基于人工智能与多源信息融合的蛋白质功能预测研究

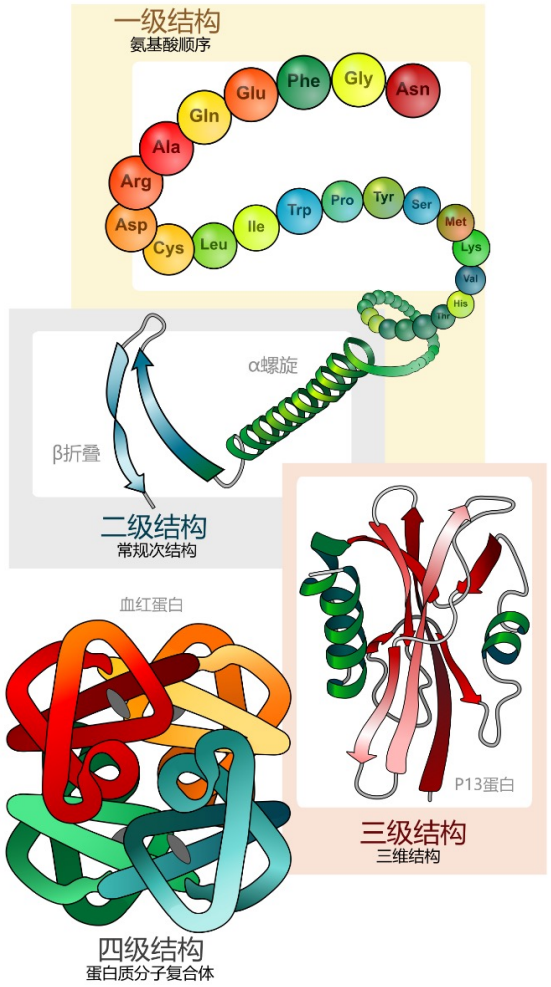
南京农业大学 人工智能学院
汇报人：朱一亨
2023年10月31日

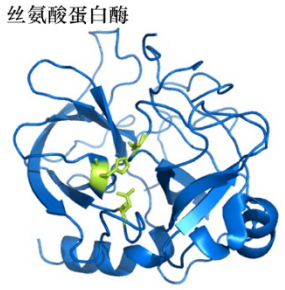
目 录

01 Part one

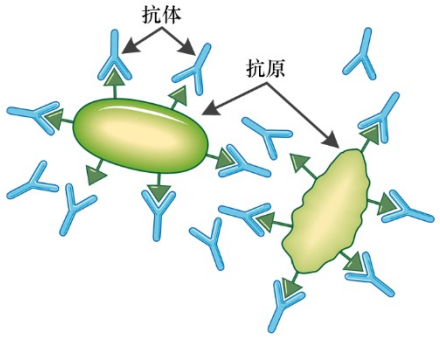
研究背景

01 蛋白质的生物功能



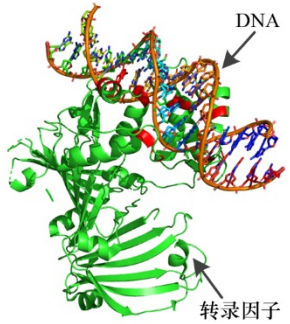


(a) 催化反应



(b) 免疫保护

(c) 运输载体



(d) 基因调控

- 识别和分析蛋白质的功能有助于解释各种生命活动现象，并阐明相关疾病的发病机理，进而指导相应的药物设计，以期推动智能医疗的发展。

- 蛋白质功能注释是后基因时代的首要任务之一。

02 蛋白质的功能注释方法

- 基因本体论 (Gene Ontology, GO)

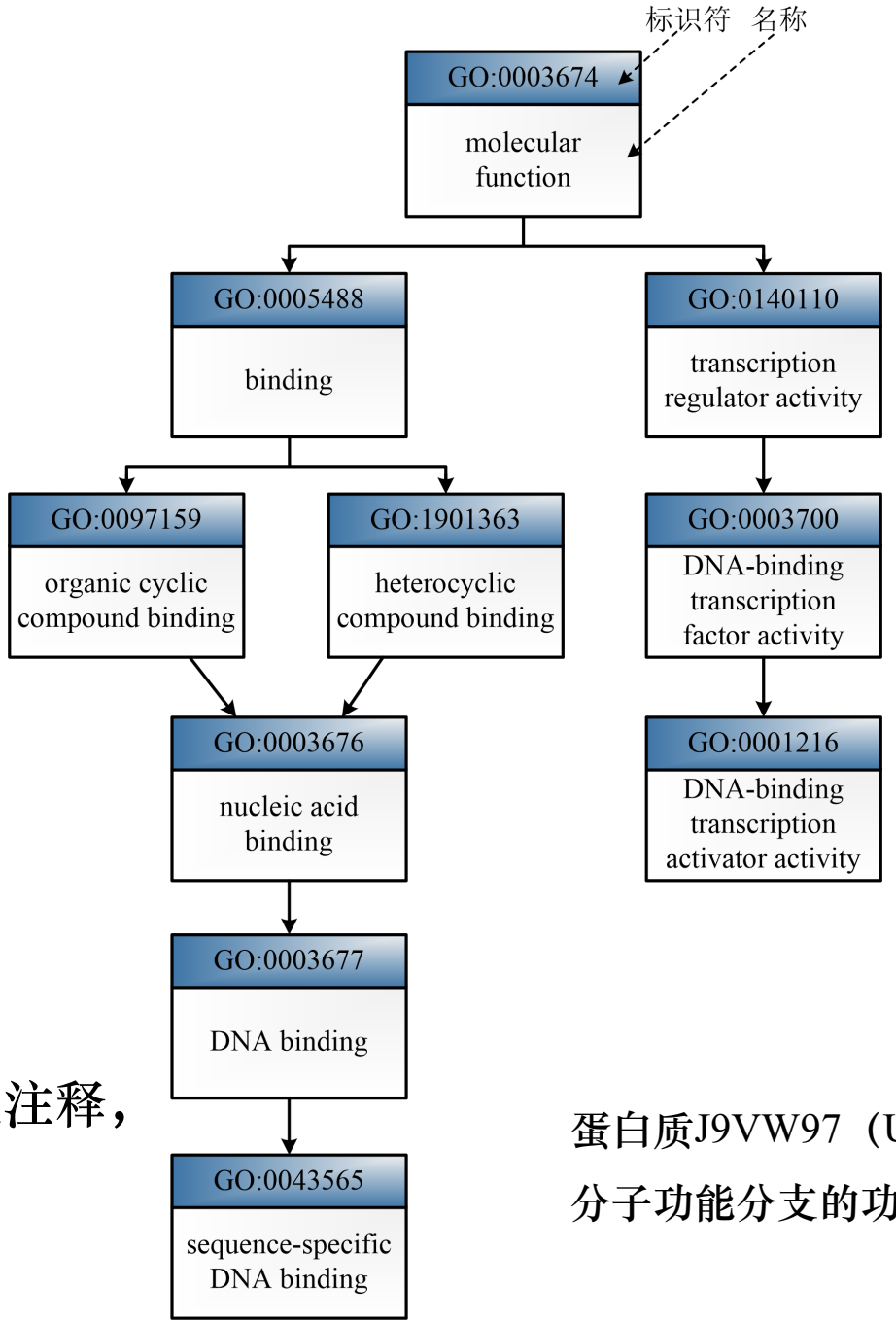
  分子功能 (Molecular Function, MF)

  生物过程 (Biological Process, BP)

  细胞组件 (Cellular Component, CC)

- 蛋白质功能注释目标

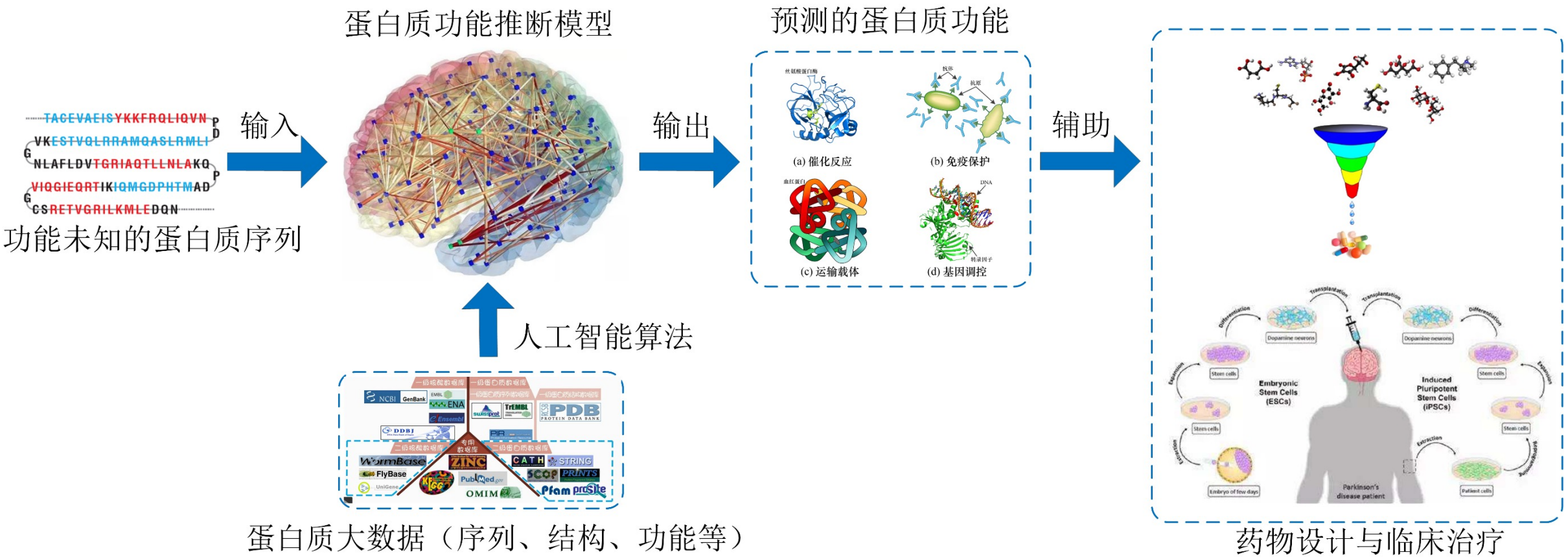
  用GO术语对蛋白质在三个分支下分别进行功能注释，形成三张有向无环图。



蛋白质J9VW97（UniProt ID）在分子功能分支的功能注释图

03 人工智能算法预测蛋白质功能

- 蛋白质功能注释最可靠的途径是生物实验，但它存在周期长、成本高等缺陷。
- 研发高效的人工智能算法来预测蛋白质功能已迫在眉睫（多标签预测任务）。



02 Part two

研究内容

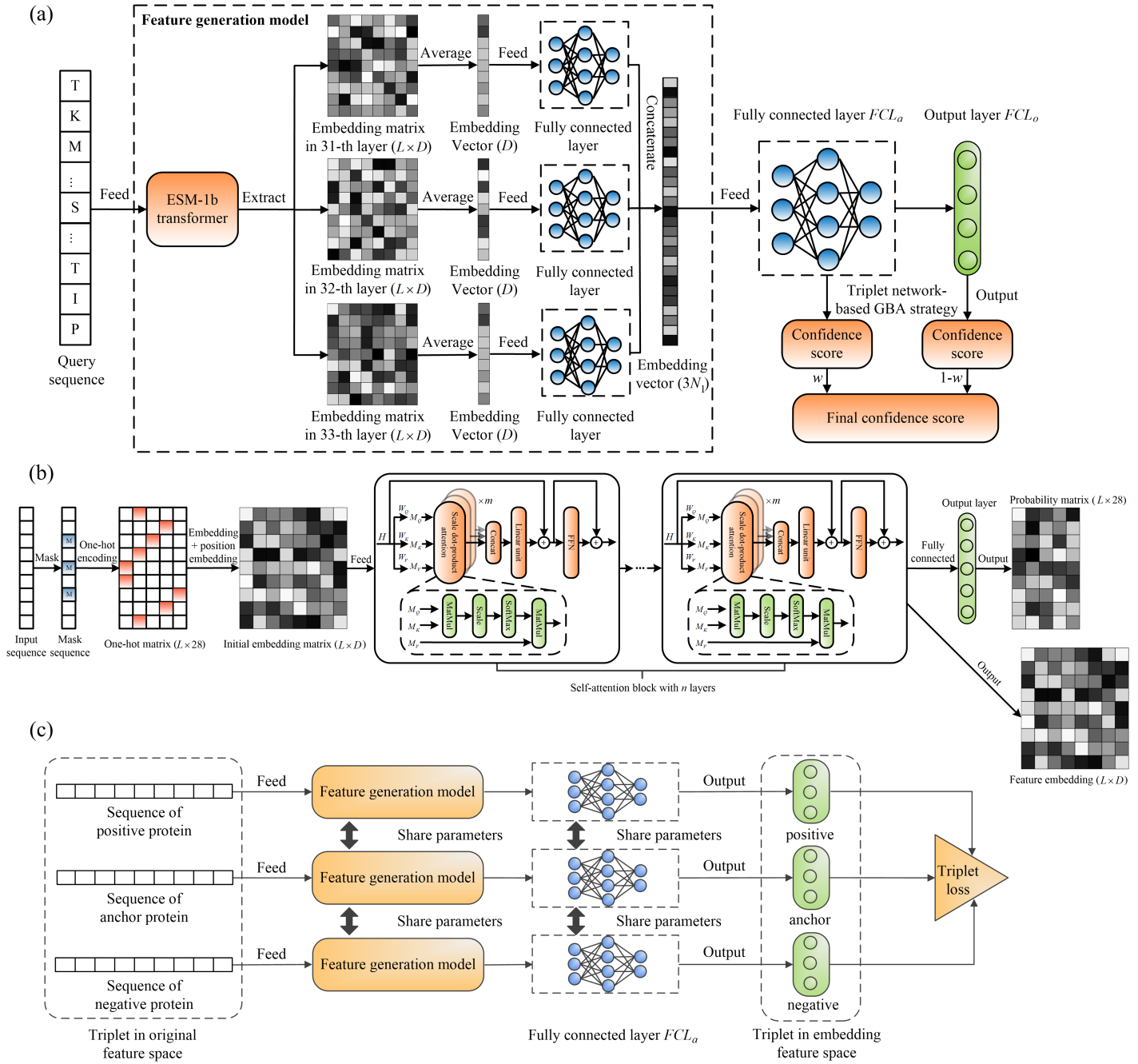
蛋白质功能预测

- 从蛋白质视角出发预测功能
- 从基因视角出发预测功能
- 从配体视角出发预测功能

从蛋白质视角出发预测功能

- 基于注意力机制与三元组神经网络的预测方法 ATGO

- 主要贡献：首次将计算机视觉领域的无监督语言模型迁移到蛋白质功能预测领域



Yi-Heng Zhu, Chengxin Zhang, Dong-Jun Yu, Yang Zhang. Integrating Unsupervised Language Model with Triplet Neural Networks for Protein Gene Ontology Prediction. **PLOS Computational Biology**. 2022, 18(12): e1010793.

Zhang Lab

UNIVERSITY OF MICHIGAN

Home | Research | COVID-19 | Services | Publications | People | Teaching | Job Opening | News | Forum | Lab Only

Online Services
- I-TASSER
- QUARK
- LOMETS
- COACH
- COFACTOR
- MetaGO
- MUSTER
- CEthreader
- SEGMER
- FG-MD
- ModRefiner
- REMO
- DEMO
- SPRING
- COTH
- BSpred
- ANGLOR
- EDock
- BSP-SLIM
- SAXSTER
- FUpred
- ThreaDom
- ThreaDomEx
- EvoDesign
- GPCR-I-TASSER
- MAGELLAN
- BindProf
- BindProfX
- SSIPe
- ResQ
- IonCom
- STRUM
- DAMpred
- TM-score

ATGO

Protein Function Prediction

ATGO is a deep learning-based algorithm for high accuracy protein Gene Ontology (GO) prediction. Starting from a query sequence, it first extracts three layers of feature embeddings from a pre-trained protein language model (ESM-1b). Next, a fully connected neural network is used to fuse the feature embeddings, which are then fed into a supervised triplet network for GO function prediction. Large-scale benchmark tests demonstrated significant advantage of ATGO on protein function annotations due to the integration of discriminative feature embeddings from attention transformer models. (view an example of ATGO prediction)

ATGO On-line Server

Input Sequence (Optional, [30,10000] residues in FASTA format)
Copy and paste your protein sequence file here (Sample input)

```
>Q9HGI3
MAYFRLYAVLLAVASSVAAVKVNPLPAPRHISWGHSGPKPLSDVSLRTERDTDDSILTNAWNRAWETIVSLEWVPAGIEA
PIPEFDEFPTSTPSASAAATRSKRANVPIQFVDVDEVEDWDADLQHGVDESYTLDAKAGSDAIDITAKTVWGALHAFTTLQ
QLVISDGNGGLILEQPVHIKDAPLYPYRGLMVDTGRNFISVRKLHEQLDGMALSKLNVLHWHLDDTQSWPVHIDAYPEM
TKDAYSARETYSHDDLRNVVAYARARGIRVIPEIDMPAHSASGWQQVDPDIVACANSWWSNDNWPLHTAVQPNPGQL
DIINPKTYEVVQDVYEELSSIFTDDWFHVGGDEIQPNCYNFSTYVTEWFQEDPSRTYNDLMQHWVDKAVPIFRSVSDSR
RLVMWEDVVLNTEHADDVPTDIVMQSWNNGLENINKLTERGYDVIVSSADFMYLDCGRGGYVTNDDRYNEQTNPDPD
TPSFNYGGIGGSWCGPYKTWQRIYNYDFTLNLTNAQAKHVIGATAPLWSEQVDDVNISNLFWPRAAALAELVWSGNRD
AKGNKRTTLFTQRILNFREYLLANGVMAATVVPKYCLQHPHACDLNYDQTVLH
```

Or upload the sequence file from your local computer
选取文件 未选择文件

Email: (mandatory, where results will be sent to)

Job ID: (optional, your given name to your job)

Run ATGO | Clear form

ATGO Download

- Download the standalone package.
- Download prediction models.
- Download benchmark datasets.

References:

- Yi-Heng Zhu, Chengxin Zhang, Dong-Jun Yu, Yang Zhang. Integrating unsupervised language model with triplet neural networks for protein gene ontology prediction. PLOS Computational Biology, 2022, 18 (12): e1010793.

ATGO result for protein E7CIP7

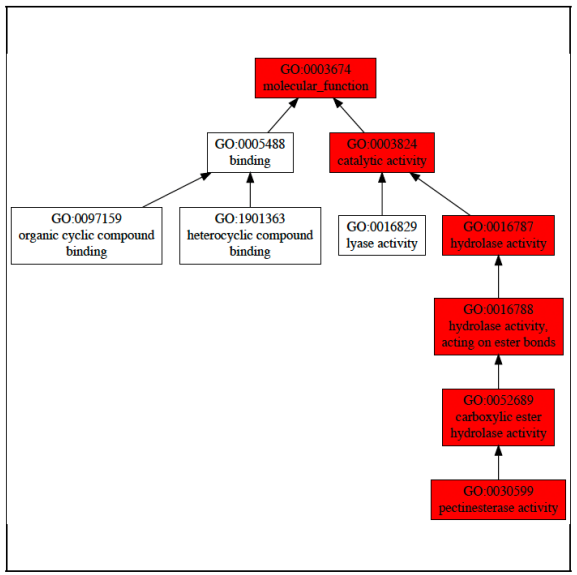
[Download result.zip for all prediction results]

User Input

```
>E7CIP7 (382 residues)
MKIIVLLLLAVVLASADQTAPGTASRPILTASESNYFTTATYLQGWSPPSISTSKADYTV
GNGYNTIQAAVNAAINTGGTTRKYIKINAGTYQEVVYIPNTKVPLTIYGGGSSPSDTLIT
LNMPAQTTPSAYKSLVGSLFNSADPAYSMYNSCASKSGTIGTSCSTVFWVKAPAVQIVNL
SIENSAKNTGDQQAVALQTNSDQIQIHNARLLGHQDTLYAGSGSSSVERSYYTNTYIEGD
IDFVFGGGSAIFESCTFYVKADRRSDTAVVFAPDTDPHKMYGYFVYKSTITGDSAWSSSK
KAYLGRAWDSGVSSSSAYVPGTSPNGQLIIKESTIDGIINTSGPWTTATSGRTYSGNNAN
SRDLNNDNYNRFWEYNNSGNGA
```

Download query sequence

Predicted Gene Ontology (GO) Terms



Molecular Function (MF)

GO term	Cscore^{GO}	Name
GO:0052689	0.982	carboxylic ester hydrolase activity
GO:0016788	0.982	hydrolase activity, acting on ester bonds
GO:0016787	0.982	hydrolase activity
GO:0003824	0.982	catalytic activity
GO:0003674	0.982	molecular_function
GO:0030599	0.935	pectinesterase activity
GO:0016829	0.027	lyase activity
GO:1901363	0.022	heterocyclic compound binding
GO:0097159	0.022	organic cyclic compound binding
GO:0005488	0.022	binding

Download full result of the above consensus prediction.

Click the graph to show a high resolution version.

(a) Cscore^{GO} is the confidence score of predicted GO terms. Cscore^{GO} values range in between [0-1]; where a higher value indicates a better confidence in predicting the function using the template.

(b) The graph shows the predicted terms within the Gene Ontology hierachy for Molecular Function. Confidently predicted terms are color coded by Cscore^{GO}:

[0.40,0.5) [0.5,0.6) [0.6,0.7) [0.7,0.8) [0.8,0.9) [0.9,1.0]

Biological Process (BP)

GO term	Cscore^{GO}	Name
GO:0008150	0.751	biological_process
GO:0071704	0.727	organic substance metabolic process
GO:0044238	0.727	primary metabolic process
GO:0008152	0.727	metabolic process

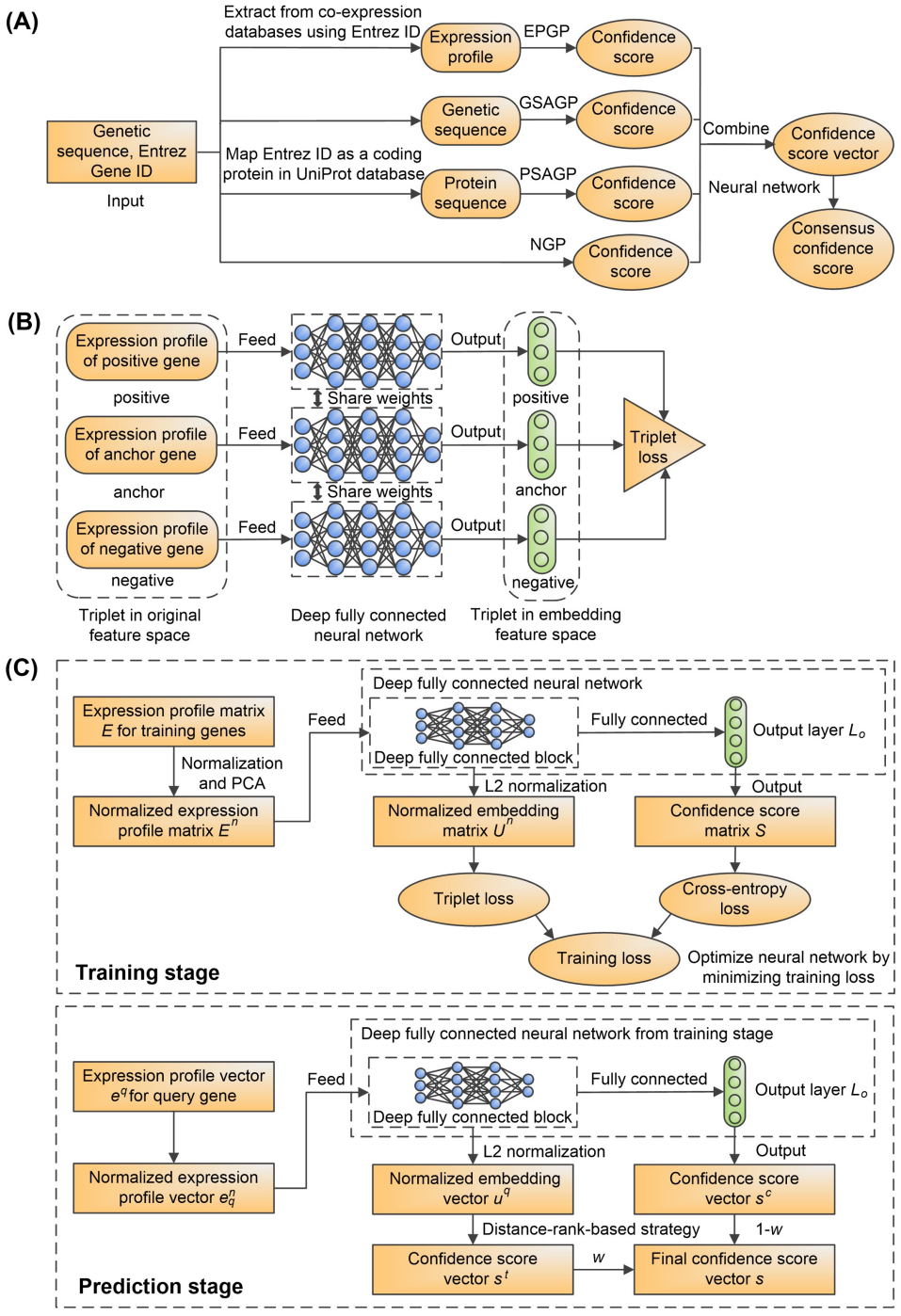
https://zhanggroup.org/ATGO/

从基因视角出发预测功能

- 基于度量学习与多源信息融合的功能预测方法 TripletGO

- 主要贡献：首次将基于基因视角的方法和基于蛋白质视角的相结合，为后续的研究开辟了新的思路

Yi-Heng Zhu, Chengxin Zhang, Yan Liu, Gilbert Omenn, Peter Freddolino, Dong-Jun Yu, Yang Zhang. TripletGO: Integrating Transcript Expression Profiles with Protein Homology Inferences for Gene Function Prediction. Genomics, **Proteomics & Bioinformatics**. 2022, 20(5): 1013-1027.



Zhang Lab

UNIVERSITY OF MICHIGAN

Home | Research | COVID-19 | Services | Publications | People | Teaching | Job Opening | News | Forum | Lab Only

Online Services
- I-TASSER
- I-TASSER-MTD
- C-I-TASSER
- CR-I-TASSER
- QUARK
- C-QUARK
- LOMETS
- MUSTER
- CEthreader
- SEGMER
- DeepFold
- DeepFoldRNA
- FoldDesign
- COFACTOR
- COACH
- MetaGO
- TripletGO
- IonCom
- FG-MD
- ModRefiner
- REMO
- DEMO
- DEMO-EM
- SPRING
- COTH
- Threpp
- PEPPI
- BSpred
- ANGLOR
- EDock
- BSP-SLIM
- SAXSTER
- FUpred
- ThreaDom
- ThreaDomEx
- EvoDesign
- BindProf
- BindProfX
- SSIPe
- GPCR-I-TASSER

TripletGO

Gene Function Prediction

TripletGO is an algorithm for predicting Gene Ontology (GO) of genes. It consists of four pipelines to detect GO terms through (1) expression profile similarity based on triplet network, (2) genetic sequence alignment, (3) protein sequence alignment, and (4) naïve probability. The final function insights are a combination of the four pipelines through neural network. (view an example of TripletGO prediction)

Triplet On-line Server

Sequence of Query Gene (Optional, [30,10000] residues in FASTA format)
Copy and paste your genetic sequence file here (Sample input)
We would suggest you provide Entrez ID for query gene, which helps to find its expression profile and coding proteins.
Entrez ID provides unique integer identifiers for genes in National Center for Biotechnology Information.

```
>839799
GGGCCTATTGGGCTGGAGCCTAGCCCATTTGTGTAGGTTGTGTTAAAACGATGTCGTTTGGCATTTCAAGTTAGG
GTTTTTTGGGGGTTTGGTTCAAGCTTCATCGTCGTCTCTCTGTCTCTTCAATTTCATTCGTTTTCTGAGATAAAAG
TGAGAGAGAAATCTAAATTCGAGAGGAGAAGTTTTAATTTTTCTGAGTTAGATTCAATGGAAGAGATCACGGAAGG
AGTTAACAACATGAACTTGGCTGTTGATACCCAGAAGAAGAATCGGATTCAAGTTTCCAACACTAAGAAACCATTG
TTCTTCTACGTCAATCTCGCCAAGAGGTACATGCAGCAGTACACTGATGTCGAATTGTCTGCACTAGGAATGGCTA
TTGCCACTGTTGTTACGGTCGCTGAGATATTGAAGAACAATGGCTTTGCTGTTGAAAAGAAGATCATGACATCGA
CTGTGGATATCAAGGATGATTCAAGGGGTCGTCCTGTGCAGAAAGCTAAGATTGAGATCACGCTTGCCAAGTCTG
AGAAGTTTGATGAACTAATGGCTGCAGCTAATGAAGAGAAGGAGGCTGCAGAAGCCCAAGAGCAAAACTAGATTG
TTTCAAGTTTTTTCTGTTCAACGATCTTATTTCTTCGTTCCCTATCTCTATCTGCTTAATTTTAAGACACTTCTATTT
CGTTAATTTTTGGTTCACTTTTTTATTTCACCTTGGATTGTGTCCTCTGTACCTCTGAGCATTTTTATTTAAAGATC
```

Or upload the sequence file from your local computer

选取文件 未选择文件

Email: (mandatory, where results will be sent to)

E-value e1 (optional, default 0.1)
The e-value for Blastn software in genetic sequence alignment

E-value e2 (optional, default 0.1)
The e-value for Blastp sofrware in protein sequence alignment

Cut-off value t1 (optional, 0.0-1.0, default 1.0)
The templates which have more than t1 seuqnece identity with the query are removed in genetic sequence alignment

Cut-off value t2 (optional, 0.0-1.0, default 1.0)
The templates which have more than t2 seuqnece identity with the query are removed in protein sequence alignment

Job ID: (optional, your given name to your job)

Run TripletGO | Clear form

TripletGO result for Gene 839799

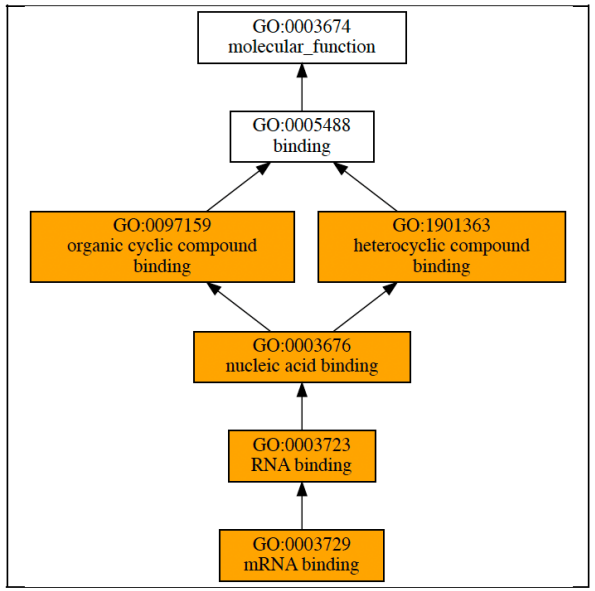
[Download result.zip for all prediction results]

User Input

```
>839799 (795 residues)
GGGCCTATTGGGCTGGAGCCTAGCCCATTTGTGTAGGTTGTGTTAAAACGATGTCGTTTG
GCATTTCAAGTTAGGGTTTTTTGGGGGTTTGGTTCAAGCTTCATCGTCGTCTCTCTGTCT
CTTCAATTTCATTCGTTTTCTGAGATAAAAGTGAGAGAGAAATCTAAATTCGAGAGGAGA
AGTTTTAATTTTTCTGAGTTAGATTCAATGGAAGAGATCACGGAAGGAGTTAACAACATG
AACTTGGCTGTTGATACCCAGAAGAAGAATCGGATTCAAGTTTCCAACACTAAGAAACCA
TTGTTCTTCTACGTCAATCTCGCCAAGAGGTACATGCAGCAGTACACTGATGTCGAATTG
TCTGCACTAGGAATGGCTATTGCCACTGTTGTTACGGTCGCTGAGATATTGAAGAACAAT
GGCTTTGCTGTTGAAAAGAAGATCATGACATCGACTGTGGATATCAAGGATGATTCAAGG
GGTCGTCCTGTGCAGAAAGCTAAGATTGAGATCACGCTTGCCAAGTCTGAGAAGTTTGAT
GAACTAATGGCTGCAGCTAATGAAGAGAAGGAGGCTGCAGAAGCCCAAGAGCAAAACTAG
ATTGTTTCAAGTTTTTTCTGTTCAACGATCTTATTTCTTCGTTCCCTATCTCTATCTGCT
TAATTTTAAGACACTTCTATTTCGTTAATTTTTGGTTCACTTTTTTATTTCACCTTGGAT
TGTGTCCTCTGTACCTCTGAGCATTTTTATTTAAAGATCGTAGGAAGTATAAAAAAGATG
GCTTCGTTGCATAAA
```

Download query sequence

Predicted Gene Ontology (GO) Terms



Molecular Function (MF)

GO term	Cscore^{GO}	Name
GO:1901363	0.886	heterocyclic compound binding
GO:0097159	0.886	organic cyclic compound binding
GO:0003676	0.884	nucleic acid binding
GO:0003723	0.877	RNA binding
GO:0003729	0.874	mRNA binding

Download full result of the above consensus prediction.

Click the graph to show a high resolution version.

(a) Cscore^{GO} is the confidence score of predicted GO terms. Cscore^{GO} values range in between [0-1]; where a higher value indicates a better confidence in predicting the function using the template.

(b) The graph shows the predicted terms within the Gene Ontology hierachy for Molecular Function. Confidently predicted terms are color coded by Cscore^{GO}:

[0.13,0.5) [0.5,0.6) [0.6,0.7) [0.7,0.8) [0.8,0.9) [0.9,1.0]

Biological Process (BP)

GO term	Cscore^{GO}	Name
GO:0009987	0.443	cellular process
GO:0008152	0.231	metabolic process
GO:0071704	0.221	organic substance metabolic

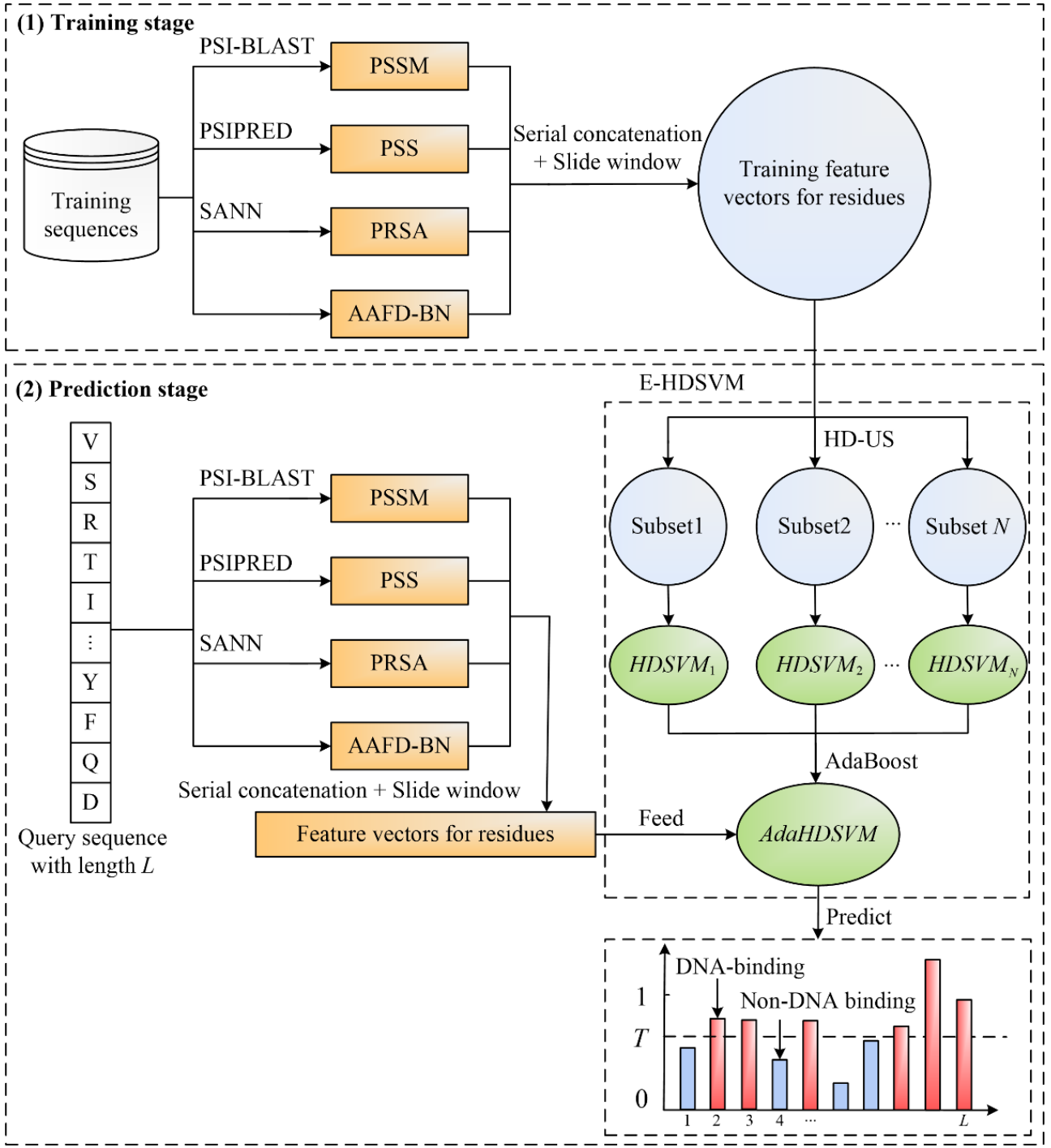
https://zhanggroup.org/TripletGO/

从配体视角出发预测功能

- 基于多粒度支持向量机集成与序列特征的蛋白质-DNA绑定位点预测方法 DNAPred

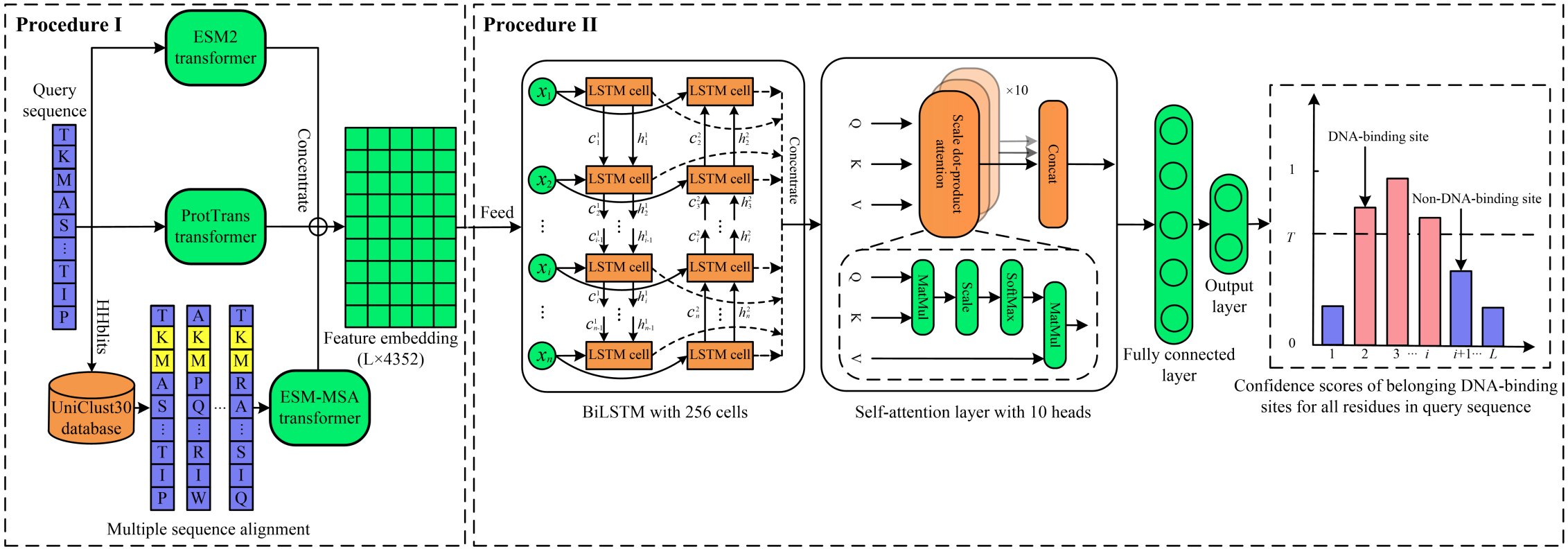
- 主要贡献：提出了新的类不平衡学习算法E-HDSVM，显著地提升了蛋白质-DNA绑定位点预测精度。

Yi-Heng Zhu, Jun Hu, Xiao-Ning Song, Dong-Jun Yu. DNAPred: Accurate Identification of DNA-binding Sites from Protein Sequence by Ensembled Hyperplane-Distance-Based Support Vector Machines. **Journal of Chemical Information and Modeling**. 2019, 59:3057-3071.



从配体视角出发预测功能

- 基于无监督语言模型与多源信息融合的蛋白质-DNA绑定位点预测方法 ULDNA
- 主要贡献：融合多种无监督蛋白质语言模型，显著地提升了蛋白质-DNA绑定位点预测精度。



Yi-Heng Zhu, Zi Liu, Zhiwei Ji, Dong-Jun Yu. ULDNA: Integrating Unsupervised Multi-Source Language Models with LSTM-Attention Network for High-Accuracy Protein-DNA Binding Site Prediction. **Briefings in Bioinformatics**. (Revision)

DNAPred: Identifying DNA-Binding Sites from Protein Sequence by Ensemble Hyperplane-Distance-Based Support Vector Machine

| Read Me | Dataset | Citation | Large-Scale Test |

Input query protein sequence(s) in FASTA format:

```
>2XTNA
MDQNEHSHWGPHAKGQCASRSELRIILVGKTGTGKSAAGNSILRKQAFESKLGS
QTLTKTCSKSQGSWGNREIVIIDTPDMFSWKDHCEALYKEVQRCYLLSAPGPHV
LLLVTQLGRYTSQDQQAAQRVKEIFGEDAMGHTIVLFTHKEDLNGGSLMDYMH
DSDNKALSKLVAACGGRICAFNNRAEGSNQDDQVKELMDCIEDLLMEKNGDHY
TNGLYSLIQRSKCGPVGSDE
```

Example | Reset Sequence(s)

Choose a prediction model

- (•) Model constructed on PDNA-543
- ( ) Model constructed on PDNA-335

Choose a threshold

- (•) Threshold 1 (*Max MCC*)
- ( ) Threshold 2 (*FPR*≈5%)
- ( ) Threshold 3 (*Sen*≈*Spe*)

Email Address (For receiving your prediction results)*

Submit | Clear All

Reference:

Yi-Heng Zhu, Jun Hu, Xiao-Ning Song and Dong-Jun Yu *. DNAPred: Identifying DNA-Binding Sites from Protein Sequence by Ensemble Hyperplane-Distance-Based Support Vector Machine. Journal of Chemical Information and Modeling, 2019.

RESULTS PAGE

Predicting Protein-DNA Binding Sites

Protein Name

2XTNA

Model constructed on Dataset

PDNA-543

Threshold

0.265 (*Max MCC*)

Prediction Summary

Number of predicted DNA-binding residues in protein **2XTNA: 4**

Specific position: **58 T 117 R 119 T 147 H**

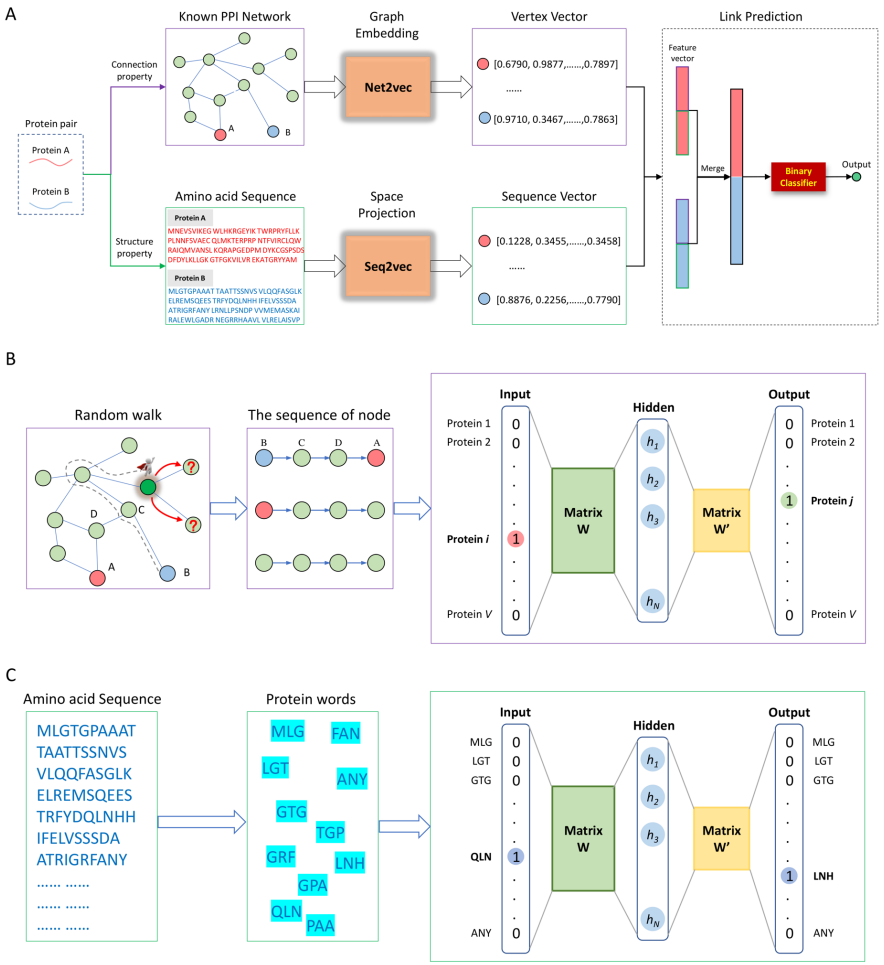
Predicted Results

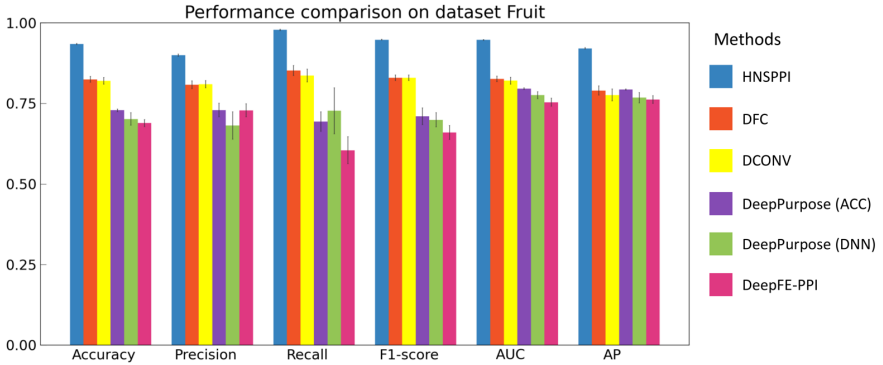
Residue #	Amino Acid Type	Probability	Binding Residue
0001	M	0.058	N
0002	D	0.028	N
0003	Q	0.024	N
0004	N	0.049	N
0005	E	0.013	N
0006	H	0.063	N
0007	S	0.008	N
0008	H	0.037	N
0009	W	0.095	N
0010	G	0.009	N
0011	P	0.019	N
0012	H	0.081	N
0013	A	0.017	N
0014	K	0.080	N
0015	G	0.006	N
0016	Q	0.013	N

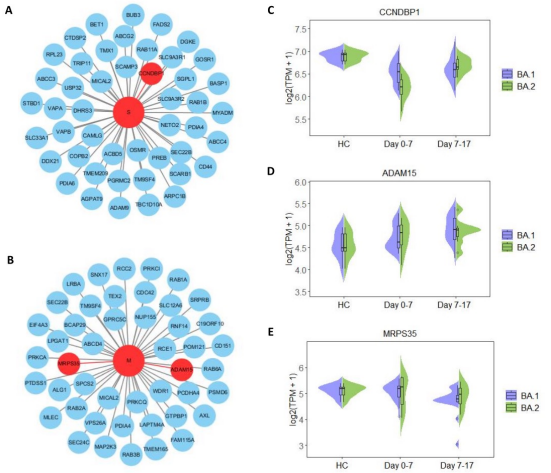
http://csbio.njust.edu.cn/bioinf/dnapred/

蛋白质-蛋白质相互作用预测

- 基于无监督语言模型与多网络融合的蛋白质相互作用预测方法HNSPPI



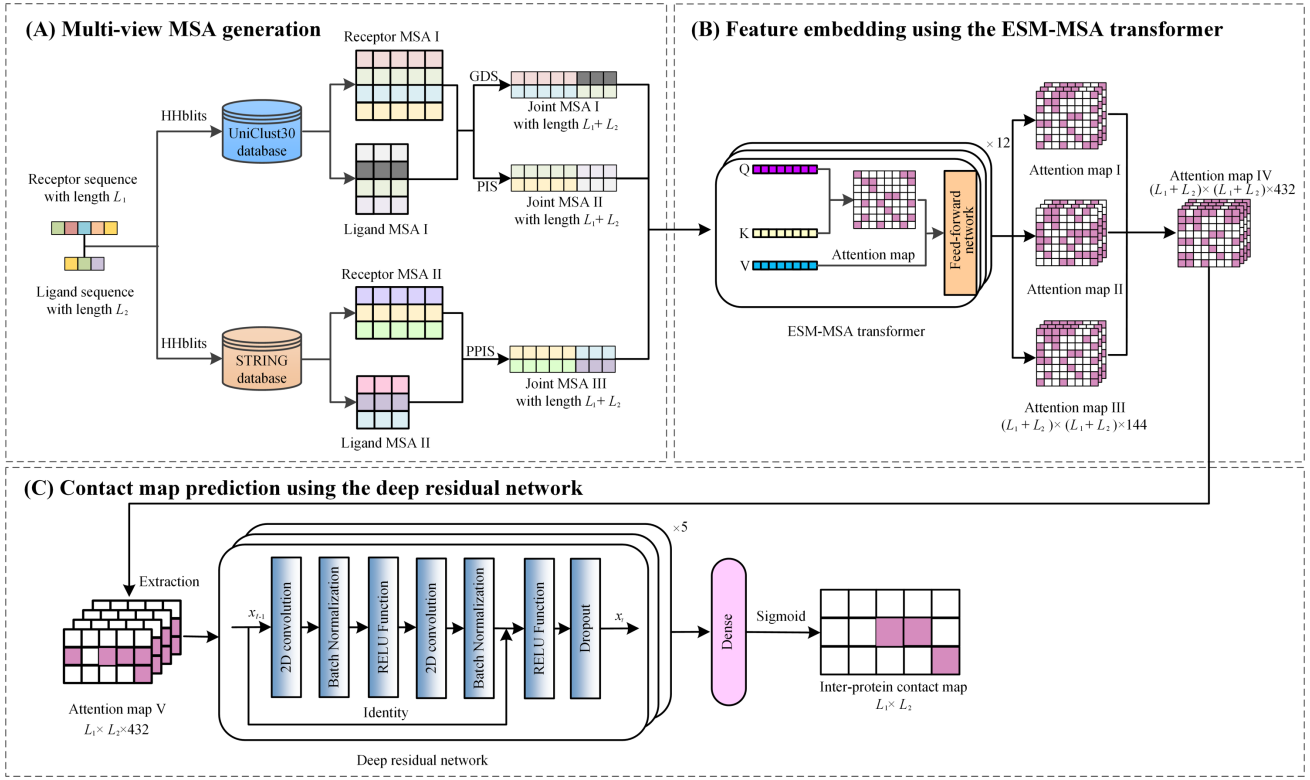


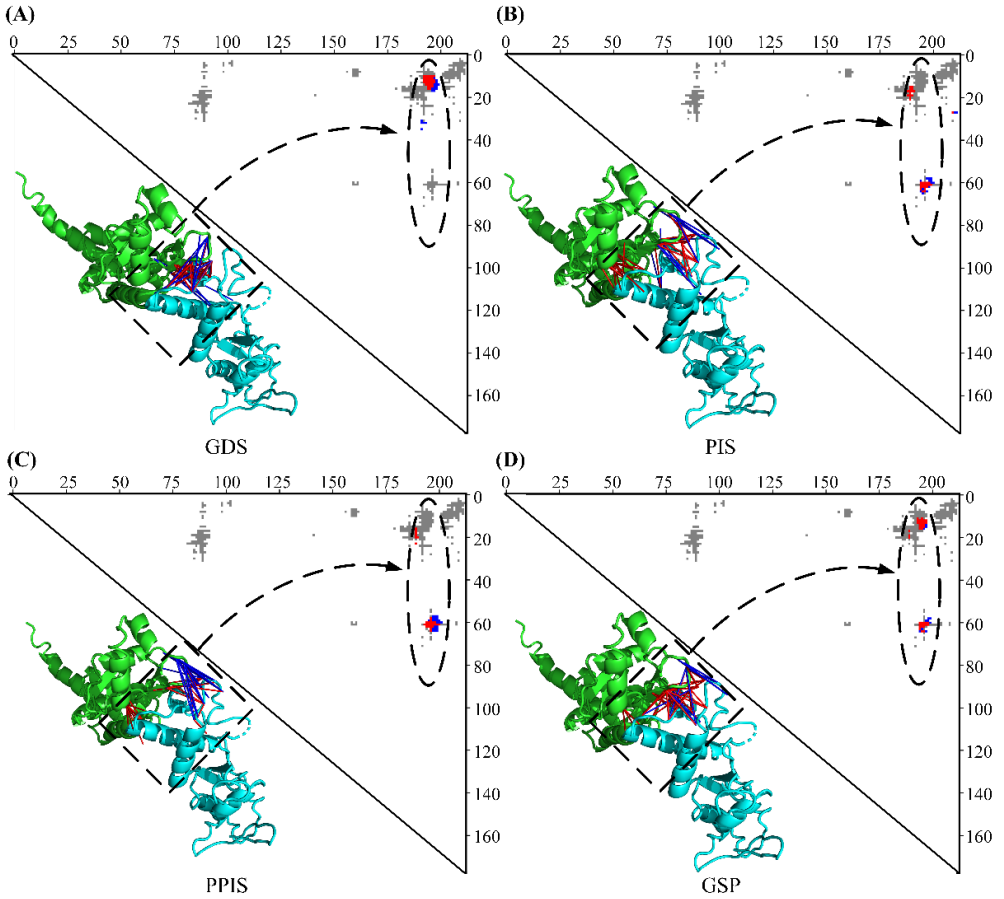


Shijie Xie, Xiaojun Xie, Xin Zhao, Fei Liu, Yiming Wang, Jihui Ping, Zhiwei Ji. HNSPPI: A hybrid computational model combing network and sequence information for predicting protein–protein interaction. **Briefings in Bioinformatics**, 2023, 24(5): bbad261.

蛋白质-蛋白质相互作用预测

➢ 基于无监督语言模型与多视角多序列联配的蛋白质-蛋白质相互作用预测方法ICCPred



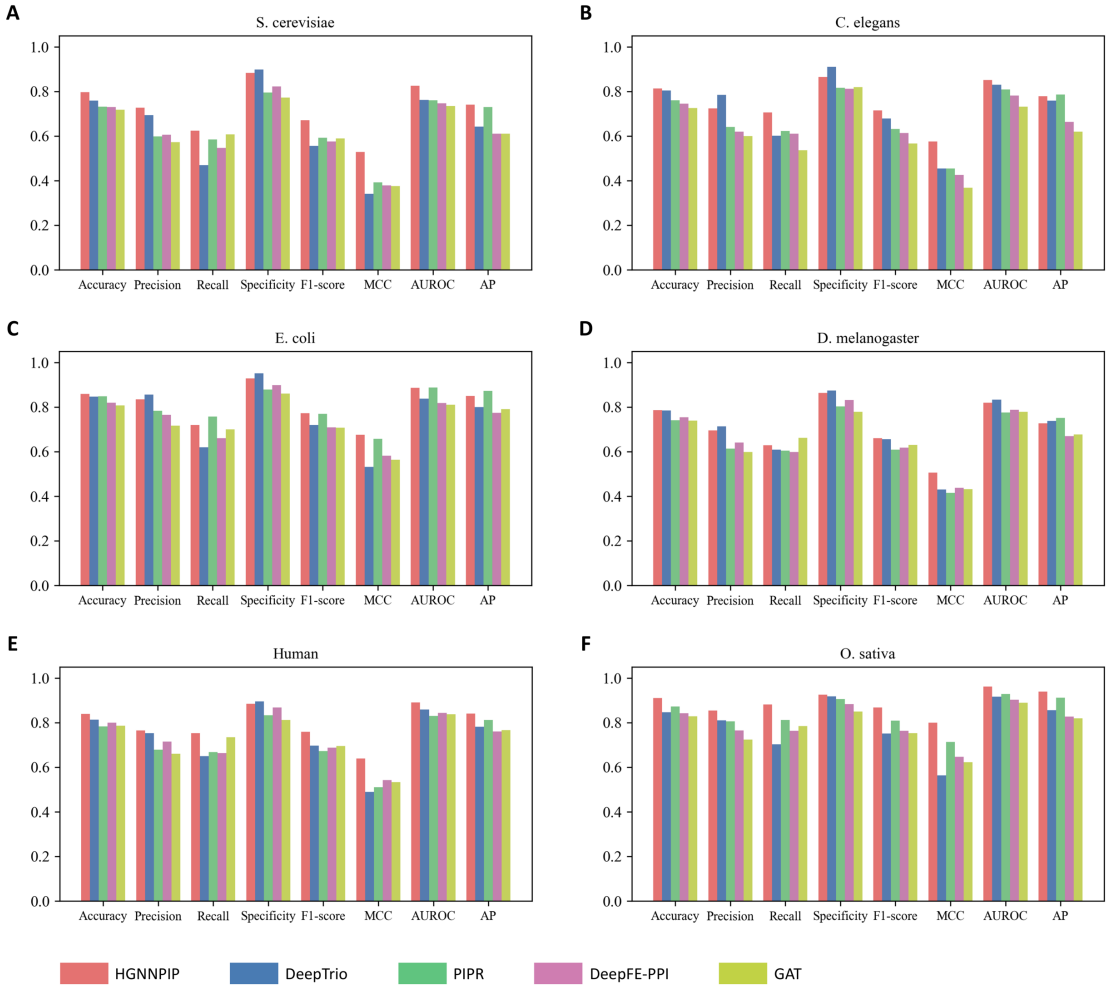


Zi Liu#, Yi-Heng Zhu#, Long-Chen Shen, Xuan Xiao, Wang-Ren Qiu, Dong-Jun Yu. Integrating Unsupervised Language Model with Multi-View Multiple Sequence Alignments for High-Accuracy Inter-Chain Contact Prediction. **Computers in Biology and Medicine**. 2023, 166: 107529

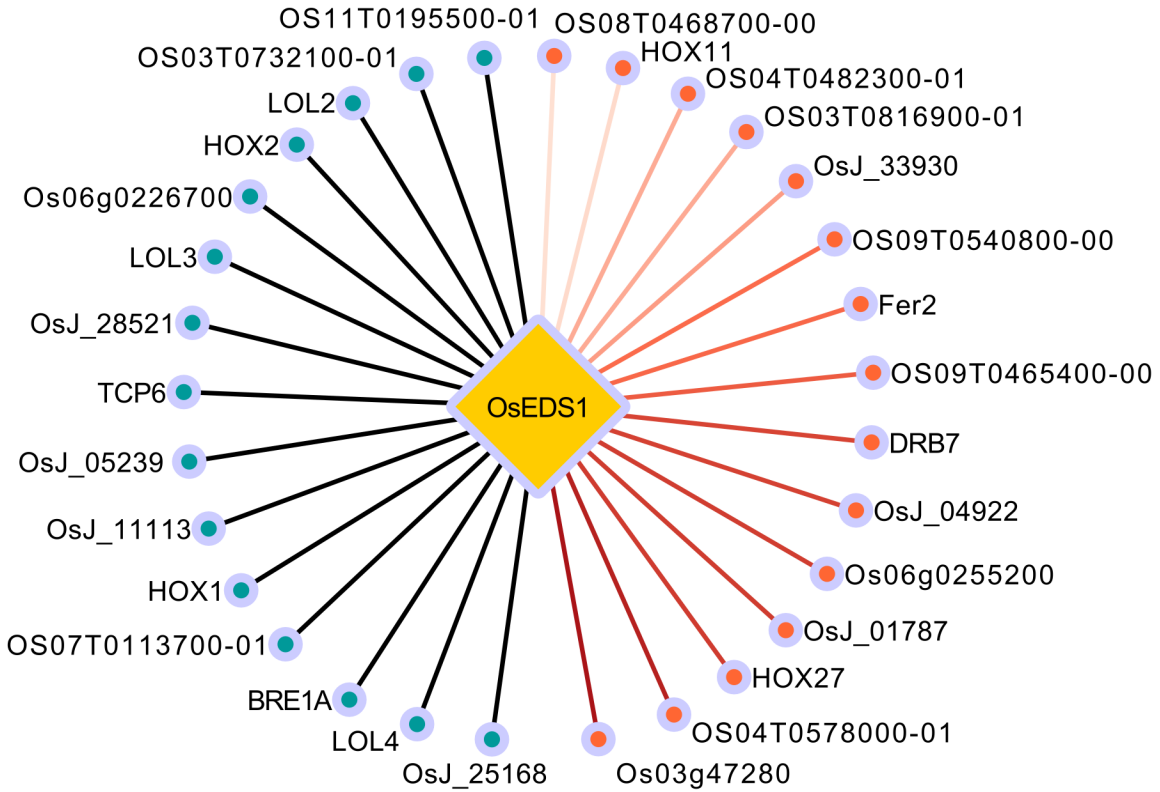
蛋白质-蛋白质相互作用预测

➢Prediction of U.virens-Rice Interactions with Graph Convolutional networks

HGNNPIP模型与4种SOTA方法的性能比较



HGNNPIP模型预测OsEDS1的作用蛋白



L Chi, J Ma, Y Wang, **Z Ji***., (unpublished)

03 Part three

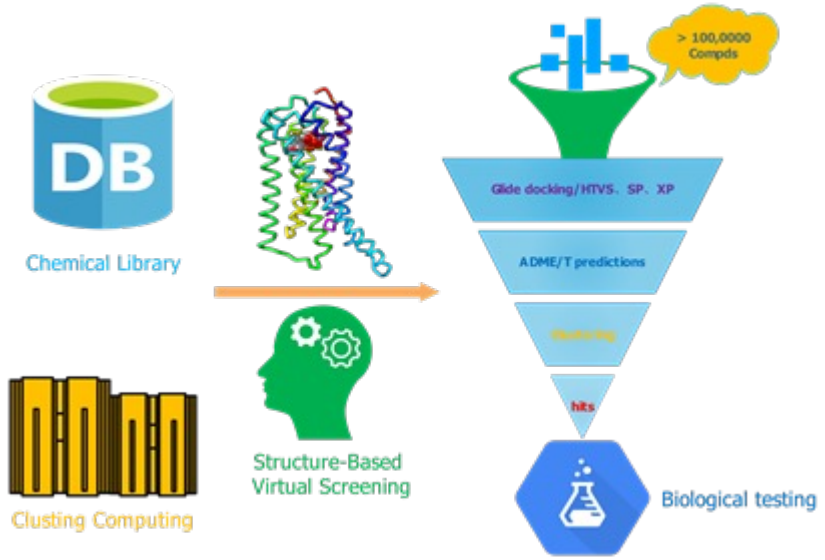
未来展望

01 蛋白质功能预测研究的应用前景

蛋白质功能预测研究有助于推动智能医疗的发展

(1) 辅助疾病分析和诊断 (推断关键致病蛋白质)

(2) 辅助药物设计（药物分子筛选）

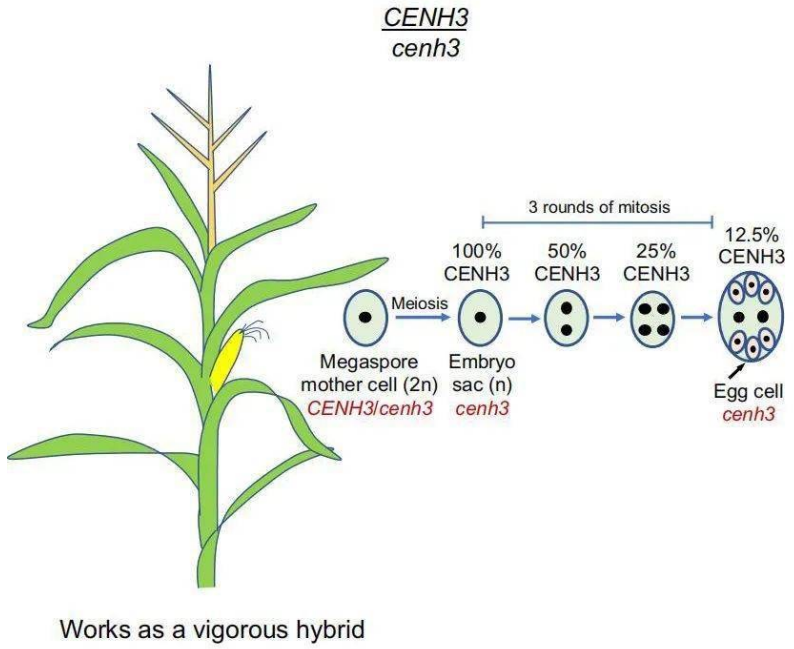


蛋白质功能预测在农业领域的应用前景

(1) 植物遗传育种

(2) 植物与微生物的相互作用

(3) 植物蛋白组学



02 科研团队信息

计智伟教授团队

- 教授1人
- 副教授1人
- 助理教授2人
- 博士研究生3人
- 硕士研究生9人

团队主页：cdsic.njau.edu.cn

研究方向

人工智能与模式识别

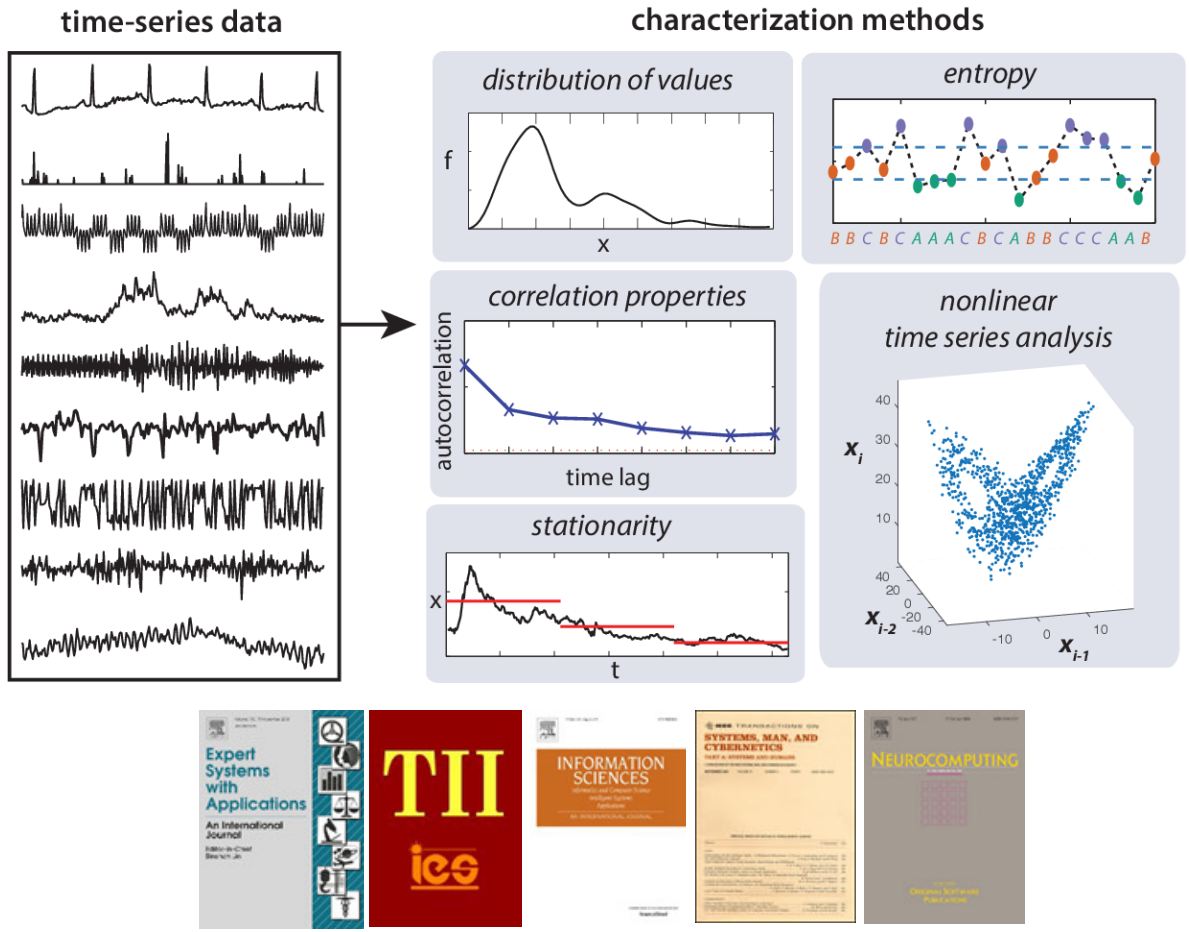
人工智能的理论及应用
大数据计算与模式识别

生物信息与系统生物学

多组学数据整合分析与计算
复杂生物系统的数学建模与预测

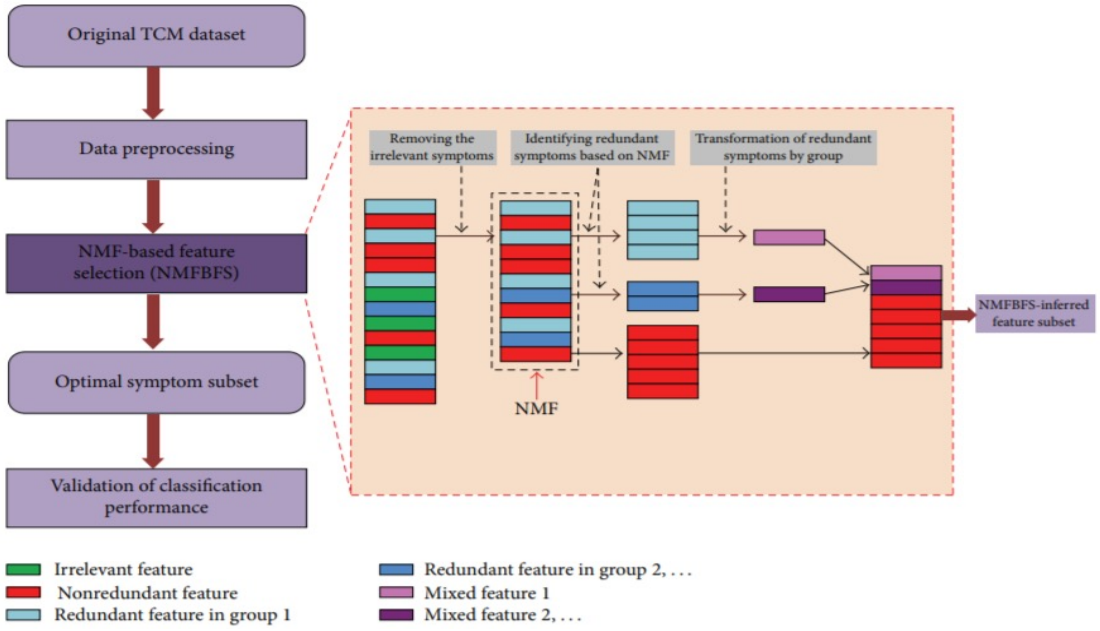
主要成果1：人工智能与模式识别领域

◆ 时间序列数据挖掘与异常模式发现



Z Ji*, Y Wang, X Xie, et al., ***Expert Systems with Applications***, 2022.
N Jin, Y Zeng, K Yan, **Z Ji**, ***IEEE Transactions on Industrial Informatics***, 2021.
M Hu, X Feng, **Z Ji***, et al., ***Information Sciences***, 2019.
K Yan#, **Z Ji#**, et al., ***IEEE Transactions on Systems, Man, and Cybernetics: Systems***, 2019.
K Yan, **Z Ji***, et al., ***Neurocomputing***, 2017.

◆ 高维复杂数据的维度约简和模型优化



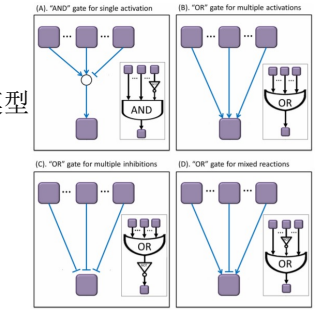
X Xie, F Xia, K Yan, H Xu, **Z Ji***, ***Plant Phenomics***, 2023.
F Xia, X Xie , S Jin, K Yan, **Z Ji***, ***Frontiers in Plant Science***, 2021.
K Yan, …, **Z Ji**., et al., ***IEEE/ACM Trans on Computational Biology and Bioinformatics,*** 2021.
X Xie, X Gu, Y Li, **Z Ji***, ***Knowledge-based Systems***, 2021.
Z Ji, ..., B Wang*, ***Computational and mathematical methods in medicine***, 2015.

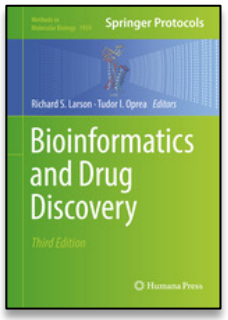
主要成果2：生物信息与系统生物学领域

◆ 创立了一套独特的生物分子网络建模方法

基于线性规划的离散时间建模方法BLP,DILP, TILP, MIP

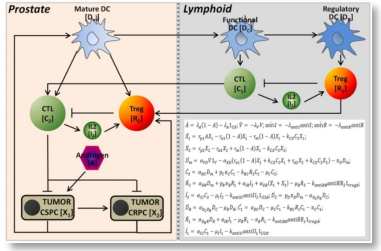
BLP模型



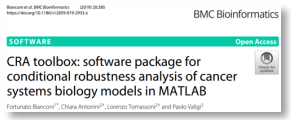


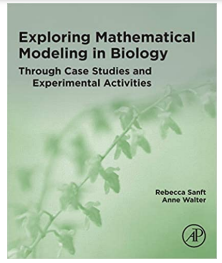
✓ **BLP作为经典模型被写入了Springer教材Methods in Molecular Biology丛书之一《Bioinformatics and Drug Discovery》(第16章第287页)**

基于微分方程的连续时间建模方法



Cell-cell interaction模型



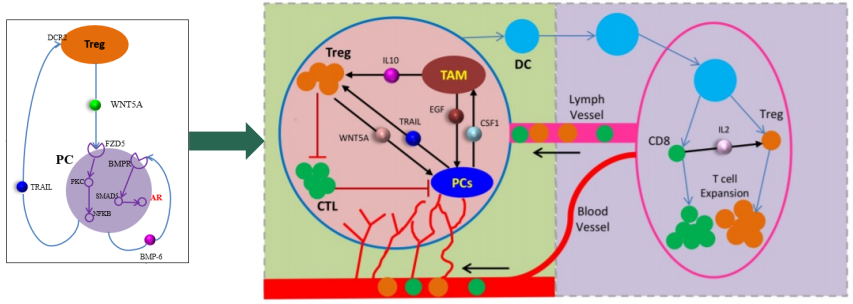


✓ 该模型被Matlab工具包CRA收录，并在BMC Bioinformatics进行长篇报道

✓ 作为经典模型被写入Elsevier教材(2020年)：《Exploring Mathematical Modeling in Biology》(第2章第54页)

◆ 创立了分子-细胞-组织的3D多尺度建模方法

发现了WNT5A调控CRPC（前列腺癌）进展的新机制
构建了面向分子-细胞-组织的多尺度3D模型HABM

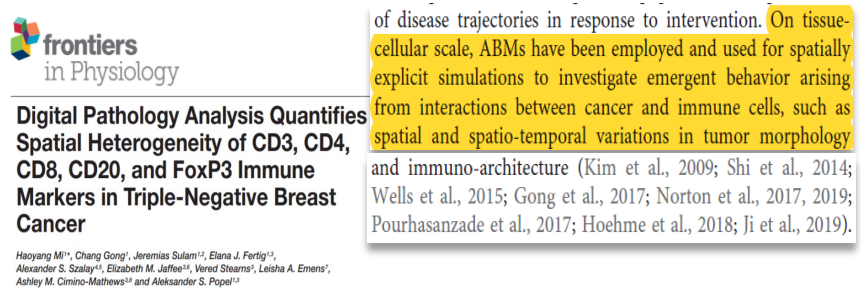



International Immunopharmacology 81 (2020) 106274
Contents lists available at ScienceDirect
International Immunopharmacology
journal homepage: www.elsevier.com/locate/intimp
Development and validation of a prognostic immune-associated gen signature in clear cell renal cell carcinoma
Chengquan Shen, Jing Liu, Jirong Wang, Xiulong Zhong, Dahai Dong, Xiaokun Yonghua Wang

"WNT5A在免疫调节中发挥重要作用"

Results: A total of 681 differentially expressed IAGs were identified and sc ...UT30, WNT5A, IRF9, AGER, PLAUR, TEK, BID) were finally selected in a IAGs signature. Survival analysis revea ...that high IAGs risk scores disease-free survival in breast cancer patients [13]. *WNT5A* belongs to the large WNT family of cysteine-rich secreted glycoproteins, which is involved in multiple signaling pathways that regulate a variety of cellular processes [14]. Ji et al. revealed that *WNT5A* could mediate the activation of Treg and TAM cells, which induced the immunosuppression during castration-resistant prostate cancer progression [15]. Chen

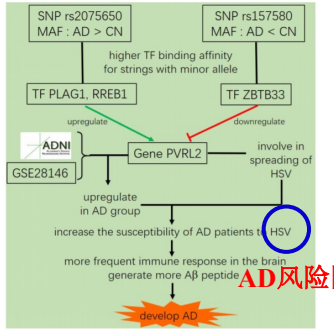
✓ **HABM是十几年来首个对肿瘤生长-免疫反应-血管生成进行3D时空建模的数学模型**


frontiers in Physiology
Digital Pathology Analysis Quantifies Spatial Heterogeneity of CD3, CD4, CD8, CD20, and FoxP3 Immune Markers in Triple-Negative Breast Cancer
Haoyang Mi, Chang Gong, Jeremias Sulam, Elana J. Fertig, Alexander S. Szalay, Elizabeth M. Jaffee, Vered Stearns, Leisha A. Emens, Ashley M. Cimino-Mathews and Aleksander S. Popel

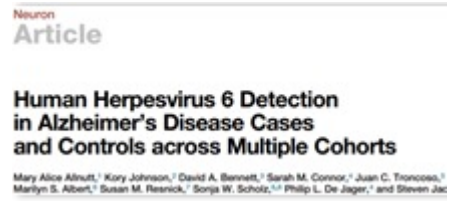
of disease trajectories in response to intervention. On tissue-cellular scale, ABMs have been employed and used for spatially explicit simulations to investigate emergent behavior arising from interactions between cancer and immune cells, such as spatial and spatio-temporal variations in tumor morphology and immuno-architecture (Kim et al., 2009; Shi et al., 2014; Wells et al., 2015; Gong et al., 2017; Norton et al., 2017, 2019; Pourhasanzade et al., 2017; Hoehme et al., 2018; Ji et al., 2019).

◆ 建立了生物组学大数据挖掘的计算框架

组学大数据挖掘发现AD潜在风险因子 (TREM2, HSV-1)



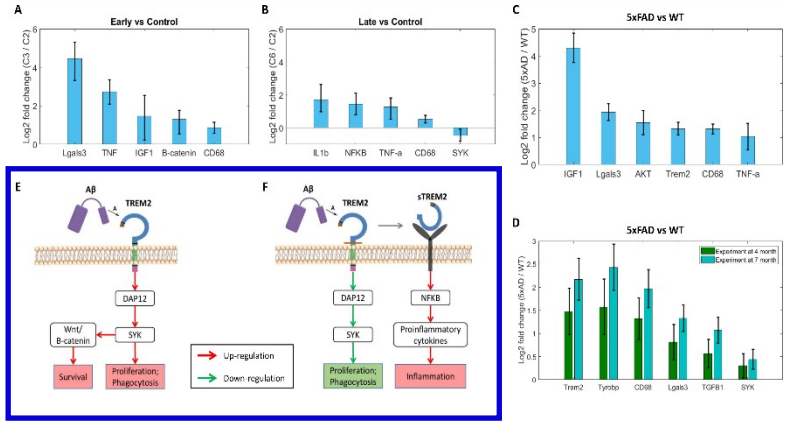
AD风险因子


Neuron
Article
Human Herpesvirus 6 Detection in Alzheimer's Disease Cases and Controls across Multiple Cohorts

是关于"HSV-1是AD风险因子"的较早报道之一

or plasma (Lövheim et al., 2018). In addition, several groups have identified overlap between AD genetic risk factors and genes affected by viral infection, such as a receptor involved in spreading HSV-1 (Liu et al., 2018) and a human leukocyte antigen (HLA) subtype associated with increased susceptibility to HHV-6A infection (Rizzo et al., 2019).

✓ **时序组学大数据挖掘，首次解析了TREM2调控Microglia表型转换的分子机制**



谢谢各位专家观看

请各位专家批评指正！