Package 'ImmuneSpaceR'

April 12, 2018

Type Package

Title A Thin Wrapper around the ImmuneSpace Database

Version 1.6.2

Date 2014-06-12

VignetteBuilder knitr

Description Provides a convenient API for accessing data sets within ImmuneSpace (www.immunespace.org), the data repository and analysis platform of the Human Immunology Project Consortium (HIPC).

biocViews DataImport, DataRepresentation, ThirdPartyClient

URL https://github.com/RGLab/ImmuneSpaceR

BugReports https://github.com/RGLab/ImmuneSpaceR/issues

License GPL-2

Suggests knitr, testthat

Imports utils, methods, data.table, httr, Rlabkey (>= 2.1.136), Biobase, pheatmap, ggplot2, scales, stats, gtools, gplots, reshape2, plotly, heatmaply (>= 0.7.0), rjson, rmarkdown, preprocessCore, parallel

Collate 'CreateConnection.R' 'ImmuneSpace.R' 'ImmuneSpaceR.R' 'ParticipantGroupFns.R' 'getDataset.R' 'getGEMatrix.R' 'netrc.R' 'quick_plot.R' 'utils.R' 'zzz.R'

RoxygenNote 6.0.1

NeedsCompilation no

Author Greg Finak [aut],
Renan Sauteraud [aut],
Mike Jiang [aut],
Gil Guday [aut],
Leo Dashevskiy [aut],
Evan Henrich [aut],
Ju Yeong Kim [aut],
Raphael Gottardo [aut],
ImmuneSpace Package Maintainer [cre, cph]

Maintainer ImmuneSpace Package Maintainer <immunespace@gmail.com>

R topics documented:

`ImmuneSpaceR-package` *A Thin Wrapper Around ImmuneSpace*

Description

ImmuneSpaceR provides a convenient API for accessing data sets within the ImmuneSpace database.

Details

Uses the Rlabkey package to connect to ImmuneSpace. Implements caching, and convenient methods for accessing data sets.

Author(s)

Greg Finak

`check_netrc` *Check netrc file*

Description

Check that there is a netrc file with a valid entry for ImmuneSpace.

Usage

```
check_netrc()
```

Details

In order to connect to ImmuneSpace, you will need a '.netrc' file in your contains a 'machine' name (hostname of ImmuneSpace), and 'login' and 'password'. See [here](https://www.labkey.org/wiki/home/Documentation/ for more information. By default `RCurl` will look for the file in your home directoty.

If no netrc is available or it is not formatted properly, `write_netrc` can be used to write one. Otherwise, when specifying login and password in `CreateConnection`, a temporary file will be created for that connection.

Value

The name of the netrc file

See Also

CreateConnection write_netrc

Examples

```
try(check_netrc())
```

`CreateConnection` *CreateConnection*

Description

Constructor for `ImmuneSpaceConnection` class

Usage

```
CreateConnection(study = NULL, login = NULL, password = NULL,
  use.data.frame = FALSE, verbose = FALSE, onTest = FALSE)
```

Arguments

`study`	A `"character"` vector naming the study.
`login`	A `"character"`. Optional argument. If there is no netrc file a temporary one can be written by passing login and password of an active ImmuneSpace account.
`password`	A `"character"`. Optional. The password for the selected login.
`use.data.frame`	A `"logical"`. If set to TRUE, the functions will return `data.frame` objects instead of `data.table`.
`verbose`	A `"logical"` whether to print the extra details for troubleshooting.
`onTest`	A `"logical"` whether to connect to the test server (https://test.immunespace.org/) instead of the production server (https://www.immunespace.org/).

Details

Instantiates an `ImmuneSpaceConnection` for `study` The constructor will try to take the values of the various 'labkey.*' parameters from the global environment. If they don't exist, it will use default values. These are assigned to 'options', which are then used by the `ImmuneSpaceConnection` class.

Value

an instance of an `ImmuneSpaceConnection`

See Also

ImmuneSpaceConnection

Examples

```
## Not run:
  # Single study
  con <- CreateConnection("SDY269")
  # Cross study
  con <- CreateConnection("")

## End(Not run)

sdy <- try(CreateConnection("SDY269"))
if(inherits(sdy, "try-error")){
  print("Read the Introduction vignette for more information on how to set up
  a .netrc file.")
}
```

`ImmuneSpaceConnection-class`
The ImmuneSpaceConnection class

Description

A connection respresents a study or a set of studies available on ImmuneSpace. It provides function to download and display the data within these studies.

Details

Uses global variables `labkey.url.base`, and `labkey.url.path`, to access a study. `labkey.url.base` should be `https://www.immunespace.org/`. `labkey.url.path` should be `/Studies/studyname`, where 'studyname' is the accession number of the study. The ImmunespaceConnection will initialize itself, and look for a `.netrc` file in `"~/"` the user's home directory. The `.netrc` file should contain a `machine`, `login`, and `password` entry to allow access to ImmuneSpace, where `machine` is the host name like "www.immunespace.org".

Value

An instance of an ImmuneSpaceConnection for a study in 'labkey.url.path'

Fields

`study` A `character`. The study accession number. Use an empty string ("") to create a connection at the project level.

`config` A `list`. Stores configuration of the connection object such as URL, path and username.

`available_datasets` A `data.table`. The table of datasets available in the connection object.

`data_cache` A `list`. Stores the data to avoid downloading the same tables multiple times.

`constants` A `list`. Used to store information regarding gene-expression data.

Methods

`addTreatment(matrixName = NULL)` Add treatment information to the phenoData of an expression matrix available in the connection object.

x: A character. The name of a expression matrix that has been downloaded from the connection.

`clear_cache()` Clear the data_cache. Remove downloaded datasets and expression matrices.

`EMNames(EM = NULL, colType = "participant_id")` Change the sampleNames of an ExpressionSet fetched by getGEMatrix using the information in the phenodData slot.

x: An ExpressionSet, as returned by getGEMatrix.

colType: A character. The type of column names. Valid options are 'expsample_accession' and 'participant_id'.

`getDataset(x, original_view = FALSE, reload = FALSE, colFilter = NULL, ...)` Get a dataset form the connection

original_view: A logical. If set tot TRUE, download the ImmPort view. Else, download the default grid view.

reload: A logical. Clear the cache. If set to TRUE, download the dataset, whether a cached version exist or not.

colFilter: A character. A filter as returned by Rlabkey's makeFilter function.

'...': Extra arguments to be passed to labkey.selectRows.

`getGEAnalysis(...)` Downloads data from the gene expression analysis results table.

'...': A list of arguments to be passed to labkey.selectRows.

`getGEFiles(files, destdir = ".", quiet = FALSE)` Download gene expression raw data files.

files: A character. Filenames as shown on the gene_expression_files dataset.

destdir: A character. The local path to store the downloaded files.

`getGEMatrix(matrixName = NULL, cohort = NULL, outputType = "summary", annotation = "latest", rel` Downloads a normalized gene expression matrix from ImmuneSpace.

'x': A 'character'. The name of the gene expression matrix to download.

'cohort': A 'character'. The name of a cohort that has an associated gene expression matrix. Note that if 'cohort' isn't NULL, then 'x' is ignored.

'outputType': one of 'raw', 'normalized' or 'summary'. If 'raw' then returns an expression matrix of non-normalized values by probe. 'normalized' returns normalized values by probe. 'summary' returns normalized values averaged by gene symbol.

'annotation': one of 'default', 'latest', or 'ImmSig'. Determines which feature annotation set is used. 'default' uses the fas from when the matrix was generated. 'latest' uses a recently updated fas based on the original. 'ImmSig' is specific to studies involved in the ImmuneSignatures project and uses the annotation from when the meta-study's manuscript was created.

'reload': A 'logical'. If set to TRUE, the matrix will be downloaded again, even if a cached cop exist in the ImmuneSpaceConnection object.

`getParticipantData(group, dataType, original_view = FALSE, ...)` returns a dataframe with ImmuneSpace data subset by groupId.

group: Use con$listParticipantGroups() to find Participant groupId or groupName.

dataType: Use con$listDatasets('datasets') to see possible dataType inputs.

`listDatasets(output = c("datasets", "expression"))` List the datasets available in the study or studies of the connection.

`listGEAnalysis()` List available gene expression analysis for the connection.

`listParticipantGroups()` returns a dataframe with all saved Participant Groups on ImmuneSpace.

`quick_plot(...)` Plots a selected dataset. This is the function used by the DataExplorer module on ImmuneSpace.

dataset: A character. The name of the dataset to plot, as displayed by the listDataset method.

normalize_to_baseline: A logical. If set to TRUE, the values are plotted as log2 fold-change from baseline.

type: A character. The type of plot. Valid choices are 'auto', 'heatmap', 'boxplot', 'lineplot', 'violinplot'. If set to 'auto', the function will select an appropriate plot type for the selected data.

filter: A filter as created by the makeFilter function from Rlabkey.

facet: The facetting for ggplot2 based plots. Valid choices are 'grid' and 'wrap'.

text_size: The size of all text elements in the plot.

legend: A character. Columns of the dataset or demographics to be added as legend on the heatmap. This argument is ignored if the plot type isn't heatmap.

show_virus_strain: A logical. Should all the virus strains be shown or should the values be averaged. Only used when dataset = 'hai'.

interactive: A logical. If set to TRUE, an interactive plot will be created. The default is FALSE.

'...': Extra argument to be passed to ggplot. e.g: shape = 'Age', color = 'Race'.

See Also

`CreateConnection` `ImmuneSpaceR-package`

Examples

```
## Not run:
  sdy269 <- CreateConnection("SDY269")
  sdy269

## End(Not run)
```

`ISpalette` *ImmuneSpace palette*

Description

Create a color gradient of the selected length that matches the ImmuneSpace theme.

Usage

```
ISpalette(n)
```

Arguments

`n`	A `numeric`. The length of the desired palette.

Value

A `character` vector colors in hexadecimal code of length `n`.

Examples

```
plot(1:10, col = ISpalette(10), cex = 10, pch = 16)
```

`loadConnection` *Save/Load an ImmuneSpaceConnection object from disk*

Description

Connection can hold a lot of data in cache. If a lot of work has been done (e.g: lots of downloaded datasets and gene-expression matrices), it can be useful to save the connection for later work or even offline use.

Usage

```
loadConnection(file)

saveConnection(con, file)
```

Arguments

`file`	The file name to be saved to or loaded from
`con`	An `ImmuneSpaceConnection`. The connection to save to file. To be loaded later using `loadConnection`.

Value

An ImmuneSpaceConnection object

Examples

```
#Sample saved connection with pre-downloaded expression matrices and datasets
saved <- system.file("extdata/saved_con.rds", package = "ImmuneSpaceR")
new_con <- loadConnection(saved)
new_con
names(new_con$data_cache)
## Not run:
  saveConnection(new_con, tempfile())

## End(Not run)
```

`template_IS` *template_IS*

Description

A HTML template for knitted reports that matches ImmuneSpace's graphic style. It is based on `html_document` from the **rmarkdown** package with css, theme, and template parameters disabled.

Usage

```
template_IS(...)
```

Arguments

...	See html_document

Details

See the documentation for html_document or the oneline documentation for additional details on using the html_document format. Compared to html_document, it:

- uses a custom css stylesheet
- does not use bootstrap themes

Value

R Markdown output format to pass to render

Examples

```
## Not run:
library(ImmuneSpaceR)
rmarkdown::render("input.Rmd", template_IS())
rmarkdown::render("input.Rmd", template_IS(toc = TRUE))

## End(Not run)
```

---

theme_IS *theme_IS*

Description

Theme that matches ImmuneSpace's graphic style. The theme modifies the background, the grid lines, the axis, and the colors used by continuous and gradient scales.

Usage

```
theme_IS(base_size = 12)
```

Arguments

base_size	A numeric. Base font size.

Details

List of modified ggplot2 elements: panel.background, panel.grid.major, panel.grid.minor, axis.ticks, axis.line.x, axis.line.y, plot.title, and strip.background.

The default scale_fill_gradient, scale_fill_continuous, scale_colour_gradient and scale_colour_contino are also replaced by a custom scale.

Value

A theme object

Examples

```
library(ggplot2)
p <- ggplot(data = mtcars) + geom_point(aes(x = mpg, y = cyl, color = hp)) + facet_grid(vs ~ am)
p + theme_IS()
```

`write_netrc` *Write a netrc file*

Description

Write a netrc file that is valid for accessing ImmuneSpace

Usage

```
write_netrc(login, password, machine = "www.immunespace.org", file = NULL)
```

Arguments

`login`	A `character`. The email address used for loging in on ImmuneSpace.
`password`	A `character`. The password associated with the login.
`machine`	A `character`. The server to connect.
`file`	A `character`. The credentials will be written into that file. If left NULL, the netrc will be written into a temporary file.

Value

A character vector containing the file paths for netrc

Examples

```
write_netrc("immunespaceuser@gmail.com", "mypassword")
```

Index