Work Report

贾宁欣

07-11-2021

XXX

Abstract

Introduction

Methods

Datasets **Table 1**

Protein representation

Architecture of XXX **Fig. 1**

Performance evaluation

Result and discussion

Select the size of the sliding window.

Performance of XXX based on different **features** and their combinations**. Table 2**

Performance comparison of XXX, … and … over **5-fold cross-validation tests** on PART-388. **Table 3** **Fig. 2**

Performance comparison of XXX and other methods **on the independent test data set**. **Table 4** **Fig. 3**

Analysis of the predicted ATP-binding residues / Postprocessing

Case study / Predictive results visualization

Setting up a web server.

Conclusion

XXX

Abstract

Introduction

localizing protein-ATP binding residues is critical to understanding the interactions between protein and ATP, which is of significant importance for both protein function analysis and drug discovery. ↵

V1.5↵

Adenosine 5'-triphosphate (ATP) is an important small molecule in cells, which is the energy source for organisms to maintain membrane transport, cellular motility, muscle contraction, signaling, replication and transcription of DNA, and various metabolic processes. For protein, ATP as a ligand interacts with it through ATP-binding residues. ATP is hydrolyzed to remove a terminal phosphate group and converted into ADP, and a large amount of energy is released at the same time. The released energy and phosphate group work together to change the protein structure at the ATP-binding residues, thereby further causing changes in protein activity. Previous studies have shown that protein-ligand binding residues tend to form spatial clusters, forming protein-ligand binding sites (pockets), which are valuable drug targets for antibacterial and anti-cancer chemotherapy. Hence, accurately localizing protein-ATP binding residues is critical to understanding the interactions between protein and ATP, which is of significant importance for both protein function analysis and drug discovery. ↵

↵

ATP 是什么↵

对于 protein，ATP 作为其 ligand 通过 ATP-binding residues 与其相互作用↵

ATP 对蛋白质起作用的方式：hydrolyze↵

绑定残基空间聚类形成 binding sites (pockets)↵

定位 protein-ligand binding residues 很重要↵

↵

D2.D3.D4 研究方法概述： ↵

实验方法↵

计算方法：↵

基于序列的↵

基于结构的↵

基于序列和结构的（本文 TM-SITE [结构] 和 I-LBR [序列]）↵

↵

D2↵

V1.1↵

Experimental methods are the most accurate techniques to identify the protein-ATP binding residues, but it relies too much on manpower and equipment resources, and is expensive and time-consuming. As the alternate techniques, computational methods have been used successfully for predicting the protein-ligand binding residues during the past decade. For the prediction of the protein-ATP binding residues, the highest Matthews correlation coefficient (MCC) value reached 0.685. According to the different types of information on

XXX

Abstract

Introduction

Methods

Datasets **Table 1**

Protein representation

Architecture of XXX

Performance evaluation

Table 1

Statistical composition of the two data sets used in this study.

Data set	$N_{pro}{}^{a}$	$N_{pos}{}^{b}$	$N_{neg}{}^{c}$	$PNratio^{d}$
PART-388	388	5657	142,086	1: 25.12
PART-TEST	41	674	14,159	1: 21.01

[a] Number of proteins.

[b] Number of ATP-binding residues.

[c] Number of non-ATP-binding residues.

[d] $PNratio = N_{pos} : N_{neg}$.

PART-388 and PART-TEST are two widely used benchmark data sets constructed in our previous work [ ].

XXX

Result and discussion

Select the size of the sliding window. #

Performance of XXX based on different **features** and their combinations**.** **Table 2**

Performance comparison of XXX, … and … over **5-fold cross-validation tests** on PART-388. **Table 3**

Performance comparison of XXX and other methods **on the independent test data set**. **Table 4**

#

A sliding window with size W (centered at the target residue) is utilized, and inspired by ATPint, we set $W = 17$ in this study.

Table 2

Performance of XXX based on different features and their combinations.

feature	Thres	Sen(%)	Spe(%)	Acc(%)	Pre(%)	MCC	F1-sco	AUC
PSFM	0.14	55.93	98.93	96.97	71.27	0.616	0.627	0.844
TMSITE	0.32	63.35	99.05	97.43	76.11	0.681	0.691	0.917
ILBR	0.38	51.93	99.27	97.11	77.09	0.619	0.621	0.892
PSFM+ TMSITE	0.28	63.95	98.96	97.37	74.57	0.677	0.688	0.884
PSFM+ILBR	0.41	55.64	99.10	97.13	74.70	0.630	0.638	0.855
TMSITE+ILBR	0.34	66.02	99.17	97.67	79.18	**0.711**	0.720	**0.932**
PSFM+TMSITE+LIBR	0.56	63.50	99.04	97.42	75.89	0.681	0.691	0.889

Note: The best results are highlighted in bold type.

XXX

Result and discussion

Select the size of the sliding window.

Performance of XXX based on different **features** and their combinations**.** **Table 2**

Performance comparison of XXX, … and … over **5-fold cross-validation tests** on PART-388. **Table 3**

Performance comparison of XXX and other methods **on the independent test data set**. **Table 4**

Table 3

Performance comparison of XXX, S-SITEatp, TM-SITEatp, ATPseq, ATPbind and DeepATPseq over 5-fold cross-validation tests on PART-388.

Predictors	Sen(%)	Spe(%)	Acc(%)	Pre(%)	MCC	AUC
S-SITEatp [a]	69.88	94.47	93.53	33.47	0.455	N/A
TM-SITEatp [a]	73.64	95.29	94.46	38.37	0.507	N/A
ATPseq [a]	57.52	98.86	97.27	66.69	0.605	0.913
ATPbind [a]	64.04	98.88	97.55	69.57	0.655	0.932
DeepATPseq [b]	52.20	99.03	97.39	69.75	0.613	
XXX			待做			

[a] Data excerpted from Ref. [ ]. (ATPbind 2018)

[b] Data excerpted from Ref. [ ]. (DeepATPseq 2021)

XXX

Result and discussion

Select the size of the sliding window.

Performance of XXX based on different **features** and their combinations**.** **Table 2**

Performance comparison of XXX, … and … over **5-fold cross-validation tests** on PART-388. **Table 3**

Performance comparison of XXX and other methods **on the independent test data set**. **Table 4**

Table 4

Performance comparison of XXX and other computational methods on the independent test data set.

Model	Sen(%)	Spe(%)	Acc(%)	Pre(%)	MCC	AUC
S-SITEatp [a]	67.51	92.65	91.51	30.41	0.416	N/A [#]
TM-SITEatp [a]	69.73	96.09	84.89	45.90	0.541	N/A
NsitePred [a]	46.74	97.70	95.39	49.22	0.456	0.852
TargetATPsite [a]	41.25	99.49	96.84	79.43	0.559	0.853
TargetS [a]	51.63	98.89	96.74	68.91	0.580	0.872
TargetSOS [a]	49.26	99.46	97.18	81.37	0.620	0.863
TargetNUCs [a]	46.88	99.66	97.26	86.81	0.627	0.856
ATPseq [a]	54.45	99.27	97.24	78.09	0.639	0.878
ATPbind [a]	62.31	98.85	97.19	72.04	0.656	0.915
DELIA [b]	62.17	98.67	97.01	69.03	0.640	
DELIA [c]				0.758	0.685	0.947
DeepATPseq [b]	57.42	99.22	97.32	77.71	0.655	
XXX	66.02	99.17	97.67	79.18	**0.711**	0.932

[a] Data excerpted from Ref. [ ]. (ATPbind 2018)

[b] Data excerpted from Ref. [ ]. (DeepATPseq 2021)

[c] Data excerpted from Ref. [ ]. (DELIA 2020)

[#] N/A, the corresponding value could not be computed.

XXX

Abstract

Introduction

Methods

Datasets **Table 1**

Protein representation

Architecture of XXX **Fig. 1**

Performance evaluation

Result and discussion

Select the size of the sliding window.

Performance of XXX based on different **features** and their combinations**. Table 2**

Performance comparison of XXX, … and … over **5-fold cross-validation tests** on PART-388. **Table 3** **Fig. 2**

Performance comparison of XXX and other methods **on the independent test data set**. **Table 4** **Fig. 3**

Analysis of the predicted ATP-binding residues / Postprocessing

Case study / Predictive results visualization

Setting up a web server.

Conclusion

Feature Works

1\ Continue to write the paper, including drawing pictures and doing experiments.

2\ Postprocessing.