

Package ‘GeomxTools’

October 14, 2021

Title NanoString GeoMx Tools

Description Tools for NanoString Technologies GeoMx Technology. Package provides functions for reading in DCC and PKC files based on an ExpressionSet derived object. Normalization and QC functions are also included.

Version 1.0.0

Encoding UTF-8

Depends R (>= 3.6), NanoStringNCTools

Imports Biobase, S4Vectors, rjson, readxl, EnvStats, reshape2, methods, utils, stats, BiocGenerics

Suggests knitr

License Artistic-2.0

Collate DccMetadata.R NanoStringGeomxSet-class.R
NanoStringGeomxSet-validity.R NanoStringGeomxSet-accessors.R
NanoStringGeomxSet-signatures.R readDccFile.R readPKCFile.R
readNanoStringGeomxSet.R utils.R

biocViews GeneExpression, Transcription, CellBasedAssays, DataImport, Transcriptomics, Proteomics, mRNAMicroarray, ProprietaryPlatforms, RNASeq, Sequencing, ExperimentalDesign, Normalization

VignetteEngine knitr::rmarkdown

VignetteBuilder knitr

RoxygenNote 7.1.1

git_url <https://git.bioconductor.org/packages/GeomxTools>

git_branch RELEASE_3_13

git_last_commit 2462584

git_last_commit_date 2021-05-19

Date/Publication 2021-10-14

Author Nicole Ortogero [cre, aut],
Zhi Yang [aut]

Maintainer Nicole Ortogero <nortogero@nanosttring.com>

R topics documented:

NanoStringGeomxSet-class	2
readDccFile	4
readNanoStringGeomxSet	5
readPKCFile	6

Index	8
--------------	----------

NanoStringGeomxSet-class

Class to Contain NanoString Spatial Expression Level Assays

Description

The NanoStringGeomxSet class extends the [ExpressionSet](#) class for NanoString GeoMx Digital Count Conversion (DCC) data.

Usage

```
NanoStringGeomxSet(assayData,
  phenoData=Biobase::annotatedDataFrameFrom(assayData, byrow=FALSE),
  featureData=Biobase::annotatedDataFrameFrom(assayData, byrow=TRUE),
  experimentData=Biobase::MIAME(),
  annotation=character(),
  protocolData=Biobase::annotatedDataFrameFrom(assayData, byrow=FALSE),
  dimLabels=c("GeneName", "SampleID"),
  signatures=SignatureSet(),
  design=NULL,
  ...)
```

Arguments

assayData	A matrix or environment containing the DCCs.
phenoData	An AnnotatedDataFrame containing the phenotypic data of areas of interest.
featureData	An AnnotatedDataFrame containing gene information; gene name, accession number, functional groups, etc.
experimentData	An optional MIAME instance with meta-data about the experiment.
annotation	A character string for the PKC file(s).
protocolData	An AnnotatedDataFrame containing meta-data about the protocol and sequencing; columns could include "FileVersion", "SoftwareVersion", "Date", "Plate_ID", "Well", "SeqSetId", "trimGaloreOpts", "flash20pts", "umiExtractOpts", "boxtie20pts", "Raw", "Trimmed", "Stitched", "Aligned", "umiQ30", "rtsQ30".
dimLabels	A character vector of length 2 that provides the column names to use as labels for the features and samples respectively in the autoplot method.
signatures	An optional SignatureSet object containing signature definitions.

design An optional one-sided formula representing the experimental design based on columns from [phenoData](#)

... Additional arguments for [ExpressionSet](#).

Value

An S4 class containing data from a NanoString GeoMx experiment

Accessing

In addition to the standard [ExpressionSet](#) accessor methods, NanoStringGeomxSet objects have the following:

`sData(object)`: extracts the data.frame containing the sample data, `cbind(pData(object), pData(protocolData(object)))`

`svarLabels(object)`: extracts the sample data column names, `c(varLabels(object), varLabels(protocolData(object)))`

`dimLabels(object)`: extracts the column names to use as labels for the features and samples.

`dimLabels(object) <-value`: replaces the dimLabels of the object.

`signatures(object)`: extracts the [SignatureSet](#) of the object.

`signatures(object) <-value`: replaces the [SignatureSet](#) of the object.

`signatureScores(object, elt="exprs")`: extracts the matrix of computed signature scores.

`design(object)`: extracts the one-sided formula representing the experimental design based on columns from [phenoData](#).

`design(object) <-value`: replaces the one-sided formula representing the experimental design based on columns from [phenoData](#).

`signatureGroups(object)`: extract the groups of [SignatureSet](#).

`signatureGroups(object) <-value`: replaces the groups of [SignatureSet](#).

Author(s)

Zhi Yang & Nicole Ortogero

See Also

[readNanoStringGeomxSet](#), [ExpressionSet](#)

Examples

```
# Create NanoStringGeomxSet from data files
datadir <- system.file("extdata", "DSP_NGS_Example_Data",
                      package="GeomxTools")
dccFiles <- dir(datadir, pattern=".dcc$", full.names=TRUE)
pkc <- unzip(zipfile = file.path(datadir, "/pkcs.zip"))
sampleAnnotationFile <- file.path(datadir, "annotations.xlsx")

dccFileColumn <- "Sample_ID"

dccSet <- readNanoStringGeomxSet(dccFiles=dccFiles,
```

```

pkcFiles=pkc,
phenoDataFile=sampleAnnotationFile,
phenoDataSheet="CW005",
phenoDataDccColName=dccFileColumn,
protocolDataColNames=c("aoi", "cell_line",
                        "roi_rep", "pool_rep",
                        "slide_rep"),
experimentDataColNames="panel",
phenoDataColPrefix="")

# Accessing sample data and column names
head(sData(dccSet))
svarLabels(dccSet)

# Accessing number of samples and features
dim(dccSet)

```

readDccFile

Read DCC File

Description

Read a NanoString GeoMx Digital Count Conversion (DCC) file.

Usage

```
readDccFile(file)
```

Arguments

`file` A character string containing the path to the DCC file.

Value

A list object with two elements:

"Header" a `data.frame` object containing the protocol and sequencing information.

"Code_Summary" a `data.frame` object containing the gene probe counts.

Author(s)

Zhi Yang & Nicole Ortogero

See Also

[readNanoStringGeomxSet](#)

Examples

```
datadir <- system.file("extdata", "DSP_NGS_Example_Data",
                      package="GeomxTools")
dccFiles <- dir(datadir, pattern=".dcc$", full.names=TRUE)
dccData <- sapply(dccFiles, readDccFile, simplify = FALSE)
```

```
readNanoStringGeomxSet
```

Read 'NanoStringGeomxSet'

Description

Create an instance of class `NanoStringGeomxSet` by reading data from NanoString GeoMx Digital Count Conversion (DCC) data.

Usage

```
readNanoStringGeomxSet(dccFiles, pkcFiles, phenoDataFile,
                      phenoDataSheet, phenoDataDccColName = "Sample_ID",
                      phenoDataColPrefix = "", protocolDataColNames = c("slide name"),
                      experimentDataColNames = c("panel"))
```

Arguments

<code>dccFiles</code>	A character vector containing the paths to the DCC files.
<code>pkcFiles</code>	An optional character string representing the path to the corresponding PKC file.
<code>phenoDataFile</code>	An optional character string representing the path to the corresponding phenotypic excel data file.
<code>phenoDataSheet</code>	An optional character string representing the excel sheet name containing the phenotypic data.
<code>phenoDataDccColName</code>	Character string identifying unique sample identifier column in <code>phenoDataFile</code> .
<code>phenoDataColPrefix</code>	An optional prefix to add to the <code>phenoData</code> column names to distinguish them from the names of <code>assayData</code> matrices, <code>featureData</code> columns, and <code>protocolData</code> columns.
<code>protocolDataColNames</code>	Character list of column names from <code>phenoDataFile</code> containing data about the experimental protocol or sequencing data.
<code>experimentDataColNames</code>	Character list of column names from <code>phenoDataFile</code> containing data about the experiment's meta-data.

Value

An instance of the `NanoStringGeomxSet` class.

Author(s)

Zhi Yang & Nicole Ortogero

See Also[NanoStringGeomxSet](#)**Examples**

```

# Data file paths
datadir <- system.file("extdata", "DSP_NGS_Example_Data",
                      package="GeomxTools")
dccFiles <- dir(datadir, pattern=".dcc$", full.names=TRUE)
pkc <- unzip(zipfile = file.path(datadir, "/pkcs.zip"))
sampleAnnotationFile <- file.path(datadir, "annotations.xlsx")

dccFileColumn <- "Sample_ID"

dccSet <- readNanoStringGeomxSet(dccFiles=dccFiles,
                                pkcFiles=pkc,
                                phenoDataFile=sampleAnnotationFile,
                                phenoDataSheet="CW005",
                                phenoDataDccColName=dccFileColumn,
                                protocolDataColNames=c("aoi", "cell_line",
                                                       "roi_rep", "pool_rep",
                                                       "slide_rep"),
                                experimentDataColNames="panel",
                                phenoDataColPrefix="")

# All data
dccSet <- readNanoStringGeomxSet(dccFiles, pkcFile = pkc,
                                phenoDataFile = sampleAnnotationFile,
                                phenoDataSheet="CW005")

varLabels(dccSet)

# All data with phenoData prefix
dccSetPhenoPrefix <- readNanoStringGeomxSet(dccFiles,
                                            pkcFile = pkc,
                                            phenoDataFile = sampleAnnotationFile,
                                            phenoDataSheet="CW005",
                                            phenoDataColPrefix = "PHENO_")

varLabels(dccSetPhenoPrefix)

```

`readPKCFile`*Read PKC File*

Description

Read a NanoString Probe Kit Configuration (PKC) file.

Usage

```
readPKCFile(file)
```

Arguments

file A character string containing the path to the PKC file.

Value

An instance of the [DataFrame](#) class containing columns:

"RTS_ID"	unique probe ID
"Gene"	gene name
"Module"	PKC name
"Negative"	negative probe
...	additional columns

Author(s)

Zhi Yang & Nicole Ortogero

See Also

[readNanoStringGeomxSet](#)

Examples

```
datadir <- system.file("extdata", "DSP_NGS_Example_Data",  
                      package="GeomxTools")  
pkc <- unzip(zipfile = file.path(datadir, "/pkcs.zip"))  
PKCData <- readPKCFile(pkc)
```

Index

- * **NanoStringGeomxSet**
 - readNanoStringGeomxSet, 5
- * **classes**
 - NanoStringGeomxSet-class, 2
- * **file**
 - readDccFile, 4
 - readNanoStringGeomxSet, 5
 - readPKCFile, 6
- * **manip**
 - readDccFile, 4
 - readNanoStringGeomxSet, 5
 - readPKCFile, 6
- * **methods**
 - NanoStringGeomxSet-class, 2

AnnotatedDataFrame, 2

class:NanoStringGeomxSet
(NanoStringGeomxSet-class), 2

coerce, ExpressionSet, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

DataFrame, 7

design, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

design<-, NanoStringGeomxSet, ANY-method
(NanoStringGeomxSet-class), 2

design<-, NanoStringGeomxSet, formula-method
(NanoStringGeomxSet-class), 2

design<-, NanoStringGeomxSet, NULL-method
(NanoStringGeomxSet-class), 2

dimLabels, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

dimLabels<-, NanoStringGeomxSet, character-method
(NanoStringGeomxSet-class), 2

ExpressionSet, 2, 3

MIAME, 2

NanoStringGeomxSet, 5, 6

NanoStringGeomxSet
(NanoStringGeomxSet-class), 2

NanoStringGeomxSet, environment-method
(NanoStringGeomxSet-class), 2

NanoStringGeomxSet, ExpressionSet-method
(NanoStringGeomxSet-class), 2

NanoStringGeomxSet, matrix-method
(NanoStringGeomxSet-class), 2

NanoStringGeomxSet, missing-method
(NanoStringGeomxSet-class), 2

NanoStringGeomxSet, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

NanoStringGeomxSet-class, 2

phenoData, 3

readDccFile, 4

readNanoStringGeomxSet, 3, 4, 5, 7

readPKCFile, 6

sData (NanoStringGeomxSet-class), 2

sData, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

show, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

signatureGroups
(NanoStringGeomxSet-class), 2

signatureGroups, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

signatures (NanoStringGeomxSet-class), 2

signatures, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

signatures<-
(NanoStringGeomxSet-class), 2

signatures<- , NanoStringGeomxSet, SignatureSet-method
(NanoStringGeomxSet-class), 2

signatureScores
(NanoStringGeomxSet-class), 2

signatureScores, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), 2

SignatureSet, [2](#), [3](#)
svarLabels (NanoStringGeomxSet-class), [2](#)
svarLabels, NanoStringGeomxSet-method
(NanoStringGeomxSet-class), [2](#)