Package 'TidyMultiqc'

May 7, 2026

Type Package

Title Converts 'MultiQC' Reports into Tidy Data Frames

Version 1.0.3

Author Michael Milton

Maintainer Michael Milton <michael.r.milton@gmail.com>

Description Provides the means to convert 'multiqc_data.json' files, produced by the wonderful 'MultiQC' tool, into tidy data frames for downstream analysis in R. This analysis might involve cohort analysis, quality control visualisation, change-point detection, statistical process control, clustering, or any other type of quality analysis.

License GPL (>= 3)

Encoding UTF-8

Imports assertthat, dplyr, jsonlite, magrittr, purrr, rlang, stringr, tibble

Suggests tidyr, testthat (>= 3.0.0), knitr, rmarkdown, ggplot2, HistDat

Config/testthat/edition 3

RoxygenNote 7.1.2

VignetteBuilder knitr

URL https://multimeric.github.io/TidyMultiqc/, https://github.com/multimeric/TidyMultiqc, https://cran.r-project.org/package=TidyMultiqc

BugReports https://github.com/multimeric/TidyMultiqc/issues

NeedsCompilation no

Repository CRAN

Date/Publication 2022-09-25 02:10:02 UTC

Contents

`TidyMultiqc-package` *TidyMultiqc: Converting MultiQC reports into tidy data frames*

Description

This package provides the means to convert `multiqc_data.json` files, produced by the wonderful MultiQC tool, into tidy data.frames for downstream analysis in R. If you are reading this manual, you should immediately stop reading this and instead refer to the documentation website at `https://multimeric.github.io/TidyMultiqc/`, which provides more accessible documentation.

`list_plots` *List the plot identifiers of all the plots in a given multiqc report*

Description

List the plot identifiers of all the plots in a given multiqc report

Usage

```
list_plots(path)
```

Arguments

`path`	The file path to the multiqc report. This should be a length 1 character vector

Details

The main use for this function is finding the plot identifiers that you will then pass into the `plots` argument of the `load_multiqc()` function. Refer to the section on "Extracting Plot Data" in the main vignette for more information.

Value

A data frame containing n rows, where n is the number of plots in the report you have provided, and two columns:

id The identifier for the plot. This is the one you should use as a name in `plot_opts`.

name The plot title. This is likely what you see in the multiqc report when you open it with your browser.

Examples

```
# Ignore this, choose your own filepath as the `filepath` variable
filepath <- system.file("extdata", "HG00096/multiqc_data.json", package = "TidyMultiqc")
# This is the actual invocation
list_plots(filepath)
```

`load_multiqc` *Loads one or more MultiQC reports into a data frame*

Description

Loads one or more MultiQC reports into a data frame

Usage

```
load_multiqc(
  paths,
  plots = NULL,
  find_metadata = function(...) {     list() },
  plot_parsers = list(),
  sections = "general"
)
```

Arguments

`paths`	A string vector of filepaths to multiqc_data.json files
`plots`	A string vector, each of which contains the ID of a plot you want to include in the output. You can use `list_plots()` to help here.
`find_metadata`	A single function that will be called with a sample name and the parsed JSON for the entire report and returns a named list of metadata fields for the sample. Refer to the vignette for an example.
`plot_parsers`	**Advanced**. A named list of custom parser functions. The names of the list should correspond to plotly plot types, such as "xy_line", and the values should be functions that return a named list of named lists. For the return value, the outer list is named by the sample ID, and the inner list is named by the name of the column. Refer to the source code for some examples.
`sections`	A string vector of zero or more sections to include in the output. Each section can be: **"plot"** Parse plot data. Note that you should also provide a list of plots via the `plots` argument **"general"** parse the general stat section **"raw"** Parse the raw data section This defaults to 'general', which tends to contain the most useful statistics

Value

A tibble (data.frame subclass) with QC data and metadata as columns, and samples as rows. Columns are named according to the respective section they belong to, and will always be listed in the following order:

`metadata.X`	This column contains metadata for this sample. By default this is only the sample ID, but if you have provided the `find_metadata` argument, there may be more columns.
`general.X`	This column contains a generally useful summary statistic for each sample
`plot.X`	This column contains a data frame of plot data for each sample. Refer to the plot parsers documentation (ie the `parse_X` functions) for more information on the output format.
`raw.X`	This column contains a raw summary statistic or value relating to each sample

See Also

`parse_xyline_plot()` `parse_bar_graph()`

Examples

```
load_multiqc(system.file("extdata", "wgs/multiqc_data.json", package = "TidyMultiqc"))
```

Index