Package 'ASpli'

April 4, 2019

Type Package
Title Analysis of alternative splicing using RNA-Seq
Version 1.5.4
Date 2017-10-12
Author Estefania Mancini, Javier Iserte, Marcelo Yanovsky and Ariel Chernomoretz
License GPL
biocViews GeneExpression, Transcription, AlternativeSplicing, Coverage, DifferentialExpression, DifferentialSplicing, TimeCourse, RNASeq, GenomeAnnotation, Sequencing, Alignment
Depends methods, grDevices, stats, utils, parallel, edgeR, limma, data.table
Imports GenomicRanges, GenomicFeatures, BiocGenerics, IRanges, GenomicAlignments, Gviz, S4Vectors, AnnotationDbi, Rsamtools, BiocStyle, igraph
Maintainer Estefania Mancini <emancini@leloir.org.ar></emancini@leloir.org.ar>
Description Integrative pipeline for the analysis of alternative splicing using RNAseq.
NeedsCompilation no
R topics documented:
ASpli-package
AS accessors
AsDiscover
ASpliAS-class
ASpliCounts
ASpliCounts-class
ASpliDu-class
ASpliFeatures-class 10 ASpliJDU-class 10
ASpliMergedReports-class
binDUreport

2 ASpli-package

X		48
	write-methods	47
	write	
	Subset ASpli objects	
	show-methods	
	readCounts	
	rds	
	plotGenomicRegions	
	plotBins	
	mergeReports	
	MergedReports accessors	
	mergeBinDUAS	
	loadBAM	
	junctionDUreport	
	JDU accessors	
	getConditions	
	filterDU	
	Features accesors	25
	Example data	24
	Examine ASpliDU objects	22
	DUreport.offset	20
	DUreport.norm	18
	DU accessors	17
	Counts accesors	15
	binGenome-methods	
	binGenome	13

Description

ASpli is an integrative and flexible package that facilitates the characterization of genome-wide changes in AS under different experimental conditions. ASpli analyzes the differential usage of introns, exons, and splice junctions using read counts, and estimates the magnitude of changes in AS by calculating differences in the percentage of exon inclusion or intron retention using splice junctions. This integrative approach allows the identification of changes in both annotated and novel AS events.

ASpli allows users to produce self-explanatory intermediate outputs, based on the aim of their analysis. A typical workflow involves parsing the genome annotation into new features called bins, overlapping read alignments against those bins, and inferring differential bin usage based on the number of reads aligning to the bins and junctions.

ASpli-package 3

Details

Package: ASpli
Type: Package
Version: 1.5.1
Date: 2018-02-22
License: GPL

Depends: methods, GenomicRanges, GenomicFeatures, edgeR, methods, BiocGenerics, IRanges, GenomicAlignments, Gvi

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky and Ariel Chernomoretz

References

- Acute effects of light on alternative splicing in light-grown plants. Photochemistry and Photobiology. Mancini, E, Sanchez, S, Romanowsky, A, Yanovsky, MJ. DOI: 10.1111/php.12550
- GEMIN2 attenuates the effects of temperature on alternative splicing and circadian rhythms in Arabidopsis thaliana. Proceedings of the National Academy of Sciences. Schlaen, RG, Mancini, E, Sanchez, SE, Perez-Santangelo, S, Rugnone, ML, Simpson, CG, Brown, JWS, Zhang, X, Chernomoretz, A, Yanovsky, MJ. DOI:10.1073/pnas.1504541112
- Genome wide comparative analysis of the effects of PRMT5 and PRMT4/CARM1 arginine methyltransferases on the Arabidopsis thaliana transcriptome. BMC Genomics. Hernando, E, Sanchez, S, Mancini, E, Yanovsky MJ. DOI:10.1186/s12864-015-1399-2
- A role for LSM genes in the regulation of circadian rhythms. Proceedings of the National Academy of Sciences. Perez Santangelo, S, Mancini, E, Francey, LJ, Schlaen, RG, Chernomoretz, A, Hogenesch, JB, Yanovsky MJ. DOI: 10.1073/pnas.1409791111
- The dengue virus NS5 protein intrudes in the cellular spliceosome and modulates splicing. PLOS Pathogens. De Maio, F., Risso, G., Iglesias, G., Shah, P, Pozzi, B., Gebhard, L., Mammi, L., Mancini, E., Yanovsky, M., Andino, R., Krogan, N., Srebrow, A. and Gamarnik, A. DOI:10.1371/journal.ppat.1005841

4 AS accessors

```
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",\\
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
# Calculate differential usage of genes, bins and junctions
         <- DUreport.norm( counts, targets )
# Calculate PSI / PIR for bins and junction.
         <- AsDiscover( counts, targets, features, bams, readLength = 100,
                         threshold = 5, cores = 1)
```

AS accessors

Accessors for ASpliAS object

Description

Methods to retrieve and set data in ASpliAS object. Setting data into an ASpliAS object is not a typical task and must be done with care, because it can affect the integrity of the object.

Usage

```
altPSI( x )
esPSI( x )
irPIR( x )
joint( x )
junctionsPIR( x )
junctionsPSI( x )
```

Arguments

Х

An ASpliAS object

Value

Returns dataframes with genomic metadata and PSI and PIR metrics

Author(s)

Estefania Mancini, Marcelo Yanovsky, Ariel Chernomoretz

AsDiscover 5

Examples

```
# Accessing data tables from an ASpliAS object
as <- aspliASexample()</pre>
ap <- altPSI(as)</pre>
ep <- esPSI(as)
ip <- irPIR(as)</pre>
j <- joint(as)</pre>
jpi <- junctionsPIR(as)</pre>
jps <- junctionsPSI(as)</pre>
# Setting data tables to an ASpliAS object
as2 <- new( 'ASpliAS' )
altPSI( as2 ) <- ap
esPSI(as2) <- ep
irPIR( as2 ) <- ip</pre>
joint( as2 ) <- j</pre>
junctionsPIR( as2 ) <- jpi</pre>
junctionsPSI( as2 ) <- jps</pre>
```

AsDiscover

Report PSI and PIR using experimental junctions

Description

Given a bin, it is possible to calculate PSI/PIR metric using junctions to estimate changes in the use of it along different conditions. PSI or PIR metric is calculated for each bin and experimental condition. The selection of which metric is used is based on the kind of splicing event associated with each bin.

Usage

```
AsDiscover( counts, targets, features, bam, readLength, threshold , cores )
```

Arguments

counts	An object of class ASpliCounts.
targets	A dataframe containing sample, bam and experimental factors as columns and samples as rows.
features	An object of class ASpliFeatures.
bam	A list with BAM files contents.

6 AsDiscover

readLength Read length of sequenced read. Default 100L

threshold Minimum number of reads supporting junctions. Default=5

cores Number of processing cores to use

Value

An object of class ASpliAS

irPIR reports: event, e1i counts (J1), ie1 counts (J2), j_within (J3), PIR by condition.

J1, J2, J3 sum of junctions (J1, J2, J3) by condition.

altPSI reports: event, J1 (start), J2 (end), J3 (exclusion), PSI. J1, J2, J3 sum of junctions

(J1, J2, J3) by condition.

esPSI reports: event, J1 (start), J2 (end), J3 (exclusion), PSI. J1, J2, J3 sum of junctions

(J1, J2, J3) by condition.

junctionsPIR PIR metric for each experimental junction using e1i and ie2 counts. Exclusion

junction is the junction itself. This output helps to discover new introns as well

as new retention events.

junctionsPSI Given a junction, it is possible to analyze if it shares start, end or both with

another junction. If so, is because there is more than one way for/of splicing.

Ratio between them along samples is reported.

Author(s)

Estefania Mancini, Marcelo Yanovsky and Ariel Chernomoretz

See Also

```
Accesors: irPIR, altPSI, esPSI, junctionsPIR, junctionsPSI
```

Export: writeAS

ASpliAS-class 7

ASpliAS-class

Class "ASpliAS"

Description

Results of PSI and PIR using experimental junctions

Slots

irPIR: Reports: event, e1i counts (J1), ie1 counts (J2), j_within (J3), PIR by condition. J1, J2, J3 sum of junctions (J1, J2, J3) by condition.

altPSI: Reports: event, J1 (start), J2 (end), J3 (exclusion), PSI. J1, J2, J3 sum of junctions (J1, J2, J3) by condition.

esPSI: Reports: event, J1 (start), J2 (end), J3 (exclusion), PSI. J1, J2, J3 sum of junctions (J1, J2, J3) by condition.

join: It is a combination of irPIR, altPSI and esPSI tables

junctionsPIR: PIR metric for each experimental junction using e1i and ie2 counts. Exclusion junction is the junction itself. This output helps to discover new introns as well as new retention events

junctionsPSI: Given a junction, it is possible to analyze if it shares start, end or both with another junction. If so, is because there is more than one way for/of splicing. Ratio between them along samples is reported.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

8 ASpliCounts-class

See Also

Methods: AsDiscover

Accesors: altPSI, irPIR, esPSI, joint, junctionsPIR, junctionsPSI

ASpliCounts

Class "ASpliCounts"

Description

Contains results of read overlaps against all feature levels summarization

Slots

```
gene.counts
exon.intron.counts
junction.counts
eli.counts
ie2.counts
gene.rd
bin.rd
```

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

ASpliCounts-class

Class "ASpliCounts"

Description

Contains results of read overlaps against all feature levels summarization

Slots

```
gene.counts: Object of class "data.frame"
exon.intron.counts: Object of class "data.frame"
junction.counts: Object of class "data.frame"
eli.counts: Object of class "data.frame"
ie2.counts: Object of class "data.frame"
gene.rd: Object of class "data.frame"
bin.rd: Object of class "data.frame"
```

ASpliDU-class 9

Methods

```
AsDiscover psi and pir metrics

countsb bin counts accesor

countseli eli counts accesor

countsig gene counts accesor

countsij junction counts accesor

DUreport_DEXSeq differential expression and usage estimation using DEXSeq

DUreport differential expression and usage estimation using DEXSeq

rdsb bin read densities accesor

rdsg gen read densities accesor

rds compute read densities on genes and bins

writeCounts Export count tables
```

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

ASpliDU-class

writeRds Export read density tables

Class "ASpliDU"

Description

Contains results of differential expression at gene level and differential usage at bin and junction level estimation using DEreport method.

Slots

genes

bins

junctions

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

10 ASpliJDU-class

ASpliFeatures-class Class "ASpliFeatures"

Description

Contains Genomic Ranges of different features extracted from a TxDb

Slots

```
genes:
bins:
junctions:
```

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

ASpliJDU-class Class "ASpliJDU"

Description

Contains results of differential expression at junction level.

Slots

```
localec
localej
anchorc
anchorj
jir
jes
jalt
contrast
```

Author(s)

Andres Rabinovich, Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

```
{\it ASpliMerged Reports-class} \\ {\it Class~"} {\it ASpliMerged Reports"}
```

Description

Contains results of differential expression at junction level.

Slots

binbased localebased anchorbased

Author(s)

Andres Rabinovich, Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

binDUreport

Differential gene expression and differential bin usage estimation

Description

Estimate differential expression at gene level and differential usage at bin level using diffSpliceDGE function from edgeR package. This is an alternative approach to DUreport. The results at gene level are the same as the results from DUreport. The results at bin level are slightly different.

12 binDUreport

Arguments

counts An object of class ASpliCounts

targets A dataframe containing sample, bam and experimental factor columns.

minGenReads Genes with at least an average of minGenReads reads for any condition are in-

cluded into the differential expression test. Bins from genes with at least an average of minGenReads reads for all conditions are included into the differen-

tial bin usage test. Default value is 10 reads.

minBinReads Bins with at least an average of minGenReads reads for any condition are in-

cluded into the differential bin usage test. Default value is 5 reads.

minRds Genes with at least an average of minRds read density for any condition are

included into the differential expression test. Bins from genes with at least an average of minRds read density for all conditions are included into the differential bin usage test. Bins with at least an average of minRds read density for any condition are included into the differential bin usage test. Default value is 0.05.

ignoreExternal Ignore Exon Bins at the beginning or end of the transcript. Default value is

TRUE

ignoreIo Ignore original introns. Default TRUE

ignoreI Ignore intron bins, test is performed only for exons. Default FALSE

contrast Define the comparison between conditions to be tested. contrast should be

a vector with length equal to the number of experimental conditions defined by targets. The values of this vector are the coefficients that will be used to weight each condition, the order of the values corresponds to the order given by getConditions function. When contrast is NULL, defaults to a vector containing -1, as the first value, 1 as the second an zero for all the remaining values, this corresponds to a pair comparison where the first condition is assumed to be a control and the second condition is the treatment, all other conditions are

ignored. Default = NULL

forceGLM Force the use of a generalized linear model to estimate differential expression.

It is not used to differential usage of bins. Default = FALSE

filterWithContrasted

A logical value specifying if bins, genes and junction will be filtered by read quantity and read density using data from those conditions that will be used in the comparison, i.e. those which coefficients in contrast argument are different from zero. The default value is FALSE, it is strongly recommended to do not

change this value.

verbose A logical value that indicates that detailed information about each step in the

analysis will be presented to the user.

Value

An ASpliDU object with results at genes, bins level.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

binGenome 13

See Also

```
edgeR, junctionDUreport Accessors: genesDE, binsDU Export: writeDU
```

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata','genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",</pre>
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
# Calculate differential usage of genes and bins
         <- binDUreport( counts, targets )</pre>
# Export results
writeDU( du = du, output.dir = "only_du" )
```

binGenome

Feature coordinates extraction

Description

Exons and introns are subdivided into new features called exon and intron bins and are then classified into exclusively exonic bins, exclusively intronic bins or alternative splicing (AS) bins .

```
binGenome(genome, geneSymbols, logTo = "ASpli_binFeatures.log" )
```

14 binGenome

Arguments

genome An object of class transcriptDb (TxDb)

geneSymbols A dataframe with symbol (common names) of TxDb genes. If geneSymbols is

NULL, gene name will be repeated

logTo Filename where to print features extraction log

Details

Exon and intron coordinates are extracted from gene annotation, only those from multi-exonic genes are saved for further evaluation. In case more than one isoform exist, some exons and introns will overlap. Exons and introns are then disjoint into new features called exon and intron bins, and then they are classified into exclusively exonic bins, exclusively intronic bind or alternative splicing bins (AS-bins), which are labeled according to which alternative splicing event are assumed to came from:

- ES: exon skipping
- IR: intron retention
- Alt5|3'ss: alternative five/three prime splicing site
- "*" (ES*, IR*, AltSS*) means this AS bin/region is involved simultaneously in more than one AS event type
- external: from the beginning or the end of a transcript

Subgenic features are labeled as follow (hypothetical GeneAAA):

- GeneAAA:E001: defines first exonic bin
- GeneAAA:I001: defines first intronic bin
- GeneAAA:Io001: defines first intron before disjoint into bins
- GeneAAA:J001: defines first junction

Junctions are defined as the last position of five prime exon (donor position) and first position of three prime exon (aceptor position). Using TxDb object, it is possible to extract annotated/known junctions. This information will be useful for the analysis of "experimental" junctions (reads aligned with gaps). Bins and junctions are labelled always in 5' to 3' sense. This notation is strand independent. It implies that bin / junction with lower numbering is always at 5'.

Value

An ASpliFeatures object. It is a list of features using GRanges format.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

featuresg, featuresb, featuresj

binGenome-methods 15

Examples

binGenome-methods

Feature coordinates extraction

Description

Feature coordinates extraction from a Transcript Database

Methods

```
signature(genome = "TxDb") An object of class transcriptDb (TxDb)
```

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
featuresg, featuresb , featuresj
```

Counts accesors

Accessors for ASpliCounts object

Description

Accessors for ASpliCounts object

16 Counts accesors

Usage

```
countsb(x)
countse1i(x)
countsg(x)
countsie2(x)
countsj(x)
rdsg(x)
rdsb(x)
```

Arguments

Х

An ASpliCounts object

Value

Returns dataframes with counts by sample and genomic metadata

Author(s)

Estefania Mancini, Marcelo Yanovsky, Ariel Chernomoretz

```
# Get data tables from an ASpliCounts object
counts <- aspliCountsExample()</pre>
cb1 <- countsb(counts)</pre>
ce1i <- countse1i(counts)</pre>
cg <- countsg(counts)</pre>
cie2 <- countsie2(counts)</pre>
cj <- countsj(counts)</pre>
rg <- rdsg(counts)
rb <- rdsb(counts)</pre>
# Set data tables to an ASpliCounts object
countsb(counts) <- cb1</pre>
countse1i(counts) <- ce1i</pre>
countsg(counts) <- cg</pre>
countsie2(counts) <- cie2</pre>
countsj(counts) <- cj</pre>
rdsg(counts) <- rg
rdsb(counts) <- rb</pre>
```

DU accessors 17

DU accessors

Accessors for ASpliDU object

Description

Accessors for ASpliDU object

Usage

```
genesDE( x )
binsDU( x )
junctionsDU( x )
```

Arguments

X

An ASpliDU object

Value

Returns dataframes with genomic metadata and logFC and pvalue

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

DUreport.norm

DUreport.norm

Differential gene expression and differential bin usage estimation

Description

Estimate differential expression at gene level and differential usage at bin level. When targets has only two conditions, and contrast is not set, the estimation of differential expression and usage is done with an exact test, otherwise is estimated using a generalized linear model.

Usage

Arguments

counts An object of class ASpliCounts

targets A data frame containing sample, bam and experimental factor columns.

minGenReads Genes with at least an average of minGenReads reads for any condition are in-

cluded into the differential expression test. Bins from genes with at least an average of minGenReads reads for all conditions are included into the differen-

tial bin usage test. Default value is 10 reads.

minBinReads Bins with at least an average of minGenReads reads for any condition are in-

cluded into the differential bin usage test. Default value is 5 reads.

minRds Genes with at least an average of minRds read density for any condition are

included into the differential expression test. Bins from genes with at least an average of minRds read density for all conditions are included into the differential bin usage test. Bins with at least an average of minRds read density for any condition are included into the differential bin usage test. Default value is 0.05.

ignoreExternal Ignore Exon Bins at the beginning or end of the transcript. Default value is

TRUE.

ignoreIo Ignore original introns. Default TRUE

ignoreI Ignore intron bins, test is performed only for exons. Default FALSE

DUreport.norm 19

contrast

Define the comparison between conditions to be tested. contrast should be a vector with length equal to the number of experimental conditions defined by targets. The values of this vector are the coefficients that will be used to weight each condition, the order of the values corresponds to the order given by getConditions function. When contrast is NULL, defaults to a vector containing -1, as the first value, 1 as the second an zero for all the remaining values, this corresponds to a pair comparison where the first condition is assumed to be a control and the second condition is the treatment, all other conditions are ignored. Default = NULL

forceGLM

Force the use of a generalized linear model to estimate differential expression and usage. Default = FALSE

filterWithContrasted

A logical value specifying if bins, genes and junction will be filtered by read quantity and read density using data from those conditions that will be used in the comparison, i.e. those which coefficients in contrast argument are different from zero. The default value is FALSE, it is strongly recommended to do not change this value.

verbose

A logical value that indicates that detailed information about each step in the analysis will be presented to the user.

threshold Default = 5

Value

An ASpliDU object with results at genes, bins level.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
edgeR, junctionDUreport Accessors: genesDE, binsDU Export: writeDU
```

20 DUreport.offset

DUreport.offset

Differential gene expression and differential bin usage estimation

Description

Estimate differential expression at gene level and differential usage at bin level. When targets has only two conditions, and contrast is not set, the estimation of differential expression and usage is done with an exact test, otherwise is estimated using a generalized linear model.

Usage

Arguments

counts An object of class ASpliCounts

targets A data frame containing sample, bam and experimental factor columns.

DUreport.offset 21

minGenReads Genes with at least an average of minGenReads reads for any condition are in-

cluded into the differential expression test. Bins from genes with at least an average of minGenReads reads for all conditions are included into the differential expression test.

tial bin usage test. Default value is 10 reads.

minBinReads Bins with at least an average of minGenReads reads for any condition are in-

cluded into the differential bin usage test. Default value is 5 reads.

minRds Genes with at least an average of minRds read density for any condition are

included into the differential expression test. Bins from genes with at least an average of minRds read density for all conditions are included into the differential bin usage test. Bins with at least an average of minRds read density for any condition are included into the differential bin usage test. Default value is 0.05.

ignoreExternal Ignore Exon Bins at the beginning or end of the transcript. Default value is

TRUE.

ignoreIo Ignore original introns. Default TRUE

ignoreI Ignore intron bins, test is performed only for exons. Default FALSE

offset Corrects bin expression using an offset matrix derived from gene expression

data. Default = FALSE

offsetAggregateMode

Choose the method to aggregate gene counts to create the offset matrix. When offsetAggregateMode is 'geneMode' and option offsetUseFitGeneX is TRUE, a generalized linear model is used to create the offset matrix. When offsetAggregateMode is 'geneMode' and option offsetUseFitGeneX is FALSE, the offset matrix is generated by adding a prior count to the gene count matrix. When offsetAggregateMode is 'binMode' a matrix from obtained from the sum of exonic bin counts, this only takes those bins that passes filters using minGenReads, minBinReads and minRds. Options:=c("geneMode", "binMode") [1],

offsetUseFitGeneX

Default= TRUE

contrast Define the comparison between conditions to be tested. contrast should be

a vector with length equal to the number of experimental conditions defined by targets. The values of this vector are the coefficients that will be used to weight each condition, the order of the values corresponds to the order given by getConditions function. When contrast is NULL, defaults to a vector containing -1, as the first value, 1 as the second an zero for all the remaining values, this corresponds to a pair comparison where the first condition is assumed to be a control and the second condition is the treatment, all other conditions are

ignored. Default = NULL

Force the use of a generalized linear model to estimate differential expression

and usage. Default = FALSE

filterWithContrasted

A logical value specifying if bins, genes and junction will be filtered by read quantity and read density using data from those conditions that will be used in the comparison, i.e. those which coefficients in contrast argument are different from zero. The default value is FALSE, it is strongly recommended to do not

change this value.

verbose A logical value that indicates that detailed information about each step in the

analysis will be presented to the user.

Value

An ASpliDU object with results at genes, bins level.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
edgeR, junctionDUreport Accessors: genesDE, binsDU Export: writeDU
```

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata', 'genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( \ "A\_C\_0.bam", \ "A\_C\_1.bam", \ "A\_C\_2.bam", \\
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
# Calculate differential usage of genes and bins
         <- DUreport.offset( counts, targets )
# Export results
writeDU( du = du, output.dir = "only_du" )
```

Examine ASpliDU objects

Examine ASpliDU objects

Description

AspliDU object may contain results of differential expression of genes, differential usage of bins and junctions, however not everything is calculated at the same or even present. Calculations for genes and bins can be done independently from junctions. Functions containsJunctions and containsGenesAndBins allow to interrogate an ASpliDU object about the kind of results it contain.

Usage

```
containsJunctions( du )
containsGenesAndBins( du )
```

Arguments

du

An ASpliDU object.

Value

A logical value that indicates that results for genes and bins, or results for junctions are available in the object.

Author(s)

Estefania Mancini, Marcelo Yanovsky, Ariel Chernomoretz

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata', 'genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",</pre>
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C','C','C','D','D','D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
```

24 Example data

```
# Create an ASpliDU object.
#
du <- DUreport.norm( counts, targets )
# Verify if du contains results for genes and bins.
containsGenesAndBins( du )
# Verify if du contains results for genes and bins.
containsJunctions( du )</pre>
```

Example data

Example Aspli objects

Description

ASpli includes functions to easily build ASpli objects, used in examples in the vignette and man pages.

Usage

```
aspliASexample()
aspliBamsExample()
aspliCountsExample()
aspliDUexample1()
aspliDUexample2()
aspliExampleBamList()
aspliExampleGTF()
aspliFeaturesExample()
aspliJunctionDUexample()
aspliTargetsExample()
```

Value

An ASpli object with example data.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

```
as <- aspliASexample()
bams <- aspliBamsExample()
counts <- aspliCountsExample()
du1 <- aspliDUexample1()
du2 <- aspliDUexample2()
bamfiles <- aspliExampleBamList()</pre>
```

Features accesors 25

```
gtffile <- aspliExampleGTF()
features <- aspliFeaturesExample()
jdu <- aspliJunctionDUexample()
targets <-aspliTargetsExample()</pre>
```

Features accesors

Accessors for ASpliFeatures object

Description

Accessors for ASpliFeatures object

Usage

```
featuresg( x )
featuresb( x )
featuresj( x )
```

Arguments

Х

An ASpliFeatures object

Value

Returns a GenomicRanges object. Function featuresg returns a GRangesList object containing exon ranges for each gene. Functions featuresb and featuresj, returns GRanges object for all bins and junctions.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

```
# Get data from an ASpliFeatures object
features <- aspliFeaturesExample()

fg <- featuresg( features )
fb <- featuresb( features )
fj <- featuresj( features )

# Set data to an ASpliFeatures object
featuresg( features ) <- fg
featuresb( features ) <- fb
featuresj( features ) <- fj</pre>
```

26 filterDU

$c:\tau$	 rDU

Filtering ASpliDU objects

Description

ASpliDU object can be filtered to retain genes, bins or junction according to their fdr corrected p-value estimated and log-fold-change.

Usage

```
filterDU(
  du ,
  what = c( 'genes','bins','junctions'),
  fdr = 1,
  logFC = 0,
  absLogFC = TRUE,
  logFCgreater = TRUE
)
```

Arguments

An ASpliDU object
A character vector that specifies the kind of features that will be filtered. Accepted values are 'genes', 'bins', 'junctions'. Multiple values can be passed at the same time. The default value is c('genes', 'bins', 'junctions')
A double value representing the maximum accepted value of fdr corrected p-value to pass the filter. The default value is 1, the neutral value for fdr filtering operation.
A double value representing the cut-off for accepted values of log-fold-change to pass the filter. The default value is 0, the neutral value for logFC filtering operation if logFCgreater and absLocFC arguments are both TRUE.
A logical value that specifies that the absolute value of log-fold-change will be used in the filter operation. The default value is TRUE.
A logical value that specifies that the log-fold-change value (or abs(log-fold-change) if absLogFC argument is TRUE) of features must be greater than the cut-off value to pass the filter. The default value is TRUE.

Value

A new ASpliDU object with the results of the filtering operations. The elements of features that were not specified to be filtered are kept from the input ASpliDU object.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

getConditions 27

See Also

DUreport.norm, DUreport.offset, junctionDUreport, binDUreport,

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata','genes.mini.gtf',</pre>
                                 package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",</pre>
                     "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
              row.names = paste0('Sample_',c(1:6)),
              bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
              factor1 = c('C','C','C','D','D','D'))
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
# Calculate differential usage of junctions only
          <- DUreport.norm( counts, targets )
# Filter by FDR
duFiltered1 <- filterDU( du, what=c('genes', 'bins'),</pre>
    fdr = 0.01)
# Filter by logFC, only those that were up-regulated
duFiltered2 <- filterDU( du, what=c('genes', 'bins'),</pre>
    logFC = log(1.5, 2), absLogFC = FALSE)
```

getConditions

Retrieve condition names from a targets data frame.

Description

Targets data frame contains experimental factors values for each sample. This function generates a simple name for each unique condition resulting from the combination of all experimental factors.

28 getConditions

The order of the conditions given by getConditions is the same that in contrast argument of DUreport.norm function.

Usage

```
getConditions( targets )
```

Arguments

targets

A dataframe containing sample, bam and experimental factors columns

Value

A character vector with the names of the conditions derived from experimental factors.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
DUreport.norm, DUreport.offset
```

```
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",\\
                                                           "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
                                       row.names = paste0('Sample_',c(1:6)),
                                       bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
                                       factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files.
# Return value is c('C', 'D') in this example.
conditions <- getConditions(targets)</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam", "A_C_3.bam", "A_C_3.bam, "A_C_3.bam
                                                           "A_D_0.bam", "A_D_1.bam", "A_D_2.bam", "A_D_3.bam" )
targets <- data.frame(</pre>
                                       row.names = paste0('Sample_',c(1:8)),
                                       bam = file.path( 'extdata', bamFileNames, package="ASpli" ),
                                       factor1 = c( 'C', 'C', 'C', 'C', 'D', 'D', 'D', 'D'),
                                        factor2 = c( 'E', 'E', 'F', 'F', 'E', 'E', 'F', 'F') )
# Load reads from bam files.
# Return value is c("C_E", "C_F", "D_E", "D_F") in this example.
conditions <- getConditions(targets)</pre>
```

JDU accessors 29

JDU accessors

Accessors for ASpliJDU object

Description

Accessors for ASpliJDU object

Usage

```
anchorc( x )
anchorj( x )
localec( x )
localej( x )
jir( x )
jes( x )
jalt( x )
```

Arguments

Χ

An ASpliJDU object

Value

Returns dataframes

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

junctionDUreport

Differential junction usage estimation

Description

Estimate differential usage at junction level. COMPLETAR

30 loadBAM

Arguments

asd An object of class ASpliAS

targets A dataframe containing sample, bam and experimental factor columns.

minAvgCounts Minimum average counts for filtering.

contrast
filterWithContrasted

Apply filter on all conditions or only on contrasts

runUniformityTest

Run uniformity test on Intron Retention

mergedBams Path to merged bams for each testing condition

maxPValForUniformityCheck

To speed up uniformity test only check junctions with pval < maxPValForUni-

formityCheck

Value

An ASpliJDU object with results of differential usage of junctions

Author(s)

Andres Rabinovich, Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
edgeR, ASpliAS Accessors: junctionsDU Export: writeDU
```

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.

#as <- new("ASpliAS")
#targets <- 1
#a <- junctionDUreportExt(as, targets)</pre>
```

loadBAM

Load BAM files

Description

Load BAM files into R session using a targets specification.

```
loadBAM(targets, cores)
```

mergeBinDUAS 31

Arguments

targets A data frame containing sample, bam and experimental factors columns

cores Number of processors to use

Value

A list of GAlignments. Each element of the list correspond to a samples BAM file.

Author(s)

Estefania Mancini, Marcelo Yanovsky, Ariel Chernomoretz

Examples

mergeBinDUAS

Differential usage of bins and PSI/PIR.

Description

This function merges the results of differential usage of bins, from an ASpliDU object, with PSI/PIR and junction information, from an ASpliAS object. Also, a delta PSI/PIR value is calculated from a contrast.

32 mergeBinDUAS

Arguments

du An object of class ASpliDU as An object of class ASpliAS

targets A data frame containing sample, bam files and experimental factor columns.

contrast Define the

Define the comparison between conditions to be tested. contrast should be a vector with length equal to the number of experimental conditions defined by targets. The values of this vector are the coefficients that will be used to weight each condition, the order of the values corresponds to the order given by getConditions function. When contrast is NULL, defaults to a vector containing -1, as the first value, 1 as the second an zero for all the remaining values, this corresponds to a pair comparison where the first condition is assumed to be a control and the second condition is the treatment, all other conditions are ignored. The default value is NULL.

Value

A data frame containing feature, event, locus, locus_overlap, symbol, gene coordinates, start of bin, end of bin, bin length, log-Fold-Change value, p-value, fdr corrected p-value, J1 inclusion junction, J1 junction counts for each sample, J2 inclusion junction, J2 junction counts for each sample, J3 exclusion junction, J3 junction counts for each sample, PSI or PIR value for each bin, and delta PSI/PIR.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
ASpliDU, ASpliAS
```

MergedReports accessors

Accessors for ASpliMergedReport object

Description

Accessors for ASpliMergedReport object

Usage

```
binbased( x )
localebased( x )
anchorbased( x )
```

Arguments

x An ASpliMergedReport object

Value

Returns dataframes

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

34 mergeReports

mergeReports

Differential junction usage estimation. COMPLETAR

Description

Estimate differential usage at junction level. COMPLETAR

Usage

Arguments

bdu An object of class ASpliDU jdu An object of class ASpliDU

Value

An ASpliMergedReports object with results of differential usage of junctions

Author(s)

Andres Rabinovich, Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

 $\verb|junctionDUreport, binDUreport, ASpliMergedReports Accessors: \verb|junctionsDUExport| writeDU| \\$

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.

#as <- new("ASpliAS")
#targets <- 1
#a <- junctionDUreportExt(as, targets)</pre>
```

plotBins 35

plotBins

Draw plots of gene counts, bin counts, PSI/PIR value, inclusion and exclusion junctions for selected bins.

Description

Creates a plot with gene counts, bin counts, PSI/PIR value, inclusion and exclusion junctions for selected bins and conditions.

Usage

```
plotBins( counts,
          as,
          bin,
          factorsAndValues,
          targets,
          main
                          = NULL,
          colors
                          = c( '#2F7955', '#79552F', '#465579',
                                '#A04935', '#752020', '#A07C35'),
          panelTitleColors = '#000000',
          panelTitleCex = 1,
                          = c(2.1, 3.1, 1.1, 1.1),
          innerMargins
          outerMargins
                          = c(0, 0, 2.4, 0),
          useBarplots
                          = NULL,
          barWidth
                          = 0.9,
          barSpacer
                          = 0.4,
          las.x
                          = 2,
          useHCColors
                          = FALSE,
          legendAtSide
                          = TRUE,
          outfolder
                          = NULL,
                          = c( 'png', 'bmp', 'jpeg', 'tiff', 'pdf')[1],
          outfileType
          deviceOpt
                          = NULL )
```

Arguments

An object of class ASpliCounts counts as An object of class ASpliAS

A character vector with the names of the bins to be plotted. bin

factorsAndValues

A list containing the factor and the values for each factor to be plotted. The order of the factors will modify how the conditions are grouped in the plot. factorsAndValues must be a named list, where the name of each element is a factor and the list element itself is a character vector of the values of this factor in the order to be plotted. See examples for more details.

targets A data frame containing sample, bam files and experimental factor columns

Main title of the plot. If NULL the bin name is used as title. main

36 plotBins

colors	A vector of character colors for lines and bar plots.
panelTitleColo	A vector of character colors for the titles of each plot panel.
panelTitleCex	Character size expansion for panel titles.
innerMargins	A numerical vector of the form c(bottom, left, top, right) which gives the size of each plot panel margins. Defaults to $c(2.1, 3.1, 1.1, 1.1)$
outerMargins	A numerical vector of the form c(bottom, left, top, right) which gives the size of margins. Defaults to $c(0,0,2.4,0)$
useBarplots	A logical value that indicates the type of plot to be used. If TRUE bar plots are used, if FALSE lines are used. If NULL the type is bar plot if there just two conditions and lines if there are more than two conditions.
barWidth	The width of the bars in bar plots. barWidth must be in $(0,1]$ range. Default value is 0.9 .
barSpacer	Fraction of barwidth used as spacer between bar plot groups. Defaule value is 0.4 .
las.x	Text orientation of x-axis labels.
useHCColors	A logical value. If TRUE panelTitleColors are not used, instead panel title are automatically chosen to have high contrast against colors.
legendAtSide	A logical value that forces panel title to be shown on the y-axis, instead of over the plot.
outfolder	Path to output folder to write plot images. Is NULL, plot are rendered on the default device
outfileType	File format of the output files used if outfolder is not NULL. Accepted values are 'png', 'jpeg', 'tiff', 'pdf'. Each value selects the graphic device of the same name. The name of the image file is the name of bin with the corresponding extension given by the chosen type

A list of named options to be passed to the graphic device selected in outfileType

Author(s)

deviceOpt

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
plotGenomicRegions,
```

plotBins 37

```
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",</pre>
                  "A_D_0.bam", "A_D_1.bam", "A_D_2.bam",
                  "B_C_0.bam", "B_C_1.bam", "B_C_2.bam",
                  "B_D_0.bam", "B_D_1.bam", "B_D_2.bam")
targets <- data.frame(</pre>
            row.names = paste0('Sample_',c(1:12)),
            bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
            # Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                       maxISize = 50000)
# Calculate differential usage of genes, bins and junctions
        <- DUreport.norm( counts, targets , contrast = c(1,-1,-1,1))
# Calculate PSI / PIR for bins and junction.
        <- AsDiscover( counts, targets, features, bams, readLength = 100,
                       threshold = 5, cores = 1)
# Plot bin data. Factor2 is the main factor for graphic representation in
# this example as it is the first in factorsAndValues argument.
# This makes a bar plot comparing four conditions, grouped by factor1.
plotBins( counts, as, 'GENE03:E002',
  factorsAndValues = list(
   factor2 = c('C', 'D'),
   factor1 = c('A', 'B')),
  las.x = 1,
  legendAtSide = TRUE,
 useHCColors = TRUE,
  targets = targets,
  barWidth = 0.95,
  innerMargins = c(2.1, 4.1, 1.1, 1.1)
# Redefine targets
targets <- data.frame(</pre>
            row.names = paste0('Sample_',c(1:12)),
            bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
            factor1 = c( 'A', 'A', 'B', 'B', 'C', 'C', 'D', 'D', 'E', 'E', 'F', 'F') )
        <- AsDiscover( counts, targets, features, bams, readLength = 100,
as
                       threshold = 5, cores = 1)
# This makes a line plot for six conditions, grouped by factor1.
```

38 plotGenomicRegions

```
plotBins( counts, as, 'GENE03:E002',
  factorsAndValues = list(
    factor1 = c('A','B','C','D','E','F') ),
  las.x = 1,
  legendAtSide = FALSE,
  targets = targets,
  innerMargins = c( 2.1, 4.1, 1.1, 1.1 ) )
```

plotGenomicRegions

Create genomic regions coverage plots

Description

Graphic representation of coverage and junctions is useful to complement the results of differential usage of bins and junction and differential expression analysis.

Function plotGenomicRegions allow to create plots for multiple conditions for bins and genes. Each individual plot can only correspond to a single gene or bin, but can contain many panels for different experimental conditions.

Usage

```
plotGenomicRegions( features, x, genomeTxDb, targets, xIsBin = TRUE,
  layout = 'auto', colors = 'auto', plotTitles = 'auto', sashimi = FALSE,
  zoomOnBins = FALSE, deviceOpt = NULL, highLightBin = TRUE, outfolder = NULL,
  outfileType = 'png', mainFontSize = 24, annotationHeight = 0.2,
  annotationCol = 'black', annotationFill = 'gray', annotationColTitle = 'black',
  preMergedBAMs = NULL, useTransparency = FALSE, tempFolder = 'tmp',
  avoidReMergeBams= FALSE, verbose = TRUE )
```

Arguments

features	An ASpliFeatures object, generated with binGenoms function.
x	A character vector with the names of bins or genes to plot. To plot into a window is recommended that the length of x be one.
genomeTxDb	A TxDb object with the annotation of reference genome.
targets	A data frame containing sample, bam files and experimental factor columns
xIsBin	A logical value that indicates if values in x corresponds to gene names or bin names.
layout	A character with value 'auto' or a character matrix with condition names arranged with the desired layout of the panels in plots. The dimensions of layout matrix, colors matrix and plotTitles matrices must be the same. Matrix can have NA values, however, the height of the panels corresponding to that column are modified to occupy the total height. The default value is 'auto'.

plotGenomicRegions 39

colors A character containing value 'auto' or containing colors strings, or a character

matrix with color names arranged with the layout specified in layout argument. The dimensions of layout matrix, colors matrix and plotTitles matrices must be

the same. The default value is 'auto'.

plotTitles A character containing value 'auto', or a character matrix with titles for each

panel arranged with the layout specified in layout argument. The dimensions of layout matrix, colors matrix and plotTitles matrices must be the same. The

default value is 'auto'.

sashimi A logical value that specifies that a sashimi plot for junctions must be included

into each panel. The default value is FALSE.

zoomOnBins A FALSE logical value or a double value between 0 and 1. If value is FALSE

then the genomic range to be plotted correspond to the complete gene, otherwise the genomic range is that the size of the bin corresponds to a fraction equals to zoomOnBins value of the total and is centered in the bin. Is used only when

xIsBin is TRUE. The default value is FALSE.

deviceOpt A named list of arguments to be passed to the graphic device used to plot. This

allow to further customization of the plot. The default value is an empty list.

highLightBin A logical value that indicates if the bin should be highlighted. The default value

is TRUE.

outfolder NULL or a character vector representing a folder path that will be used to save

the plot images. If the folder doesn't exists it is created. If NULL, the plot is made in a window. The default value is NULL. For each bin, a single image file is generated. The name of the file is the name of the bin, added with a '.gr.' string, and the file extension at the end. If the name of the bin contains invalid character for a file name, those will be replaced by an underscore character.

outfileType A character value the specifies the file format of the plot to be created. Is used

only when outfolder is not NULL. Accepted values are 'png', 'jpeg', 'bmp', 'tiff', 'pdf'. Each value is the graphic device used to create the image. The

default value is 'png'.

mainFontSize A numeric value specifying the size of the main title. The default value is 24.

annotationHeight

A double value specifying the proportion of the total height used to represent

the gene model. The default value is 0.2.

annotationCol A character value that specifies the color of the borders of bars in gene model

representation. The default value is 'black'.

annotationFill A character value that specifies the color of the filling of bars in gene model

representation. The default value is 'gray'.

 $annotation {\tt ColTitle}$

A character value that specifies the color of text in gene model representation.

The default value is 'black'.

preMergedBAMs A one column data frame that associates a condition, specified in the row name,

with a character value representing the path of bam file with the reads for that condition. This bam file is typically generated by merging the bam files of all replicates for that condition. The default value is NULL, this specifies that not merged bam files are used, instead on-the-fly read extraction and merging is

done from the bam files specified in the targets argument.

plotGenomicRegions

useTransparency

A logical value that specifies if transparency will be used to generate the plots, this leads to better looking plot, however not all graphic devices support transparency. The default value is FALSE.

tempFolder

40

A character value specifying the path to store intermediate files produced while extracting and merging reads from bam files. It is only used when preMerged-BAMs arguments is NULL. The files created are not automatically removed after plotting because can be reused to create a new plot of the same genes or bins with different graphic options. The default value is 'tmp', that means a new 'tmp' folder in the current working folder.

avoidReMergeBams

A logical value specifying that extraction and merging of bam files will be avoided. This is only meaningful when the extraction and merging of the same set of genes and bins was done in the previous execution of plotGenomicRegions function. The default value is FALSE.

verbose

A logical value specifying that detailed information about the execution will be informed to the user. The default value is TRUE.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
Devices, pdf, png, bmp, jpeg, tiff
```

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata','genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D'),
             stringsAsFactors = FALSE )
# Plot a single bin to a window
plotGenomicRegions(
  features,
  'GENE01:E002',
```

rds 41

```
genomeTxDb,
  targets,
  sashimi = FALSE,
  colors = '#AA4444',
  annotationHeight = 0.1,
  tempFolder = 'tmp',
  verbose = TRUE ,
  avoidReMergeBams = FALSE,
 useTransparency = FALSE )
# plot two bins to pdf files.
plotGenomicRegions(
  features, c( 'GENE01:E002', 'GENE02:E002'),
  genomeTxDb,
  targets,
  layout = matrix( c( 'C', 'D'), ncol = 1),
  colors = matrix( c( '#663243', '#363273'), ncol = 1),
 plotTitles = matrix( c( 'C condition', 'D condition'), ncol = 1),
  sashimi = FALSE,
  mainFontSize = 12,
  annotationHeight = 0.1,
  tempFolder = 'tmp',
  verbose = TRUE ,
  avoidReMergeBams = FALSE,
 useTransparency = TRUE,
  outfolder = '.',
  outfileType = 'pdf',
  deviceOpt = list( height = 6, width = 5, paper = 'a4r' ) )
```

rds

Read density of gene and bins

Description

Read density of gene and bins is the quotient between the number of read counts and the length of the feature. The results are appended into an ASpliCounts object that must be given as argument. The explicit calculation of read densities is usually not required because is automatically performed by readCounts function.

Usage

```
rds( counts, targets )
```

Arguments

counts An ASpliCounts object

targets A data frame containing sample, bam and experimental factors columns

42 readCounts

Value

An ASpliCounts object containing read densities of genes and bins.

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata','genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",</pre>
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
# Calculates read densities
counts <- rds( counts, targets )</pre>
```

 ${\sf readCounts}$

Summarize read overlaps

Description

Summarize read overlaps against all feature levels

Usage

readCounts 43

Arguments

features	An object of class ASpliFeatures. It is a list of GRanges at gene, bin and junction level
bam	List of GAlignments objects corresponding to bam files of samples.
targets	A dataframe containing sample, bam and experimental factors columns
readLength	Read length of sequenced library. It is used for compute E1I and IE2 read summarization
maxISize	maximum intron expected size. Junctions longer than this size will be dicarded
cores	Number of cores to use. Default is 1
minAnchor	Minimum percentage of read that should be aligned to an exon-intron boundary

Value

An object of class ASpliCounts. Each slot is a dataframe containing features metadata and read counts. Summarization is reported at gene, bin, junction and intron flanking regions (E1I, IE2).

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

Accesors: countsg, countsb, countsj, countse1i, countsie2, rdsg, rdsb, Export: writeCounts

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata', 'genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam", \\
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
```

44 Subset ASpli objects

```
# Export data
writeCounts( counts, output.dir = "only_counts" )
```

show-methods

Display a summary of data contained in ASpliObjects

Description

Display a summary of data contained in ASpliObjects

Details

Display a summary of data contained in ASpliObjects

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

```
Subset ASpli objects
                         Subset ASpli objects
```

Description

ASpli provides utility functions to easy subset ASpliCounts objects, ASpliAS objects, targets data frame and lists GAlignments generated with loadBAM function. The subset can be done selecting some of the experimental conditions or samples names (but not both).

Usage

```
subset( x, ... )
subsetBams( x, targets, select )
subsetTargets( targets, select, removeRedundantExpFactors )
```

Arguments

An ASpliCount or ASpliAS object for subset function, or list of GAlignments Х

for subsetrBams function.

targets A dataframe containing sample, bam and experimental factor columns.

select A character vector specifying the conditions or samples to be kept after subset

operation. It's assumed that condition names are different from sample names.

removeRedundantExpFactors

When sub-setting the targets data frame, one or more experimental factors can have only one value. If this argument is TRUE those experimental factors are absent in the resulting target data frame.

Subsetting ASpliCounts and ASpliAS objects sub subset method requires a targets argument and a select argument with the same specifications that the arguments with the same name in subsetBams and subsetTargets functions

Subset ASpli objects 45

Value

A data frame similar to x (or targets for subsetTargets) with only the containing only the selected elements

Author(s)

Estefania Mancini, Marcelo Yanovsky, Ariel Chernomoretz

Examples

```
# Create a transcript DB from gff/gtf annotation file.
# Warnings in this examples can be ignored.
library(GenomicFeatures)
genomeTxDb <- makeTxDbFromGFF( system.file('extdata', 'genes.mini.gtf',</pre>
                                package="ASpli") )
# Create an ASpliFeatures object from TxDb
features <- binGenome( genomeTxDb )</pre>
# Define bam files, sample names and experimental factors for targets.
bamFileNames <- c( "A_C_0.bam", "A_C_1.bam", "A_C_2.bam",</pre>
                    "A_D_0.bam", "A_D_1.bam", "A_D_2.bam")
targets <- data.frame(</pre>
             row.names = paste0('Sample_',c(1:6)),
             bam = system.file( 'extdata', bamFileNames, package="ASpli" ),
             factor1 = c( 'C', 'C', 'C', 'D', 'D', 'D') )
# Load reads from bam files
bams <- loadBAM( targets )</pre>
# Read counts from bam files
counts <- readCounts( features, bams, targets, cores = 1, readLength = 100,</pre>
                         maxISize = 50000)
# Create ASpliAS object
         <- AsDiscover( counts, targets, features, bams, readLength = 100,
                         threshold = 5, cores = 1)
# Define selection
select <- c('Sample_1', 'Sample_2', 'Sample_4', 'Sample_5')</pre>
# Subset target
targets2 <- subsetTargets( targets, select )</pre>
# Subset bams
bams2 <- subsetBams( bams, targets, select )</pre>
# Subset ASpliCounts object
counts2 <- subset( counts, targets, select )</pre>
# Subset ASpliAS object
```

46 write

```
as2 <- subset( as, targets, select )
```

write Write results

Description

Export tab delimited files in structured output

Usage

```
writeCounts(counts, output.dir="counts")
writeRds(counts, output.dir="rds")
writeDU(du, output.dir="du")
writeAS(as, output.dir="as")
writeAll(counts, du, as, output.dir="output")
```

Arguments

counts An ASpliCounts object

as An ASpliAS object

du An ASpliDU object

output.dir Name of output folder (new or existing)

Value

Tab delimited files are exported in a tidy manner into output folder

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

```
AsDiscover, binGenome, DUreport.norm, DUreport.offset
```

write-methods 47

write-methods Write results

Description

Export tab delimited files in structured output

Details

Tab delimited files are exported in a tidy manner into output folder

Author(s)

Estefania Mancini, Javier Iserte, Marcelo Yanovsky, Ariel Chernomoretz

See Also

AsDiscover, binGenome, DUreport.norm, DUreport.offset

Index

*Topic alternative splicing, RNA-seq,	aspliBamsExample (Example data), 24
junctions	ASpliCounts, 8
ASpli-package, 2	ASpliCounts-class, 8
	aspliCountsExample (Example data), 24
altPSI, 6, 8	ASpliDU, 32
altPSI (AS accessors), 4	ASpliDU (ASpliDU-class), 9
altPSI, ASpliAS-method (ASpliAS-class), 7	ASpliDU-class, 9
altPSI<- (AS accessors), 4	aspliDUexample1 (Example data), 24
altPSI<-,ASpliAS,data.frame-method	aspliDUexample2 (Example data), 24
(ASpliAS-class), 7	aspliExampleBamList(Example data), 24
anchorbased (MergedReports accessors),	aspliExampleGTF (Example data), 24
33	ASpliFeatures-class, 10
anchorbased,ASpliMergedReports-method	aspliFeaturesExample (Example data), 24
(ASpliMergedReports-class), 11	ASpliJDU (ASpliJDU-class), 10
anchorbased<- (MergedReports	ASpliJDU-class, 10
accessors), 33	aspliJunctionDUexample (Example data),
$anchorbased \verb <, ASpliMergedReports , data.frame$	-method 24
(ASpliMergedReports-class), 11	ASpliMergedReports, <i>34</i>
anchorc (JDU accessors), 29	ASpliMergedReports
anchorc,ASpliJDU-method	(ASpliMergedReports-class), 11
(ASpliJDU-class), 10	ASpliMergedReports-class, 11
anchorc<- (JDU accessors), 29	aspliTargetsExample (Example data), 24
anchorc<-,ASpliJDU,data.frame-method	
(ASpliJDU-class), 10	binbased (MergedReports accessors), 33
anchorj (JDU accessors), 29	binbased,ASpliMergedReports-method
anchorj,ASpliJDU-method	(ASpliMergedReports-class), 11
(ASpliJDU-class), 10	binbased<- (MergedReports accessors), 33
anchorj<-(JDU accessors), 29	<pre>binbased<-,ASpliMergedReports,data.frame-method</pre>
anchorj<-,ASpliJDU,data.frame-method	(ASpliMergedReports-class), 11
(ASpliJDU-class), 10	binDUreport, 11, 27, 34
AS accessors, 4	binDUreport,ASpliCounts-method
AsDiscover, 5, 8, 46, 47	(ASpliCounts-class), 8
AsDiscover, ASpliCounts-method	binGenome, 13, 46, 47
(ASpliCounts-class), 8	binGenome,TxDb-method
ASpli (ASpli-package), 2	(binGenome-methods), 15
ASpli-package, 2	binGenome-methods, 15
ASpliAS, 30, 32	binsDU, 13, 19, 22
ASpliAS (ASpliAS-class), 7	binsDU (DU accessors), 17
ASpliAS-class, 7	binsDU, ASpliDU-method (ASpliDU-class), 9
aspliASexample (Example data), 24	binsDU<- (DU accessors), 17

INDEX 49

11. 80 1180	5
binsDU<-,ASpliDU-method	Devices, 40
(ASpliDU-class), 9	DU accessors, 17
bmp, 40	DUreport.norm, 18, 27, 28, 46, 47
	DUreport.norm, ASpliCounts-method
containsGenesAndBins (Examine ASpliDU	(ASpliCounts-class), 8
objects), 22	DUreport.offset, 20, 27, 28, 46, 47
containsGenesAndBins,ASpliDU-method	DUreport.offset,ASpliCounts-method
(ASpliDU-class), 9	(ASpliCounts-class), 8
containsJunctions (Examine ASpliDU	
objects), 22	edgeR, 13, 19, 22, 30
containsJunctions, ASpliDU-method	esPSI, 6, 8
(ASpliDU-class), 9	esPSI (AS accessors), 4
Counts accesors, 15	esPSI,ASpliAS-method(ASpliAS-class),7
countsb, 43	esPSI<- (AS accessors), 4
countsb (Counts accesors), 15	esPSI<-,ASpliAS,data.frame-method
	(ASpliAS-class), 7
countsb, ASpliCounts-method	Examine ASpliDU objects, 22
(ASpliCounts-class), 8	Example data, 24
countsb<- (Counts accesors), 15	,
countsb<-,ASpliCounts,data.frame-method	Features accesors, 25
(ASpliCounts-class), 8	featuresb, <i>14</i> , <i>15</i>
countseli, 43	featuresb (Features accesors), 25
countseli (Counts accesors), 15	featuresb, ASpliFeatures-method
countse1i,ASpliCounts-method	(ASpliFeatures-class), 10
(ASpliCounts-class), 8	featuresb<- (Features accesors), 25
countseli<- (Counts accesors), 15	featuresb<-,ASpliFeatures,GRanges-method
<pre>countse1i<-,ASpliCounts,data.frame-method</pre>	(ASpliFeatures-class), 10
(ASpliCounts-class), 8	featuresg, 14, 15
countsg, 43	
countsg (Counts accesors), 15	featuresg (Features accesors), 25
countsg, ASpliCounts-method	featuresg, ASpliFeatures -method
(ASpliCounts-class), 8	(ASpliFeatures-class), 10
countsg<- (Counts accesors), 15	featuresg<- (Features accesors), 25
countsg<-,ASpliCounts,data.frame-method	featuresg<-,ASpliFeatures,GRangesList-method
(ASpliCounts-class), 8	(ASpliFeatures-class), 10
	featuresj, <i>14</i> , <i>15</i>
countsie2,43	featuresj (Features accesors), 25
countsie2 (Counts accesors), 15	featuresj, ASpliFeatures-method
countsie2, ASpliCounts-method	(ASpliFeatures-class), 10
(ASpliCounts-class), 8	featuresj<- (Features accesors), 25
countsie2<- (Counts accesors), 15	featuresj<-,ASpliFeatures,GRanges-method
<pre>countsie2<-,ASpliCounts,data.frame-method</pre>	(ASpliFeatures-class), 10
(ASpliCounts-class), 8	filterDU, 26
countsj, 43	filterDU,ASpliDU-method
countsj (Counts accesors), 15	(ASpliDU-class), 9
countsj,ASpliCounts-method	
(ASpliCounts-class), 8	genesDE, 13, 19, 22
countsj<- (Counts accesors), 15	genesDE (DU accessors), 17
countsj<-,ASpliCounts,data.frame-method	<pre>genesDE, ASpliDU-method (ASpliDU-class),</pre>
(ASpliCounts-class), 8	9

50 INDEX

<pre>genesDE<- (DU accessors), 17 genesDE<-,ASpliDU,data.frame-method</pre>	<pre>junctionsPIR (AS accessors), 4 junctionsPIR, ASpliAS-method</pre>
(ASpliDU-class), 9	(ASpliAS-class), 7
getConditions, 27	junctionsPIR<- (AS accessors), 4
getconditions, 27	junctionsPIR<-,ASpliAS,data.frame-method
irPIR, 6, 8	(ASpliAS-class), 7
irPIR (AS accessors), 4	junctionsPSI, 6 , 8
irPIR, ASpliAS-method (ASpliAS-class), 7	junctionsPSI (AS accessors), 4
irPIR<- (AS accessors), 4	junctionsPSI, ASpliAS-method
<pre>irPIR<-,ASpliAS,data.frame-method</pre>	(ASpliAS-class), 7
(ASpliAS-class), 7	junctionsPSI<- (AS accessors), 4
(**************************************	junctionsPSI<-,ASpliAS,data.frame-method
jalt (JDU accessors), 29	(ASpliAS-class), 7
<pre>jalt, ASpliJDU-method (ASpliJDU-class),</pre>	(ASPITAS-CIASS), /
10	1 IDAM 20
jalt<- (JDU accessors), 29	loadBAM, 30
<pre>jalt<-,ASpliJDU,data.frame-method</pre>	localebased (MergedReports accessors),
(ASpliJDU-class), 10	33
JDU accessors, 29	localebased, ASpliMergedReports-method
jes (JDU accessors), 29	(ASpliMergedReports-class), 11
jes, ASpliJDU-method (ASpliJDU-class), 10	localebased<- (MergedReports
jes<- (JDU accessors), 29	accessors), 33
jes<-,ASpliJDU,data.frame-method	localebased<-,ASpliMergedReports,data.frame-method
(ASpliJDU-class), 10	(ASpliMergedReports-class), 11
jir (JDU accessors), 29	localec(JDU accessors), 29
jir, ASpliJDU-method (ASpliJDU-class), 10	localec,ASpliJDU-method
jir<- (JDU accessors), 29	(ASpliJDU-class), 10
jir<-,ASpliJDU,data.frame-method	localec<- (JDU accessors), 29
(ASpliJDU-class), 10	localec<-,ASpliJDU,data.frame-method
joint, 8	(ASpliJDU-class), 10
joint (AS accessors), 4	localej(JDU accessors), 29
<pre>joint,ASpliAS-method (ASpliAS-class), 7</pre>	localej,ASpliJDU-method
joint<- (AS accessors), 4	(ASpliJDU-class), 10
<pre>joint<-,ASpliAS,data.frame-method</pre>	localej<-(JDU accessors), 29
(ASpliAS-class), 7	<pre>localej<-,ASpliJDU,data.frame-method</pre>
jpeg, 40	(ASpliJDU-class), 10
junctionDUreport, <i>13</i> , <i>19</i> , <i>22</i> , <i>27</i> , <i>29</i> , <i>34</i>	
junctionDUreport, ASpliAS-method	mergeBinDUAS, 31
(ASpliAS-class), 7	mergeBinDUAS,ASpliDU,ASpliAS-method
junctionsDU, <i>30</i> , <i>34</i>	(ASpliDU-class), 9
junctionsDU (DU accessors), 17	MergedReports accessors, 33
junctionsDU, ASpliDU-method	mergedReports, ASpliMergedReports-method
(ASpliDU-class), 9	(ASpliMergedReports-class), 11
junctionsDU<- (DU accessors), 17	mergeReports, 34
junctionsDU<-,ASpliDU,data.frame-method	mergeReports, ASpliDU-method
(ASpliDU-class), 9	(ASpliDU-class), 9
junctionsDUreport,ASpliJDU-method	, ,
(ASpliJDU-class), 10	pdf, 40
junctionsPIR, 6, 8	plotBins, 35
, -, -	1 7 7 7

INDEX 51

plotBins,ASpliCounts-method	writeAll(write),46
(ASpliCounts-class), 8	writeAll, ANY-method (write-methods), 47
plotGenomicRegions, 36, 38	writeAll,ASpliCounts-method
plotGenomicRegions,ASpliFeatures-method	(ASpliCounts-class), 8
(ASpliFeatures-class), 10	writeAS, 6
png, 40	writeAS(write),46
	writeAS, ASpliAS-method (ASpliAS-class)
rds, 41	,
rds, ASpliCounts-method	writeAS-methods (write-methods), 47
(ASpliCounts-class), 8	writeCounts, 43
rdsb, 43	writeCounts (write), 46
rdsb (Counts accesors), 15	writeCounts,ASpliCounts-method
rdsb,ASpliCounts-method	(ASpliCounts-class), 8
(ASpliCounts-class), 8	writeCounts-methods (write-methods), 47
rdsb<- (Counts accesors), 15	writeDU, <i>13</i> , <i>19</i> , <i>22</i> , <i>30</i> , <i>34</i>
rdsb<-,ASpliCounts,data.frame-method	writeDU(write),46
(ASpliCounts-class), 8	writeDU, ASpliDU-method (ASpliDU-class)
rdsg, <i>43</i>	9
rdsg (Counts accesors), 15	writeDU-methods (write-methods), 47
rdsg,ASpliCounts-method	writeRds(write),46
(ASpliCounts-class), 8	writeRds,ASpliCounts-method
rdsg<- (Counts accesors), 15	(ASpliCounts-class), 8
rdsg<-,ASpliCounts,data.frame-method	writeRds-methods (write-methods), 47
(ASpliCounts-class), 8	
readCounts, 42	
readCounts,ASpliFeatures-method	
(ASpliFeatures-class), 10	
show, ASpliAS-method (show-methods), 44	
show, ASpliCounts-method (show-methods),	
44	
show, ASpliDU-method (show-methods), 44	
show, ASpliFeatures-method	
(show-methods), 44	
show, ASpliMergedReports-method	
(show-methods), 44	
show-methods, 44	
subset (Subset ASpli objects), 44	
Subset ASpli objects, 44	
subset, ASpliAS-method (ASpliAS-class), 7	
subset, ASpliCounts-method	
(ASpliCounts-class), 8	
subsetBams (Subset ASpli objects), 44	
subsetTargets (Subset ASpli objects), 44	
tiff, 40	
write,46	
write, 40 write-methods, 47	
WI I CO IIIC CITOUS, T/	