Package ‘tuberculosis’

May 16, 2024

Title  Tuberculosis Gene Expression Data for Machine Learning

Description  The tuberculosis R/Bioconductor package features tuberculosis gene expression data for machine learning. All human samples from GEO that did not come from cell lines, were not taken postmortem, and did not feature recombination have been included. The package has more than 10,000 samples from both microarray and sequencing studies that have been processed from raw data through a hyper-standardized, reproducible pipeline.

biocViews  ExperimentHub, ExpressionData, Homo_sapiens_Data, MicroarrayData, SequencingData, ReproducibleResearch

Version  1.10.0

License  Artistic-2.0

Depends  R (>= 4.1.0), SummarizedExperiment

Imports  AnnotationHub, ExperimentHub, S4Vectors, dplyr, magrittr, purrr, rlang, stringr, tibble, tidyr

Suggests  BiocStyle, ggplot2, hrbrthemes, knitr, readr, rmarkdown, scater, usethis, utils

URL  https://github.com/schifferl/tuberculosis

BugReports  https://github.com/schifferl/tuberculosis/issues

Encoding  UTF-8

Roxygen  list(markdown = TRUE)

RoxygenNote  7.3.1

VignetteBuilder  knitr

git_url  https://git.bioconductor.org/packages/tuberculosis
git_branch  RELEASE_3_19
git_last_commit  ebf01f6
git_last_commit_date  2024-04-30

Repository  Bioconductor 3.19

Date/Publication  2024-05-16

Author  Lucas Schiffer [aut, cre] (<https://orcid.org/0000-0003-3628-0326>)

Maintainer  Lucas Schiffer <schiffer.lucas@gmail.com>
To find or get tuberculosis gene expression data, users will use the `tuberculosis` function. The `dryrun` argument allows users to test a query prior to returning resources. When `dryrun = TRUE`, the function will print the names of matching resources as a message and return them invisibly as a character vector. When `dryrun = FALSE`, the function will either download resources from ExperimentHub or load them from the user’s local cache. If a resource has multiple creation dates, the most recent is selected by default; add a date to override this behavior.

**Usage**

`tuberculosis(pattern, dryrun = TRUE)`

**Arguments**

- `pattern`  
  regular expression pattern to look for in the titles of resources available in tuberculosis; `"."` will return all resources
- `dryrun`  
  if `TRUE` (the default), a character vector of resource names is returned invisibly; if `FALSE`, a list of resources is returned

**Value**

- if `dryrun = TRUE`, a character vector of resource names is returned invisibly; if `dryrun = FALSE`, a list of resources is returned

**Examples**

`tuberculosis("GSE103147")`
Index

tuberculosis, 2