2022 7 116 2022 8 15

——

Recursion Benevolent AI

Benevolent AI

Baricitinib [2]

Insilico Medicine 3CL

[2] Baricitinib Janus(JAK1/JAK2) Incyte Eli Lilly 2017 Olumiant

ADMET

AI +

AI +

AI

Ribonucleic Acid Sequencing RNA-seq

High-Throughput Sequencing HTS

+^ V
ÝH •
Â
© D
1-2

National

Center for Biotechnology Information NCBI

Gene Expression O

1-2

PubChem

ChEMBL

Absorption, Distribution, Metabolism and Excretion ADME

DrugBank

LINCS L1000 PDB

Computer Aided Drug Design CADD ML DL

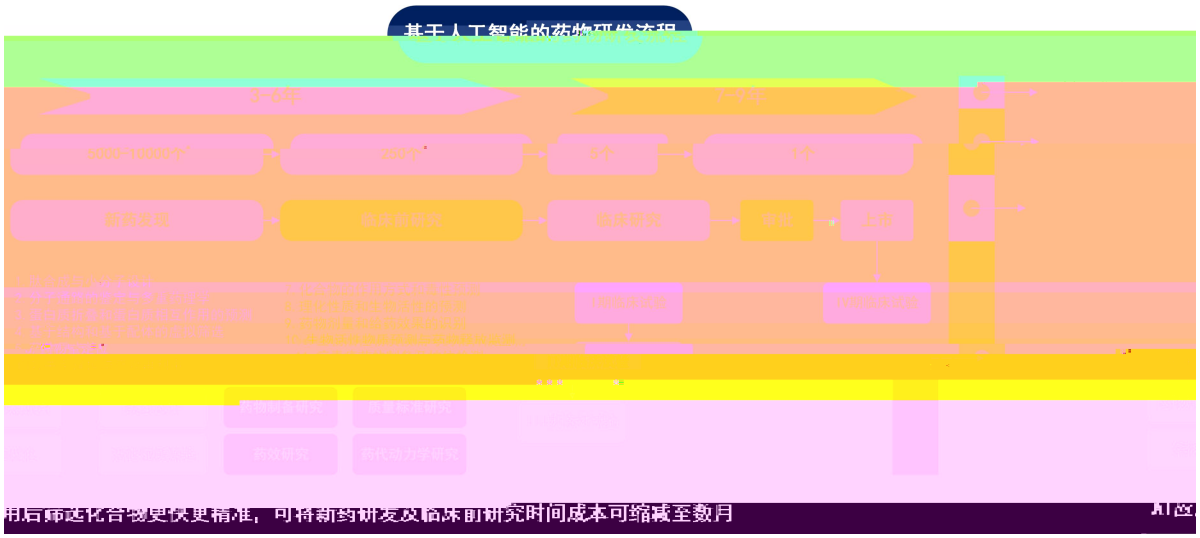
Quantitative Structure Activity Relationship QSAR

ML

①

②

2-1



2-1

2 50

2-2

Yan [2] 2020 DL

Antimicrobial peptides AMPs Deep-AmPEP30

AmPEP30 Convolutional Neural Network, CNN

DeoxyriboNucleic Acid

DNA AMP

—

AMPs

Zhavoronkov [3]

GENTRL

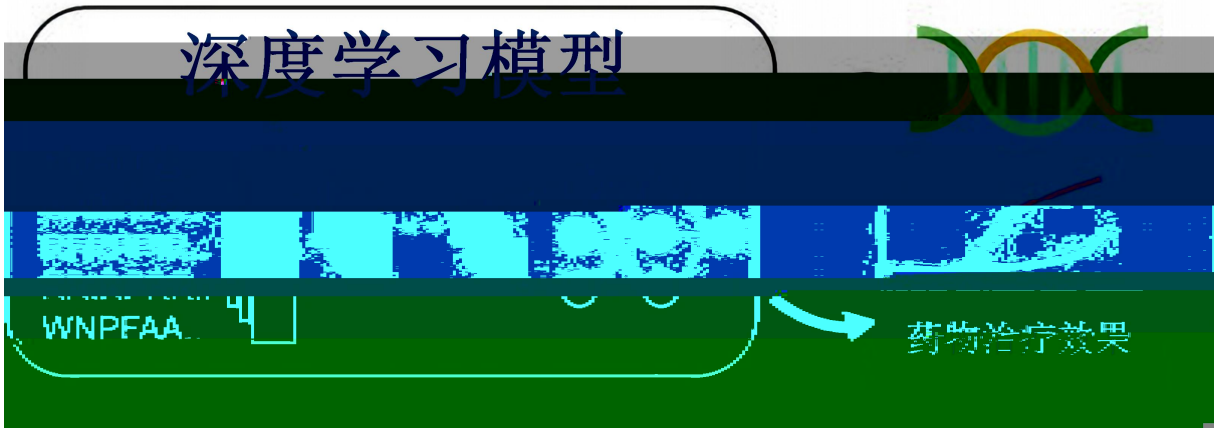
https://github.com/insilicomedicine/GENTRL

DDR1 McCloskey [4] DNA

DNA-Encoded small molecule Libraries DEL

ML

Xing [5] XGBoost



2-2

—

—

"

— Protein-Protein Interaction, PPI

2-4

Bayesian Network BN PPI

Gene Ontology GO

PPI BN

PCA

—

[7]

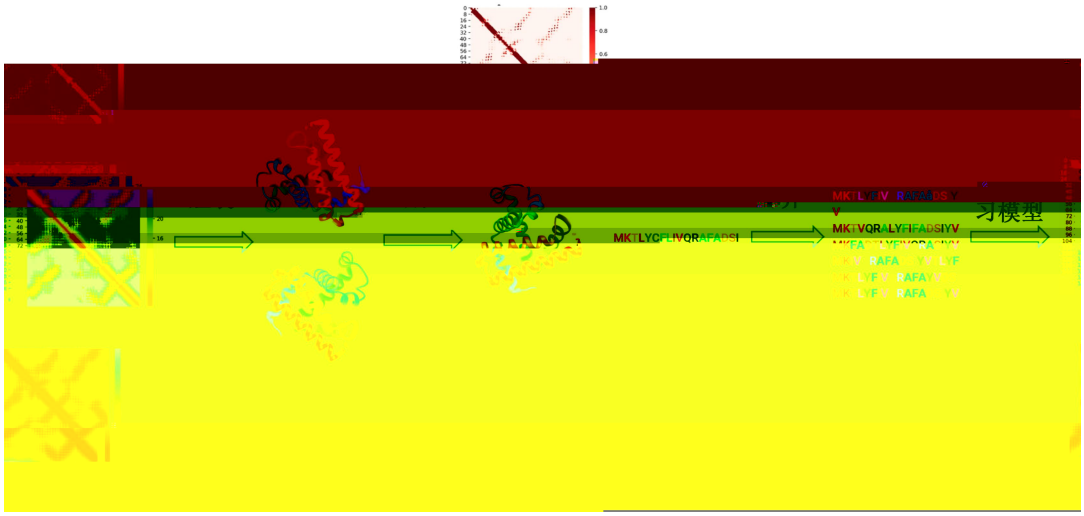
K—

PPI

Gene Ontology GO

PPI

[7] —



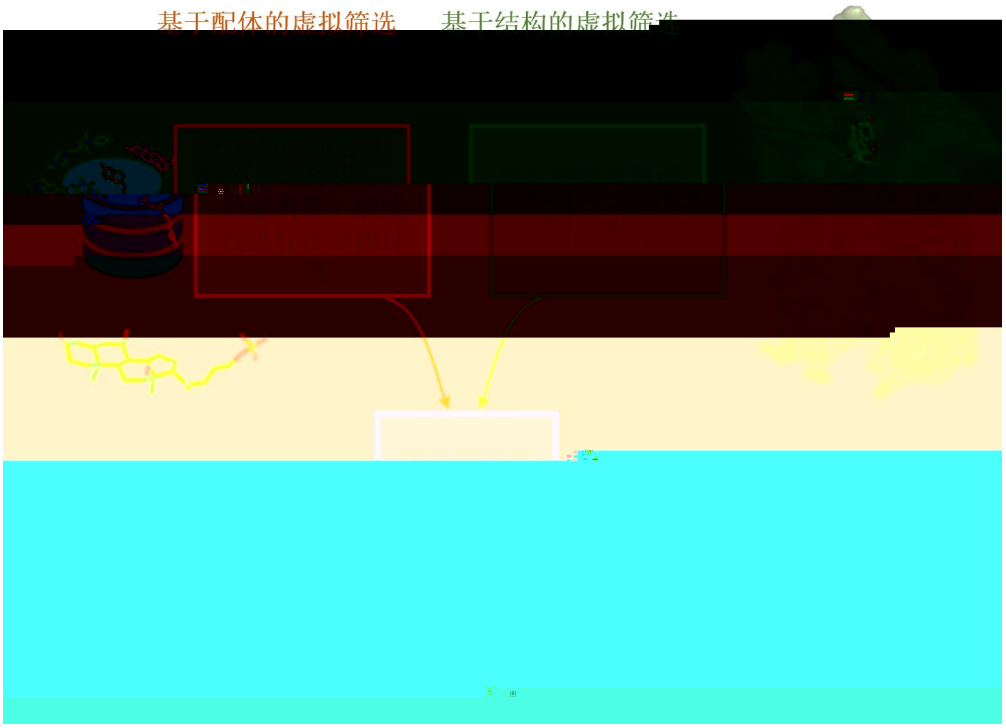
2-4

Virtual Screening VS

CADD

ML VS

ML VS VS



2-5

VS 2-5 VS

Structural Based Virtual Screening SBVS VS

Ligand Based Virtual Screening LBVS

SBVS AI ML

[8] NNScore CScore SVR-SCORE ID-SCORE

SBVS SBVS

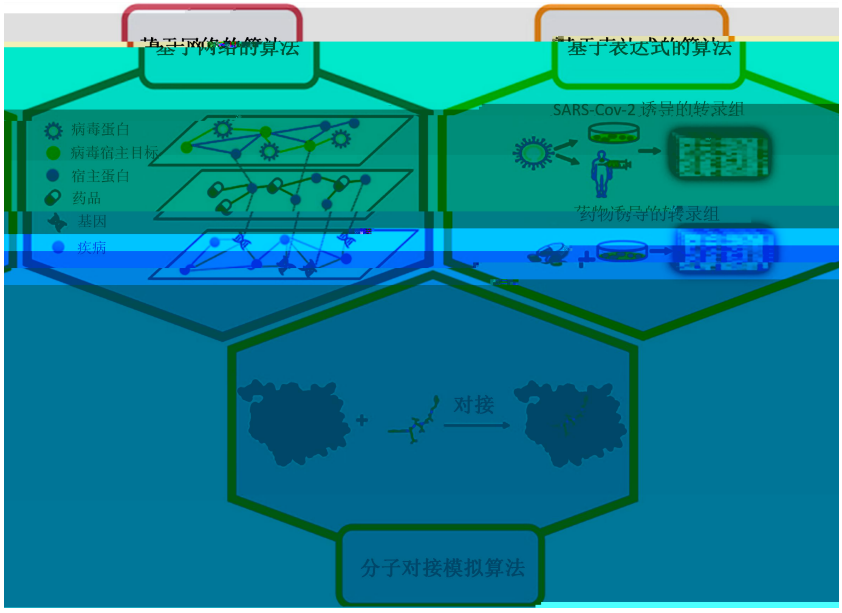
—

LBVS SwissSimilarity[9]

METADOCK[10] HybridSim-V[11]S AutoDock Bias[12]

1.06

Logistic



2-6 AI ML

COVID-19

2 SARS-CoV-2

COVID-19

COVID-19

AI ML COVID-19

2-6 2-

1

2-1 AI ML

AI ML		
[13]	SARS-CoV-2	https://github.com/Murali-group/SARS-CoV-2-network-analysis
[14]	-	
[15]	MT-DTI	
[16]	-	
[17]		https://github.com/ChengF-Lab/CoV-KGE
[18]		https://github.com/yejinjkim/drug-repurposing-graph
[17]		https://github.com/FangpingWan/NeoDTI
[19]		https://github.com/FangpingWan/CoV-DTI
[20]	-	https://github.com/Barabasi-Lab/COVID-19
- InfinityPhenotype[21]		

		https://[illegible]hub.com/[illegible]lab/covid19_re[illegible]rposing
[illegible]2]		https://g[illegible]ub.com/[illegible]993/DeepCE
[24]		
[14]		
		https://github.com/ekraka/SSnet
[26]		

[27,28]

S

[illegible]

[illegible] [13]

(Regularized Laplacian RL) [illegible]

FDA SARS-CoV-2

[29] Zhu [21] AI InfinityPhenotype

FDA

Liquiritin I

COVID-19

[30]

AI ML Nguyen

[24] MathPose MathDL

SARS-CoV-2 3CL —

MathPose

MathDL

COVID-19 15

AI ML [31–33]

2-7

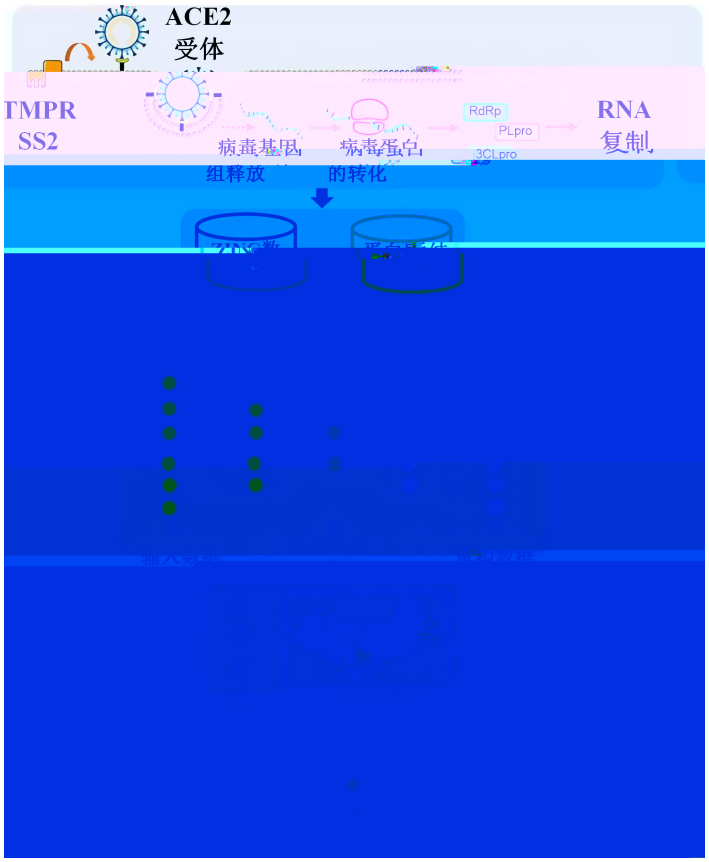
2-2

[34] Zhavoronkov [35]

COVID-19

Srinivasan [36]

COVID-19



2-7 AI ML

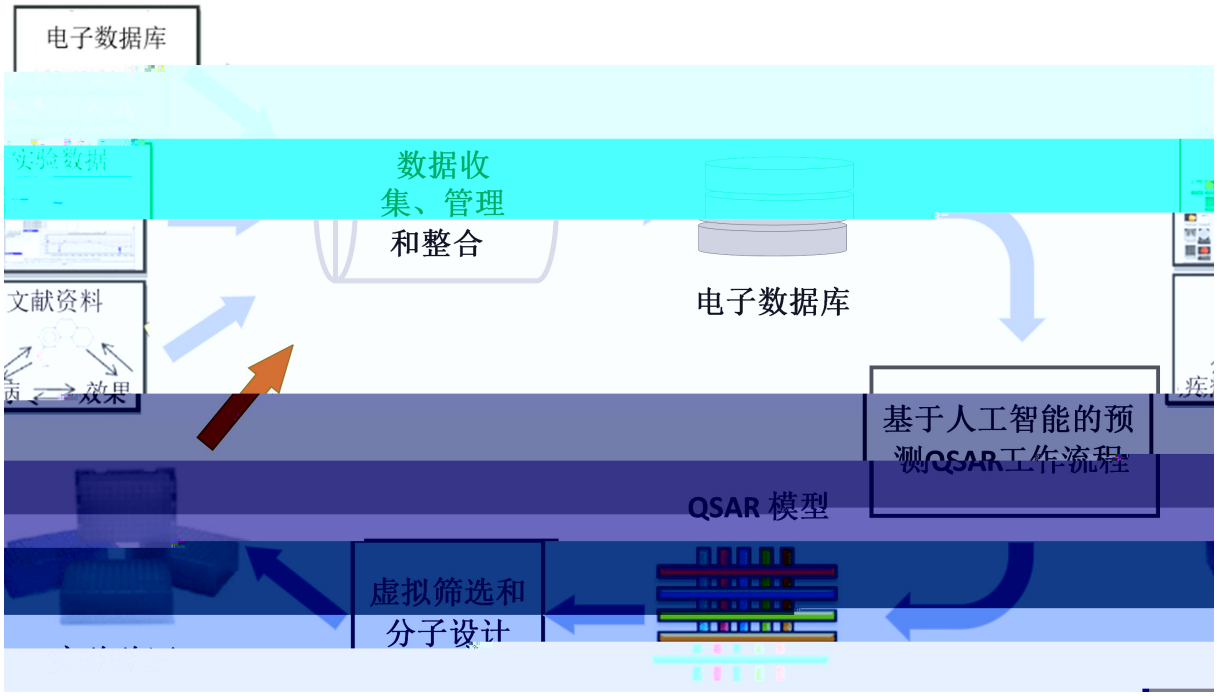
2-2 AI ML

AI ML		
[35]		
Q [37]		https://github.com/tbwxmu/2019-nCov
[38]		

Hooshmand [45]

16 HCoV

12



2-8

2-9

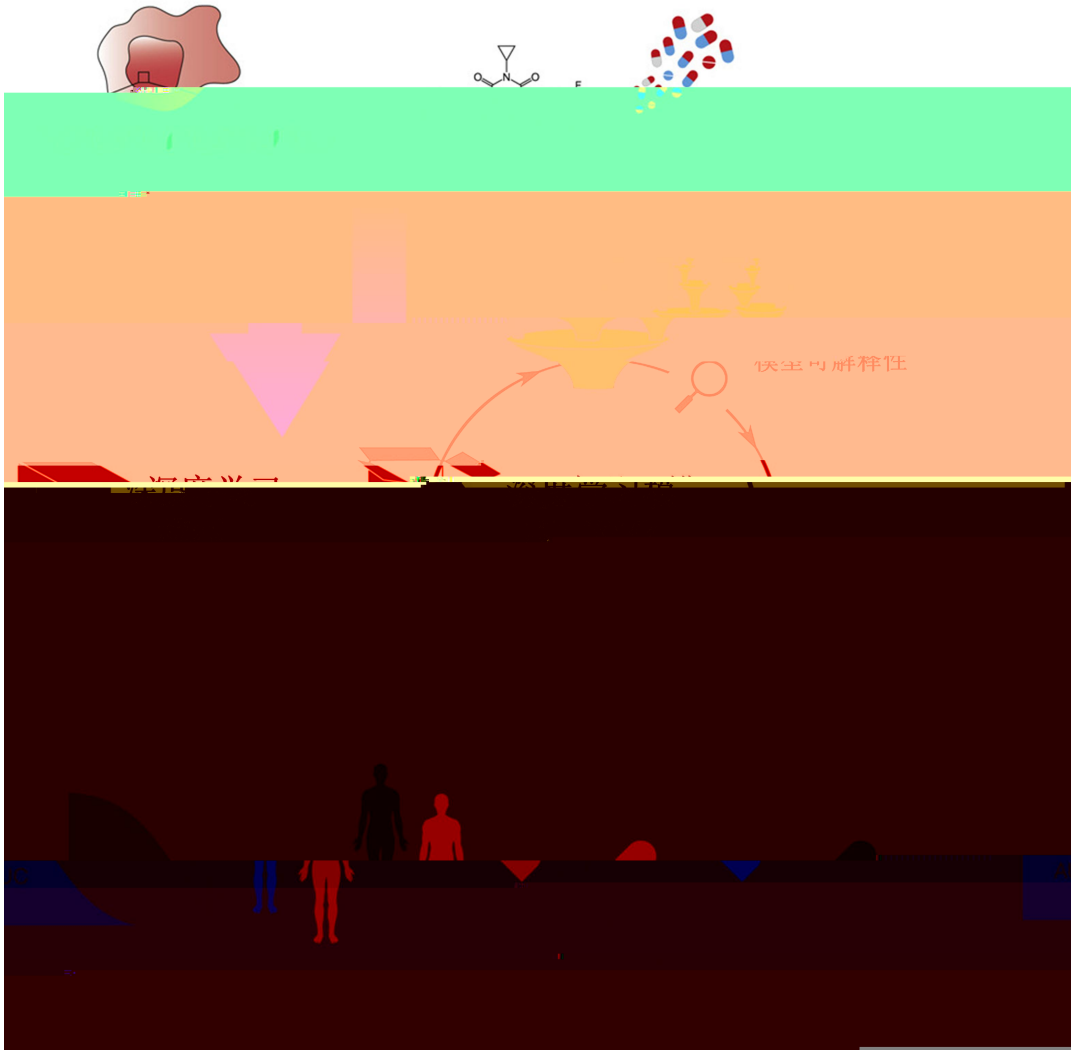
Web LimTox pkCSM admetSAR Toxtree DeepTox[46] PrOCTOR[47]

Robledo-Cadena [48]

PrOCTOR

2,576

16 2

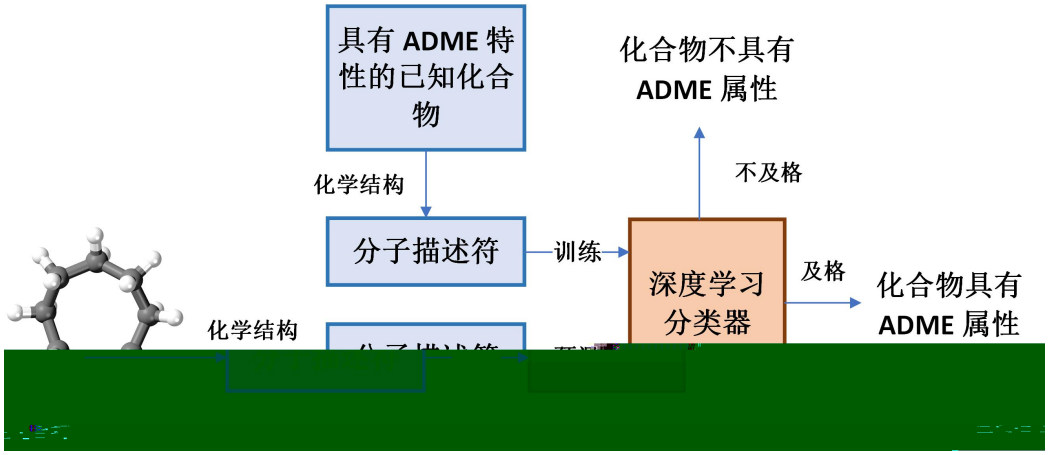


2-9

2-10 SMILES

Coulomb matrices

Deep Neural Networks DNN



2-10

ChemMapper

ML DL

[49] KronRLS SimBoost DeepDTA Padme

ML

DL 2-11 Shen

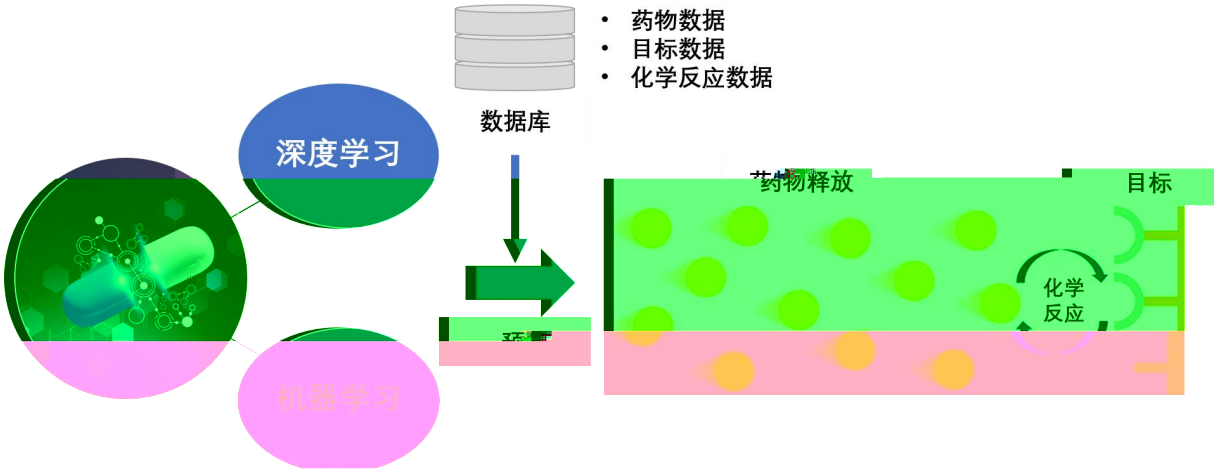
[50] AI-PRS

AI-PRS

Wu DL RF WDL-RF (https://zhanglab.ccmb.med.umich.edu/WDL-RF/)

G G Protein-Coupled Receptors GPCRs Cichonska [54]

pairwiseMKL https://github.com/aalto-ics-kepaco [55]



2-12

2-13 2-3 COVID-19

[56] Fast [57]

MARIA NetMHCPan4

SARS-CoV-2 T B 405 T

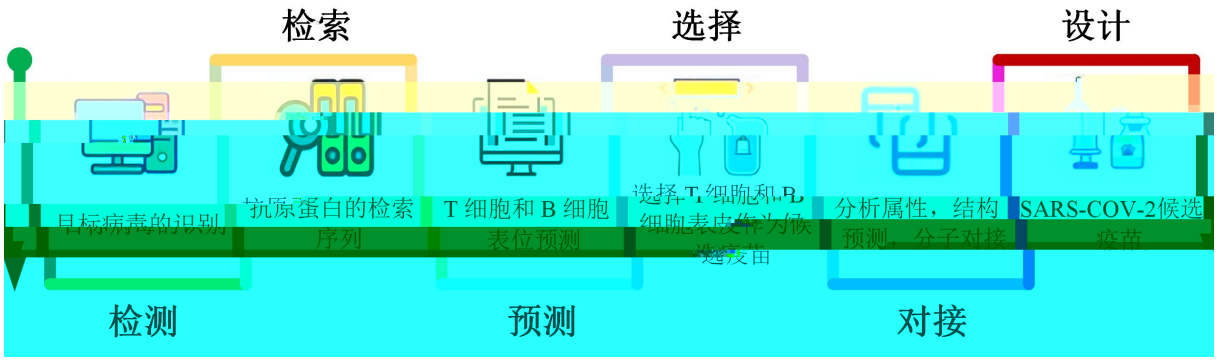
MHC-I MHC-II

S B

COVID-19

2-3 AI ML

AI ML		
[57]	SARS-CoV-2 T B	
XGBoost[58]		
[59]	SARS-CoV-2 HLA	
[60]		https://github.com/zikunyang/DCVST



2-13

Prachar [59] 174 SARS-CoV-2 11 HLA SARS-CoV-2 HLA SARS-CoV-2

COVID-19 Yang [60]

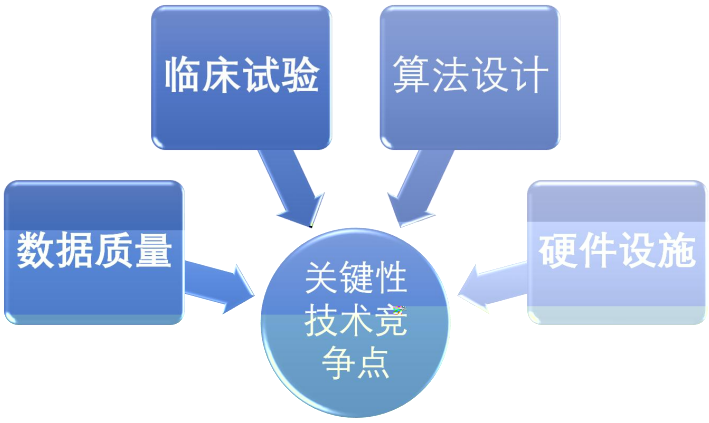
DeepVacPred

DeepVacPred 694aa 16 B

82 CTL 89 HTL SARS-CoV-2 RNA

3.5

1



3-1

—

—

—

3

Graphic Processing Unit GPU

Tensor Processing Unit TPU

Central Processing Unit CPU GPU

Field Programmable Gate Array FPGA

AI/ML

2014

Generative Adversarial Networks GANs

2014 2015 AI

AI Exscientia

Insilico Medicine

AI

AI

2016 2017 AI

AI

2018 AI

AI Exscientia Atomwise Recursion Insilico Medicine

AI AI

3-1

3-1

Benevolent AI				Baricitinib
Exscientia				Exscientia AI

Atom

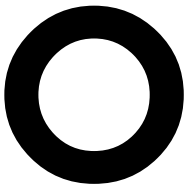
Recursion

NME

2019 Recursion

Takeda 60

6



Insilico
Medicine

7

IND-enabling COVID-19

3CL

MAT2A USP1 “

”

0

1 ά

				AI
				PandaOmics Chemistry42
				“ ”
				AI

AI 2019

AI Benevolent AI Atomwise Recursion

3-2

AI

3-2 AI

Benevolent AI	2019 9 17 - BenevolentAI 9000 2018 4 19 - BenevolentAI 1.15 2015 8 26 - BenevolentAI 8700
Exscientia	Exscientia 9 4.744 2022 1 10 IPO 2022 1 10 IPO - Exscientia 1 2021 7 8 Grant - Exscientia 150 2021 4 20 D - Exscientia 2.25 * https://www.crunchbase.com/organization/exscientia/company_financials
Atomwise	2020 09 17 D - 9000 - MSD Partners Octave Group Shumway Capital TpTfOCV Partners OMX Ventures Avenir Growth 2019 08 22 C 5300 – Shumway Capital TPTF OCV Partners 2018 03 1 B 3600 - Shumway Capital Morgan Noble Avenir Growth Capital
Recursion	2019 7 15 Recursion Pharmaceuticals 1.21 C Baillie Gifford Scottish Mortgage Investment Trust Intermountain Ventures University of Minnesota Texas Tech University System
Insilico Medicine	Insilico Medicine 10 3.063 2022 1 11 Venture - Series Unknown 2022 1 11 - Insilico Medicine 2021 6 22 C - Insilico Medicine 2.55 2020 4 20 B -Insilico Medicine * https://www.crunchbase.com/organization/insilico-medicine/company_financials
	2 2021 2 19 SciClone Pharmaceuticals 1.34 2021 2 19 - SciClone Pharmaceuticals 1.34

	2016 12 21 -Healenvoy 300
	1 2000 2020 7 9 2020 7 9 2000
	2018 7 3 SIG 2020 3 16 1000 A
	2021 08 4 D IMO Ventures Neumann Capital Artisan Partners Artisan Partners 2020 9 3.188 C 2 Mirae Asset IMO Ventures Parkway SIG

2000

3-3 AI

	AI DrugSpaceX ModelArts		
	80%		
	CPCR CPCR		
	DNA 2021 10 MindSpore AlphaFold2 2 3		
	1.5		

	2019 11 FDA		

Artificial Intelligence AI

— **Protein-Protein Inte**

[5] Xing G , Liang L , Deng C , et al. Activity Prediction of Small Molecule Inhibitors for Antirheumatoid Arthritis Targets Based on Artificial Intelligence[J]. ACS Combinatorial Science, 2020, 22(12): 873–886.

[6] Gu S , Lai L H . Associating 197 Chinese herbal medicine with drug targets and diseases using the similarity ensemble approach[J]. Acta Pharmacologica Sinica, 2019, 41(3): 7.

[7] Zhu-Hong, You, Ying-Ke, et al. Prediction of protein-protein interactions from amino acid sequences with ensemble extreme learning machines and principal component analysis[J]. BMC Bioinformatics, 2013, 14(10): S10.

[8] Serafim, Mateus Sa Magalhaes, et al. The application of machine learning techniques to innovative antibacterial discovery and development. [J]. Expert opinion on drug discovery, England: 2020, 15(10): 1165 – 1180.

[9] Zoete V , Daina A , Bovigny C , et al. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening[J]. Journal of Chemical Information & Modeling, 2016:1399 – 1404.

[10] Imbernon B , JM Cecilia, Perez-Sanchez H , et al. METADOCK: A parallel metaheuristic schema for virtual screening methods[J]. Experimental Mechanics, 2018, 32(6):789-803.

[11] Hongjian L , Kwong-S. L , Man-H. W , et al. USR-VS: a web server for large-scale prospective virtual screening using ultrafast shape recognition techniques[J]. Nucleic Acids Research, 2016(W1):W436-W441.

[12] Pablo A J , Modenutti C P , Demian A , et al. AutoDock Bias: improving binding mode prediction and virtual screening using known protein-ligand interactions[J]. Bioinformatics, 2019, 35(19): 3836 – 3838.

[13] Law J N , N Tasnina, Kshirsagar M , et al. Identifying human interactors of SARS-CoV-2 proteins and drug targets for COVID-19 using network-based label propagation[J]. arXiv preprint arXiv:2006.01968, 2020.

[22] Belyaeva A , Cammarata L , Radhakrishnan A , et al. Causal network models of SARS-CoV-2 expression and aging to identify candidates for drug repurposing[J]. Nature Communications, 2021, 12(1): 1024.

[23] Pham, Thai-Hoang, et al. A deep learning framework for high-throughput mechanism-driven phenotype compound screening and its application to COVID-19 drug repurposing[J]. Nature machine intelligence, 2021, 3(3): 247 – 257.

[24] Nguyen D D , Gao K , Chen J , et al. Potentially highly potent drugs for 2019-nCoV[J]. bioRxiv, Cold Spring Harbor Laboratory, 2020.

[25] Karki N K , Verma N , Trozzi F , et al. Predicting Potential SARS-COV-2 Drugs - In Depth Drug Database Screening Using Deep Neural Network Framework SSnet, Classical Virtual Screening and Docking. 2020, 22(4).

[26] Mohapatra S , Nath P , Chatterjee M , et al. Repurposing therapeutics for COVID-19: Rapid prediction of commercially available drugs through machine learning and docking[J]. PLoS ONE, 2020, 15(11):e0241543.

[27] Liu L , Zhang L R , Dao F Y , et al. A computational framework for identifying the transcription factors involved in enhancer-promoter loop formation.[J]. Molecular therapy. Nucleic acids, 2021, 23: 347 – 354.

[28] Liu L , Li Q Z , Jin W , et al. Revealing Gene Function and Transcription Relationship by Reconstructing Gene-Level Chromatin Interaction.[J]. Computational and structural biotechnology journal, 2019, 17: 195 – 205.

[29] Killick R, Ballard C, Doherty P, et al. Transcription-based drug repurposing for COVID-19.[J]. Virus research, 2020, 290: 198176.

[30] Zulfiqar H , Masoud M S , Yang H , et al. Screening of Prospective Plant Compounds as H1R and CL1R Inhibitors and Its Antiallergic Efficacy through Molecular Docking Approach[J]. Computational and Mathematical Methods in Medicine, 2021, 2021(1):1-9.

[31] Hasan, Md Mehedi, et al. NeuroPred-FRL: an interpretable prediction model for identifying neuropeptide using feature representation learning.[J]. Briefings in bioinformatics, England: 2021, 22(6).

[32] Charoenkwan P , Nantasenamat C , Hasan M M , et al. BERT4Bitter: a bidirectional encoder representations from transformers (BERT)-based model for improving the prediction of bitter peptides[J]. Bioinformatics, 2021.

[33] Phasit C , Wararat C , Chanin N , et al. StackIL6: a stacking ensemble model for improving the prediction of IL-6 inducing peptides[J]. Briefings in Bioinformatics, 2021(6):6.

[34] Xu B , Liu D , Wang Z , et al. Multi-substrate selectivity based on key loops and non-homologous domains: new insight into ALKBH family[J]. Cellular and molecular life sciences : CMLS, 78(1):129-141.

[35] Zhavoronkov A , Aladinskiy V A , Zhebrak A , et al. Potential 2019-nCoV 3C-like protease inhibitors designed using generative deep learning approaches[J]. 2020.

[36] Srinivasan S , Batra R , Chan H , et al. Artificial Intelligence-Guided De Novo Molecular Design Targeting COVID-19[J]. ACS Omega, 2021, 6(19):12557-12566.

[37] Tang B , He F , Liu D , et al. AI-aided design of novel targeted covalent inhibitors against SARS-CoV-2[J]. BioRxiv, Cold Spring Harbor Laboratory, 2020.

[38] Ton A T , Gentile F , Hsing M , et al. Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds.[J]. Molecular informatics, 2020, 39(8): e2000028.

[39] Chenthamarakshan V , Das P , Hoffman S C , et al. Cogmol: target-specific and selective drug design for COVID-19 using deep generative models[J]. arXiv preprint arXiv:2004.01215, 2020.

[40] Bung N , Krishnan S R , Bulusu G , et al. De novo design of new chemical entities for SARS-CoV-2 using artificial intelligence[J]. Future Medicinal Chemistry, 2021(2): 575 – 585.

[41] Wang S , Sun Q , Xu Y , et al. A transferable deep learning approach to fast screen potential antiviral drugs against SARS-CoV-2[J]. Briefings in bioinformatics, Oxford University Press, 2021.

[42] Ambure P , Halder A K , H González-Díaz, et al. QSAR-Co: An Open Source Software for Developing Robust Multi-tasking or Multi-target Classification-Based QSAR Models[J]. Journal of Chemical Information and Modeling, 2019, 59(6): 2538 – 2544.

[43] Karpov P , Godin G , Tetko I V . Transformer-CNN: Swiss knife for QSAR modeling and interpretation[J]. Journal of Cheminformatics, 2020, 12(1): 17.

[44] Luo H , Li M , Wang S , et al. Computational Drug Repositioning using Low-Rank Matrix Approximation and Randomized Algorithms[J]. Bioinformatics(11): 1904 – 1912.

[45] Hooshmand S A , · Mohadeseh, Ghobadi Z , et al. A multimodal deep learning-based drug repurposing approach for treatment of COVID-19[J]. Molecular Diversity, 2020:3: 1717 – 1730.

[46] Mayr, Andreas, et al. DeepTox: toxicity prediction using deep learning[J]. Frontiers in Environmental Science, Frontiers, 2016, 3: 80.

[47] Kaitlyn M, Gayvert, Neel S, et al. A Data-Driven Approach to Predicting Successes and Failures of Clinical Trials[J]. Cell Chemical Biology, 2016, 23(10): 1294–1301.

[48] Gilvary C , Elkhader J , Madhukar N , et al. A machine learning and network framework to discover new indications for small molecules[J]. PLoS Computational Biology, 2020, 16(8):e1008098.

[49] Zelik R , H Ztürk, A Zgür, et al. ChemBoost: A Chemical Language Based Approach for Protein – Ligand Binding Affinity Prediction[J]. Molecular Informatics, 2020, 40(5): 2000212.

[50] Harnessing Artificial Intelligence to Optimize Long - Term Maintenance Dosing for Antiretroviral - Naive Adults with HIV - 1 Infection[J]. Advanced Therapeutics, 2020, 3(4): 1900114.

[51] Tang J , Liu R , Zhang Y L , et al. Application of Machine-Learning Models to Predict Tacrolimus Stable Dose in Renal Transplant Recipients[J]. Scientific Reports, 2017, 7:42192.

[52] Hu YH, Tai CT, Tsai CF, Huang MW. Improvement of Adequate Digoxin Dosage: An Application of Machine Learning Approach.[J]. Journal of healthcare engineering, 2018, 2018: 3948245.

[53] Imai S , Takekuma Y , Miyai T , et al. A New Algorithm Optimized for Initial Dose Settings of Vancomycin Using Machine Learning[J]. Biological & Pharmaceutical Bulletin, 2020, 43(1):188-193.

[54] Jiansheng W, Qiuming Z , Weijian W , et al. WDL-RF: predicting bioactivities of ligand molecules acting with G protein-coupled receptors by combining weighted deep learning and random forest[J]. Bioinformatics, 2018, 2(13):13.

[55] Cichonska, Anna, et al. Learning with multiple pairwise kernels for drug bioactivity prediction. [J]. Bioinformatics, 2018, 34(13): i509 – i518.

[56] Wang J , Wang H , Wang X , et al. Predicting Drug-target Interactions via FM-DNN Learning[J]. Current Bioinformatics, 2020, 15(1): 68 – 76.

[57] Fast E, Chen B. Potential T-cell and B-cell epitopes of 2019-nCoV[J]. BioRxiv, Cold Spring Harbor Laboratory, 2020.

[58] Ong E , Mei U W , Huffman A , et al. COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning. 2020, 11: 1581.

[59] Prachar M , Justesen S , Steen-Jensen D B , et al. Identification and validation of 174 COVID-19 vaccine candidate epitopes reveals low performance of common epitope prediction tools[J]. Scientific Reports, 2020, 10(1): 20465.

[60] Yang Z , Bogdan P , Nazarian S . An in silico deep learning approach to multi-epitope vaccine design: a SARS-CoV-2 case study[J]. Scientific Reports, 2021, 11(1): 3238.

zhuysh@pbcsf.tsinghua.edu.cn

0 1

“ ”

/

“

”

1

2

/

3

/

4

/

Exscientia Atomwise Recursion Insilico Medicine

“ ”