Package ‘SummarizedExperiment’

April 29, 2020

Title SummarizedExperiment container

Description The SummarizedExperiment container contains one or more assays, each represented by a matrix-like object of numeric or other mode. The rows typically represent genomic ranges of interest and the columns represent samples.

Version 1.18.0

Encoding UTF-8

Author Martin Morgan, Valerie Obenchain, Jim Hester, Hervé Pagès

Maintainer Bioconductor Package Maintainer <maintainer@bioconductor.org>

biocViews Genetics, Infrastructure, Sequencing, Annotation, Coverage, GenomeAnnotation

Depends R (>= 3.2), methods, GenomicRanges (>= 1.33.6), Biobase, DelayedArray (>= 0.3.20)

Imports utils, stats, tools, Matrix, BiocGenerics (>= 0.15.3), S4Vectors (>= 0.25.14), IRanges (>= 2.21.6), GenomeInfoDb (>= 1.13.1)

Suggests annotate, AnnotationDbi, hgu95av2.db, GenomicFeatures, TxDb.Hsapiens.UCSC.hg19.knownGene, BiocStyle, knitr, markdown, digest, jsonlite, rhdf5, HDF5Array (>= 1.7.5), airway, RUnit, testthat

VignetteBuilder knitr

License Artistic-2.0

Collate Assays-class.R SummarizedExperiment-class.R
        RangedSummarizedExperiment-class.R intra-range-methods.R
        inter-range-methods.R coverage-methods.R findOverlaps-methods.R
        nearest-methods.R makeSummarizedExperimentFromExpressionSet.R
        makeSummarizedExperimentFromDataFrame.R
        makeSummarizedExperimentFromLoom.R readKallisto.R zzz.R

git_url https://git.bioconductor.org/packages/SummarizedExperiment

git_branch RELEASE_3_11

git_last_commit ebaaf2

git_last_commit_date 2020-04-27

Date/Publication 2020-04-28
**Assays-class**

**R topics documented:**

- Assays-class .................................................. 2
- coverage-methods ............................................ 4
- findOverlaps-methods ....................................... 5
- inter-range-methods .......................................... 7
- intra-range-methods ......................................... 8
- makeSummarizedExperimentFromDataFrame ..................... 10
- makeSummarizedExperimentFromExpressionSet .................... 12
- makeSummarizedExperimentFromLoom .......................... 13
- nearest-methods ............................................. 14
- RangedSummarizedExperiment-class ............................. 16
- readKallisto .................................................. 20
- SummarizedExperiment-class .................................. 22

**Index**

<table>
<thead>
<tr>
<th>Assays-class</th>
<th>Assays objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assays-class</td>
<td>Assays objects</td>
</tr>
</tbody>
</table>

**Description**

The Assays virtual class and its methods provide a formal abstraction of the assays slot of **SummarizedExperiment** objects.

SimpleAssays and ShallowSimpleListAssays are concrete subclasses of Assays with the former being currently the default implementation of Assays objects. Other implementations (e.g. disk-based) could easily be added.

Note that these classes are not meant to be used directly by the end user and the material in this man page is aimed at package developers.

**Details**

Assays objects have a list-like semantics with elements having matrix- or array-like semantics (e.g., dim, dimnames).

The Assays API consists of:

- (a) The Assays() constructor function.
- (b) Lossless back and forth coercion from/to SimpleList. The coercion method from SimpleList doesn’t need (and should not) validate the returned object.
- (c) length, names, `names<--`, getListElement, setListElement, dim, [`, `[<--`, rbind, cbind.

An Assays concrete subclass needs to implement (b) (required) plus, optionally any of the methods in (c).

**IMPORTANT:**

1. Nobody in the Assays hierarchy is allowed to inherit from SimpleList because of the conflicting semantic of [.

---

**Index**

<table>
<thead>
<tr>
<th>Assays-class</th>
<th>Assays objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assays-class</td>
<td>Assays objects</td>
</tr>
</tbody>
</table>

---

**Description**

The Assays virtual class and its methods provide a formal abstraction of the assays slot of **SummarizedExperiment** objects.

SimpleAssays and ShallowSimpleListAssays are concrete subclasses of Assays with the former being currently the default implementation of Assays objects. Other implementations (e.g. disk-based) could easily be added.

Note that these classes are not meant to be used directly by the end user and the material in this man page is aimed at package developers.

**Details**

Assays objects have a list-like semantics with elements having matrix- or array-like semantics (e.g., dim, dimnames).

The Assays API consists of:

- (a) The Assays() constructor function.
- (b) Lossless back and forth coercion from/to SimpleList. The coercion method from SimpleList doesn’t need (and should not) validate the returned object.
- (c) length, names, `names<--`, getListElement, setListElement, dim, [`, `[<--`, rbind, cbind.

An Assays concrete subclass needs to implement (b) (required) plus, optionally any of the methods in (c).

**IMPORTANT:**

1. Nobody in the Assays hierarchy is allowed to inherit from SimpleList because of the conflicting semantic of [.

---

**Index**

<table>
<thead>
<tr>
<th>Assays-class</th>
<th>Assays objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assays-class</td>
<td>Assays objects</td>
</tr>
</tbody>
</table>
2. Methods that return a modified Assays object (a.k.a. endomorphisms), that is, as well as replacement methods names<-, setListElement, and [<-, must respect the copy-on-change contract. With objects that don’t make use of references internally, the developer doesn’t need to take any special action for that because it’s automatically taken care of by R itself. However, for objects that do make use of references internally (e.g. environments, external pointers, pointer to a file on disk, etc...), the developer needs to be careful to implement endomorphisms with copy-on-change semantics. This can easily be achieved (and is what the default methods for Assays objects do) by performing a full (deep) copy of the object before modifying it instead of trying to modify it in-place. However note that this full (deep) copy can be very expensive and is actually not necessary in order to achieve copy-on-change semantics: it’s enough (and often preferable for performance reasons) to copy only the parts of the object that need to be modified.

Assays has currently 3 implementations which are formalized by concrete subclasses SimpleAssays, ShallowSimpleListAssays, and AssaysInEnv. SimpleAssays is the default (prior to SummarizedExperiment 1.15.4, ShallowSimpleListAssays was the default). AssaysInEnv is a broken alternative to ShallowSimpleListAssays that does NOT respect the copy-on-change contract. It is only provided for illustration purposes (see source file Assays-class.R for the details).

A little more detail about ShallowSimpleListAssays: a small reference class hierarchy (not exported from the GenomicRanges name space) defines a reference class ShallowData with a single field data of type ANY, and a derived class ShallowSimpleListAssays that specializes the type of data as SimpleList, and contains=c("ShallowData","Assays"). The assays slot of a SummarizedExperiment object contains an instance of ShallowSimpleListAssays.

Author(s)

Martin Morgan and Hervé Pagès

See Also

• SummarizedExperiment objects.

• SimpleList objects in the S4Vectors package.

Examples

```r
## DIRECT MANIPULATION OF Assays OBJECTS
m1 <- matrix(runif(24), ncol=3)
m2 <- matrix(runif(24), ncol=3)
a <- Assays(SimpleList(m1, m2))
a

as(a, "SimpleList")
length(a)
getListElement(a, 2)
dim(a)

b <- a[-4, 2]
b
length(b)
getListElement(b, 2)
dim(b)
```
coverage-methods

Coverage of a RangedSummarizedExperiment object

Description

This man page documents the coverage method for RangedSummarizedExperiment objects.

Usage

## S4 method for signature 'RangedSummarizedExperiment'
coverage(x, shift=0L, width=NULL, weight=1L,
    method=c("auto", "sort", "hash"))

Arguments

x A RangedSummarizedExperiment object.

shift, width, weight, method
See ?coverage in the GenomicRanges package.
details
This method operates on the rowRanges component of the RangedSummarizedExperiment object, which can be a GenomicRanges or GRangesList object.

More precisely, on RangedSummarizedExperiment object x, coverage(x,...) is equivalent to coverage(rowRanges(x),...).

See ?coverage in the GenomicRanges package for the details of how coverage operates on a GenomicRanges or GRangesList object.

value
See ?coverage in the GenomicRanges package.

see also
- RangedSummarizedExperiment objects.
- The coverage man page in the GenomicRanges package where the coverage methods for GenomicRanges and GRangesList objects are documented.

Examples

nrows <- 20; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
rowRanges <- GRanges(rep(c("chr1", "chr2"), c(5, 15)),
    IRanges(sample(1000L, 20), width=100),
    strand=Rle(c("+", "-"), c(12, 8)),
    seqlengths=c(chr1=1800, chr2=1300))
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3),
    row.names=LETTERS[1:6])
rse <- SummarizedExperiment(assays=SimpleList(counts=counts),
    rowRanges=rowRanges, colData=colData)
cvg <- coverage(rse)
cvg
stopifnot(identical(cvg, coverage(rowRanges(rse))))

findOverlaps-methods Finding overlapping ranges in RangedSummarizedExperiment objects

description
This man page documents the findOverlaps methods for RangedSummarizedExperiment objects. RangedSummarizedExperiment objects also support countOverlaps, overlapsAny, and subsetByOverlaps thanks to the default methods defined in the IRanges package and to the findOverlaps methods defined in this package and documented below.
Usage

```r
## S4 method for signature 'RangedSummarizedExperiment,Vector'
findOverlaps(query, subject,
  maxgap=-1L, minoverlap=0L,
  type=c("any", "start", "end", "within", "equal"),
  select=c("all", "first", "last", "arbitrary"),
  ignore.strand=FALSE)
## S4 method for signature 'Vector,RangedSummarizedExperiment'
findOverlaps(query, subject,
  maxgap=-1L, minoverlap=0L,
  type=c("any", "start", "end", "within", "equal"),
  select=c("all", "first", "last", "arbitrary"),
  ignore.strand=FALSE)
```

Arguments

- `query, subject` One of these two arguments must be a `RangedSummarizedExperiment` object.
- `maxgap, minoverlap, type` See `?findOverlaps` in the `GenomicRanges` package.
- `select, ignore.strand` See `?findOverlaps` in the `GenomicRanges` package.

Details

These methods operate on the `rowRanges` component of the `RangedSummarizedExperiment` object, which can be a `GenomicRanges` or `GRangesList` object.

More precisely, if any of the above functions is passed a `RangedSummarizedExperiment` object thru the query and/or subject argument, then it behaves as if `rowRanges(query)` and/or `rowRanges(subject)` had been passed instead.

See `?findOverlaps` in the `GenomicRanges` package for the details of how `findOverlaps` and family operate on `GenomicRanges` and `GRangesList` objects.

Value

See `?findOverlaps` in the `GenomicRanges` package.

See Also

- `RangedSummarizedExperiment` objects.
- The `findOverlaps` man page in the `GenomicRanges` package where the `findOverlaps` family of methods for `GenomicRanges` and `GRangesList` objects is documented.

Examples

```r
nrows <- 20; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
rowRanges <- GRanges(rep(c("chr1", "chr2"), c(5, 15)),
  IRanges(sample(1000L, 20), width=100),
  strand=Rle(c("+", "-"), c(12, 8)))
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3),
  row.names=LETTERS[1:6])
rse0 <- SummarizedExperiment(assays=SimpleList(counts=counts),
```
inter-range-methods

```r
rse1 <- shift(rse0, 100)

hits <- findOverlaps(rse0, rse1)

stopifnot(identical(hits, findOverlaps(rowRanges(rse0), rowRanges(rse1))))
stopifnot(identical(hits, findOverlaps(rse0, rowRanges(rse1))))
stopifnot(identical(hits, findOverlaps(rowRanges(rse0), rse1)))
```

---

**Description**

This man page documents the *inter range transformations* that are supported on `RangedSummarizedExperiment` objects.

**Usage**

```r
## S4 method for signature 'RangedSummarizedExperiment'
isDisjoint(x, ignore.strand=FALSE)

## S4 method for signature 'RangedSummarizedExperiment'
disjointBins(x, ignore.strand=FALSE)
```

**Arguments**

- `x` A `RangedSummarizedExperiment` object.
- `ignore.strand` See `?isDisjoint` in the `GenomicRanges` package.

**Details**

These transformations operate on the `rowRanges` component of the `RangedSummarizedExperiment` object, which can be a `GenomicRanges` or `GRangesList` object.

More precisely, any of the above functions performs the following transformation on `RangedSummarizedExperiment` object `x`:

```r
f(rowRanges(x), ...)
```

where `f` is the name of the function and `...` any additional arguments passed to it.

See `?isDisjoint` in the `GenomicRanges` package for the details of how these transformations operate on a `GenomicRanges` or `GRangesList` object.

**Value**

See `?isDisjoint` in the `GenomicRanges` package.

**See Also**

- `RangedSummarizedExperiment` objects.
- The `isDisjoint` man page in the `GenomicRanges` package where *inter range transformations* of a `GenomicRanges` or `GRangesList` object are documented.
**Examples**

```r
nrows <- 20; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
rowRanges <- GRanges(rep(c("chr1", "chr2"), c(5, 15)), 
                      IRanges(sample(1000L, 20), width=100), 
                      strand=Rle(c("+", "-"), c(12, 8)))
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3), 
                      row.names=LETTERS[1:6])
rse0 <- SummarizedExperiment(assays=SimpleList(counts=counts), 
                             rowRanges=rowRanges, colData=colData)
rse1 <- shift(rse0, 99*start(rse0))
isDisjoint(rse0) # FALSE
isDisjoint(rse1) # TRUE
bins0 <- disjointBins(rse0)
bins0
stopifnot(identical(bins0, disjointBins(rowRanges(rse0))))

bins1 <- disjointBins(rse1)
bins1
stopifnot(all(bins1 == bins1[1]))
```

---

**Description**

This man page documents the *intra range transformations* that are supported on *RangedSummarizedExperiment* objects.

**Usage**

```r
## S4 method for signature 'RangedSummarizedExperiment'
shift(x, shift=0L, use.names=TRUE)

## S4 method for signature 'RangedSummarizedExperiment'
narrow(x, start=NA, end=NA, width=NA, use.names=TRUE)

## S4 method for signature 'RangedSummarizedExperiment'
resize(x, width, fix="start", use.names=TRUE, 
       ignore.strand=FALSE)

## S4 method for signature 'RangedSummarizedExperiment'
flank(x, width, start=TRUE, both=FALSE, 
      use.names=TRUE, ignore.strand=FALSE)

## S4 method for signature 'RangedSummarizedExperiment'
promoters(x, upstream=2000, downstream=200)

## S4 method for signature 'RangedSummarizedExperiment'
```
 restrict(x, start=NA, end=NA, keep.all.ranges=FALSE, 
       use.names=TRUE)

## S4 method for signature 'RangedSummarizedExperiment'
trim(x, use.names=TRUE)

Arguments

  x

  A RangedSummarizedExperiment object.

  shift, use.names

  See ?shift in the IRanges package.

  start, end, width, fix

  See ?shift in the IRanges package.

  ignore.strand, both

  See ?shift in the IRanges package.

  upstream, downstream

  See ?shift in the IRanges package.

  keep.all.ranges

  See ?shift in the IRanges package.

Details

These transformations operate on the rowRanges component of the RangedSummarizedExperiment object, which can be a GenomicRanges or GRangesList object.

More precisely, any of the above functions performs the following transformation on RangedSummarizedExperiment object x:

rowRanges(x) <- f(rowRanges(x), ...)

where f is the name of the function and ... any additional arguments passed to it.

See ?shift in the IRanges package for the details of how these transformations operate on a GenomicRanges or GRangesList object.

See Also

• RangedSummarizedExperiment objects.

• The shift man page in the IRanges package where intra range transformations of a GenomicRanges or GRangesList object are documented.

Examples

nrows <- 20; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
rowRanges <- GRanges(rep(c("chr1", "chr2"), c(5, 15)),
                   IRanges(sample(1000L, 20), width=100),
                   strand=Rle(c("+", "-"), c(12, 8)))
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3),
                   row.names=LETTERS[1:6])
rse0 <- SummarizedExperiment(assays=SimpleList(counts=counts),
                          rowRanges=rowRanges, colData=colData)
rse1 <- shift(rse0, 1)
stopifnot(identical(rowRanges(rse1), shift(rowRanges(rse0), 1)))

se2 <- narrow(rse0, start=10, end=-15)
stopifnot(identical(rowRanges(se2), narrow(rowRanges(rse0), start=10, end=-15)))

se3 <- resize(rse0, width=75)
stopifnot(identical(rowRanges(se3), resize(rowRanges(rse0), width=75)))

se4 <- flank(rse0, width=20)
stopifnot(identical(rowRanges(se4), flank(rowRanges(rse0), width=20)))

se5 <- restrict(rse0, start=200, end=700, keep.all.ranges=TRUE)
stopifnot(identical(rowRanges(se5), restrict(rowRanges(rse0), start=200, end=700, keep.all.ranges=TRUE)))

---

**makeSummarizedExperimentFromDataFrame**

*Make a RangedSummarizedExperiment from a data.frame or DataFrame*

**Description**

makeSummarizedExperimentFromDataFrame uses data.frame or DataFrame column names to create a GRanges object for the rowRanges of the resulting SummarizedExperiment object. It requires that non-range data columns be coercible into a numeric matrix for the SummarizedExperiment constructor. All columns that are not part of the row ranges attribute are assumed to be experiment data; thus, keeping metadata columns will not be supported. Note that this function only returns SummarizedExperiment objects with a single assay.

If metadata columns are to be kept, one can first construct the row ranges attribute by using the makeGRangesFromDataFrame function and subsequently creating the SummarizedExperiment.

**Usage**

```r
makeSummarizedExperimentFromDataFrame(df, 
...,
  seqinfo = NULL,
  starts.in.df.are.0based = FALSE)
```
makeSummarizedExperimentFromDataFrame

Arguments

df 
A data.frame or DataFrame object. If not, then the function first tries to turn df into a data frame with as.data.frame(df).

... Additional arguments passed on to makeGRangesFromDataFrame

seqinfo 
Either NULL, or a Seqinfo object, or a character vector of seqlevels, or a named numeric vector of sequence lengths. When not NULL, it must be compatible with the genomic ranges in df i.e. it must include at least the sequence levels represented in df.

starts.in.df.are.0based 
TRUE or FALSE (the default). If TRUE, then the start positions of the genomic ranges in df are considered to be 0-based and are converted to 1-based in the returned GRanges object. This feature is intended to make it more convenient to handle input that contains data obtained from resources using the “0-based start” convention. A notorious example of such resource is the UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables).

Value

A RangedSummarizedExperiment object with rowRanges and a single assay

Author(s)

M. Ramos

See Also

• makeGRangesFromDataFrame

Examples

```r
## BASIC EXAMPLES

# Note that rownames of the data.frame are also rownames of the result
df <- data.frame(chr="chr2", start = 11:15, end = 12:16,
                 strand = c("+", "-", "+", "A", "."), expr0 = 3:7,
                 expr1 = 8:12, expr2 = 12:16,
                 row.names = paste0("GENE", letters[5:1]))

df

exRSE <- makeSummarizedExperimentFromDataFrame(df)

exRSE

assay(exRSE)

rowRanges(exRSE)
```
**Description**

Coercion between `RangedSummarizedExperiment` and `ExpressionSet` is supported in both directions.

For going from `ExpressionSet` to `RangedSummarizedExperiment`, the `makeSummarizedExperimentFromExpressionSet` function is also provided to let the user control how to map features to ranges.

**Usage**

```r
makeSummarizedExperimentFromExpressionSet(from,
mapFun=naiveRangeMapper,
...)
```

```r
## range mapping functions
naiveRangeMapper(from)
probeRangeMapper(from)
geneRangeMapper(txDbPackage, key = "ENTREZID")
```

**Arguments**

- `from` An `ExpressionSet` object.
- `mapFun` A function which takes an `ExpressionSet` object and returns a `GRanges`, or `GRangesList` object which corresponds to the genomic ranges used in the `ExpressionSet`. The `rownames` of the returned `GRanges` are used to match the `featureNames` of the `ExpressionSet`. The `naiveRangeMapper` function is used by default.
- `...` Additional arguments passed to `mapFun`.
- `txDbPackage` A character string with the Transcript Database to use for the mapping.
- `key` A character string with the Gene key to use for the mapping.

**Value**

`makeSummarizedExperimentFromExpressionSet` takes an `ExpressionSet` object as input and a **range mapping function** that maps the features to ranges. It then returns a `RangedSummarizedExperiment` object that corresponds to the input.

The range mapping functions return a `GRanges` object, with the `rownames` corresponding to the `featureNames` of the `ExpressionSet` object.

**Author(s)**

Jim Hester, james.f.hestergmail.com
makeSummarizedExperimentFromLoom

See Also

- RangedSummarizedExperiment objects.
- ExpressionSet objects in the Biobase package.
- TxDb objects in the GenomicFeatures package.

Examples

```r
## ---------------------------------------------------------------------
## GOING FROM ExpressionSet TO SummarizedExperiment
## ---------------------------------------------------------------------

data(sample.ExpressionSet, package="Biobase")

# naive coercion
makeSummarizedExperimentFromExpressionSet(sample.ExpressionSet)
as(sample.ExpressionSet, "RangedSummarizedExperiment")
as(sample.ExpressionSet, "SummarizedExperiment")

# using probe range mapper
makeSummarizedExperimentFromExpressionSet(sample.ExpressionSet, probeRangeMapper)

# using the gene range mapper
se <- makeSummarizedExperimentFromExpressionSet(
  sample.ExpressionSet,
  geneRangeMapper("TxDb.Hsapiens.UCSC.hg19.knownGene")
)
se
rowData(se) # duplicate row names

## ---------------------------------------------------------------------
## GOING FROM SummarizedExperiment TO ExpressionSet
## ---------------------------------------------------------------------

example(RangedSummarizedExperiment) # to create 'rse'
rse
as(rse, "ExpressionSet")
```

makeSummarizedExperimentFromLoom

Make a SummarizedExperiment from a `.loom` hdf5 file

Description

makeSummarizedExperimentFromLoom represents a `.loom` file as a SummarizedExperiment. The `/matrix` and `/layers` are represented as HDF5Array objects; row and column attributes are parsed to DataFrame. Optionally, row or column attributes can be specified as row and and column names.

Usage

```r
makeSummarizedExperimentFromLoom(file, 
  rownames_attr = NULL, 
  colnames_attr = NULL)
```
nearest-methods

Arguments

- **file**: The path (as a single character string) to the HDF5 file where the dataset is located.
- **rownames_attr**: The name of the row attribute to be used as row names.
- **colnames_attr**: The name of the column attribute to be used as column names.

Value

A *SummarizedExperiment* object with row and column data and one or more assays.

Author(s)

Martin Morgan

See Also

[http://loompy.org/loompy-docs/format/index.html](http://loompy.org/loompy-docs/format/index.html) for a specification of the .loom format.

Examples

```r
## BASIC EXAMPLE

file <- system.file(
    package="SummarizedExperiment", "extdata", "example.loom"
)
se <- makeSummarizedExperimentFromLoom(file)
se
assay(se)
metadata(se)
```

### Description

This man page documents the nearest methods and family (i.e. precede, follow, distance, and distanceToNearest methods) for *RangedSummarizedExperiment* objects.

### Usage

```r
## S4 method for signature 'RangedSummarizedExperiment,ANY'
precede(x, subject, select=c("arbitrary", "all"),
    ignore.strand=FALSE)
## S4 method for signature 'ANY,RangedSummarizedExperiment'
precede(x, subject, select=c("arbitrary", "all"),
    ignore.strand=FALSE)
## S4 method for signature 'RangedSummarizedExperiment,ANY'
follow(x, subject, select=c("arbitrary", "all"),
```
nearest-methods

```r
## S4 method for signature 'ANY,RangedSummarizedExperiment'
follow(x, subject, select=c("arbitrary", "all"),
ignore.strand=FALSE)

## S4 method for signature 'RangedSummarizedExperiment,ANY'
nearest(x, subject, select=c("arbitrary", "all"), ignore.strand=FALSE)

## S4 method for signature 'ANY,RangedSummarizedExperiment'
nearest(x, subject, select=c("arbitrary", "all"), ignore.strand=FALSE)

## S4 method for signature 'RangedSummarizedExperiment,ANY'
distance(x, y, ignore.strand=FALSE, ...)

## S4 method for signature 'ANY,RangedSummarizedExperiment'
distance(x, y, ignore.strand=FALSE, ...)

## S4 method for signature 'RangedSummarizedExperiment,ANY'
distanceToNearest(x, subject, ignore.strand=FALSE, ...)

## S4 method for signature 'ANY,RangedSummarizedExperiment'
distanceToNearest(x, subject, ignore.strand=FALSE, ...)
```

### Arguments

- `x, subject`
  One of these two arguments must be a `RangedSummarizedExperiment` object.
- `select, ignore.strand`
  See `?nearest` in the `GenomicRanges` package.
- `y`
  For the distance methods, one of `x` or `y` must be a `RangedSummarizedExperiment` object.
- `...`
  Additional arguments for methods.

### Details

These methods operate on the `rowRanges` component of the `RangedSummarizedExperiment` object, which can be a `GenomicRanges` or `GRangesList` object.

More precisely, if any of the above functions is passed a `RangedSummarizedExperiment` object thru the `x, subject,` and/or `y` argument, then it behaves as if `rowRanges(x), rowRanges(subject),` and/or `rowRanges(y)` had been passed instead.

See `?nearest` in the `GenomicRanges` package for the details of how nearest and family operate on GenomicRanges and GRangesList objects.

### Value

See `?nearest` in the GenomicRanges package.

### See Also

- `RangedSummarizedExperiment` objects.
- The `nearest` man page in the GenomicRanges package where the nearest family of methods for GenomicRanges and GRangesList objects is documented.
Examples

```r
nrows <- 20; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
rowRanges <- GRanges(rep(c("chr1", "chr2"), c(5, 15)),
                       IRanges(sample(1000L, 20), width=100),
                       strand=Rle(c("+", "-"), c(12, 8)))
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3),
                      row.names=LETTERS[1:6])
rse0 <- SummarizedExperiment(assays=SimpleList(counts=counts),
                             rowRanges=rowRanges, colData=colData)
```

```r
rse1 <- shift(rse0, 100)
```

```r
res <- nearest(rse0, rse1)
```

```r
stopifnot(identical(res, nearest(rowRanges(rse0), rowRanges(rse1))))
```

```r
res <- nearest(rse0, rowRanges(rse1))
```

```r
res <- nearest(rse0) # missing subject
```

```r
res <- nearest(rse0, rowRanges(rse0))
```

```r
hits <- nearest(rse0, rse1, select="all")
```

```r
stopifnot(identical(hits, nearest(rowRanges(rse0), rowRanges(rse1), select="all")))
```

```r
stopifnot(identical(hits, nearest(rse0, rowRanges(rse1), select="all")))
```

```r
stopifnot(identical(hits, nearest(rowRanges(rse0), rse1, select="all")))
```

RangedSummarizedExperiment-class

RangedSummarizedExperiment objects

Description

The RangedSummarizedExperiment class is a matrix-like container where rows represent ranges of interest (as a GRanges or GRangesList object) and columns represent samples (with sample data summarized as a DataFrame). A RangedSummarizedExperiment contains one or more assays, each represented by a matrix-like object of numeric or other mode.

RangedSummarizedExperiment is a subclass of SummarizedExperiment and, as such, all the methods documented in ?SummarizedExperiment also work on a RangedSummarizedExperiment object. The methods documented below are additional methods that are specific to RangedSummarizedExperiment objects.
Usage

## Constructor

```r
SummarizedExperiment(assays=SimpleList(),
                      rowData=NULL, rowRanges=GRangesList(),
                      colData=DataFrame(),
                      metadata=list())
```

## Accessors

```r
rowRanges(x, ...)
rowRanges(x, ...) <- value
```

## Subsetting

## S4 method for signature 'RangedSummarizedExperiment'

```r
subset(x, subset, select, ...)
```

## rowRanges access

## see 'GRanges compatibility', below

Arguments

- **assays**
  A list or `SimpleList` of matrix-like elements, or a matrix-like object (e.g. an ordinary matrix, a data frame, a `DataFrame` object from the `S4Vectors` package, a `sparseMatrix` derivative from the `Matrix` package, a `DelayedMatrix` object from the `DelayedArray` package, etc...). All elements of the list must have the same dimensions, and dimension names (if present) must be consistent across elements and with the row names of `rowRanges` and `colData`.

- **rowData**
  A `DataFrame` object describing the rows. Row names, if present, become the row names of the `SummarizedExperiment` object. The number of rows of the `DataFrame` must equal the number of rows of the matrices in `assays`.

- **rowRanges**
  A `GRanges` or `GRangesList` object describing the ranges of interest. Names, if present, become the row names of the `SummarizedExperiment` object. The length of the `GRanges` or `GRangesList` must equal the number of rows of the matrices in `assays`. If `rowRanges` is missing, a `SummarizedExperiment` instance is returned.

- **colData**
  An optional `DataFrame` describing the samples. Row names, if present, become the column names of the `RangedSummarizedExperiment`.

- **metadata**
  An optional `list` of arbitrary content describing the overall experiment.

- **x**
  A `RangedSummarizedExperiment` object. The `rowRanges` setter will also accept a `SummarizedExperiment` object and will first coerce it to `RangedSummarizedExperiment` before it sets `value` on it.

- **...**
  Further arguments to be passed to or from other methods.

- **value**
  A `GRanges` or `GRangesList` object.

- **subset**
  An expression which, when evaluated in the context of `rowRanges(x)`, is a logical vector indicating elements or rows to keep: missing values are taken as false.

- **select**
  An expression which, when evaluated in the context of `colData(x)`, is a logical vector indicating elements or rows to keep: missing values are taken as false.
Details

The rows of a RangedSummarizedExperiment object represent ranges (in genomic coordinates) of interest. The ranges of interest are described by a GRanges or a GRangesList object, accessible using the rowRanges function, described below. The GRanges and GRangesList classes contain sequence (e.g., chromosome) name, genomic coordinates, and strand information. Each range can be annotated with additional data; this data might be used to describe the range or to summarize results (e.g., statistics of differential abundance) relevant to the range. Rows may or may not have row names; they often will not.

Constructor

RangedSummarizedExperiment instances are constructed using the SummarizedExperiment function with arguments outlined above.

Accessors

In the following code snippets, x is a RangedSummarizedExperiment object.

rowRanges(x), rowRanges(x) <- value: Get or set the row data. value is a GenomicRanges object. Row names of value must be NULL or consistent with the existing row names of x.

GRanges compatibility (rowRanges access)

Many GRanges and GRangesList operations are supported on RangedSummarizedExperiment objects, using rowRanges.

Supported operations include: pcompare, duplicated, end, end<-, granges, is.unsorted, match, mcols, mcols<-, order, ranges, ranges<-, rank, seqinfo, seqinfo<-, seqnames, sort, start, start<-, strand, strand<-, width, width<-. See also ?shift, ?isDisjoint, ?coverage, ?findOverlaps, and ?nearest for more GRanges compatibility methods.

Not all GRanges operations are supported, because they do not make sense for RangedSummarizedExperiment objects (e.g., length, name, as.data.frame, c, splitAsList), involve non-trivial combination or splitting of rows (e.g., disjoin, gaps, reduce, unique), or have not yet been implemented (Ops, map, window, window<-).

Subsetting

In the code snippets below, x is a RangedSummarizedExperiment object.

subset(x, subset, select): Create a subset of x using an expression subset referring to columns of rowRanges(x) (including ‘seqnames’, ‘start’, ‘end’, ‘width’, ‘strand’, and names(rowData(x))) and / or select referring to column names of colData(x).

Extension

RangedSummarizedExperiment is implemented as an S4 class, and can be extended in the usual way, using contains="RangedSummarizedExperiment" in the new class definition.

Author(s)

Martin Morgan, mtmorgan@fhcrc.org
### Examples

```r	nrows <- 200; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
rowRanges <- GRanges(rep(c("chr1", "chr2"), c(50, 150)),
                     IRanges(floor(runif(200, 1e5, 1e6)), width=100),
                     strand=sample(c("+", "-"), 200, TRUE),
                     feature_id=sprintf("ID%03d", 1:200))
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3),
                     row.names=LETTERS[1:6])
rse <- SummarizedExperiment(assays=SimpleList(counts=counts),
                           rowRanges=rowRanges, colData=colData)
rse
dim(rse)
dimnames(rse)
assayNames(rse)
head(assay(rse))
assays(rse) <- endoapply(assay(rse), asinh)
head(assay(rse))

rowRanges(rse)
rowData(rse) # same as `mcols(rowRanges(rse))`
colData(rse)

rse[, rse$Treatment == "ChIP"]

## cbind() combines objects with the same ranges but different samples:
rse1 <- rse
rse2 <- rse1[, 1:3]
colnames(rse2) <- letters[1:ncol(rse2)]
cmb1 <- cbind(rse1, rse2)
dim(cmb1)
dimnames(cmb1)

## rbind() combines objects with the same samples but different ranges:
rse1 <- rse
rse2 <- rse1[1:50, ]
rownames(rse2) <- letters[1:nrow(rse2)]
cmb2 <- rbind(rse1, rse2)
dim(cmb2)
dimnames(cmb2)

## Coercion to/from SummarizedExperiment:
se0 <- as(rse, "SummarizedExperiment")
se0
as(se0, "RangedSummarizedExperiment")

## Setting rowRanges on a SummarizedExperiment object turns it into a
## RangedSummarizedExperiment object:
```

### See Also
- `SummarizedExperiment-class`
- `shift, isDisjoint, coverage, findOverlaps, and nearest` for more GRanges compatibility methods.
- `GRanges` objects in the GenomicRanges package.
se <- se0
rowRanges(se) <- rowRanges
se # RangedSummarizedExperiment

## Sanity checks:
stopifnot(identical(assays(se0), assays(rse)))
stopifnot(identical(dim(se0), dim(rse)))
stopifnot(identical(dimnames(se0), dimnames(rse)))
stopifnot(identical(rowData(se0), rowData(rse)))
stopifnot(identical(colData(se0), colData(rse)))

---

**readKallisto**

**Input kallisto or kallisto bootstrap results.**

**Description**

readKallisto inputs several kallisto output files into a single SummarizedExperiment instance, with rows corresponding to estimated transcript abundance and columns to samples. readKallistoBootstrap inputs kallisto bootstrap replicates of a single sample into a matrix of transcript x bootstrap abundance estimates.

**Usage**

```r
readKallisto(files,
  json = file.path(dirname(files), "run_info.json"),
  h5 = any(grepl("\.h5$", files)), what = KALLISTO_ASSAYS,
  as = c("SummarizedExperiment", "list", "matrix"))
```

```r
readKallistoBootstrap(file, i, j)
```

**Arguments**

- **files** character() paths to kallisto `abundance.tsv` output files. The assumption is that files are organized in the way implied by kallisto, with each sample in a distinct directory, and the directory containing files `abundance.tsv`, `run_info.json`, and perhaps `abundance.h5`.
- **json** character() vector of the same length as `files` specifying the location of JSON files produced by kallisto and containing information on the run. The default assumes that json files are in the same directory as the corresponding abundance file.
- **h5** character() vector of the same length as `files` specifying the location of HDF5 files produced by kallisto and containing bootstrap estimates. The default assumes that HDF5 files are in the same directory as the corresponding abundance file.
- **what** character() vector of kallisto per-sample outputs to be input. See KALLISTO_ASSAYS for available values.
- **as** character(1) specifying the output format. See Value for additional detail.
- **file** character(1) path to a single HDF5 output file.
- **i, j** integer() vector of row (i) and column (j) indexes to input.
Value

A `SummarizedExperiment`, `list`, or `matrix`, depending on the value of argument `as`; by default a `SummarizedExperiment`. The `as="SummarizedExperiment"` `rowData(se)` the length of each transcript; `colData(se)` includes summary information on each sample, including the number of targets and bootstraps, the kallisto and index version, the start time and operating system call used to create the file. `assays()` contains one or more transcript x sample matrices of parameters estimated by kallisto (see `KALLISTO_ASSAYS`).

`as="list"` return value contains information similar to `SummarizedExperiment` with row, column and assay data as elements of the list without coordination of row and column annotations into an integrated data container. `as="matrix"` returns the specified assay as a simple `R` matrix.

Author(s)

Martin Morgan martin.morgan@roswellpark.org

References

http://pachterlab.github.io/kallisto software for quantifying transcript abundance.

Examples

```r
outputs <- system.file(package="SummarizedExperiment", "extdata", "kallisto")
files <- dir(outputs, pattern="abundance.tsv", full=TRUE, recursive=TRUE)
stopifnot(all(file.exists(files)))

## default: input 'est_counts'
(se <- readKallisto(files, as="SummarizedExperiment"))
str(readKallisto(files, as="list"))
str(readKallisto(files, as="matrix"))

## available assays
KALLISTO_ASSAYS
## one or more assay
readKallisto(files, what=c("tpm", "eff_length"))

## alternatively: read hdf5 files
files <- sub(".tsv", ".h5", files, fixed=TRUE)
readKallisto(files)

## input all bootstraps
xx <- readKallistoBootstrap(files[1])
ridx <- head(which(rowSums(xx) != 0), 3)
cidx <- c(1:5, 96:100)
xx[ridx, cidx]

## selective input of rows (transcripts) and/or bootstraps
readKallistoBootstrap(files[1], i=c(ridx, rev(ridx)), j=cidx)
```
SummarizedExperiment-class

SummarizedExperiment objects

Description

The SummarizedExperiment class is a matrix-like container where rows represent features of interest (e.g. genes, transcripts, exons, etc...) and columns represent samples (with sample data summarized as a DataFrame). A SummarizedExperiment object contains one or more assays, each represented by a matrix-like object of numeric or other mode.

Note that SummarizedExperiment is the parent of the RangedSummarizedExperiment class which means that all the methods documented below also work on a RangedSummarizedExperiment object.

Usage

## Constructor

# See ?RangedSummarizedExperiment for the constructor function.

## Accessors

assayNames(x, ...)
assayNames(x, ...) <- value
assays(x, withDimnames=TRUE, ...)
assays(x, withDimnames=TRUE, ...) <- value
assay(x, i, withDimnames=TRUE, ...)
assay(x, i, withDimnames=TRUE, ...) <- value
rowData(x, use.names=TRUE, ...)
rowData(x, ...) <- value
colData(x, ...)
colData(x, ...) <- value
#dim(x)
#dimnames(x)
#dimnames(x) <- value

## Quick colData access

## S4 method for signature 'SummarizedExperiment'
x$name
## S4 replacement method for signature 'SummarizedExperiment'
x$name <- value
## S4 method for signature 'SummarizedExperiment,ANY,missing'
x[[i, j, ...]]
## S4 replacement method for signature 'SummarizedExperiment,ANY,missing'
x[[i, j, ...]] <- value

## Subsetting
Arguments

x A SummarizedExperiment object.

... For assay, arguments in ... are forwarded to assays.

For cbind, rbind, ... contains SummarizedExperiment objects to be combined.

For other accessors, ignored.

value An object of a class specified in the S4 method signature or as outlined in ‘Details’.

i, j For assay, assay<-, i is an integer or numeric scalar; see ‘Details’ for additional constraints.

For [.SummarizedExperiment, [.SummarizedExperiment<-, i, j are subscripts that can act to subset the rows and columns of x, that is the matrix elements of assays.

For [[,SummarizedExperiment, [[<-,SummarizedExperiment, i is a scalar index (e.g., character(1)) or integer(1)) into a column of colData.

name A symbol representing the name of a column of colData.

withDimnames A logical(1), indicating whether the dimnames of the SummarizedExperiment object should be applied (i.e. copied) to the extracted assays. More precisely, setting withDimnames=FALSE in the getter returns the assays as-is whereas setting withDimnames=FALSE return them with possibly modified dimnames.

Setting withDimnames=FALSE in the setter (assays<-) is required when the dimnames on the supplied assays are not identical to the dimnames on the SummarizedExperiment object; it does not influence actual assignment of dimnames to assays (they’re always stored as-is).

Note that

assays(x, withDimnames=FALSE) <- assays(x, withDimnames=FALSE)

is guaranteed to always work and be a no-op. This is not the case if withDimnames=TRUE is used or if withDimnames is not specified.
use.names

Like `mcols(x)`, by default `rowData(x)` propagates the rownames of `x` to the returned `DataFrame` object (note that for a `SummarizedExperiment` object, the rownames are also the names i.e. `rownames(x)` is always the same as `names(x)`). Setting `use.names=FALSE` suppresses this propagation i.e. it returns a `DataFrame` object with no rownames. Use this when `rowData(x)` fails, which can happen when the rownames contain NAs (because the rownames of a `SummarizedExperiment` object can contain NAs, but the rownames of a `DataFrame` object cannot).

drop

A logical(1), ignored by these methods.

deparse.level

See ?base::cbind for a description of this argument.

subset

An expression which, when evaluated in the context of `rowData(x)`, is a logical vector indicating elements or rows to keep: missing values are taken as false.

select

An expression which, when evaluated in the context of `colData(x)`, is a logical vector indicating elements or rows to keep: missing values are taken as false.

BACKEND

NULL (the default), or a single string specifying the name of the backend. When the backend is set to NULL, each element of `assays(x)` is realized in memory as an ordinary array by just calling `as.array` on it.

Details

The `SummarizedExperiment` class is meant for numeric and other data types derived from a sequencing experiment. The structure is rectangular like a matrix, but with additional annotations on the rows and columns, and with the possibility to manage several assays simultaneously.

The rows of a `SummarizedExperiment` object represent features of interest. Information about these features is stored in a `DataFrame` object, accessible using the function `rowData`. The `DataFrame` must have as many rows as there are rows in the `SummarizedExperiment` object, with each row of the `DataFrame` providing information on the feature in the corresponding row of the `SummarizedExperiment` object. Columns of the `DataFrame` represent different attributes of the features of interest, e.g., gene or transcript IDs, etc.

Each column of a `SummarizedExperiment` object represents a sample. Information about the samples are stored in a `DataFrame`, accessible using the function `colData`, described below. The `DataFrame` must have as many rows as there are columns in the `SummarizedExperiment` object, with each row of the `DataFrame` providing information on the sample in the corresponding column of the `SummarizedExperiment` object. Columns of the `DataFrame` represent different sample attributes, e.g., tissue of origin, etc. Columns of the `DataFrame` can themselves be annotated (via the `mcols` function). Column names typically provide a short identifier unique to each sample.

A `SummarizedExperiment` object can also contain information about the overall experiment, for instance the lab in which it was conducted, the publications with which it is associated, etc. This information is stored as a list object, accessible using the `metadata` function. The form of the data associated with the experiment is left to the discretion of the user.

The `SummarizedExperiment` container is appropriate for matrix-like data. The data are accessed using the `assays` function, described below. This returns a `SimpleList` object. Each element of the list must itself be a matrix (of any mode) and must have dimensions that are the same as the dimensions of the `SummarizedExperiment` in which they are stored. Row and column names of each matrix must either be NULL or match those of the `SummarizedExperiment` during construction. It is convenient for the elements of `SimpleList` of assays to be named.

Constructor

`SummarizedExperiment` instances are constructed using the `SummarizedExperiment` function documented in ?`RangedSummarizedExperiment`. 

Accessors

In the following code snippets, \( x \) is a SummarizedExperiment object.

\[ \text{assays}(x), \text{assays}(x) \leftarrow \text{value} \]: Get or set the assays. \( \text{value} \) is a list or SimpleList, each element of which is a matrix with the same dimensions as \( x \).

\[ \text{assay}(x, i), \text{assay}(x, i) \leftarrow \text{value} \]: A convenient alternative (to \( \text{assays}(x)[[i]], \text{assays}(x)[[i]] \leftarrow \text{value} \)) to get or set the \( i \)th (default first) assay element. \( \text{value} \) must be a matrix of the same dimension as \( x \), and with dimension names NULL or consistent with those of \( x \).

\[ \text{assayNames}(x), \text{assayNames}(x) \leftarrow \text{value} \]: Get or set the names of \( \text{assay}() \) elements.

\[ \text{rowData}(x, \text{use.names}=\text{TRUE}), \text{rowData}(x) \leftarrow \text{value} \]: Get or set the row data. \( \text{value} \) is a DataFrame object.

\[ \text{colData}(x), \text{colData}(x) \leftarrow \text{value} \]: Get or set the column data. \( \text{value} \) is a DataFrame object. Row names of \( \text{value} \) must be NULL or consistent with the existing column names of \( x \).

\[ \text{metadata}(x), \text{metadata}(x) \leftarrow \text{value} \]: Get or set the experiment data. \( \text{value} \) is a list with arbitrary content.

\[ \text{dim}(x) \]: Get the dimensions (features of interest x samples) of the SummarizedExperiment.

\[ \text{dimnames}(x), \text{dimnames}(x) \leftarrow \text{value} \]: Get or set the dimension names. \( \text{value} \) is usually a list of length 2, containing elements that are either NULL or vectors of appropriate length for the corresponding dimension. \( \text{value} \) can be NULL, which removes dimension names. This method implies that \( \text{rownames}, \text{rownames}<-, \text{colnames}, \text{colnames}<- \) are all available.

Subsetting

In the code snippets below, \( x \) is a SummarizedExperiment object.

\[ x[i,j], x[i,j] \leftarrow \text{value} \]: Create or replace a subset of \( x \). \( i, j \) can be numeric, logical, character, or missing. \( \text{value} \) must be a SummarizedExperiment object with dimensions, dimension names, and assay elements consistent with the subset \( x[i,j] \) being replaced.

\[ \text{subset}(x, \text{subset}, \text{select}) \]: Create a subset of \( x \) using an expression \( \text{subset} \) referring to columns of \( \text{rowData}(x) \) and / or \( \text{select} \) referring to column names of \( \text{colData}(x) \).

Additional subsetting accessors provide convenient access to \( \text{colData} \) columns

\[ x\$name, x\$name \leftarrow \text{value} \] Access or replace column name in \( x \).

\[ x[[i, \ldots]], x[[i, \ldots]] \leftarrow \text{value} \] Access or replace column \( i \) in \( x \).

Combining

In the code snippets below, \( \ldots \) are SummarizedExperiment objects to be combined.

\[ \text{cbind}(\ldots) \]: \( \text{cbind} \) combines objects with the same features of interest but different samples (columns in \( \text{assays} \)). The colnames in \( \text{colData} \) (SummarizedExperiment) must match or an error is thrown. Duplicate columns of \( \text{rowData} \) (SummarizedExperiment) must contain the same data. Data in \( \text{assays} \) are combined by name matching; if all assay names are NULL matching is by position. A mixture of names and NULL throws an error.

\( \text{metadata} \) from all objects are combined into a list with no name checking.
rbind(...): rbind combines objects with the same samples but different features of interest (rows in assays). The colnames in rowData(SummarizedExperiment) must match or an error is thrown. Duplicate columns of colData(SummarizedExperiment) must contain the same data.

Data in assays are combined by name matching; if all assay names are NULL matching is by position. A mixture of names and NULL throws an error.

metadata from all objects are combined into a list with no name checking.

Implementation and Extension

This section contains advanced material meant for package developers.

SummarizedExperiment is implemented as an S4 class, and can be extended in the usual way, using contains="SummarizedExperiment" in the new class definition.

In addition, the representation of the assays slot of SummarizedExperiment is as a virtual class Assays. This allows derived classes (contains="Assays") to implement alternative requirements for the assays, e.g., backed by file-based storage like NetCDF or the ff package, while re-using the existing SummarizedExperiment class without modification. See Assays for more information.

Author(s)

Martin Morgan, mttmorgan@fhcrc.org

See Also

- RangedSummarizedExperiment objects.
- DataFrame, SimpleList, and Annotated objects in the S4Vectors package.
- The metadata and mcols accessors in the S4Vectors package.
- saveHDF5SummarizedExperiment and loadHDF5SummarizedExperiment in the HDF5Array package for saving/loading an HDF5-based SummarizedExperiment object to/from disk.
- The realize generic function in the DelayedArray package for more information about on-disk realization of objects carrying delayed operations.

Examples

nrows <- 200; ncols <- 6
counts <- matrix(runif(nrows * ncols, 1, 1e4), nrows)
colData <- DataFrame(Treatment=rep(c("ChIP", "Input"), 3),
                      row.names=LETTERS[1:6])
se0 <- SummarizedExperiment(assays=SimpleList(counts=counts),
                            colData=colData)
se0
dim(se0)
dimnames(se0)
assayNames(se0)
head(assay(se0))
assays(se0) <- endoapply(assays(se0), asinh)
head(assay(se0))
rowData(se0)
colData(se0)

se0[, se0$Treatment == "ChIP"]
subset(se0, select = Treatment == "ChIP")
## cbind() combines objects with the same features of interest but different samples:
se1 <- se0
se2 <- se1[,1:3]
colnames(se2) <- letters[seq_len(ncol(se2))]
cmb1 <- cbind(se1, se2)
dim(cmb1)
dimnames(cmb1)

## rbind() combines objects with the same samples but different features of interest:
se1 <- se0
se2 <- se1[1:50,]
rownames(se2) <- letters[seq_len(nrow(se2))]
cmb2 <- rbind(se1, se2)
dim(cmb2)
dimnames(cmb2)

## ON-DISK REALIZATION
setRealizationBackend("HDF5Array")
cmb3 <- realize(cmb2)
assay(cmb3, withDimnames=FALSE) # an HDF5Matrix object
Index

* file
  readKallisto, 20

* manip
  makeSummarizedExperimentFromExpressionSet, 12

* methods
  coverage-methods, 4
  findOverlaps-methods, 5
  inter-range-methods, 7
  intra-range-methods, 8
  nearest-methods, 14

* utilities
  coverage-methods, 4
  findOverlaps-methods, 5
  inter-range-methods, 7
  intra-range-methods, 8
  nearest-methods, 14
  [, Assays, ANY, ANY, ANY-method (Assays-class), 2
  [, Assays, ANY-method (Assays-class), 2
  [, SummarizedExperiment, ANY, ANY, ANY-method (SummarizedExperiment-class), 22
  [, SummarizedExperiment, ANY-method (SummarizedExperiment-class), 22
  <<- SummarizedExperiment-method (SummarizedExperiment-class), 22
  acbind, Matrix-method (Assays-class), 2
  Annotated, 26
  arbind, Matrix-method (Assays-class), 2
  assay (SummarizedExperiment-class), 22
  assay, SummarizedExperiment, character-method (SummarizedExperiment-class), 22
  assay, SummarizedExperiment, missing-method (SummarizedExperiment-class), 22
  assay, SummarizedExperiment, numeric-method (SummarizedExperiment-class), 22
  assay<-, SummarizedExperiment-class), 22
  assay<-, SummarizedExperiment, character-method (SummarizedExperiment-class), 22
  assay<-, SummarizedExperiment, missing-method (SummarizedExperiment-class), 22
  assay<-, SummarizedExperiment, numeric-method (SummarizedExperiment-class), 22
  assayNames (SummarizedExperiment-class), 22
  assayNames, SummarizedExperiment-method (SummarizedExperiment-class), 22
  assayNames<-, SummarizedExperiment-character-method (SummarizedExperiment-class), 22
  Assays, 26
  Assays (Assays-class), 2
  assays (SummarizedExperiment-class), 22
  assays, SummarizedExperiment-method (SummarizedExperiment-class), 22

28
INDEX

`dimnames`, `SummarizedExperiment-method` (SummarizedExperiment-class), 22
`dimnames<-`, `RangedSummarizedExperiment,list-method` (RangedSummarizedExperiment-class), 16
`dimnames<-`, `SummarizedExperiment,list-method` (SummarizedExperiment-class), 22
`dimnames<-`, `SummarizedExperiment,NULL-method` (SummarizedExperiment-class), 22
`disjointBins` (inter-range-methods), 7
`disjointBins`, `RangedSummarizedExperiment-method` (inter-range-methods), 7
`distance` (nearest-methods), 14
`distance`, `ANY`, `RangedSummarizedExperiment-method` (nearest-methods), 14
`distance`, `RangedSummarizedExperiment`, `ANY-method` (nearest-methods), 14
`distance`, `RangedSummarizedExperiment`, `RangedSummarizedExperiment-method` (nearest-methods), 14
`duplicated`, 18
`duplicated`, `RangedSummarizedExperiment-method` (RangedSummarizedExperiment-class), 16
`elementMetadata`, `RangedSummarizedExperiment-method` (RangedSummarizedExperiment-class), 16
`elementMetadata<-`, `RangedSummarizedExperiment-method` (RangedSummarizedExperiment-class), 16
end, 18
end, `RangedSummarizedExperiment-method` (RangedSummarizedExperiment-class), 16
end<-, `RangedSummarizedExperiment-method` (RangedSummarizedExperiment-class), 16
`ExpressionSet`, 12, 13
`featureNames`, 12
`findOverlaps`, 6, 18, 19
`findOverlaps` (findOverlaps-methods), 5
`findOverlaps`, `RangedSummarizedExperiment`, `RangedSummarizedExperiment-method` (findOverlaps-methods), 5
`findOverlaps`, `RangedSummarizedExperiment`, `Vector-method` (findOverlaps-methods), 5
`findOverlaps`, `Vector`, `RangedSummarizedExperiment-method` (findOverlaps-methods), 5
`findOverlaps-methods`, 5
`flank` (intra-range-methods), 8
`flank`, `RangedSummarizedExperiment-method` (intra-range-methods), 8
follow (nearest-methods), 14
follow, `ANY`, `RangedSummarizedExperiment-method` (nearest-methods), 14
follow, `RangedSummarizedExperiment`, `ANY-method` (nearest-methods), 14
follow, `RangedSummarizedExperiment`, `RangedSummarizedExperiment-method` (nearest-methods), 14
`getGeneRangeMapper` (makeSummarizedExperimentFromExpressionSet), 12
`getGeneRangeMapper`, `RangedSummarizedExperiment-method` (makeSummarizedExperimentFromExpressionSet), 12
`GenomicRanges`, 5–7, 9, 15
`getListElement`, `Assays-method` (Assays-class), 2
`getListElement`, `AssaysInEnv-method` (Assays-class), 2
`GRanges`, 10–12, 16–19
`GRanges`, 18
`GRangesList`, 5–7, 9, 12, 15–18
`inter-range-methods`, 7
`intra-range-methods`, 8
is.unsorted, 18
is.unsorted, `RangedSummarizedExperiment-method` (RangedSummarizedExperiment-class), 16
isDisjoint, 7, 18, 19
isDisjoint (inter-range-methods), 7
isDisjoint, `RangedSummarizedExperiment-method` (inter-range-methods), 7
`KALLISTO_ASSAYS` (readKallisto), 20
length, `Assays-method` (Assays-class), 2
length, `AssaysInEnv-method` (Assays-class), 2
length, `SummarizedExperiment-method` (SummarizedExperiment-class), 22
loadHDF5SummarizedExperiment, 26
makeGRangesFromDataFrame, 10, 11
makeSummarizedExperimentFromDataFrame, 10
makeSummarizedExperimentFromExpressionSet, 12
makeSummarizedExperimentFromLoom, 13
match, 18
mcols, 18, 24, 26
mcols, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
mcols<-, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
metadata, 26
naiveRangeMapper
(probeRangeMapper
(probeRangeMapper
(probeRangeMapper
(name, Assays-method (Assays-class), 2
name, AssaysInEnv-method (Assays-class), 2
name, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
name, SummarizedExperiment-method (SummarizedExperiment-class), 22
name<-, Assays-method (Assays-class), 2
name<-, AssaysInEnv-method (Assays-class), 2
name<-, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
name<-, SummarizedExperiment-method (SummarizedExperiment-class), 22
narrow (intra-range-methods), 8
narrow, RangedSummarizedExperiment-method (intra-range-methods), 8
nearest, 15, 18, 19
nearest (nearest-methods), 14
nearest, ANY, RangedSummarizedExperiment-method (nearest-methods), 14
nearest, RangedSummarizedExperiment, ANY-method (nearest-methods), 14
nearest, RangedSummarizedExperiment, RangedSummarizedExperiment-method (nearest-methods), 14
nearest-methods, 14
order, 18
order, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
pcompare, 18
pcompare, ANY, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
pcompare, RangedSummarizedExperiment, ANY-method (RangedSummarizedExperiment-class), 16
pcompare, RangedSummarizedExperiment, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
precede (nearest-methods), 14
precede, ANY, RangedSummarizedExperiment-method (nearest-methods), 14
precede, RangedSummarizedExperiment, ANY-method (nearest-methods), 14
promoters (intra-range-methods), 8
promoters, RangedSummarizedExperiment-method (intra-range-methods), 8
RangedSummarizedExperiment, 4–9, 11–15, 22, 24, 26
RangedSummarizedExperiment
(ranges, 18
ranges, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
ranges<-, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
rank, 18
rank, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
rbind, Assays-method (Assays-class), 2
rbind, SummarizedExperiment-method (SummarizedExperiment-class), 22
readKallisto, 20
readKallistoBootstrap (readKallisto), 20
realize, 26
realize, SummarizedExperiment-method (SummarizedExperiment-class), 22
resize (intra-range-methods), 8
resize, RangedSummarizedExperiment-method (intra-range-methods), 8
restrict (intra-range-methods), 8
restrict, RangedSummarizedExperiment-method (intra-range-methods), 8
rowData (SummarizedExperiment-class), 22
rowData, SummarizedExperiment-method (SummarizedExperiment-class), 22
rowData <- (SummarizedExperiment-class), 22
rownames, 12
rowRanges (RangedSummarizedExperiment-class), 16
rowRanges, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
rowRanges, SummarizedExperiment-method (SummarizedExperiment-class), 16
rowRanges <- (RangedSummarizedExperiment-class), 16
rowRanges <-, RangedSummarizedExperiment, NULL-method (RangedSummarizedExperiment-class), 16
rowRanges <-, SummarizedExperiment, GenomicRanges-method (RangedSummarizedExperiment-class), 16
rowRanges <-, SummarizedExperiment, GRangesList-method (RangedSummarizedExperiment-class), 16
rowRanges <-, SummarizedExperiment, NULL-method (RangedSummarizedExperiment-class), 16
saveHDF5SummarizedExperiment, 26
Seqinfo, 11
seqinfo, 18
seqinfo, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
seqinfo <-, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
seqnames, 18
seqnames, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
setListElement, Assays-method (Assays-class), 2
setListElement, AssaysInEnv-method (Assays-class), 2
ShallowData (Assays-class), 2
ShallowData-class (Assays-class), 2
ShallowSimpleListAssays (Assays-class), 2
ShallowSimpleListAssays-class (Assays-class), 2
shift, 9, 18, 19
shift (intra-range-methods), 8
shift, RangedSummarizedExperiment-method (intra-range-methods), 8
show, SummarizedExperiment-method (SummarizedExperiment-class), 22
SimpleAssays (Assays-class), 2
SimpleAssays-class (Assays-class), 2
SimpleList, 2, 3, 24, 26
sort, 18
sort, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
sparseMatrix, 17
split, RangedSummarizedExperiment, ANY-method (RangedSummarizedExperiment-class), 16
split, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
start, 18
start, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
start <-, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
strand, 18
strand, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
strand <-, RangedSummarizedExperiment, ANY-method (RangedSummarizedExperiment-class), 16
subset, RangedSummarizedExperiment-method (RangedSummarizedExperiment-class), 16
subset, SummarizedExperiment-method (SummarizedExperiment-class), 22
```plaintext
Index

SummarizedExperiment, 2, 3, 10, 14, 16, 17
SummarizedExperiment
(RangedSummarizedExperiment-class), 16
SummarizedExperiment-class, 19, 22

trim, RangedSummarizedExperiment-method
(intra-range-methods), 8

TxDb, 13

updateObject, Assays-method
(Assays-class), 2
updateObject, RangedSummarizedExperiment-method
(RangedSummarizedExperiment-class), 16
updateObject, SummarizedExperiment-method
(SummarizedExperiment-class), 22

width, 18
width, RangedSummarizedExperiment-method
(RangedSummarizedExperiment-class), 16

width<-, RangedSummarizedExperiment-method
(RangedSummarizedExperiment-class), 16
```