Package ‘DNAcycP2’

April 24, 2025

Title DNA Cyclizability Prediction

Version 1.1.0

Description This package performs prediction of intrinsic cyclizability of of every 50-bp subsequence in a DNA sequence. The input could be a file either in FASTA or text format. The output will be the C-score, the estimated intrinsic cyclizability score for each 50 bp sequences in each entry of the sequence set.

Depends R (>= 4.4.0)

Imports basilisk, reticulate

License Artistic-2.0

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

StagedInstall yes

BinaryFiles inst/extdata/irlstm/saved_model.pb, inst/extdata/irlstm_smooth/saved_model.pb

biocViews NeuralNetwork, StructuralPrediction

VignetteBuilder knitr

Suggests knitr, rmarkdown, BiocGenerics, RUnit, tinytest, BiocStyle, Biostrings

LazyData false

URL `https://github.com/jipingw/DNAcycP2`

BugReports `https://github.com/jipingw/DNAcycP2`

git_url https://git.bioconductor.org/packages/DNAcycP2

git_branch devel

git_last_commit 8f14169

git_last_commit_date 2025-04-15

Repository Bioconductor 3.22

Date/Publication 2025-04-24

Author Ji-Ping Wang [aut, cre] (ORCID: <`https://orcid.org/0000-0002-8398-939X`>)

Maintainer Ji-Ping Wang <`jzwang@northwestern.edu`>

Contents

`cycle` *Predict Cyclizability*

Description

This predicts cyclizability for a set of sequences.

Usage

```
cycle(sequences, smooth, save_path_prefix = "")
```

Arguments

`sequences`	A list or vector of sequences
`smooth`	Whether to predict smoothed C0 (DNAcycP2) or original C0 (DNAcycP)
`save_path_prefix`	Base path for output files. If it is an empty string, the output files will not be saved (default="")

Details

Optionally, saves output files (use argument 'save_path_prefix')

Value

A list of predictions for each input sequence.

Examples

```
# Example usage of cycle
cycle(c("ACTGCTAGTCACTGCTAGTCACTGCTAGTCACTGCTAGTCACTGCTAGTC"), smooth=TRUE)
# where sequences is a list/vector of sequences
```

`cycle_fasta` *Predict Cyclizability*

Description

This predicts cyclizability for all subsequences of length 50bp from a .fasta input file.

Usage

```
cycle_fasta(
  file_path,
  smooth,
  n_cores = 1,
  chunk_length = 1e+05,
  save_path_prefix = ""
)
```

Arguments

`file_path`	.fasta input file path
`smooth`	Whether to predict smoothed C0 (DNAcycP2) or original C0 (DNAcycP)
`n_cores`	Number of cores to use for parallel processing (default=1)
`chunk_length`	Length of sequence that each core will predict on at a given time. (default=100000)
`save_path_prefix`	Base path for output files. If it is an empty string, the output files will not be saved (default="")

Details

Optionally, saves output files (use argument 'save_path_prefix')

Value

A list of predictions for each ID in the .fasta file.

Each list item has the following columns: position, c_score_norm (predictions on a normalized scale), and c_score_unnorm (predictions on an unnormalized scale).

Each list item is named "cycle_id" corresponding to the fasta id

Examples

```
# Create a temporary file
temp_file <- tempfile(fileext = ".fasta")
writeLines(">1", temp_file)
writeLines("ACTGCTAGTCACTGCTAGTCACTGCTAGTCACTGCTAGTCACTGCTAGTC", temp_file)

# Example usage of cycle_fasta
cycle_fasta(temp_file, smooth=TRUE)

# Cleanup
unlink(temp_file)
```

Index