

Package ‘AmyloGram’

July 21, 2025

Type Package
Title Prediction of Amyloid Proteins
Version 1.1
LazyData TRUE
Date 2017-10-11
Description Predicts amyloid proteins using random forests trained on the
n-gram encoded peptides. The implemented algorithm can be accessed from
both the command line and shiny-based GUI.
License GPL-3
URL <https://github.com/michbur/AmyloGram>
BugReports <https://github.com/michbur/AmyloGram/issues>
RoxxygenNote 6.0.1
Depends R (>= 3.0.0)
Imports biogram, ranger, seqinr, shiny
Repository CRAN
NeedsCompilation no
Author Michal Burdukiewicz [cre, aut],
Piotr Sobczyk [ctb],
Stefan Roediger [ctb]
Maintainer Michal Burdukiewicz <michalburdukiewicz@gmail.com>
Date/Publication 2017-10-11 14:46:15 UTC

Contents

AmyloGram-package	2
AmyloGram_gui	2
AmyloGram_model	3
is_protein	3
pep424	4
predict.ag_model	4

print.ag_model	5
read_txt	5
spec_sens	6

Index	7
--------------	----------

AmyloGram-package	<i>Prediction of amyloids</i>
-------------------	-------------------------------

Description

Amyloids are proteins associated with the number of clinical disorders (e.g., Alzheimer’s, Creutzfeldt-Jakob’s and Huntington’s diseases). Despite their diversity, all amyloid proteins can undergo aggregation initiated by 6- to 15-residue segments called hot spots. Henceforth, amyloids form unique, zipper-like beta-structures, which are often harmful. To find the patterns defining the hot spots, we developed our novel predictor of amyloidogenicity AmyloGram, based on random forests.

Details

AmyloGram is available as R function ([predict.ag_model](#)) or shiny GUI ([AmyloGram_gui](#)).
The package is enriched with the benchmark data set [pep424](#).

Author(s)

Maintainer: Michal Burdukiewicz <michalburdukiewicz@gmail.com>

References

Burdukiewicz MJ, Sobczyk P, Roediger S, Duda-Madej A, Mackiewicz P, Kotulska M. (2017) *Amyloidogenic motifs revealed by n-gram analysis*. Scientific Reports 7 <https://doi.org/10.1038/s41598-017-13210-9>

AmyloGram_gui	<i>AmyloGram Graphical User Interface</i>
---------------	---

Description

Launches graphical user interface that predicts presence of amyloids.

Usage

AmyloGram_gui()

Warning

Any ad-blocking software may cause malfunctions.

AmyloGram_model	<i>Random forest model of amyloid proteins</i>
-----------------	--

Description

Random forest grown using the ranger package with additional information.

Format

A list of length three: random forest, a vector of important n-grams and the best-performing encoding.

See Also

[ranger](#)

is_protein	<i>Protein test</i>
------------	---------------------

Description

Checks if an object is a protein (contains letters from one-letter amino acid code).

Usage

```
is_protein(object)
```

Arguments

object character vector where each elements represent one amino acid.

Value

TRUE or FALSE.

 pep424

 pep424 data set

Description

Benchmark dataset for PASTA 2.0. 5 sequences shorter than 6 amino acids (1% of the original dataset) were removed.

Usage

```
pep424
```

Format

a list of 424 peptides (class [SeqFastaAA](#)).

Source

Walsh, I., Seno, F., Tosatto, S.C.E., and Trovato, A. (2014). *PASTA 2.0: an improved server for protein aggregation prediction*. Nucleic Acids Research gku399.

 predict.ag_model

 Predict amyloids

Description

Recognizes amyloids using AmyloGram algorithm.

Usage

```
## S3 method for class 'ag_model'
predict(object, newdata, ...)
```

Arguments

object	ag_model object.
newdata	list of sequences (for example as given by read.fasta).
...	further arguments passed to or from other methods.

Examples

```
data(AmyloGram_model)
data(pep424)
predict(AmyloGram_model, pep424[17])
```

print.ag_model	<i>Print AmyloGram object</i>
----------------	-------------------------------

Description

Prints ag_model objects.

Usage

```
## S3 method for class 'ag_model'  
print(x, ...)
```

Arguments

x	ag_model object.
...	further arguments passed to or from other methods.

Examples

```
data(AmyloGram_model)  
print(AmyloGram_model)
```

read_txt	<i>Read sequences from .txt file</i>
----------	--------------------------------------

Description

Read sequence data saved in text file.

Usage

```
read_txt(connection)
```

Arguments

connection	a connection to the text (.txt) file.
------------	---

Details

The input file should contain one or more amino acid sequences separated by empty line(s).

Value

a list of sequences. Each element has class [SeqFastaAA](#). If connection contains no characters, function prompts warning and returns NULL.

spec_sens	<i>Specificity/sensitivity balance</i>
-----------	--

Description

Sensitivity, specificity and Matthew's Correlation Coefficient of AmyloGram for different cutoffs computed on pep424 dataset.

Usage

spec_sens

Format

a data frame with four columns and 99 rows.

Source

Walsh, I., Seno, F., Tosatto, S.C.E., and Trovato, A. (2014). *PASTA 2.0: an improved server for protein aggregation prediction*. Nucleic Acids Research gku399.

Index

* **datasets**

- AmyloGram_model, [3](#)
- pep424, [4](#)
- spec_sens, [6](#)

* **manip**

- read_txt, [5](#)

* **package**

- AmyloGram-package, [2](#)

AmyloGram (AmyloGram-package), [2](#)

AmyloGram-package, [2](#)

AmyloGram_gui, [2](#), [2](#)

AmyloGram_model, [3](#)

connection, [5](#)

is_protein, [3](#)

pep424, [2](#), [4](#)

predict.ag_model, [2](#), [4](#)

print.ag_model, [5](#)

ranger, [3](#)

read.fasta, [4](#)

read_txt, [5](#)

SeqFastaAA, [4](#), [5](#)

spec_sens, [6](#)