Package 'yriMulti'

April 14, 2020

Title support for expression, methylation, DHS, VCF for YRI

Version 0.16.0

Author VJ Carey <stvjc@channing.harvard.edu>

Description expression, methylation, DHS for YRI

Suggests erma, BiocStyle, knitr, rmarkdown, gQTLstats (>= 1.9.2),
doParallel, geuvPack, knitcitations, bibtex

Imports GenomicFiles (>= 1.13.6), VariantAnnotation (>= 1.23.1),
gQTLBase, SummarizedExperiment, GenomicRanges, dsQTL,
GenomeInfoDb

Depends Homo.sapiens, geuvPack, MultiAssayExperiment (>= 1.3.14)

Maintainer VJ Carey <stvjc@channing.harvard.edu>

License Artistic-2.0

LazyLoad yes

VignetteBuilder knitr

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

git_branch RELEASE_3_10

git_last_commit 2505f95

git_last_commit_date 2019-10-29

Date/Publication 2020-04-14

R topics documented:

---

yriMulti-package   *support for expression, methylation, DHS, VCF for YRI*

Description

expression, methylation, DHS for YRI

Details

The DESCRIPTION file: This package was not yet installed at build time.

Index: This package was not yet installed at build time.

This is a relatively experimental package that is used to exercise aspects of MultiAssayExperiment and gQTLstats. The primary innovations are illustration of VcfStack class of GenomicFiles as a distributed on-disk (or in-cloud) resource for assay elements representing genotypes, and adoption of MultiAssayExperiment to unite relatively uncommonly integrated assays such as DnaseI hypersensitivity and 450k methylation.

Author(s)

VJ Carey <stvjc@channing.harvard.edu>

Maintainer: VJ Carey <stvjc@channing.harvard.edu>

Examples

```
# see the vignette
```

---

banovichSE   *Illumina 450k methlyation assay applied to Yoruba cell lines*

Description

Illumina 450k methlyation assay applied to Yoruba cell lines

Usage

```
data("banovichSE")
```

Format

The format is:
Formal class 'RangedSummarizedExperiment' [package "SummarizedExperiment"] with 6 slots
..@ rowRanges :Formal class 'GRanges' [package "GenomicRanges"] with 6 slots
.. .. ..@ seqnames :Formal class 'Rle' [package "S4Vectors"] with 4 slots
.. ..@ values : Factor w/ 24 levels "chr1","chr2",..: 16 1 8 14 1 15 19 3 12 15 ...
.. ..@ lengths : int [1:301401] 1 1 1 1 1 1 1 1 1 1 ...
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()

.. .. ..@ ranges :Formal class 'IRanges' [package "IRanges"] with 6 slots
.. ..@ start : int [1:329469] 53468112 91194674 42263294 69341139 230560793 23034447 54695678 128902377 124086477 59785306 ...
.. ..@ width : int [1:329469] 2 2 2 2 2 2 2 2 2 2 ...
.. ..@ NAMES : chr [1:329469] "cg00000029" "cg00000165" "cg00000236" "cg00000289" ...
.. ..@ elementType : chr "integer"
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()
.. .. ..@ strand :Formal class 'Rle' [package "S4Vectors"] with 4 slots
.. ..@ values : Factor w/ 3 levels "+","-","*": 3
.. ..@ lengths : int 329469
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()
.. .. ..@ elementMetadata:Formal class 'DataFrame' [package "S4Vectors"] with 6 slots
.. ..@ rownames : NULL
.. ..@ nrows : int 329469
.. ..@ listData :List of 10
.. ..$ addressA : chr [1:329469] "14782418" "12637463" "12649348" "18766346" ...
.. ..$ addressB : chr [1:329469] "" "" "" "" ...
.. ..$ channel :Formal class 'Rle' [package "S4Vectors"] with 4 slots
.. ..@ values : Factor w/ 3 levels "Both","Grn","Red": 1 3 1 2 1 2 1 3 1 2 ...
.. ..@ lengths : int [1:140453] 5 1 2 1 3 1 1 1 2 1 ...
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()
.. ..$ platform :Formal class 'Rle' [package "S4Vectors"] with 4 slots
.. ..@ values : Factor w/ 2 levels "BOTH","HM450": 2 1 2 1 2 1 2 1 2 1 ...
.. ..@ lengths : int [1:35429] 38 1 31 1 41 1 19 1 8 1 ...
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()
.. ..$ percentGC : num [1:329469] 0.46 0.48 0.5 0.44 0.42 0.8 0.6 0.58 0.78 0.54 ...
.. ..$ sourceSeq :Formal class 'DNAStringSet' [package "Biostrings"] with 5 slots
.. ..@ pool :Formal class 'SharedRaw_Pool' [package "XVector"] with 2 slots
.. ..@ xp_list :List of 1
.. ..$:<externalptr>
.. ..@ .link_to_cached_object_list:List of 1
.. ..$:<environment: 0x7f87a4476990>
.. ..@ ranges :Formal class 'GroupedIRanges' [package "XVector"] with 7 slots
.. ..@ group : int [1:329469] 1 1 1 1 1 1 1 1 1 1 ...
.. ..@ start : int [1:329469] 1 151 201 251 401 451 551 651 701 851 ...
.. ..@ width : int [1:329469] 50 50 50 50 50 50 50 50 50 50 ...
.. ..@ NAMES : NULL
.. ..@ elementType : chr "integer"
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()
.. ..@ elementType : chr "DNAString"
.. ..@ elementMetadata: NULL
.. ..@ metadata : list()
.. ..$ probeType :Formal class 'Rle' [package "S4Vectors"] with 4 slots
.. ..@ values : Factor w/ 3 levels "cg","ch","rs": 1 2
.. ..@ lengths : int [1:2] 327286 2183
.. ..@ elementMetadata: NULL

```
.. .. .. .. .. .. .. ..@ metadata : list()
.. .. .. .. .. ..$ probeStart : chr [1:329469] "53468112" "91194626" "42263246" "69341139" ...
.. .. .. .. .. ..$ probeEnd : chr [1:329469] "53468161" "91194675" "42263295" "69341188" ...
.. .. .. .. .. ..$ probeTarget: num [1:329469] 5.35e+07 9.12e+07 4.23e+07 6.93e+07 2.31e+08 ...
.. .. .. .. .. ..@ elementType : chr "ANY"
.. .. .. .. .. ..@ elementMetadata: NULL
.. .. .. .. .. ..@ metadata : list()
.. .. ..@ seqinfo :Formal class 'Seqinfo' [package "GenomeInfoDb"] with 4 slots
.. .. .. .. .. ..@ seqnames : chr [1:24] "chr1" "chr2" "chr3" "chr4" ...
.. .. .. .. .. ..@ seqlengths : int [1:24] 249250621 243199373 198022430 191154276 180915260
171115067 159138663 146364022 141213431 135534747 ...
.. .. .. .. .. ..@ is_circular: logi [1:24] FALSE FALSE FALSE FALSE FALSE FALSE ...
.. .. .. .. .. ..@ genome : chr [1:24] "hg19" "hg19" "hg19" "hg19" ...
.. .. .. ..@ metadata :List of 1
.. .. .. ..$ :Formal class 'DataFrame' [package "S4Vectors"] with 6 slots
.. .. .. .. .. ..@ rownames : NULL
.. .. .. .. .. ..@ nrows : int 13
.. .. .. .. .. ..@ listData :List of 2
.. .. .. .. .. .. ..$ name : chr [1:13] "Db type" "Supporting package" "data_nrow" "Db created by" ...
.. .. .. .. .. .. ..$ value: chr [1:13] "FeatureDb" "GenomicFeatures" "487173" "GenomicFeatures
package from Bioconductor" ...
.. .. .. .. .. ..@ elementType : chr "ANY"
.. .. .. .. .. ..@ elementMetadata: NULL
.. .. .. .. .. ..@ metadata : list()
..@ colData :Formal class 'DataFrame' [package "S4Vectors"] with 6 slots
.. .. ..@ rownames : chr [1:64] "NA18498" "NA18499" "NA18501" "NA18502" ...
.. .. ..@ nrows : int 64
.. .. ..@ listData :List of 35
.. .. ..$ title : Factor w/ 64 levels "GM18489","GM18498",..: 2 3 4 5 11 12 13 14 19 20 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ geo_accession : Factor w/ 64 levels "GSM1383567","GSM1383568",..: 1 2 3 4 5 6 7 8 9
10 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ status : Factor w/ 1 level "Public on May 09 2014": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ submission_date : Factor w/ 1 level "May 08 2014": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ last_update_date : Factor w/ 1 level "Aug 12 2014": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ type : Factor w/ 1 level "genomic": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ channel_count : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ source_name_ch1 : Factor w/ 1 level "HapMap_Yoruba": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ organism_ch1 : Factor w/ 1 level "Homo sapiens": 1 1 1 1 1 1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ characteristics_ch1 : Factor w/ 2 levels "gender: Female",..: 2 1 2 1 2 1 2 1 1 2 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. .. ..$ characteristics_ch1.1 : Factor w/ 1 level "cell type: Lymphoblastoid Cell Line": 1 1 1 1 1
1 1 1 1 1 ...
.. .. .. .. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
```

.. ..$ molecule_ch1 : Factor w/ 1 level "genomic DNA": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ extract_protocol_ch1 : Factor w/ 1 level "DNA was extracted from lymphoblastoid Cell Line": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ label_ch1 : Factor w/ 1 level "cy3 and cy5": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ label_protocol_ch1 : Factor w/ 1 level "Standard illumina protocols": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ taxid_ch1 : Factor w/ 1 level "9606": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ hyb_protocol : Factor w/ 1 level "The samples were bisulphite-converted and hybridized to the Infinium HumanMethylation450 BeadChip at the University of Chicago "| __truncated__: 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ scan_protocol : Factor w/ 1 level "not provided": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ description : Factor w/ 64 levels "Sample 1","Sample 10",..: 1 12 23 34 45 56 62 63 64 2 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ data_processing : Factor w/ 1 level "To ensure high data quality, probes were mapped to a bisulfite converted genome and only uniquely mapped probes were retained. "| __truncated__: 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ data_processing.1 : Factor w/ 1 level "Normalized data: Normalized average beta (PCs removed).": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ data_processing.2 : Factor w/ 1 level "Un-normalized data: Unmethylated and methylated signal intensities and detection p-value.": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ platform_id : Factor w/ 1 level "GPL13534": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_name : Factor w/ 1 level "Nicholas,E,Banovich": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_laboratory : Factor w/ 1 level "Gilad": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_department : Factor w/ 1 level "Human Genetics": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_institute : Factor w/ 1 level "University of Chicago": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_address : Factor w/ 1 level "920 E. 58th Street, CLSC 317": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_city : Factor w/ 1 level "Chicago": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_state : Factor w/ 1 level "IL": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_zip.postal_code: Factor w/ 1 level "60453": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ contact_country : Factor w/ 1 level "USA": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ supplementary_file : Factor w/ 1 level "NONE": 1 1 1 1 1 1 1 1 1 1 ...

.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...

.. ..$ data_row_count : Factor w/ 1 level "329469": 1 1 1 1 1 1 1 1 1 1 ...
.. ..- attr(*, "names")= chr [1:64] "V2" "V3" "V4" "V5" ...
.. ..$ naid : chr [1:64] "NA18498" "NA18499" "NA18501" "NA18502" ...
.. .. ..@ elementType : chr "ANY"
.. .. ..@ elementMetadata: NULL
.. .. ..@ metadata : list()
..@ assays :Reference class 'ShallowSimpleListAssays' [package "GenomicRanges"] with 1 field
.. ..$ data: NULL
.. ..and 14 methods.
..@ NAMES : NULL
..@ elementMetadata:Formal class 'DataFrame' [package "S4Vectors"] with 6 slots
.. .. ..@ rownames : NULL
.. .. ..@ nrows : int 329469
.. .. ..@ listData : Named list()
.. .. ..@ elementType : chr "ANY"
.. .. ..@ elementMetadata: NULL
.. .. ..@ metadata : list()
..@ metadata : list()

Examples

```
data(banovichSE)
```

---

mexGR   *combine expression data for a gene with nearby methylation data*

Description

combine expression data for a gene with nearby methylation data,

Usage

```
mexGR(methSE, exprSE, gradius = 0, symbol = "ORMDL3", etagname = "gene_name")
```

Arguments

methSE	SummarizedExperiment instance
exprSE	SummarizedExperiment instance
gradius	number of base pairs around 'gene' to search
symbol	string selecting a gene
etagname	metadata column in which symbol can be found to select gene

Value

GRanges instance with assay results in mcols

Examples

```
data(geuFPKM)
data(banovichSE)
m1 = mexGR(banovichSE, geuFPKM, gradius=5000)
ii = bindelms(geuFPKM, banovichSE)
```

---

pwplot   *pairwise plot for components of MultiAssayExperiment*

Usage

```
pwplot(fmla1, fmla2, mae, ytx = force, xtx = force, ...)
```

Arguments

fmla1   a formula with one name in each of lhs and rhs. these are the names of the
            experiments to be extracted

fmla2   a formula with lhs the name of a feature in rowData(lhs(fmla1)) and rhs the
            name of a feature in rowData(rhs(fmla1)). these are the features to be plotted

mae   instance of MultiAssayExperiment

ytx   function to transform lhs(fmla2)

xtx   function to transform rhs(fmla2)

...   passed to plot

Value

runs plot

Author(s)

VJ Carey <stvjc@channing.harvard.edu>

Examples

```
data(demoMAE)
pwplot(geuvRNAseq~yri450k, ENSG00000139618.9~cg20073910,
    demoMAE, ytx=log, main="demo")
```

Index