Computational method for predicting Self-Interactions Protein using Recurrent Neural Network from Protein Evolutionary Information

CURRENT STATUS: POSTED

Ji-Yong An

Corresponding Author

ajy@cumt.edu.cn

DOI:
10.21203/rs.2.23742/v1

SUBJECT AREAS
General Cell Biology & Physiology

KEYWORDS
SIPs, Recurrent Neural Network, Scale Invariant Feature Transform, PSSM
Abstract
Self-interactions Protein (SIPs) play crucial roles in biological activities of organisms. Many high-throughput methods can be used to identify SIPs. However, these methods are both time-consuming and expensive. How to develop effective computational approaches for identifying SIPs is a challenging task. In the paper, we presented a novelty computational method called RRN-SIFT, which combines the Recurrent Neural Network (RNN) with Scale Invariant Feature Transform (SIFT) to predict SIPs based on protein evolutionary information. The main advantage of the proposed RNN-SIFT model is that it used SIFT for extracting key feature by exploring the evolutionary information embedded in PSI-BLAST-constructed position specific scoring matrix (PSSM) and employed RNN classifier to carry out classification based on extracted features. Extensive experiments show that the RRN-SIFT obtained average accuracy of 94.34% and 97.12% on yeast and human dataset. We also compared our performance with the Back Propagation Neural Network (BPNN), the state-of-the-art support vector machine (SVM) and other exiting methods. By comparing with experimental results, the performance of RNN-SIFT is significantly better than those of the BPNN, SVM and other previous methods in the domain. Therefore, we can come to the conclusion that the proposed RNN-SIFT model is useful tools and can execute incredibly well for predicting SIPs, as well as other bioinformatics tasks. In order to facilitate widely studies and encourage future proteomics research, a freely available web server called RNN-SIFT-SIPs was developed, and is available at http://219.219.62.123:8888/RNNSIFT/ and includes source code and SIPs datasets.

Full-text
Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures
Figure 1
the schematic of a PSSM

Figure 2
the technology roadmap of the proposed method
Figure 3

the structure of RNN

Figure 4

the prediction flowchart of RNN-SIFT
Comparison of ROC curves between RNN, BPNN and SVM on yeast dataset.
Comparison of ROC curves between RNN, BPNN and SVM on human dataset.

Figure 6