Package ‘rRDPData’

April 8, 2025

Title Databases for the Default RDP Classifier

Description The package provides the data for the RDP Classifier 2.14 released in August 2023. It contains the latest bacterial and archaeal taxonomy training set No. 19 as described in Wang Q, Cole JR. 2024. Updated RDP taxonomy and RDP Classifier for more accurate taxonomic classification. Microbiol Resour Announc 0:e01063-23. <doi.org/10.1128/mra.01063-23>

Version 1.27.0

Date 2024-03-26

biocViews SequencingData, MicrobiomeData, RNASeqData

Bugreports https://github.com/mhahsler/rRDP/issues

URL `https://github.com/mhahsler/rRDP`

Depends rRDP

License GPL-2

git_url https://git.bioconductor.org/packages/rRDPData

git_branch devel

git_last_commit bc56232

git_last_commit_date 2024-10-29

Repository Bioconductor 3.21

Date/Publication 2025-04-08

Author Michael Hahsler [aut, cre] (ORCID: <`https://orcid.org/0000-0003-2716-1405`>), Nagar Anurag [aut]

Maintainer Michael Hahsler <`mhahsler@lyle.smu.edu`>

Contents

16SrRNA 2

Index **4**

`16SrRNA` *RDP Classifier Model for RDP 16S rRNA Trainingset No 19*

Description

Model data for he RDP Classifier 2.14 released in August 2023. The classifier is trained with the latest bacterial and archaeal taxonomy training set No. 19. See Wang and Cole (2024).

16SrRNA is the default model data used by the classifier in package ‘rRDP‘.

Details

`rdp()` creates a default classifier using this data set.

The model data is stored in `system.file("extdata", package="rRDPData")` and consists of several files in proprietary format used by the RDP software. Since RDP is a naive Bayes classifier the data mainly contain conditional word (8-mer) probabilities.

References

RDP Classifier `http://sourceforge.net/projects/rdp-classifier/`

Qiong W. and Cole J.R. Updated RDP taxonomy and RDP Classifier for more accurate taxonomic classification, Microbial Ecology, Announcement, 4 March 2024. doi:10.1128/mra.0106323

Examples

```
### Use the RDP with the default classifier model data
cl_16S <- rdp()
cl_16S

seq <- readRNAStringSet(system.file("examples/RNA_example.fasta",
 package="rRDP"))

## shorten names
names(seq) <-  sapply(strsplit(names(seq), " "), "[", 1)
seq

## use RDP for classification
pred <- predict(cl_16S, seq)
pred

attr(pred, "confidence")

### Query a different model

## List all available models
dir(system.file("extdata", package="rRDPData"))

cl_fungus <- rdp(system.file("extdata", "fungallsu", package="rRDPData"))
```

```
cl_fungus

pred <- predict(cl_fungus, seq)
pred

attr(pred, "confidence")
```

Index

∗ **model**
    `16SrRNA`, 2
`16S` *(*`16SrRNA`*)*, 2
`16SrRNA`, 2

`fungalits_unite` *(*`16SrRNA`*)*, 2
`fungalits_warcup` *(*`16SrRNA`*)*, 2
`fungallsu` *(*`16SrRNA`*)*, 2