Package: AntAngioCOOL (via r-universe)

September 2, 2024

Type Package

Title Anti-Angiogenic Peptide Prediction

Version 1.2

Date 2016-08-01

Author Babak Khorsand <khorsand@yahoo.com>

Maintainer Javad Zahiri <zahiri@modares.ac.ir>

Description Machine learning based package to predict anti-angiogenic peptides using heterogeneous sequence descriptors. 'AntAngioCOOL' exploits five descriptor types of a peptide of interest to do prediction including: pseudo amino acid composition, k-mer composition, k-mer composition (reduced alphabet), physico-chemical profile and atomic profile. According to the obtained results, 'AntAngioCOOL' reached to a satisfactory performance in anti-angiogenic peptide prediction on a benchmark non-redundant independent test dataset.

License GPL-2

RoxygenNote 5.0.1

Depends caret,rJava,RWeka,rpart, R(>= 2.10.0)

NeedsCompilation no

Repository CRAN

Date/Publication 2016-08-01 14:01:06

Contents

AntAngioCOOL 2

Index **4**

AntAngioCOOL *AntAngioCOOL*

Description

Machine learning based package to predict anti-angiogenic peptides using heterogeneous sequence descriptors.

Usage

```
AntAngioCOOL(Input_Sequence, Classifier = 1, SF = FALSE, AF = FALSE)
```

Arguments

Input_Sequence	A peptide sequence
Classifier	1 if you want to get the prediction from model with maximum Accuracy (according to the independent test reasults), 2 if maximum Sensivity is desired and 3 if maximum Specefity is desired.
SF	if True then all 2343 selected features (out of 175062 features) that had been used for prediction will be returned.
AF	if True then all 175062 extracted features will be returned.

Details

AntAngioCOOL is a machine learning based package to predict anti-angiogenic peptides using heterogeneous sequence descriptors.

This package consists of three different predictors according to the obtained performances on independent test set: sensitive model, specific model and accurate model. These models have been build using the gold standard dataset that published by Ramaprasad et al. (Ettayapuram Ramaprasad et al., 2015).

Four different features have been used to encode peptides:

1- Pseudo Amino Acid Composition (PseAAC) that has been used effectively in predicting cell penetrating peptides (Chen, Chu, Huang, Kong, & Cai, 2015). Despite the simple amino acid composition, PseAAC considers the sequence-order information of the peptide.

2- K-mer composition that shows the fraction of all possible subsequences with length k in the given peptide. To compute k-mer composition features, reduced amino acid alphabet that proposed by Zahiri et al (Zahiri et al., 2014) has been exploited: the 20 alphabet of amino acids have been reduced to a new alphabet with size 8 according to 544 physicochemical and biochemical indices that extracted from AAIndex database (Kawashima et al., 2008) (C1=A, E, C2=I, L, F, M, V, C3=N, D, T, S, C4=G, C5=P, C6=R, K, Q, H, C7=Y, W, C8=C). We have computed k-mer composition for k=2,3,4 for each peptide.

3- Physico-chemical profile : To compute these features, 544 different physico-chemical indices have been extracted from AAINDEX (Kawashima et al., 2008). To remove redundant indices, a subset of indices with correlation coefficient less than 0.8 and greater than -0.8 has been selected. This feature type for 5 amino acids of N- termini (5-NT) and C-termini (5-CT).

4- Atomic profile : A 50-dimentional feature vector has been used to encode each peptide according to its atomic properties (frequency of five types of atoms: C, H, N, O, S). For details of atomic composition for each 20 natural amino acid see Kumar et al., 2015.

References

1- Chen, L., Chu, C., Huang, T., Kong, X., & Cai, Y.-D. (2015). Prediction and analysis of cell-penetrating peptides using pseudo-amino acid composition and random forest models. Amino Acids, 47(7), 1485-93. http://doi.org/10.1007/s00726-015-1974-5

2- Ettayapuram Ramaprasad, A. S., Singh, S., Gajendra P. S, R., Venkatesan, S., Brem, S., Cotran, R., . Stephens, R. (2015). AntiAngioPred: A Server for Prediction of Anti-Angiogenic Peptides. PLOS ONE, 10(9), e0136990. http://doi.org/10.1371/journal.pone.0136990

3- Kawashima, S., Pokarowski, P., Pokarowska, M., Kolinski, A., Katayama, T., & Kanehisa, M. (2008). AAindex: amino acid index database, progress report 2008. Nucleic Acids Research, 36(Database issue), D202-5. http://doi.org/10.1093/nar/gkm998

3- Kumar, R., Chaudhary, K., Singh Chauhan, J., Nagpal, G., Kumar, R., Sharma, M., & Raghava, G. P. S. (2015). An in silico platform for predicting, screening and designing of antihypertensive peptides. Scientific Reports, 5, 12512. http://doi.org/10.1038/srep12512

4- Zahiri, J., Mohammad-Noori, M., Ebrahimpour, R., Saadat, S., Bozorgmehr, J. H., Goldberg, T., & Masoudi-Nejad, A. (2014). LocFuse: Human protein-protein interaction prediction via classifier fusion using protein localization information. Genomics, 104(6), 496-503.

Value

If Predicted class (Anti-angiogenic/ Not anti-angiogenic) of the input peptide and a subset of descriptors upon request.

Author(s)

Babak Khorsand

Examples

```
AntAngioCOOL("AAPFLECQGN",2,SF=TRUE)
```

Index

AntAngioCOOL, 2