

# Package ‘RTCGAToolbox’

January 26, 2026

**Type** Package

**Title** A new tool for exporting TCGA Firehose data

**Version** 2.41.0

**Description** Managing data from large scale projects such as The Cancer Genome Atlas (TCGA) for further analysis is an important and time consuming step for research projects. Several efforts, such as Firehose project, make TCGA pre-processed data publicly available via web services and data portals but it requires managing, downloading and preparing the data for following steps. We developed an open source and extensible R based data client for Firehose pre-processed data and demonstrated its use with sample case studies. Results showed that RTCGAToolbox could improve data management for researchers who are interested with TCGA data. In addition, it can be integrated with other analysis pipelines for following data analysis.

**License** GPL-2

**Depends** R (>= 4.3.0)

**Imports** BiocGenerics, data.table, DelayedArray, GenomicRanges, Seqinfo, httr, methods, RaggedExperiment, RCurl, RJSONIO, rvest, S4Vectors, stats, stringr, SummarizedExperiment, TCGAutils, utils

**Suggests** BiocStyle, Homo.sapiens, knitr, readr, rmarkdown

**biocViews** DifferentialExpression, GeneExpression, Sequencing

**URL** <http://mksamur.github.io/RTCGAToolbox/>

**BugReports** <https://github.com/mksamur/RTCGAToolbox/issues>

**VignetteBuilder** knitr

**RoxygenNote** 7.3.2

**Roxygen** list(markdown = TRUE)

**Collate** 'RTCGAToolbox-Class.R' 'RTCGAToolbox.R' 'utils.R'  
'biocExtract.R' 'data.R' 'getBroadSubtypes.R'  
'getFirehoseAnalyzeDates.R' 'getFirehoseData.R'  
'getFirehoseDatasets.R' 'getFirehoseRunningDates.R'  
'getGISTICPeaks.R' 'getLinks.R' 'getMutationRate.R'  
'makeSummarizedExperimentFromGISTIC.R' 'selectType.R'

**Encoding** UTF-8  
**Date** 2024-12-18  
**git\_url** <https://git.bioconductor.org/packages/RTCGAToolbox>  
**git\_branch** devel  
**git\_last\_commit** d55f300  
**git\_last\_commit\_date** 2025-10-29  
**Repository** Bioconductor 3.23  
**Date/Publication** 2026-01-25  
**Author** Mehmet Samur [aut],  
Marcel Ramos [aut, cre] (ORCID:  
<<https://orcid.org/0000-0002-3242-0582>>),  
Ludwig Geistlinger [ctb]  
**Maintainer** Marcel Ramos <[marcel.ramos@sph.cuny.edu](mailto:marcel.ramos@sph.cuny.edu)>

Contents

|                                              |    |
|----------------------------------------------|----|
| accmini . . . . .                            | 3  |
| biocExtract . . . . .                        | 3  |
| CorResult-class . . . . .                    | 5  |
| DGEResult-class . . . . .                    | 5  |
| FirehoseCGHArray-class . . . . .             | 5  |
| FirehoseData-class . . . . .                 | 6  |
| FirehoseGISTIC-class . . . . .               | 7  |
| FirehoseMethylationArray-class . . . . .     | 8  |
| FirehosemRNAArray-class . . . . .            | 8  |
| getBroadSubtypes . . . . .                   | 9  |
| getData . . . . .                            | 9  |
| getFirehoseAnalyzeDates . . . . .            | 10 |
| getFirehoseData . . . . .                    | 11 |
| getFirehoseDatasets . . . . .                | 13 |
| getFirehoseRunningDates . . . . .            | 14 |
| getGISTICPeaks . . . . .                     | 14 |
| getLinks . . . . .                           | 15 |
| getMutationRate . . . . .                    | 16 |
| hg19.ucsc.gene.locations . . . . .           | 17 |
| makeSummarizedExperimentFromGISTIC . . . . . | 18 |
| RTCGAToolbox . . . . .                       | 19 |
| selectType . . . . .                         | 19 |
| showResults . . . . .                        | 20 |
| showResults,CorResult-method . . . . .       | 21 |
| showResults,DGEResult-method . . . . .       | 21 |

---

|         |                                                               |
|---------|---------------------------------------------------------------|
| accmini | <i>A subset of the Adrenocortical Carcinoma (ACC) dataset</i> |
|---------|---------------------------------------------------------------|

---

### Description

See the `acc_sample.R` script to see how the data was generated. This dataset contains real data from the The Cancer Genome Atlas for the pipeline run date and GISTIC analysis date of 2016-01-28.

### Usage

```
data("accmini", package = "RTCGAToolbox")
```

### Format

A `FirehoseData` data object

---

|             |                                                                                                  |
|-------------|--------------------------------------------------------------------------------------------------|
| biocExtract | <i>Extract and convert data from a <code>FirehoseData</code> object to a Bioconductor object</i> |
|-------------|--------------------------------------------------------------------------------------------------|

---

### Description

This function processes data from a [FirehoseData](#) object. Raw data is converted to a conventional Bioconductor object. The function returns either [SummarizedExperiment](#) or [RaggedExperiment](#). In cases where there are multiple platforms in a data type, an attempt to consolidate datasets will be made based on matching dimension names. For ranged data, this functionality is provided with more control as part of the `RaggedExperiment` features. See the [RaggedExperiment-class](#) for more details.

### Usage

```
biocExtract(
  object,
  type = c("clinical", "RNASeqGene", "RNASeq2Gene", "miRNASeqGene", "RNASeq2GeneNorm",
    "CNASNP", "CNVSNP", "CNASeq", "CNACGH", "Methylation", "Mutation", "mRNAArray",
    "miRNAArray", "RPPAArray", "GISTIC", "GISTICA", "GISTICT", "GISTICP"),
  ...
)
```

### Arguments

|                     |                                                                                                                                                                                                                              |
|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>object</code> | A <code>FirehoseData</code> object from which to extract data.                                                                                                                                                               |
| <code>type</code>   | The type of data to extract from the "FirehoseData" object, see type section.                                                                                                                                                |
| <code>...</code>    | Additional arguments passed to lower level functions that convert tabular data into Bioconductor object such as <code>.makeRangedSummarizedExperimentFromDataFrame</code> or <code>.makeRaggedExperimentFromDataFrame</code> |

## Details

A typical additional argument for this function passed down to lower level functions is the `names.field` which indicates the row names in the data. By default, it is the "Hugo\_Symbol" column in the internal code that converts `data.frames` to `SummarizedExperiment` representations (via the `.makeSummarizedExperimentFromInternal` internal function).

## Value

Either [SummarizedExperiment](#) or [RaggedExperiment](#).

## type

Choices include the following:

- `clinical`: Get the clinical data slot
- `RNASeqGene`: `RNASeqGene`, `RNASeq v1`
- `RNASeqGene`: `RNASeq2Gene`, `RNASeq v2`
- `RNASeq2GeneNorm`: `RNASeq v2 Normalized`
- `miRNASeqGene`: `micro RNA SeqGene`
- `CNASNP`: `Copy Number Alteration`
- `CNVSNP`: `Copy Number Variation`
- `CNASeq`: `Copy Number Alteration`
- `CNACGH`: `Copy Number Alteration`
- `Methylation`: `Methylation`
- `mRNAArray`: `Messenger RNA`
- `miRNAArray`: `micro RNA`
- `RPPAArray`: `Reverse Phase Protein Array`
- `Mutation`: `Mutations`
- `GISTICA`: `GISTIC v2 ('AllByGene' only)`
- `GISTICT`: `GISTIC v2 ('ThresholdedByGene' only)`
- `GISTICP`: `GISTIC v2 ('Peaks' only)`
- `GISTIC`: `GISTIC v2 scores, probabilities, and peaks`

## Author(s)

Marcel Ramos <marcel.ramos@sph.cuny.edu>

## Examples

```
data(accmini)
biocExtract(accmini, "RNASeq2Gene")
biocExtract(accmini, "miRNASeqGene")
biocExtract(accmini, "RNASeq2GeneNorm")
biocExtract(accmini, "CNASNP")
```

```

biocExtract(accmini, "CNVSNP")
biocExtract(accmini, "Methylation")
biocExtract(accmini, "Mutation")
biocExtract(accmini, "RPPAArray")
biocExtract(accmini, "GISTIC")

```

---

|                 |                                                                                             |
|-----------------|---------------------------------------------------------------------------------------------|
| CorResult-class | <i>An S4 class to store correlations between gene expression level and copy number data</i> |
|-----------------|---------------------------------------------------------------------------------------------|

---

**Description**

An S4 class to store correlations between gene expression level and copy number data

**Slots**

Dataset A cohort name  
 Correlations Results data frame

---

|                 |                                                                  |
|-----------------|------------------------------------------------------------------|
| DGEResult-class | <i>An S4 class to store differential gene expression results</i> |
|-----------------|------------------------------------------------------------------|

---

**Description**

An S4 class to store differential gene expression results

**Slots**

Dataset Dataset name  
 Toptable Results data frame

---

|                        |                                                     |
|------------------------|-----------------------------------------------------|
| FirehoseCGHArray-class | <i>An S4 class to store data from CGA platforms</i> |
|------------------------|-----------------------------------------------------|

---

**Description**

An S4 class to store data from CGA platforms

**Slots**

Filename Platform name  
 DataMatrix A data frame that stores the CGH data.

---

FirehoseData-class      *An S4 class to store main data object from clinent function.*

---

## Description

An S4 class to store main data object from clinent function.

## Usage

```
## S4 method for signature 'FirehoseData'
show(object)

## S4 method for signature 'FirehoseData'
getData(object, type, platform)

## S4 method for signature 'FirehoseGISTIC'
getData(object, type, platform)

## S4 method for signature 'ANY'
getData(object, type, platform)

## S4 method for signature 'FirehoseData'
updateObject(object, ..., verbose = FALSE)

## S4 method for signature 'FirehoseData'
selectType(object, dataType)
```

## Arguments

|          |                                                                                                                                                                      |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| object   | A FirehoseData object                                                                                                                                                |
| type     | A data type to be extracted                                                                                                                                          |
| platform | An index for data types that may come from multiple platforms (such as mRNAArray), for GISTIC data, one of the options: 'AllByGene', 'ThresholdedByGene', or 'Peaks' |
| ...      | additional arguments for updateObject                                                                                                                                |
| verbose  | logical (default FALSE) whether to print extra messages                                                                                                              |
| dataType | An available data type, see object show method                                                                                                                       |

## Methods (by generic)

- show(FirehoseData): show method
- getData(FirehoseData): Get a matrix or data.frame from FirehoseData
- getData(FirehoseGISTIC): Get GISTIC data from FirehoseData
- getData(ANY): Default method for getting data from FirehoseData

- `updateObject(FirehoseData)`: Update an old RTCGAToolbox FirehoseData object to the most recent API
- `selectType(FirehoseData)`: Extract data type

## Slots

**Dataset** A cohort name  
**runDate** Standard data run date from [getFirehoseRunningDates](#)  
**gistic2Date** Analyze running date from [getFirehoseAnalyzeDates](#)  
**clinical** clinical data frame  
**RNASeqGene** Gene level expression data matrix from RNAseq  
**RNASeq2Gene** Gene level expression data matrix from RNAseqV2  
**RNASeq2GeneNorm** Gene level expression data matrix from RNAseqV2 (RSEM)  
**miRNASeqGene** miRNA expression data from matrix smallRNAseq  
**CNASNP** A data frame to store somatic copy number alterations from SNP array platform  
**CNVSNP** A data frame to store germline copy number variants from SNP array platform  
**CNASeq** A data frame to store somatic copy number alterations from sequencing platform  
**CNACGH** A list that stores FirehoseCGHArray object for somatic copy number alterations from CGH platform  
**Methylation** A list that stores FirehoseMethylationArray object for methylation data  
**mRNAArray** A list that stores FirehosemRNAArray object for gene expression data from microarray  
**miRNAArray** A list that stores FirehosemRNAArray object for miRNA expression data from microarray  
**RPPAArray** A list that stores FirehosemRNAArray object for RPPA data  
**Mutation** A data frame for mutation information from sequencing data  
**GISTIC** A FirehoseGISTIC object to store processed copy number data  
**BarcodeUUID** A data frame that stores the Barcodes, UUIDs and Short sample identifiers

---

**FirehoseGISTIC-class** *An S4 class to store processed copy number data. (Data processed by using GISTIC2 algorithm)*

---

## Description

An S4 class to store processed copy number data. (Data processed by using GISTIC2 algorithm)

## Usage

```
## S4 method for signature 'FirehoseGISTIC'
isEmpty(x)

## S4 method for signature 'FirehoseGISTIC'
updateObject(object, ..., verbose = FALSE)
```

**Arguments**

|         |                                                         |
|---------|---------------------------------------------------------|
| x       | A FirehoseGISTIC class object                           |
| object  | A FirehoseGISTIC object                                 |
| ...     | additional arguments for updateObject                   |
| verbose | logical (default FALSE) whether to print extra messages |

**Methods (by generic)**

- isEmpty(FirehoseGISTIC): check whether the FirehoseGISTIC object has data in it or not
- updateObject(FirehoseGISTIC): Update an old FirehoseGISTIC object to the most recent API

**Slots**

Dataset Cohort name  
 AllByGene A data frame that stores continuous copy number  
 ThresholdedByGene A data frame for discrete copy number data  
 Peaks A data frame storing GISTIC peak data. See [getGISTICPeaks](#).

---

FirehoseMethylationArray-class

*An S4 class to store data from methylation platforms*

---

**Description**

An S4 class to store data from methylation platforms

**Slots**

Filename Platform name  
 DataMatrix A data frame that stores the methylation data.

---

FirehosemRNAArray-class

*An S4 class to store data from array (mRNA, miRNA etc.) platforms*

---

**Description**

An S4 class to store data from array (mRNA, miRNA etc.) platforms

**Slots**

Filename Platform name  
 DataMatrix A data matrix that stores the expression data.

---

|                  |                                                                           |
|------------------|---------------------------------------------------------------------------|
| getBroadSubtypes | <i>Download expression-based cancer subtypes from the Broad Institute</i> |
|------------------|---------------------------------------------------------------------------|

---

**Description**

Obtain the mRNA expression clustering results from the Broad Institute for a specific cancer code (see [getFirehoseDatasets](#)).

**Usage**

```
getBroadSubtypes(dataset, clust.alg = c("CNMF", "ConsensusPlus"))
```

**Arguments**

|           |                                                                                   |
|-----------|-----------------------------------------------------------------------------------|
| dataset   | A TCGA cancer code, e.g. "OV" for ovarian cancer                                  |
| clust.alg | The selected cluster algorithm, either "CNMF" or "ConsensusPlus" (default "CNMF") |

**Value**

A data.frame of cluster and silhouette values

**Author(s)**

Ludwig Geistlinger

**Examples**

```
co <- getBroadSubtypes("COAD", "CNMF")
head(co)
```

---

|         |                                              |
|---------|----------------------------------------------|
| getData | <i>Extract data from FirehoseData object</i> |
|---------|----------------------------------------------|

---

**Description**

A go-to function for getting top level information from a [FirehoseData](#) object. Available datatypes for a particular object can be seen by entering the object name in the console ('show' method).

**Usage**

```
getData(object, type, platform)
```

**Arguments**

|          |                                                                                                                                                             |
|----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| object   | A <a href="#">FirehoseData</a> object                                                                                                                       |
| type     | A data type to be extracted                                                                                                                                 |
| platform | An index for data types that may come from multiple platforms (such as mRNAArray), for GISTIC data, one of the options: 'AllByGene' or 'Thresholded-ByGene' |

**Value**

Returns matrix or data.frame depending on data type

**Examples**

```
data(accmini)
getData(accmini, "clinical")
getData(accmini, "RNASeq2GeneNorm")
getData(accmini, "Methylation", 1)[1:4]
```

---

`getFirehoseAnalyzeDates`*Get data analyze dates.*

---

**Description**

getFirehoseAnalyzeDates returns the character vector for analyze release dates.

**Usage**

```
getFirehoseAnalyzeDates(last = NULL)
```

**Arguments**

|      |                                      |
|------|--------------------------------------|
| last | To list last n dates. (Default NULL) |
|------|--------------------------------------|

**Value**

A character vector for dates.

**Examples**

```
getFirehoseAnalyzeDates(last=2)
```

---

|                 |                                       |
|-----------------|---------------------------------------|
| getFirehoseData | <i>Get data from Firehose portal.</i> |
|-----------------|---------------------------------------|

---

## Description

getFirehoseData returns FirehoseData object that stores TCGA data.

## Usage

```
getFirehoseData(
  dataset,
  runDate = "20160128",
  gistic2Date = "20160128",
  RNASeqGene = FALSE,
  RNASeq2Gene = FALSE,
  clinical = TRUE,
  miRNASeqGene = FALSE,
  miRNASeqGeneType = c("read_count", "reads_per_million_miRNA_mapped", "cross-mapped"),
  RNASeq2GeneNorm = FALSE,
  CNASNP = FALSE,
  CNVSNP = FALSE,
  CNASeq = FALSE,
  CNACGH = FALSE,
  Methylation = FALSE,
  Mutation = FALSE,
  mRNAArray = FALSE,
  miRNAArray = FALSE,
  RPPAArray = FALSE,
  GISTIC = FALSE,
  RNASeqNorm = "raw_count",
  RNASeq2Norm = c("normalized_counts", "RSEM_normalized_log2", "raw_counts",
    "scaled_estimate"),
  forceDownload = FALSE,
  destdir = .setCache(),
  fileSizeLimit = 500,
  getUUIDs = FALSE,
  ...
)
```

## Arguments

|             |                                                                                                   |
|-------------|---------------------------------------------------------------------------------------------------|
| dataset     | A cohort disease code. TCGA cancer codes can be obtained via <a href="#">getFirehose-Datasets</a> |
| runDate     | Standard data run dates. Date list can be accessible via <a href="#">getFirehoseRunning-Dates</a> |
| gistic2Date | Analysis run date for GISTIC obtained via <a href="#">getFirehoseAnalyzeDates</a>                 |

|                  |                                                                                                                                                                                             |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RNASeqGene       | Logical (default FALSE) RNAseq TPM data.                                                                                                                                                    |
| RNASeq2Gene      | Logical (default FALSE) RNAseq v2 (RSEM processed) data; see RNAseqNorm argument.                                                                                                           |
| clinical         | Logical (default TRUE) clinical data.                                                                                                                                                       |
| miRNASeqGene     | Logical (default FALSE) smallRNAseq data.                                                                                                                                                   |
| miRNASeqGeneType | Character (default "read_count") Indicate which type of data should be pulled from the miRNASeqGene data. Must be one of "reads_per_million_miRNA_mapped", "read_count", or "cross-mapped". |
| RNASeq2GeneNorm  | Logical (default FALSE) RNAseq v2 (RSEM processed) data.                                                                                                                                    |
| CNASNP           | Logical (default FALSE) somatic copy number alterations data from SNP array.                                                                                                                |
| CNVSNP           | Logical (default FALSE) germline copy number variants data from SNP array.                                                                                                                  |
| CNASeq           | Logical (default FALSE) somatic copy number alterations data from sequencing.                                                                                                               |
| CNACGH           | Logical (default FALSE) somatic copy number alterations data from CGH.                                                                                                                      |
| Methylation      | Logical (default FALSE) methylation data.                                                                                                                                                   |
| Mutation         | Logical (default FALSE) mutation data from sequencing.                                                                                                                                      |
| mRNAArray        | Logical (default FALSE) mRNA expression data from microarray.                                                                                                                               |
| miRNAArray       | Logical (default FALSE) miRNA expression data from microarray.                                                                                                                              |
| RPPAArray        | Logical (default FALSE) RPPA data                                                                                                                                                           |
| GISTIC           | logical (default FALSE) processed copy number data                                                                                                                                          |
| RNAseqNorm       | RNAseq data normalization method. (Default raw_count)                                                                                                                                       |
| RNAseq2Norm      | RNAseq v2 data normalization method. (Default normalized_count or one of RSEM_normalized_log2, raw_count, scaled_estimate)                                                                  |
| forceDownload    | A logic (Default FALSE) key to force download RTCGAToolbox every time. By default if you download files into your working directory once than RTCGAToolbox using local files next time.     |
| destdir          | Directory in which to store the resulting downloaded file. Defaults to a cache directory given by RTCGAToolbox:::setCache().                                                                |
| fileSizeLimit    | Files that are larger than set value (megabyte) won't be downloaded (Default: 500)                                                                                                          |
| getUUIDs         | Logical key to get UUIDs from barcode (Default: FALSE)                                                                                                                                      |
| ...              | Additional arguments to pass down.                                                                                                                                                          |

## Details

This is a main client function to download data from Firehose TCGA portal.

To avoid unnecessary downloads, we use `tools::R_user_dir("RTCGAToolbox", "cache")` to set the default `destdir` parameter to the cached directory. To get the actual default directory, one can run `RTCGAToolbox:::setCache()`.

**Value**

A FirehoseData data object that stores data for selected data types.

**See Also**

[getLinks](#), <https://gdac.broadinstitute.org/>

**Examples**

```
# Sample Dataset
data(accmini)
accmini
## Not run:
BRCAdata <- getFirehoseData(dataset="BRCA",
runDate="20140416",gistic2Date="20140115",
RNASeqGene=TRUE,clinical=TRUE,mRNAArray=TRUE,Mutation=TRUE)

## End(Not run)
```

---

|                     |                                           |
|---------------------|-------------------------------------------|
| getFirehoseDatasets | <i>Get a list of TCGA disease cohorts</i> |
|---------------------|-------------------------------------------|

---

**Description**

getFirehoseDatasets returns a character vector of TCGA disease codes. A reference table can be seen at <https://gdac.broadinstitute.org/>.

**Usage**

```
getFirehoseDatasets()
```

**Value**

A character string

**See Also**

<https://gdac.broadinstitute.org/>

**Examples**

```
getFirehoseDatasets()
```

---

```
getFirehoseRunningDates
```

*Get standard data running dates.*

---

### Description

getFirehoseRunningDates returns the character vector for standard data release dates.

### Usage

```
getFirehoseRunningDates(last = NULL)
```

### Arguments

last                      To list last n dates. (Default NULL)

### Value

A character vector for dates.

### Examples

```
getFirehoseRunningDates()
getFirehoseRunningDates(last=2)
```

---

```
getGISTICPeaks
```

*Download GISTIC2 peak-level data from the Firehose pipeline*

---

### Description

Access GISTIC2 level 4 copy number data through [gdac.broadinstitute.org](http://gdac.broadinstitute.org)

### Usage

```
getGISTICPeaks(object, peak = c("wide", "narrow", "full"), rm.chrX = TRUE)
```

### Arguments

object                    A FirehoseData GISTIC type object  
peak                        The peak type, select from "wide", "narrow", "full".  
rm.chrX                    (logical default TRUE) Whether to remove observations in the X chromosome

### Value

A data.frame of peak values

**Author(s)**

Ludwig Geistlinger

**Examples**

```
co <- getFirehoseData("COAD", clinical = FALSE, GISTIC = TRUE)
peaks <- getGISTICPeaks(co, "wide")
class(peaks)
head(peaks)[1:6]
```

getLinks

*Get resource links from inputs***Description**

This function provides a reference to the resources downloaded from the GDAC Firehose pipeline. Based on the input, the function returns a URL location to the resource if there exists one.

**Usage**

```
getLinks(
  dataset,
  data_date = "20160128",
  RNASeqGene = FALSE,
  RNASeq2Gene = FALSE,
  clinical = FALSE,
  miRNASeqGene = FALSE,
  RNASeq2GeneNorm = FALSE,
  RNASeq2Norm = c("normalized_counts", "RSEM_normalized_log2", "raw_counts",
    "scaled_estimate"),
  CNASNP = FALSE,
  CNVSNP = FALSE,
  CNASeq = FALSE,
  CNACGH = FALSE,
  Methylation = FALSE,
  Mutation = FALSE,
  mRNAArray = FALSE,
  miRNAArray = FALSE,
  RPPAArray = FALSE,
  GISTIC = FALSE
)
```

**Arguments**

dataset      A cohort disease code. TCGA cancer codes can be obtained via [getFirehose-Datasets](#)

|                 |                                                                                                                            |
|-----------------|----------------------------------------------------------------------------------------------------------------------------|
| data_date       | Either a runDate or analysisDate typically entered in getFirehoseData                                                      |
| RNASeqGene      | Logical (default FALSE) RNAseq TPM data.                                                                                   |
| RNASeq2Gene     | Logical (default FALSE) RNAseq v2 (RSEM processed) data; see RNAseqNorm argument.                                          |
| clinical        | Logical (default TRUE) clinical data.                                                                                      |
| miRNASeqGene    | Logical (default FALSE) smallRNAseq data.                                                                                  |
| RNASeq2GeneNorm | Logical (default FALSE) RNAseq v2 (RSEM processed) data.                                                                   |
| RNASeq2Norm     | RNAseq v2 data normalization method. (Default normalized_count or one of RSEM_normalized_log2, raw_count, scaled_estimate) |
| CNASNP          | Logical (default FALSE) somatic copy number alterations data from SNP array.                                               |
| CNVSNP          | Logical (default FALSE) germline copy number variants data from SNP array.                                                 |
| CNASeq          | Logical (default FALSE) somatic copy number alterations data from sequencing.                                              |
| CNACGH          | Logical (default FALSE) somatic copy number alterations data from CGH.                                                     |
| Methylation     | Logical (default FALSE) methylation data.                                                                                  |
| Mutation        | Logical (default FALSE) mutation data from sequencing.                                                                     |
| mRNAArray       | Logical (default FALSE) mRNA expression data from microarray.                                                              |
| miRNAArray      | Logical (default FALSE) miRNA expression data from microarray.                                                             |
| RPPAArray       | Logical (default FALSE) RPPA data                                                                                          |
| GISTIC          | logical (default FALSE) processed copy number data                                                                         |

**Value**

A character URL to a dataset location

**Examples**

```
getLinks("BRCA", CNASeq = TRUE)
```

---

|                 |                                                                  |
|-----------------|------------------------------------------------------------------|
| getMutationRate | <i>Make a table for mutation rate of each gene in the cohort</i> |
|-----------------|------------------------------------------------------------------|

---

**Description**

Make a table for mutation rate of each gene in the cohort

**Usage**

```
getMutationRate(dataObject)
```

**Arguments**

dataObject      This must be FirehoseData object.

**Value**

Returns a data table

**Examples**

```
data(accmini)
mutRate <- getMutationRate(dataObject=accmini)
mutRate <- mutRate[order(mutRate[,2],decreasing = TRUE),]
head(mutRate)
```

---

hg19.ucsc.gene.locations

*Gene coordinates for circle plot.*

---

**Description**

A dataset containing the gene coordinates The variables are as follows:

**Format**

A data frame with 28454 rows and 5 variables

**Details**

- GeneSymbol: Gene symbols
- Chromosome: Chromosome name
- Strand: Gene strand on chromosome
- Start: Gene location on chromosome
- End: Gene location on chromosome

---

```
makeSummarizedExperimentFromGISTIC
```

*Create a SummarizedExperiment from FireHose GISTIC*

---

## Description

Use the output of `getFirehoseData` to create a [SummarizedExperiment](#). This can be done for three types of data, G-scores threshold by gene, copy number by gene, and copy number by peak regions.

## Usage

```
makeSummarizedExperimentFromGISTIC(
  gistic,
  dataType = c("AllByGene", "ThresholdedByGene", "Peaks"),
  rownameCol = "Gene.Symbol",
  ...
)
```

## Arguments

|                         |                                                                                                                                                                                                                 |
|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>gistic</code>     | A <a href="#">FirehoseGISTIC</a> object                                                                                                                                                                         |
| <code>dataType</code>   | character(1) One of "ThresholdedByGene" (default), "AllByGene", or "Peaks"                                                                                                                                      |
| <code>rownameCol</code> | character(1) The name of the column in the data to use as rownames in the data matrix (default: 'Gene.Symbol'). The row names are only set when the column name is found in the data and all values are unique. |
| <code>...</code>        | Additional arguments passed to 'getGISTICPeaks'.                                                                                                                                                                |

## Value

A `SummarizedExperiment` object

## Author(s)

L. Geistlinger, M. Ramos

## Examples

```
co <- getFirehoseData("COAD", clinical = FALSE, GISTIC = TRUE,
  destdir = tempdir())
makeSummarizedExperimentFromGISTIC(co, "AllByGene")
```

**Description**

Managing data from large-scale projects (such as The Cancer Genome Atlas (TCGA) for further analysis is an important and time consuming step for research projects. Several efforts, such as the Firehose project, make TCGA pre-processed data publicly available via web services and data portals, but this information must be managed, downloaded and prepared for subsequent steps. We have developed an open source and extensible R based data client for pre-processed data from the Firehose, and demonstrate its use with sample case studies. Results show that our RTCGAToolbox can facilitate data management for researchers interested in working with TCGA data. The RTCGAToolbox can also be integrated with other analysis pipelines for further data processing.

**Details**

The main function you're likely to need from RTCGAToolbox is [getFirehoseData](#). Otherwise refer to the vignettes to see how to use the RTCGAToolbox

**Author(s)**

Mehmet Kemal Samur

**See Also**

Useful links:

- <http://mksamur.github.io/RTCGAToolbox/>
- Report bugs at <https://github.com/mksamur/RTCGAToolbox/issues>

**Description**

An accessor function for the [FirehoseData](#). An argument will specify the data type to return See [FirehoseData](#) for more details.

**Usage**

```
selectType(object, dataType)
```

**Arguments**

|          |                             |
|----------|-----------------------------|
| object   | A FirehoseData class object |
| dataType | A data type, see details.   |

Details

- clinical: Get the clinical data slot
- RNASeqGene: RNASeqGene
- RNASeq2GeneNorm: Normalized
- miRNASeqGene: micro RNA SeqGene
- CNASNP: Copy Number Alteration
- CNVSNP: Copy Number Variation
- CNASeq: Copy Number Alteration
- CNACGH: Copy Number Alteration
- Methylation: Methylation
- mRNAArray: Messenger RNA
- miRNAArray: micro RNA
- RPPAArray: Reverse Phase Protein Array
- Mutation: Mutations
- GISTIC: GISTIC v2 scores and probabilities

Value

The data type element of the FirehoseData object

---

|             |                                                  |
|-------------|--------------------------------------------------|
| showResults | <i>Export toptable or correlation data frame</i> |
|-------------|--------------------------------------------------|

---

Description

Export toptable or correlation data frame

Usage

showResults(object)

Arguments

object                   A [DGEResult](#) or [CorResult](#) object

Value

Returns toptable or correlation data frame

Examples

data(accmini)

---

`showResults,CorResult-method`*Export toptable or correlation data frame*

---

**Description**

Export toptable or correlation data frame

**Usage**

```
## S4 method for signature 'CorResult'  
showResults(object)
```

**Arguments**

object            A [DGEResult](#) or [CorResult](#) object

**Value**

Returns correlation results data frame

**Examples**

```
data(accmini)
```

---

`showResults,DGEResult-method`*Export toptable or correlation data frame*

---

**Description**

Export toptable or correlation data frame

**Usage**

```
## S4 method for signature 'DGEResult'  
showResults(object)
```

**Arguments**

object            A [DGEResult](#) or [CorResult](#) object

**Value**

Returns toptable for DGE results

**Examples**

```
data(accmini)
```

# Index

- \* **datasets**
  - accmini, [3](#)
- accmini, [3](#)
- biocExtract, [3](#)
- CorResult, [20](#), [21](#)
- CorResult-class, [5](#)
- DGEResult, [20](#), [21](#)
- DGEResult-class, [5](#)
- FirehoseCGHArray-class, [5](#)
- FirehoseData, [3](#), [9](#), [10](#), [19](#)
- FirehoseData-class, [6](#)
- FirehoseGISTIC, [18](#)
- FirehoseGISTIC-class, [7](#)
- FirehoseMethylationArray-class, [8](#)
- FirehosemRNAArray-class, [8](#)
- getBroadSubtypes, [9](#)
- getData, [9](#)
- getData, ANY-method (FirehoseData-class), [6](#)
- getData, FirehoseData-method (FirehoseData-class), [6](#)
- getData, FirehoseGISTIC-method (FirehoseData-class), [6](#)
- getFirehoseAnalyzeDates, [7](#), [10](#), [11](#)
- getFirehoseData, [11](#), [19](#)
- getFirehoseDatasets, [9](#), [11](#), [13](#), [15](#)
- getFirehoseRunningDates, [7](#), [11](#), [14](#)
- getGISTICPeaks, [8](#), [14](#)
- getLinks, [13](#), [15](#)
- getMutationRate, [16](#)
- hg19.ucsc.gene.locations, [17](#)
- isEmpty, FirehoseGISTIC-method (FirehoseGISTIC-class), [7](#)
- makeSummarizedExperimentFromGISTIC, [18](#)
- RaggedExperiment, [3](#), [4](#)
- RaggedExperiment-class, [3](#)
- RTCGAToolbox, [19](#)
- RTCGAToolbox-package (RTCGAToolbox), [19](#)
- selectType, [19](#)
- selectType, FirehoseData-method (FirehoseData-class), [6](#)
- show, FirehoseData-method (FirehoseData-class), [6](#)
- showResults, [20](#)
- showResults, CorResult, CorResult-method (showResults, CorResult-method), [21](#)
- showResults, CorResult-method, [21](#)
- showResults, DGEResult, DGEResult-method (showResults, DGEResult-method), [21](#)
- showResults, DGEResult-method, [21](#)
- SummarizedExperiment, [3](#), [4](#), [18](#)
- updateObject, FirehoseData-method (FirehoseData-class), [6](#)
- updateObject, FirehoseGISTIC-method (FirehoseGISTIC-class), [7](#)