Package: OpEnCAST (via r-universe)

May 9, 2026

Type Package

Title Optimized Ensemble Model for C and A Methylation Search in Plant

Version 0.1.1

Author Abhik Sarkar [aut, cre], Dipro Sinha [aut], Sneha Murmu [aut], Md Yeasin [aut], Dwijesh Chandra Mishra [aut], Sunil Archak [aut]

Maintainer Abhik Sarkar <`abhik.iasri@gmail.com`>

Description DNA methylation is an important epigenetic process that regulates gene activity through chemical modifications of DNA without changing its sequence. 'OpEnCAST' is a plant-specific ensemble-based prediction package that identifies 4mC, 5mC and 6mA methylation sites directly from DNA sequences. It combines multiple machine learning algorithms trained on monocot (Oryza sp.) and dicot (Arabidopsis sp.) reference models to deliver accurate predictions. This methodology is being inspired by the ensemble algorithm for methylation prediction developed by Wang et al. (2022) <doi:10.1186/s12859-022-04756-1>.

Imports Biostrings, seqinr, stringr, tibble, entropy, ftrCOOL, stats

Suggests caret, kernlab, ranger, xgboost, gbm

Encoding UTF-8

License GPL-3

RoxygenNote 7.3.3

NeedsCompilation no

Config/pak/sysreqs libicu-dev zlib1g-dev

Repository https://abhiksarkariasri.r-universe.dev

Date/Publication 2026-01-08 19:20:22 UTC

RemoteUrl https://github.com/cran/OpEnCAST

RemoteRef HEAD

RemoteSha fd03bdcf8c7ac6386bd4087903f0d9627338b850

Contents

`Dicot_MethPred` *DNA Methylation Prediction in Dicot Plants*

Description

Predicting sequences with DNA methylation sites like 4mC or 6mA based on Arabidopsis as reference model.

Usage

```
Dicot_MethPred(fasta_file_path, Reference = "Arabidopsis")
```

Arguments

`fasta_file_path`	Sequence file (.fasta format)
`Reference`	Arabidopsis as Reference Model for Dicot plants

Value

Methylation Status: Sequences with their probable DNA methylation state such as 4mC, 6mA or Non Methhylated.

References

Lv, H., Dao, F. Y., Zhang, D., Guan, Z. X., Yang, H., Su, W., ... & Lin, H. (2020). iDNA-MS: an integrated computational tool for detecting DNA modification sites in multiple genomes. Iscience, 23(4).

Examples

```
library(OpEnCAST)
data<-system.file("exdata/test.fasta", package = "OpEnCAST")
pred<-Dicot_MethPred(fasta_file_path=data, Reference="Arabidopsis")
```

`Monocot_MethPred` *DNA Methylation Prediction in Monocot Plants*

Description

Predicting sequences with DNA methylation sites like 5mC or 6mA based on Rice as reference model.

Usage

```
Monocot_MethPred(fasta_file_path, Reference = "Rice")
```

Arguments

`fasta_file_path`	Sequence file (.fasta format)
`Reference`	Rice as Reference Model for Monocot plants

Value

Methylation Status: Sequences with their probable DNA methylation state such as 5mC, 6mA or Non Methhylated.

References

Lv, H., Dao, F. Y., Zhang, D., Guan, Z. X., Yang, H., Su, W., ... & Lin, H. (2020). iDNA-MS: an integrated computational tool for detecting DNA modification sites in multiple genomes. Iscience, 23(4).

Examples

```
library(OpEnCAST)
data<-system.file("exdata/test.fasta", package = "OpEnCAST")
pred<-Monocot_MethPred(fasta_file_path=data, Reference="Rice")
```

Index