



DRUG REPOSITIONING

AI in DRUG DISCOVERY II
2020. 11. 11

Korean AI Center For Drug Discovery and Development

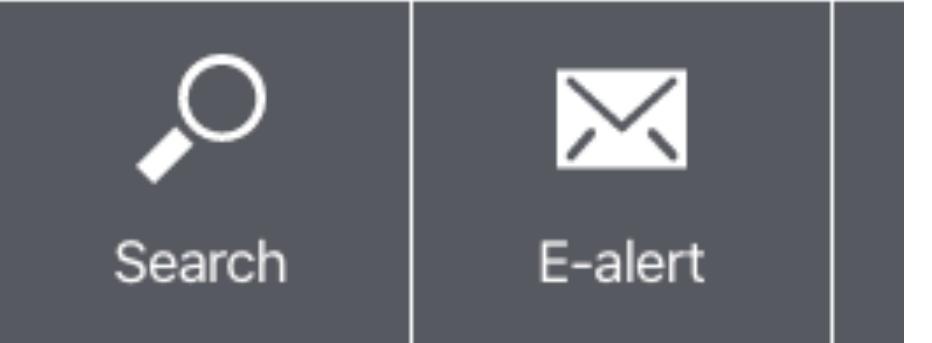
Erkhembayar J. Ph.D



DRUG REPOSITIONING

“The most fruitful basis for the discovery of a new drug is to start with an old drug”

Nobel Prize-winning pharmacologist Sir James Black



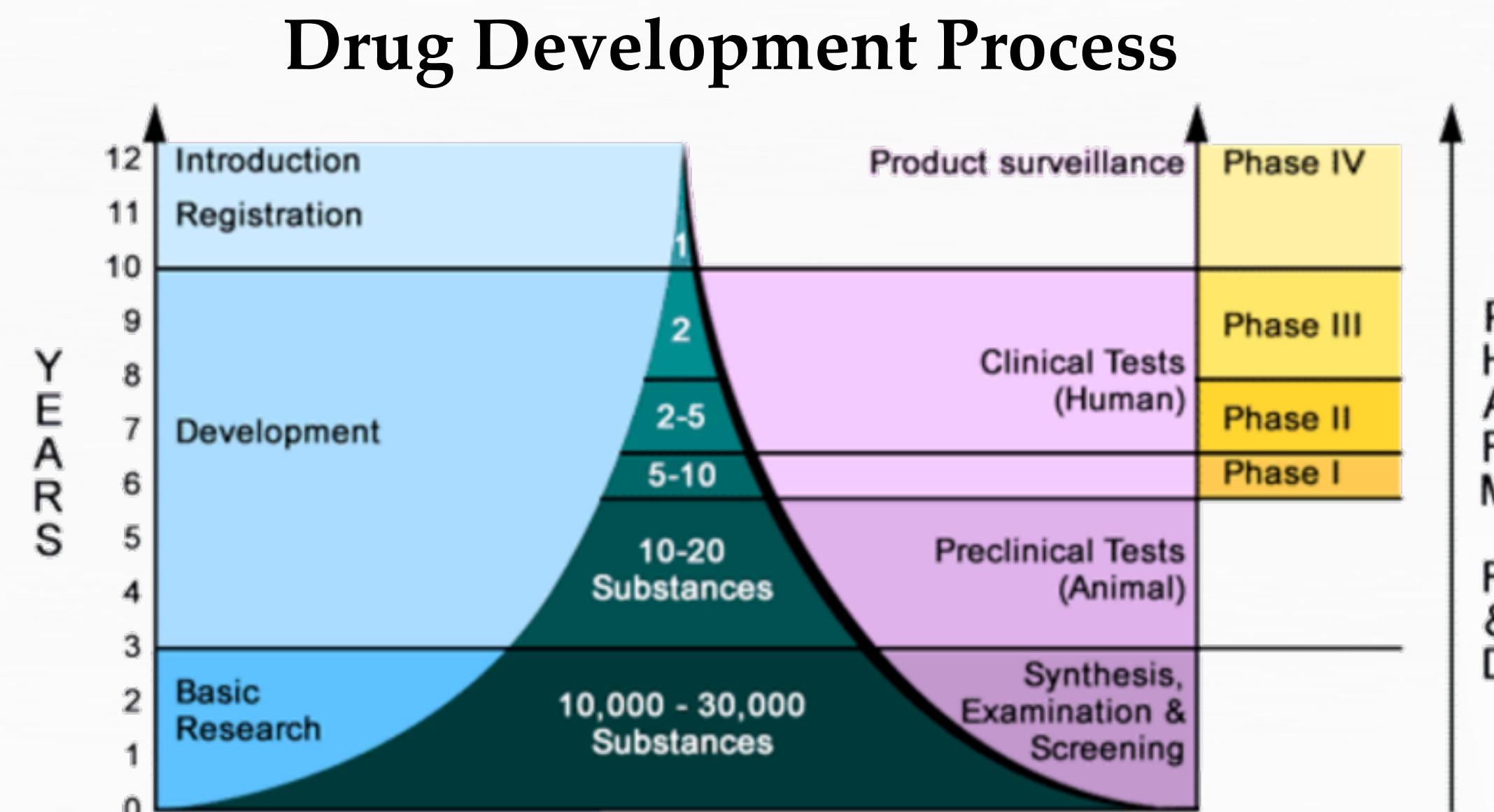
NEWS · 27 FEBRUARY 2020

Coronavirus puts drug repurposing on the fast track

Existing antivirals and knowledge gained from the SARS and MERS outbreaks gain traction as the fastest route to fight the current coronavirus epidemic.

Drug Repositioning

- A method for identifying and discovering new uses for existing drugs



It takes **too long** and **costs too much**
to bring new drugs to market

~ 15 years, US\$ 1 billions for one drug



nature

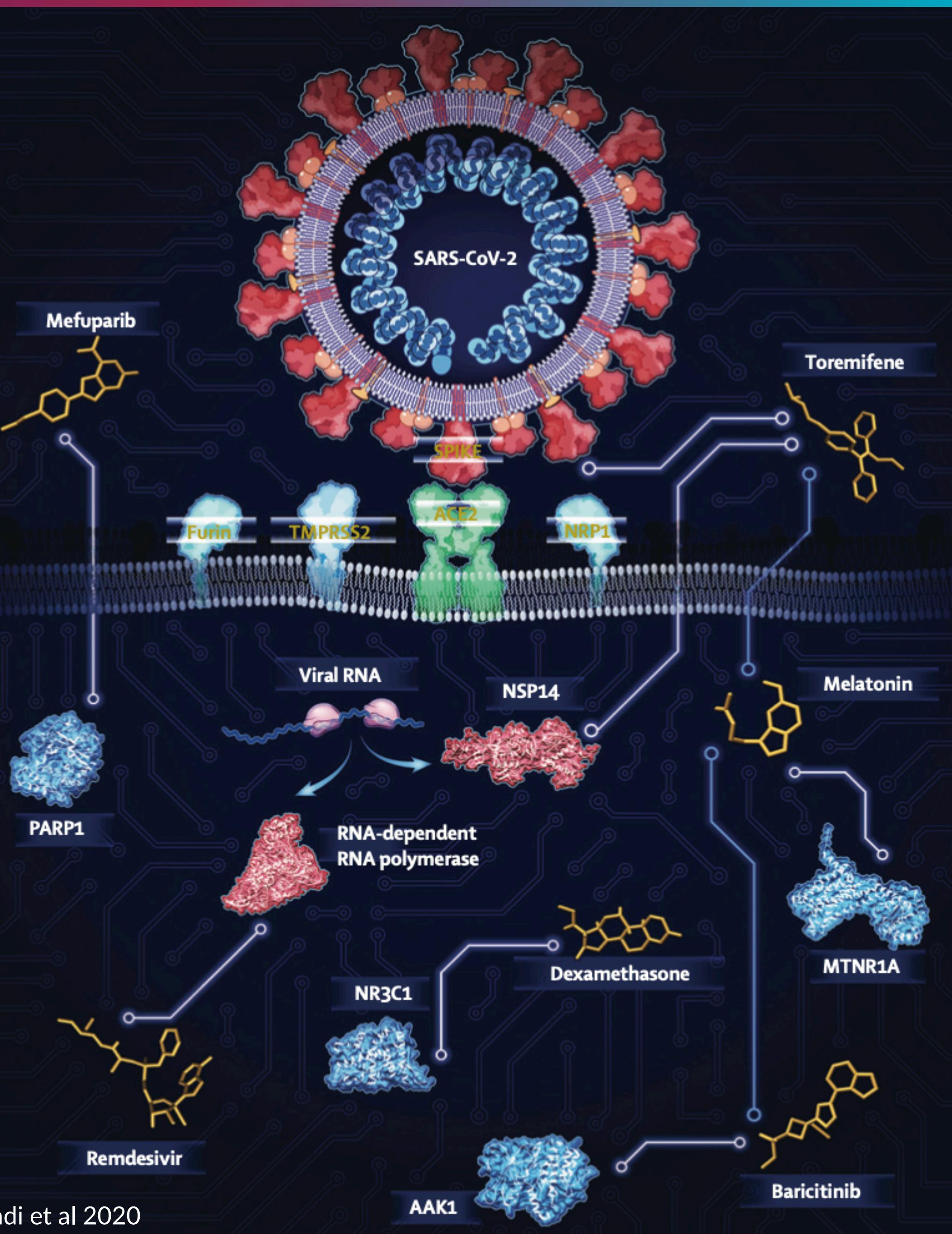
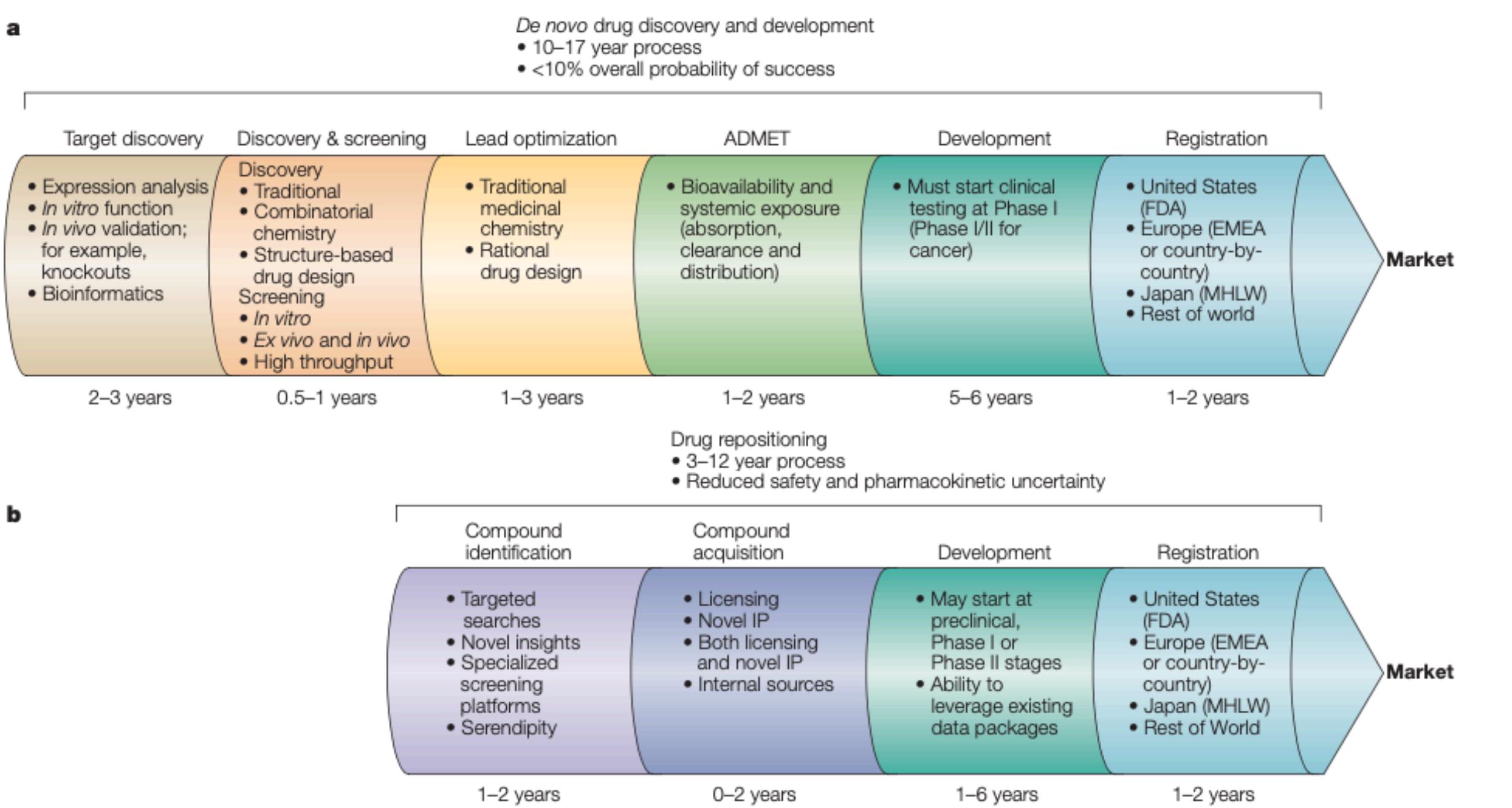
New uses for old drugs

2007

Curtis R. Chong¹ & David J. Sullivan, Jr²

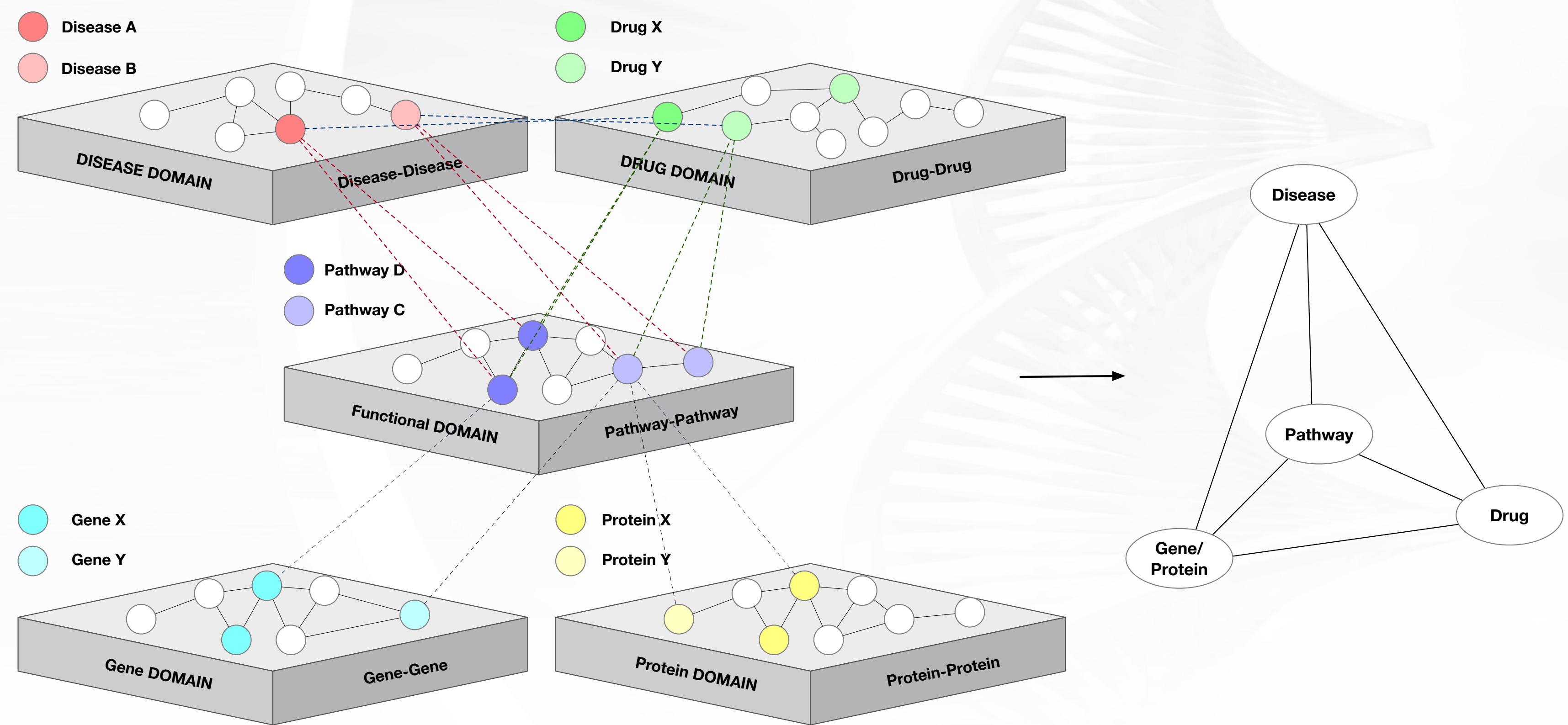
Advantages

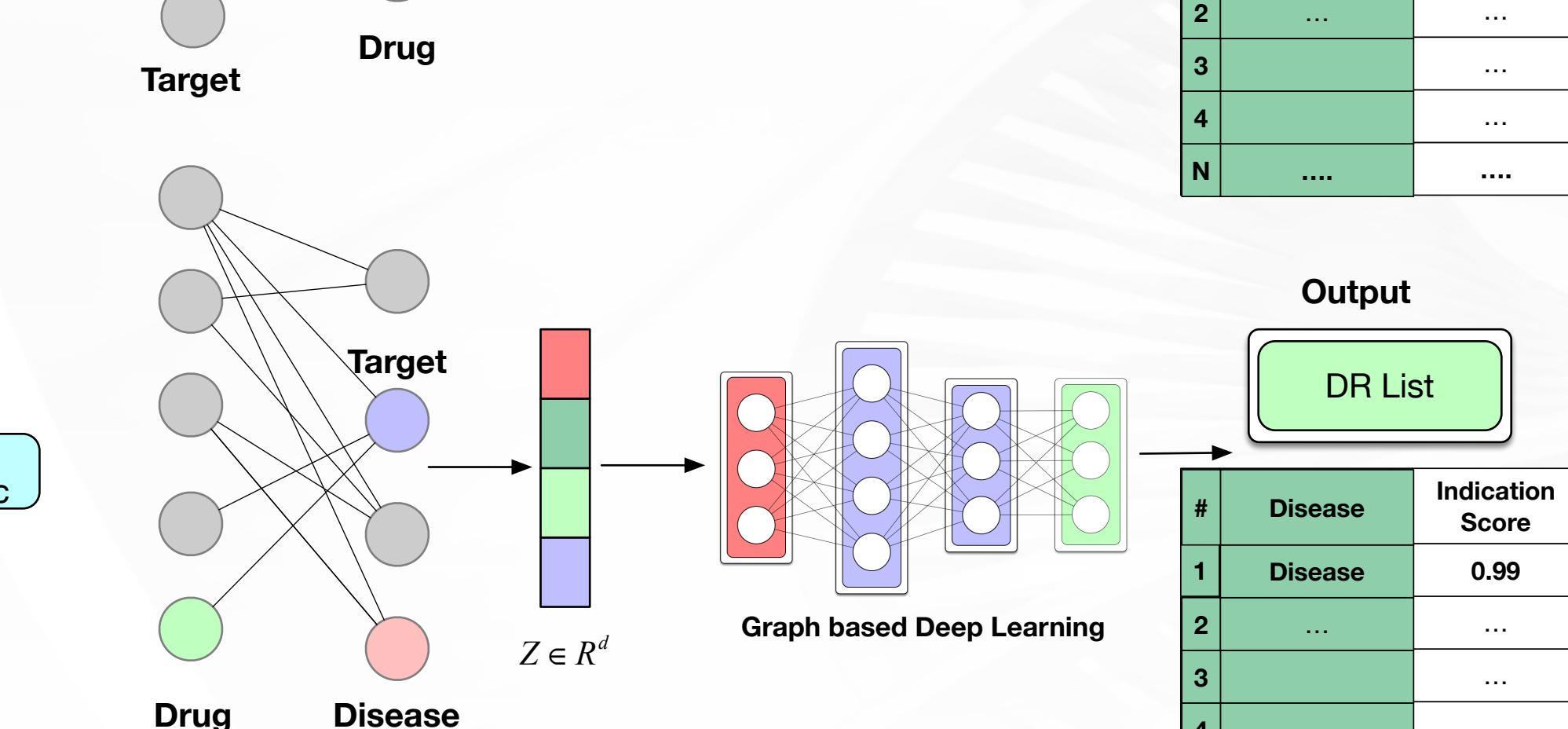
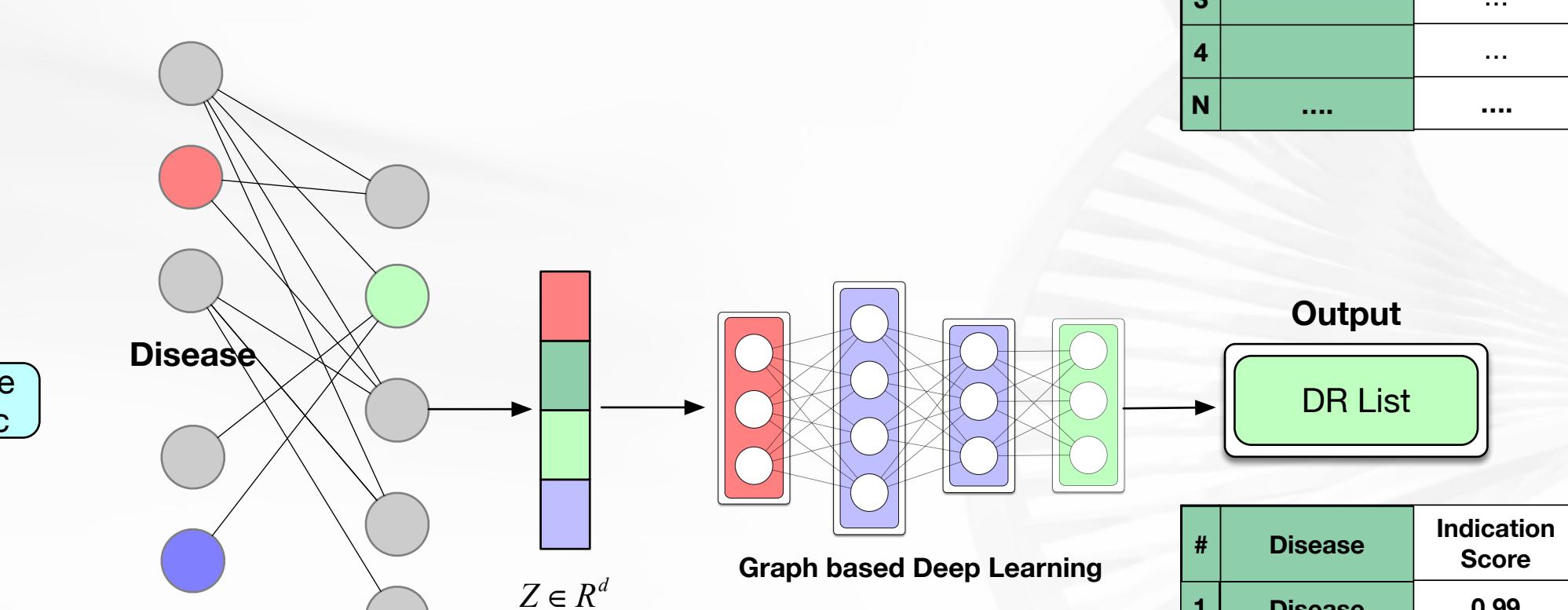
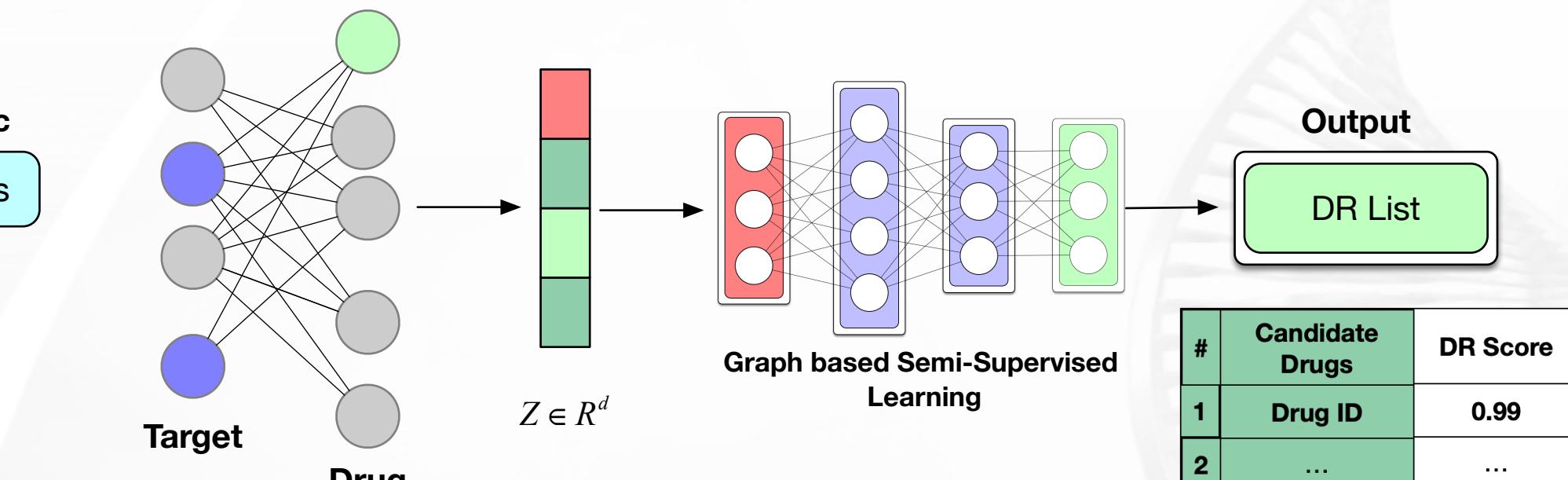
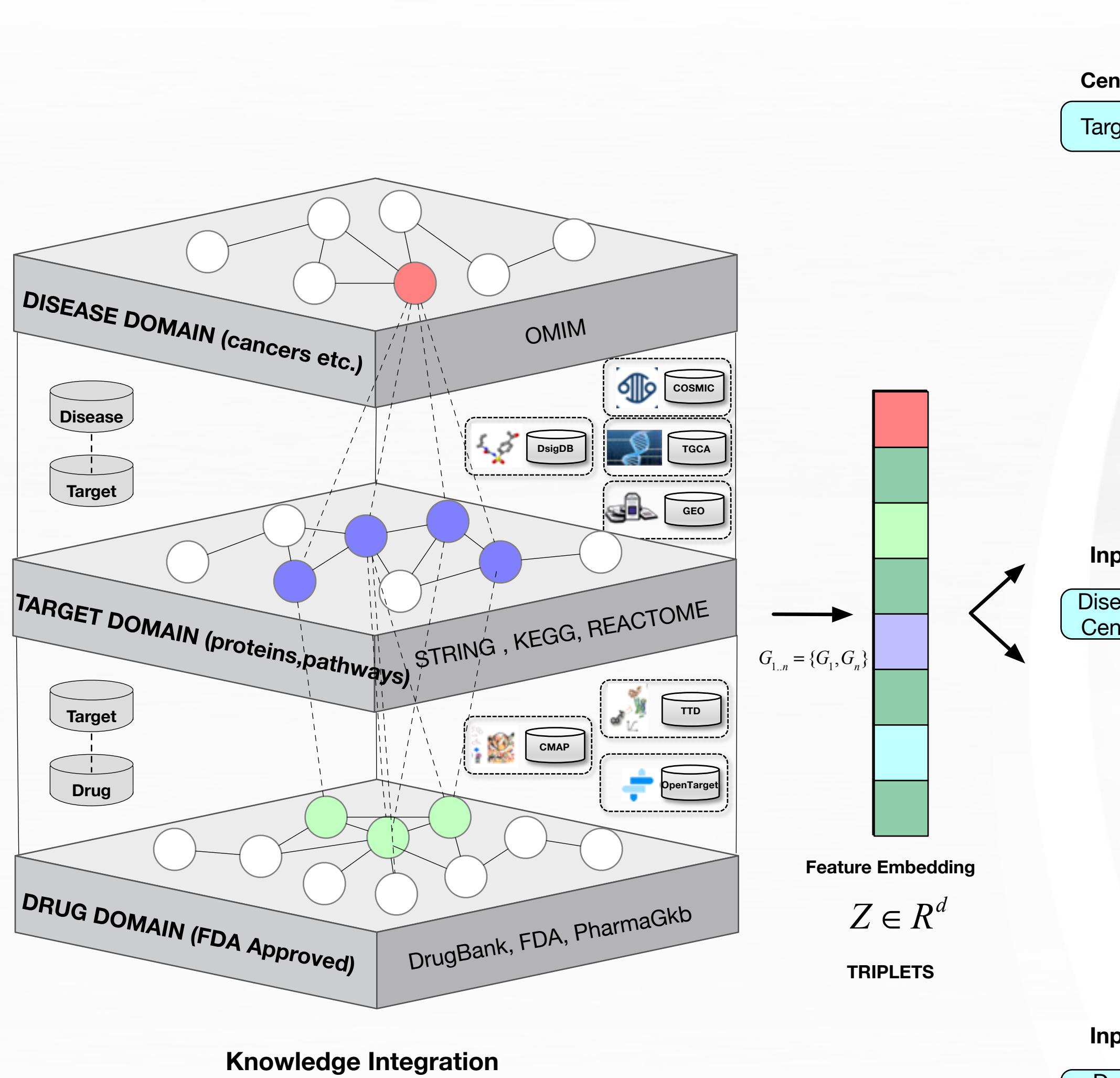
- Opportunity:
 - Reduced development timelines and overall cost
 - **Fast response to pandemic/epidemic (e.g: COVID19,)**



Traditional Method

- A classic way to repurpose drugs is through network medicine, which includes the construction of medical knowledge graphs
 - Target-centric
 - Disease-centric
 - Drug-centric





#	Candidate Drugs	DR Score
1	Drug ID	0.99
2
3
4
N

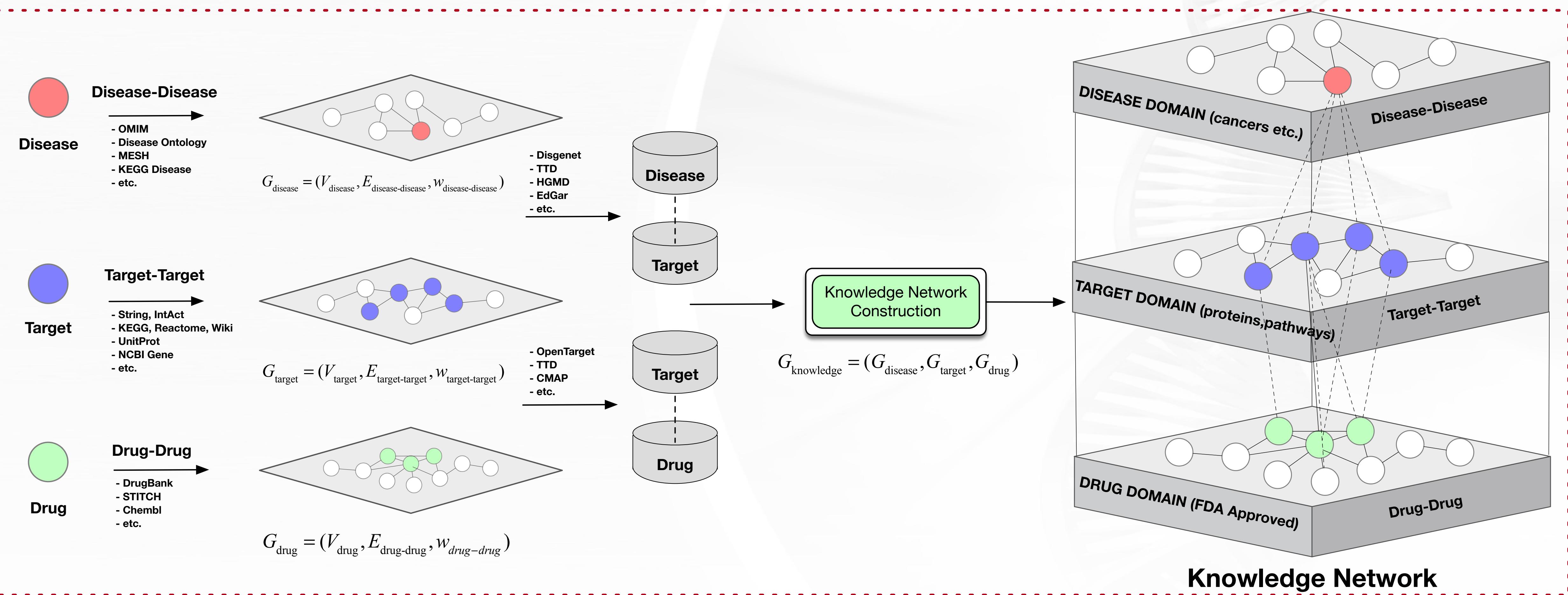
#	Disease	Indication Score
1	Disease	0.99
2
3
4
N

#	Disease	Indication Score
1	Disease	0.99
2
3
4
N

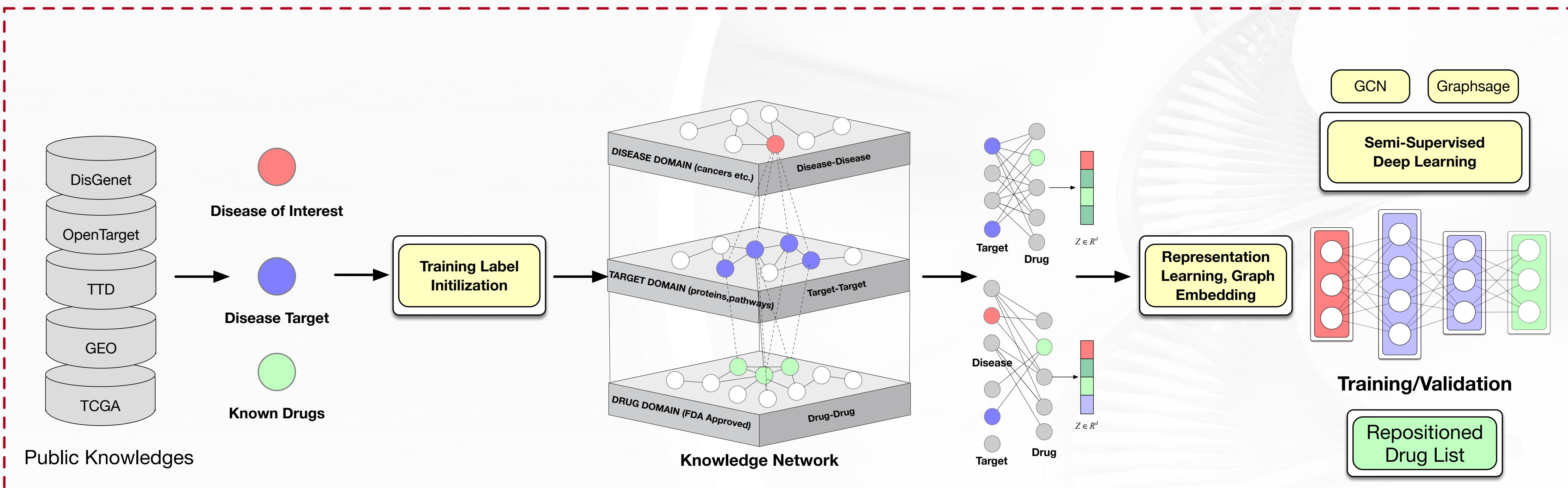
#	Drug	Indication Score
1	Drug ID	0.99
2
3
4
N

Initial Step

1) Knowledge Network Construction



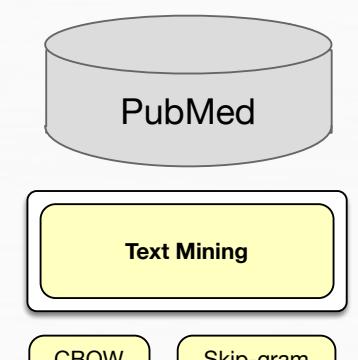
Prior Knowledge of Biomedical Entities



Validation (in silico)

a) Literature-based Validation

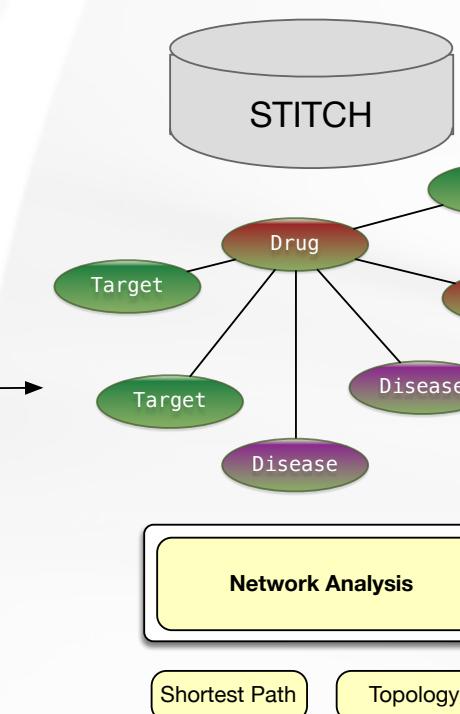
#	Candidate Drugs
1	Drug ID
2	...
3	...
4	...
N



#	Literature Validation Score
1	0.86
2	...
3	...
4	...
N

b) MOA-based Validation

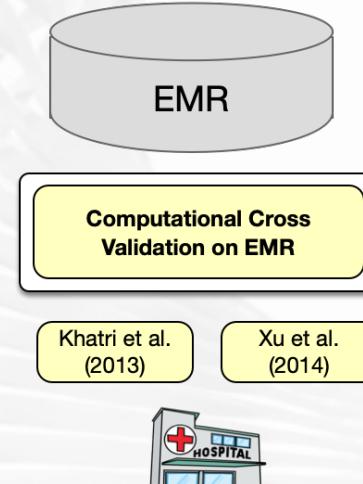
#	Candidate Drugs
1	Drug ID
2	...
3	...
4	...
N



#	Computational Validation Score
1	0.86
2	...
3	...
4	...
N

c) Real World Patient EMR-based Validation

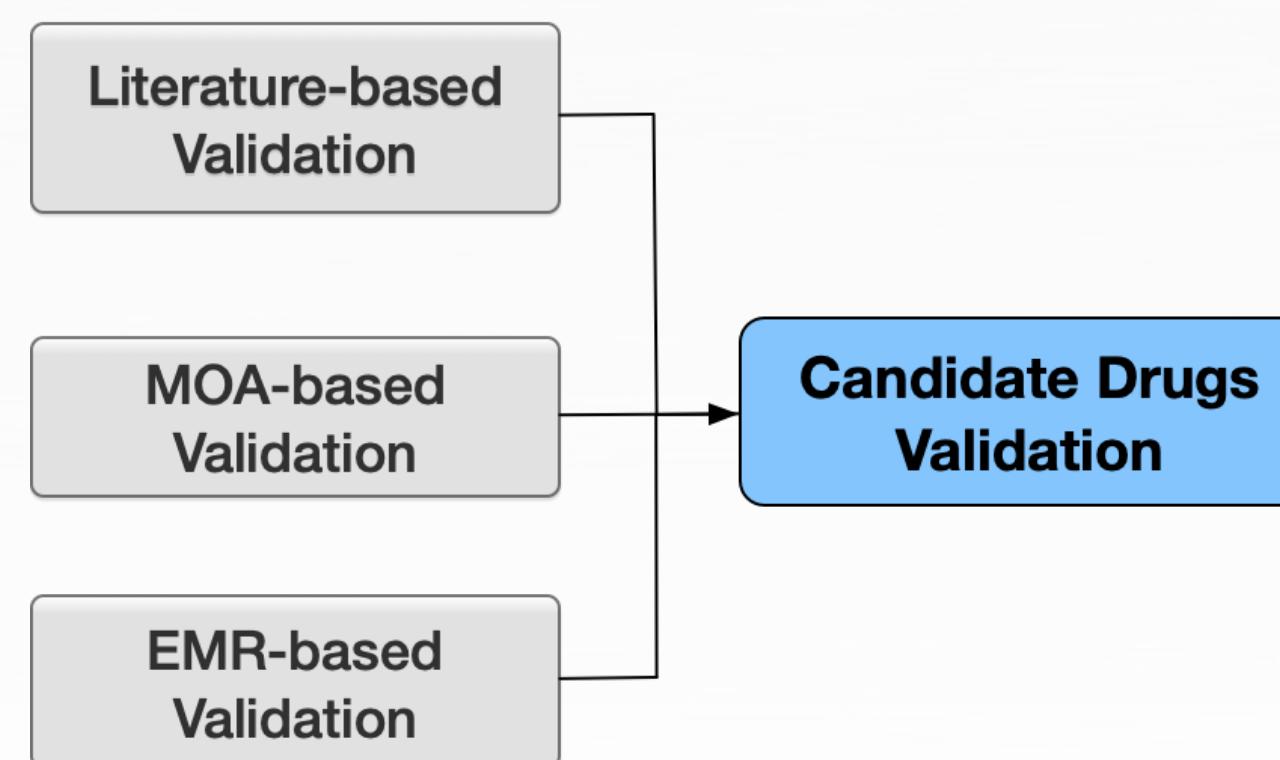
#	Candidate Drugs
1	Drug ID
2	...
3	...
4	...
N



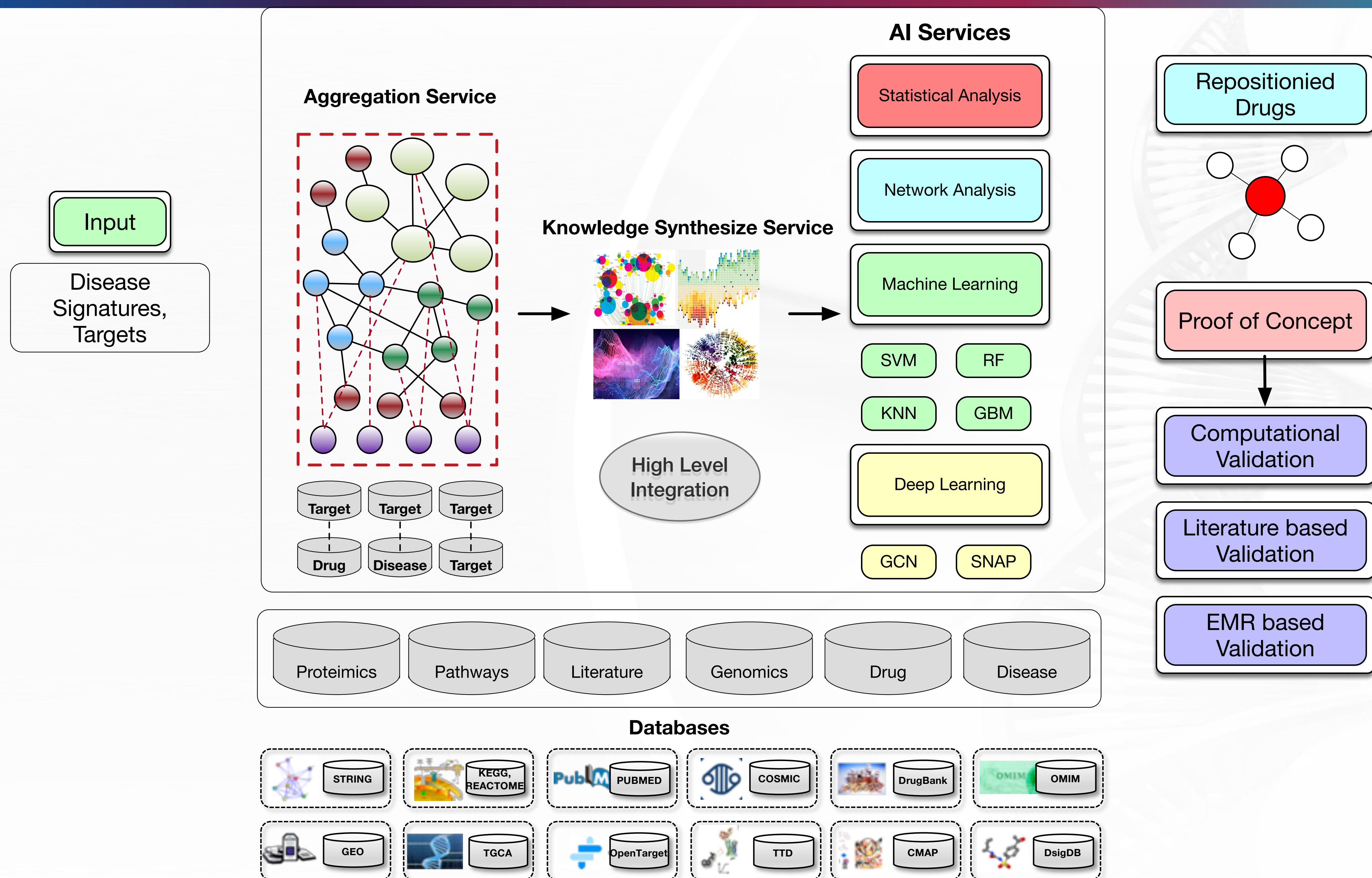
#	Computational Validation Score
1	0.83
2	...
3	...
4	...
N

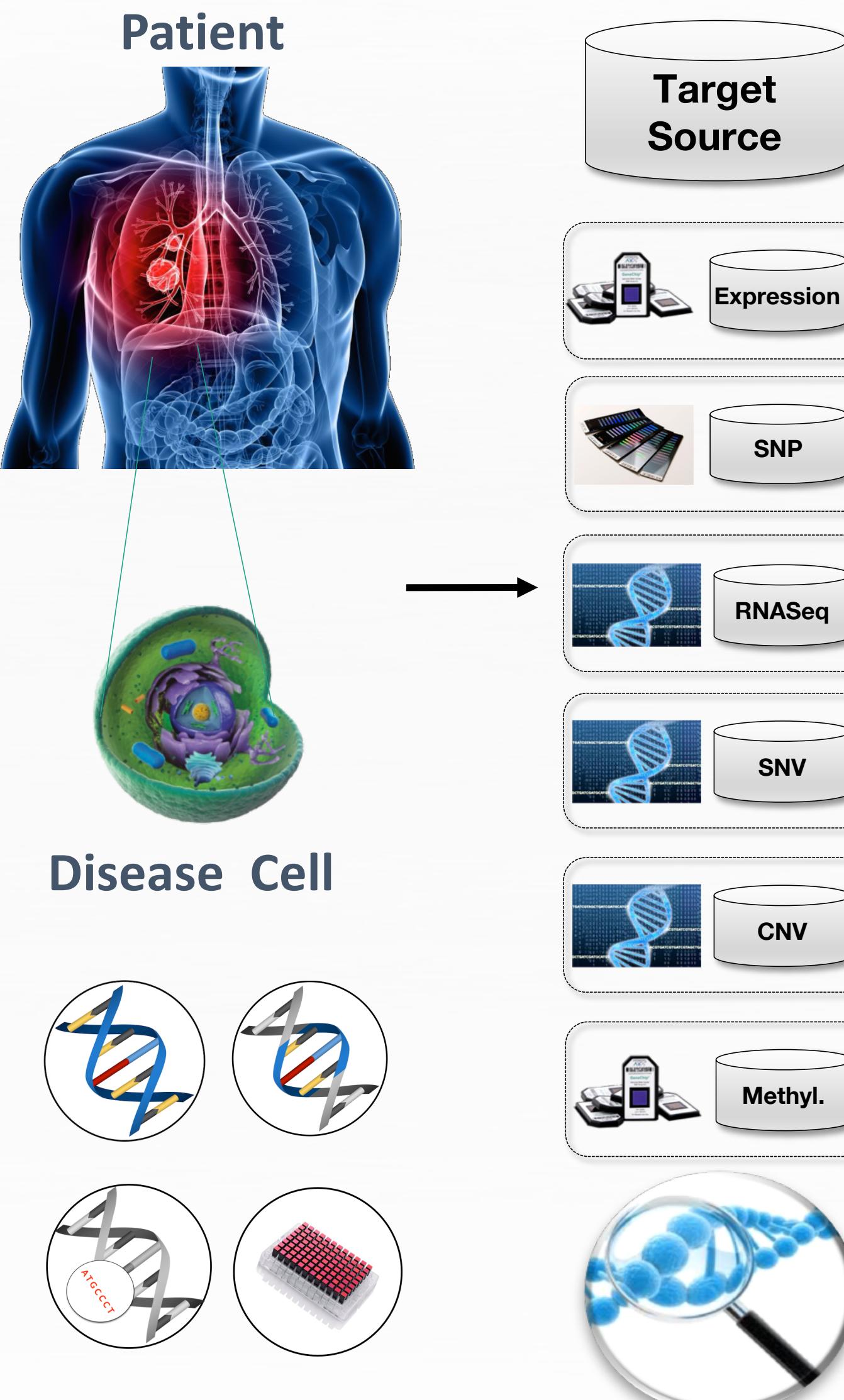
Yonsei Severance Hospital

Candidate Drugs Validation Score

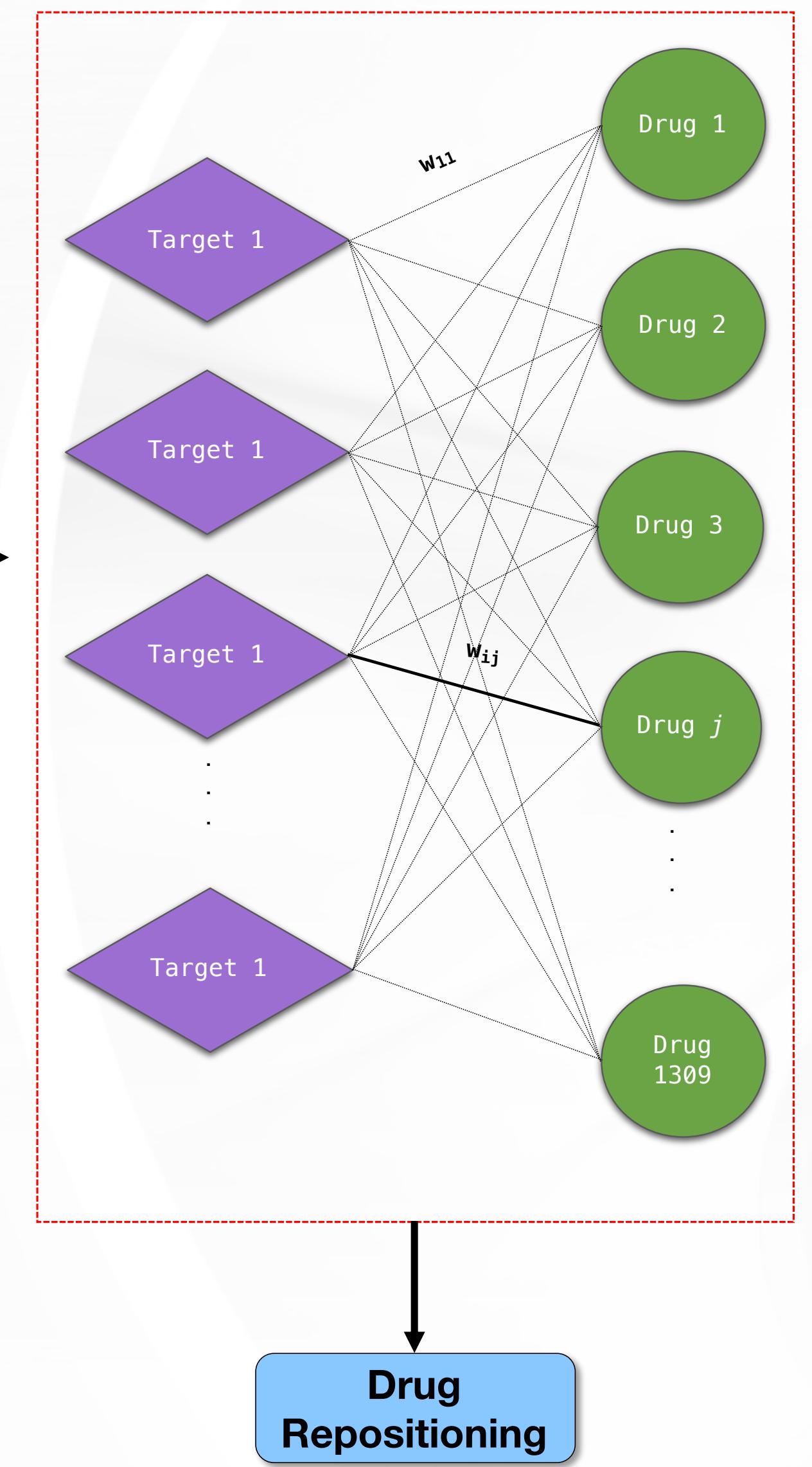


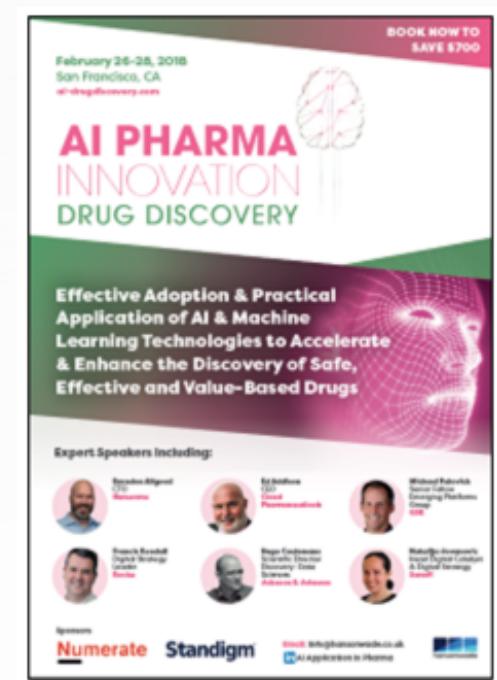
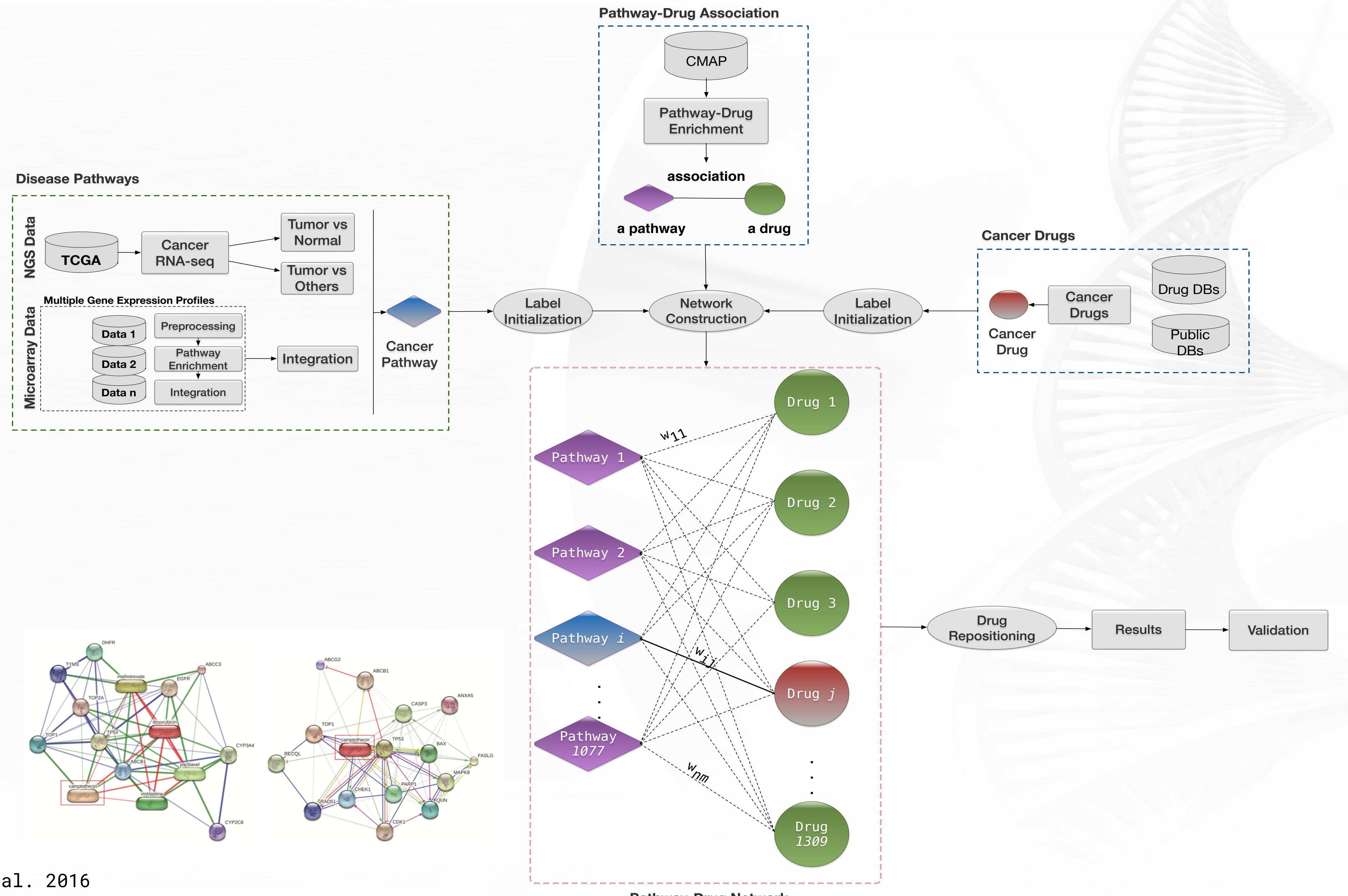
#	Candidate Drugs	Literature-based Validation Score	MOA-based Validation Score	EMR-based Validation Score	Candidate Drug Validation Score
1	Drug ID	0.86	0.99	0.83	0.88
2
3
4
N





EMBEDDED TRIPLET





2018: AI PHARM: Keynote - TRENDLIST

Erkhembayar J. et al. 2016

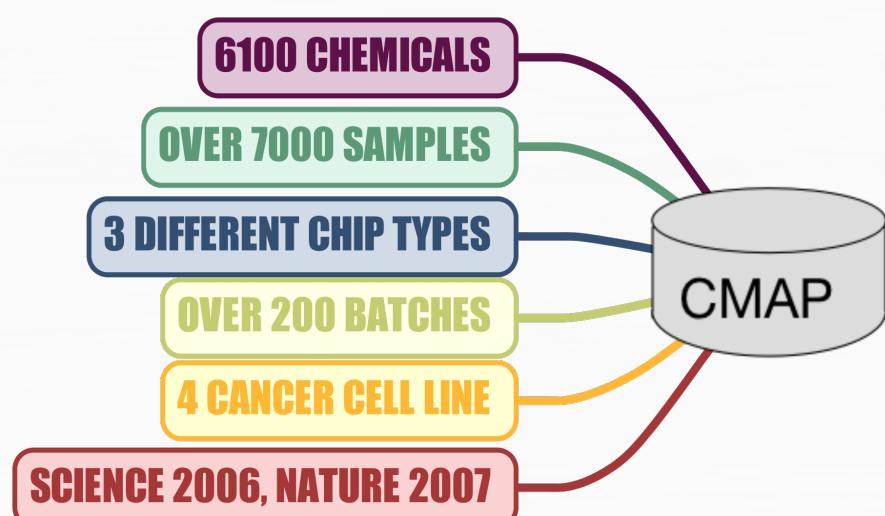
Erkhembayar J. KAICD

Pathway Drug Network ADVANCE TRAINING

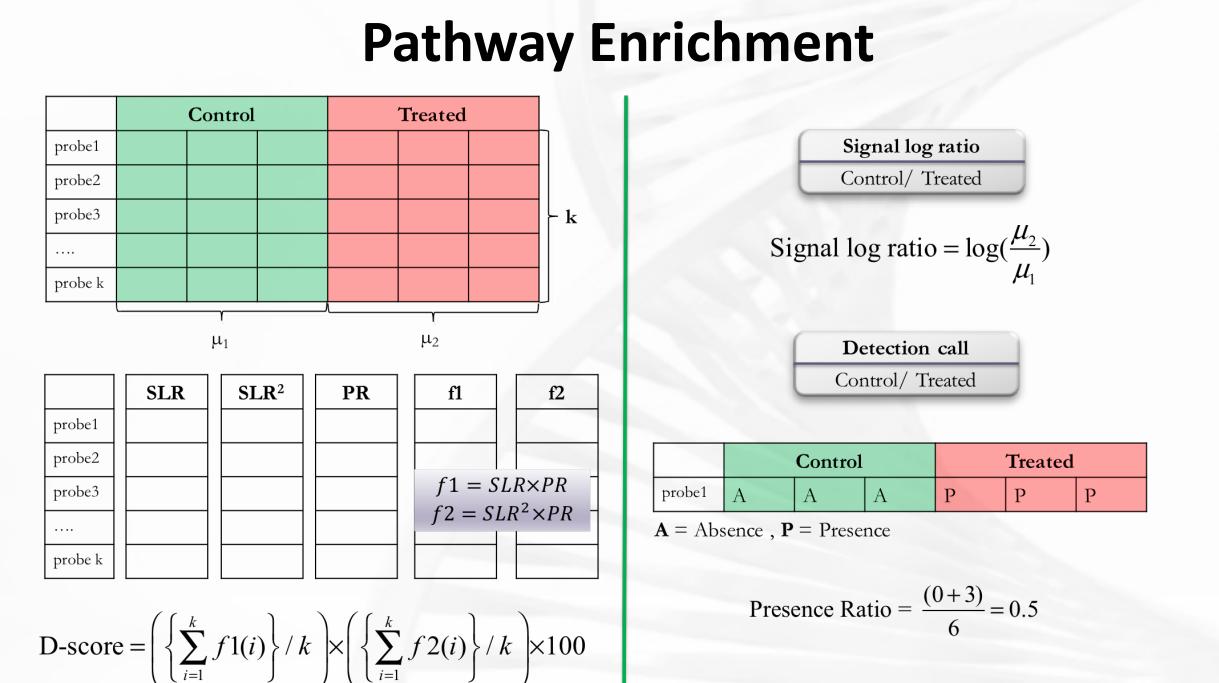
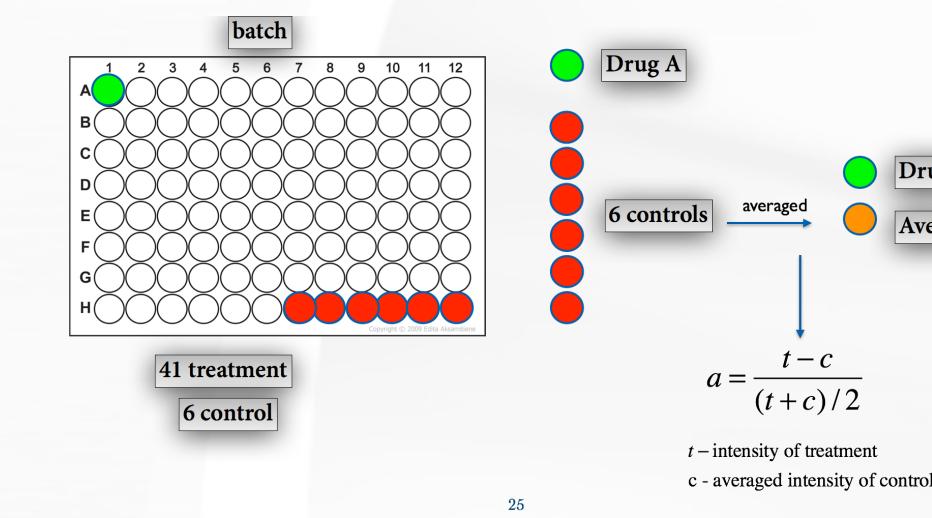
Database

- CMAP was mostly used for DR, but now LINCS

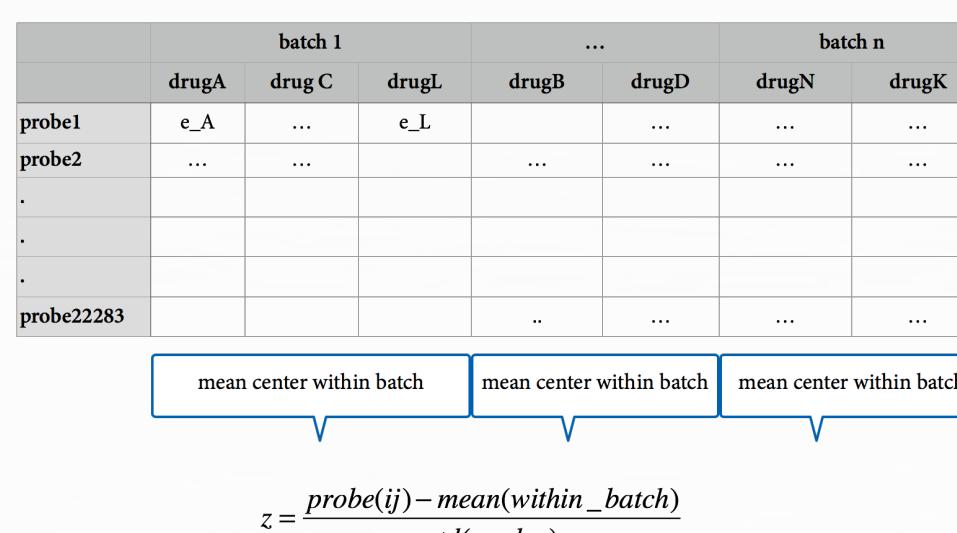
	HT_HG-U133A	HG-U133A	HT_HG-U133A_EA	Total
Treated	5242	674	184	6100
Control	787	133	36	956
Total	6029	807	220	7056



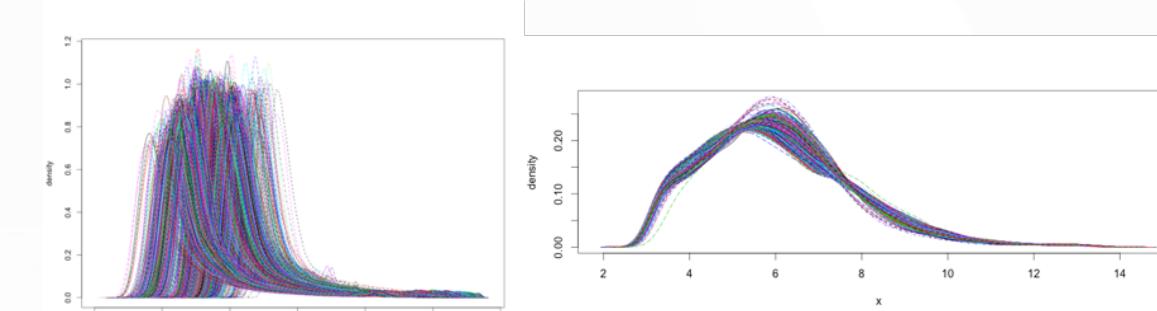
- ✓ HT-HG-U133A platform
- ✓ 106 batches
- ✓ 3906 samples
- ✓ 1309 drugs



- Data Preprocessing → Gene Expression Matrix → Pathway Enrichment → Drug-Pathway Enrichment Matrix
- ✓ Each cel file is are normalised by RMA
 - ✓ 222277 probes converted 19930 gene symbol
 - ✓ D-score enrichment
 - ✓ GSEA
 - ✓ Each Pathway enrichment scores for each drug



Batch effect removal



	Drug 1	Drug 2	Drug 3	Drug 1309
Pathway 1	Enrichment score				
Pathway 2					
Pathway 3					
...					
Pathway 1077					

Drug-Pathway Matrix

Zhou et al. *Cell Discovery* (2020)6:14
<https://doi.org/10.1038/s41421-020-0153-3>

ARTICLE

Cell Discovery
 www.nature.com/celldisc

Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2

 Yadi Zhou¹, Yuan Hou¹, Jiayu Shen¹, Yin Huang¹, William Martin¹ and Feixiong Cheng^{1,2,3}
Abstract

Human coronaviruses (HCoVs), including severe acute respiratory syndrome coronavirus (SARS-CoV) and 2019 novel coronavirus (2019-nCoV, also known as SARS-CoV-2), lead global epidemics with high morbidity and mortality. However, there are currently no effective drugs targeting 2019-nCoV/SARS-CoV-2. Drug repurposing, representing as an effective drug discovery strategy from existing drugs, could shorten the time and reduce the cost compared to de novo drug discovery. In this study, we present an integrative, antiviral drug repurposing methodology implementing a systems pharmacology-based network medicine platform, quantifying the interplay between the HCoV-host interactome and drug targets in the human protein–protein interaction network. Phylogenetic analyses of 15 HCoV whole genomes reveal that 2019-nCoV/SARS-CoV-2 shares the highest nucleotide sequence identity with SARS-CoV (79.7%). Specifically, the envelope and nucleocapsid proteins of 2019-nCoV/SARS-CoV-2 are two evolutionarily conserved regions, having the sequence identities of 96% and 89.6%, respectively, compared to SARS-CoV. Using network proximity analyses of drug targets and HCoV-host interactions in the human interactome, we prioritize 16 potential anti-HCoV repurposable drugs (e.g., melatonin, mercaptopurine, and sirolimus) that are further validated by enrichment analyses of drug-gene signatures and HCoV-induced transcriptomics data in human cell lines. We further identify three potential drug combinations (e.g., sirolimus plus dactinomycin, mercaptopurine plus melatonin, and toremifene plus emodin) captured by the “Complementary Exposure” pattern: the targets of the drugs both hit the HCoV-host subnetwork but target separate neighborhoods in the human interactome network. In summary, this study offers powerful network-based methodologies for rapid identification of candidate repurposable drugs and potential drug combinations targeting 2019-nCoV/SARS-CoV-2.

Introduction

Coronaviruses (CoVs) typically affect the respiratory tract of mammals, including humans, and lead to mild to severe respiratory tract infections¹. In the past two decades, two highly pathogenic human CoVs (HCoVs), including severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), emerging from animal reservoirs, have led to global epidemics with high morbidity and

mortality². For example, 8098 individuals were infected and 774 died in the SARS-CoV pandemic, which cost the global economy with an estimated \$30 to \$100 billion^{3,4}. According to the World Health Organization (WHO), as of November 2019, MERS-CoV has had a total of 2494 diagnosed cases causing 858 deaths, the majority in Saudi Arabia². In December 2019, the third pathogenic HCoV, named 2019 novel coronavirus (2019-nCoV/SARS-CoV-2), as the cause of coronavirus disease 2019 (abbreviated as COVID-19)⁵, was found in Wuhan, China. As of 24 February 2020, there have been over 79,000 cases with over 2600 deaths for the 2019-nCoV/SARS-CoV-2 outbreak worldwide; furthermore, human-to-human transmission has occurred among close contacts⁶. However, there are currently no effective medications against 2019-

 Correspondence: Feixiong Cheng (chengf@ccf.org)

¹Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA

²Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, OH 44195, USA

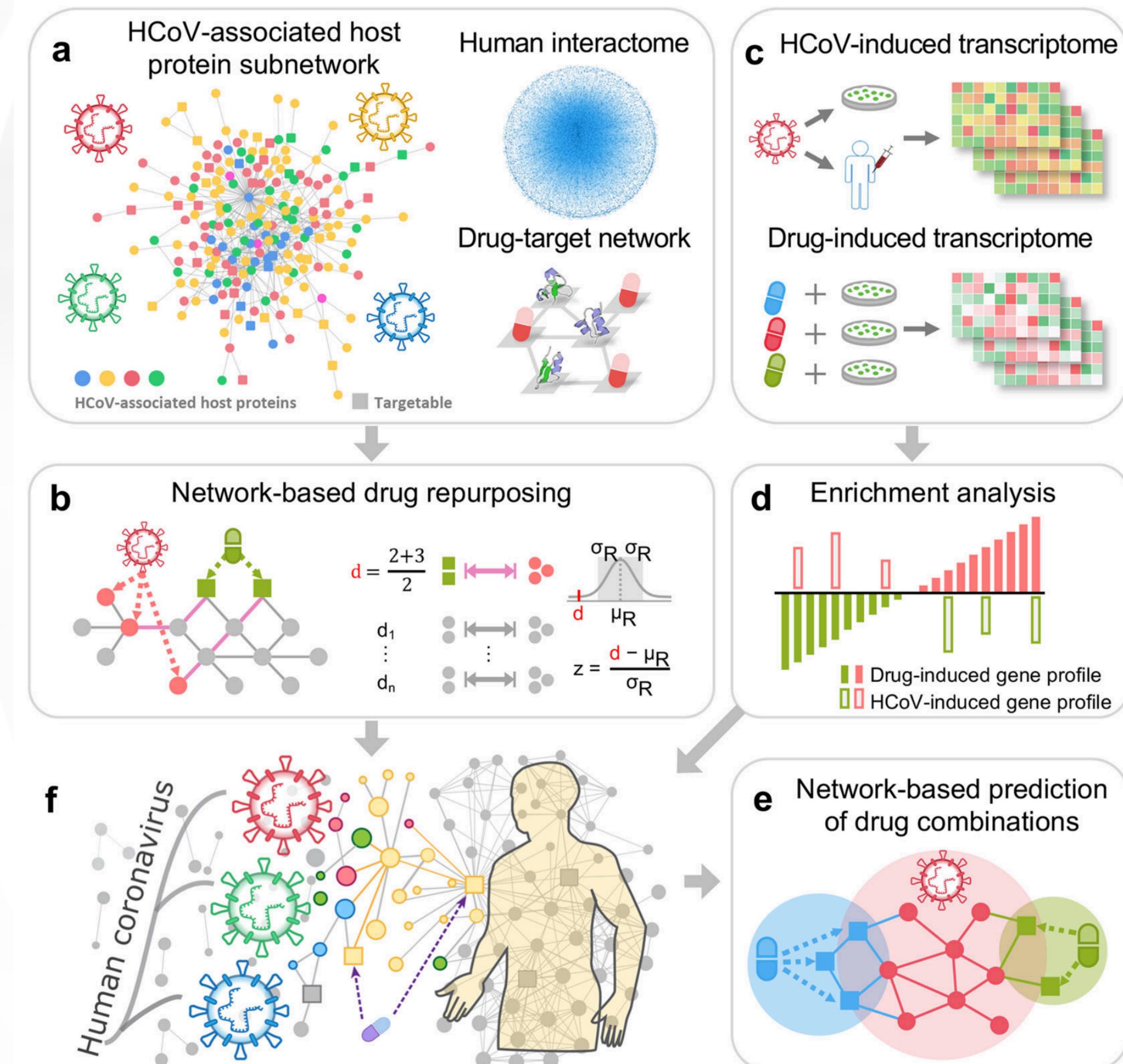
Full list of author information is available at the end of the article

These authors contributed equally: Yadi Zhou, Yuan Hou

© The Author(s) 2020

 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Nature / Cell Discovery, 16 March 2020



bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.11.986836>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

A data-driven drug repositioning framework discovered a potential therapeutic agent targeting COVID-19

Yiyue Ge^{1,2,†}, Tingzhong Tian^{1,2,†}, Suling Huang^{3,†}, Fangping Wan^{1,†}, Jingxin Li^{2,†}, Shuya Li¹, Hui Yang¹¹, Lixiang Hong¹, Nian Wu¹, Enming Yuan¹, Lili Cheng⁴, Yipin Lei¹¹, Hantao Shu¹, Xiaolong Feng^{6,7}, Ziyuan Jiang⁵, Ying Chi², Xiling Guo², Lumbiao Cui², Liang Xiao¹⁰, Zeng Li¹⁰, Chunhao Yang³, Zehong Miao³, Haidong Tang⁴, Ligong Chen⁴, Haimian Zeng¹¹, Dan Zhao^{1,*}, Fengcai Zhu^{2,8,*}, Xiaokun Shen^{10,*}, Jianyang Zeng^{1,9,*}

¹Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, 100084, China.

²NHC Key laboratory of Enteric Pathogenic Microbiology, Jiangsu Provincial Center for Diseases Control and Prevention, Nanjing, Jiangsu Province, 210009, China.

³Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China.

⁴School of Pharmaceutical Sciences, Beijing Advanced Innovation Center for Structural Biology, Tsinghua University, Beijing, 100084, China.

⁵Department of Automation, Tsinghua University, Beijing, 100084, China.

⁶School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, Hubei Province, 430074, China.

⁷Institute of Pathology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, 430030, China.

⁸Center for Global Health, Nanjing Medical University, Nanjing, Jiangsu Province, 210009, China.

⁹MOE Key Laboratory of Bioinformatics, Tsinghua University, Beijing, 100084, China.

¹⁰Convalife (Shanghai) Co., Ltd., Shanghai, 201203, China.

¹¹Silexon AI Technology Co., Ltd., Nanjing, Jiangsu Province, 210033, China.

[†]These authors contributed equally to this work.

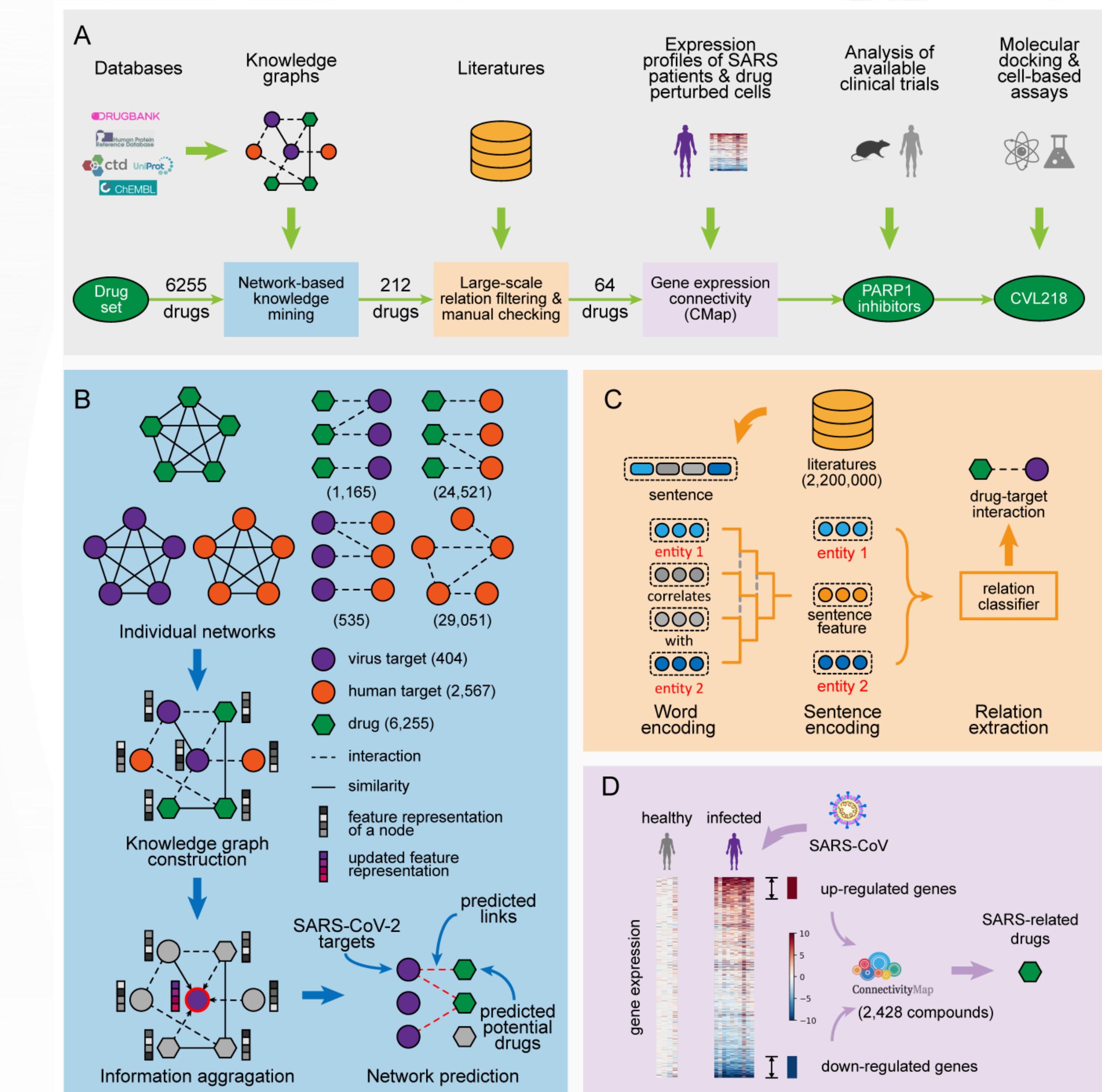
^{*}Corresponding authors.

Abstract

The global spread of SARS-CoV-2 requires an urgent need to find effective therapeutics for the treatment of COVID-19. We developed a data-driven drug repositioning framework, which applies both machine learning and statistical analysis approaches to systematically integrate and mine large-scale knowledge graph, literature and transcriptome data to discover the potential drug candidates against SARS-CoV-2. The retrospective study using the past SARS-CoV and MERS-CoV data demonstrated that our machine learning based method can successfully predict effective drug candidates

Email addresses: zhaodan2018@tsinghua.edu.cn (Dan Zhao), jszfc@vip.sina.com (Fengcai Zhu), steve.shen@convalife.com (Xiaokun Shen), zengjy321@tsinghua.edu.cn (Jianyang Zeng)

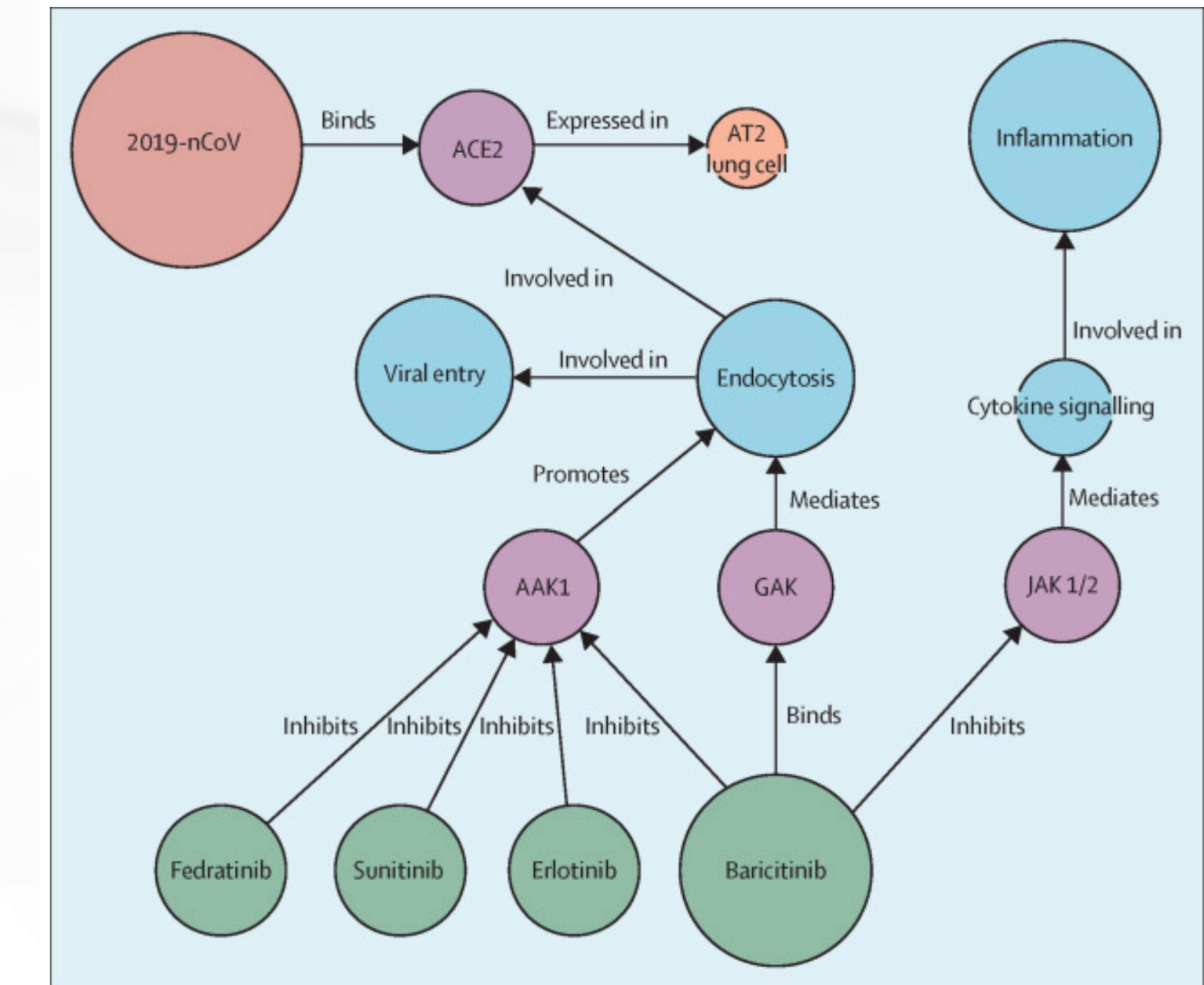
March 11, 2020



▪ BenevolentAI



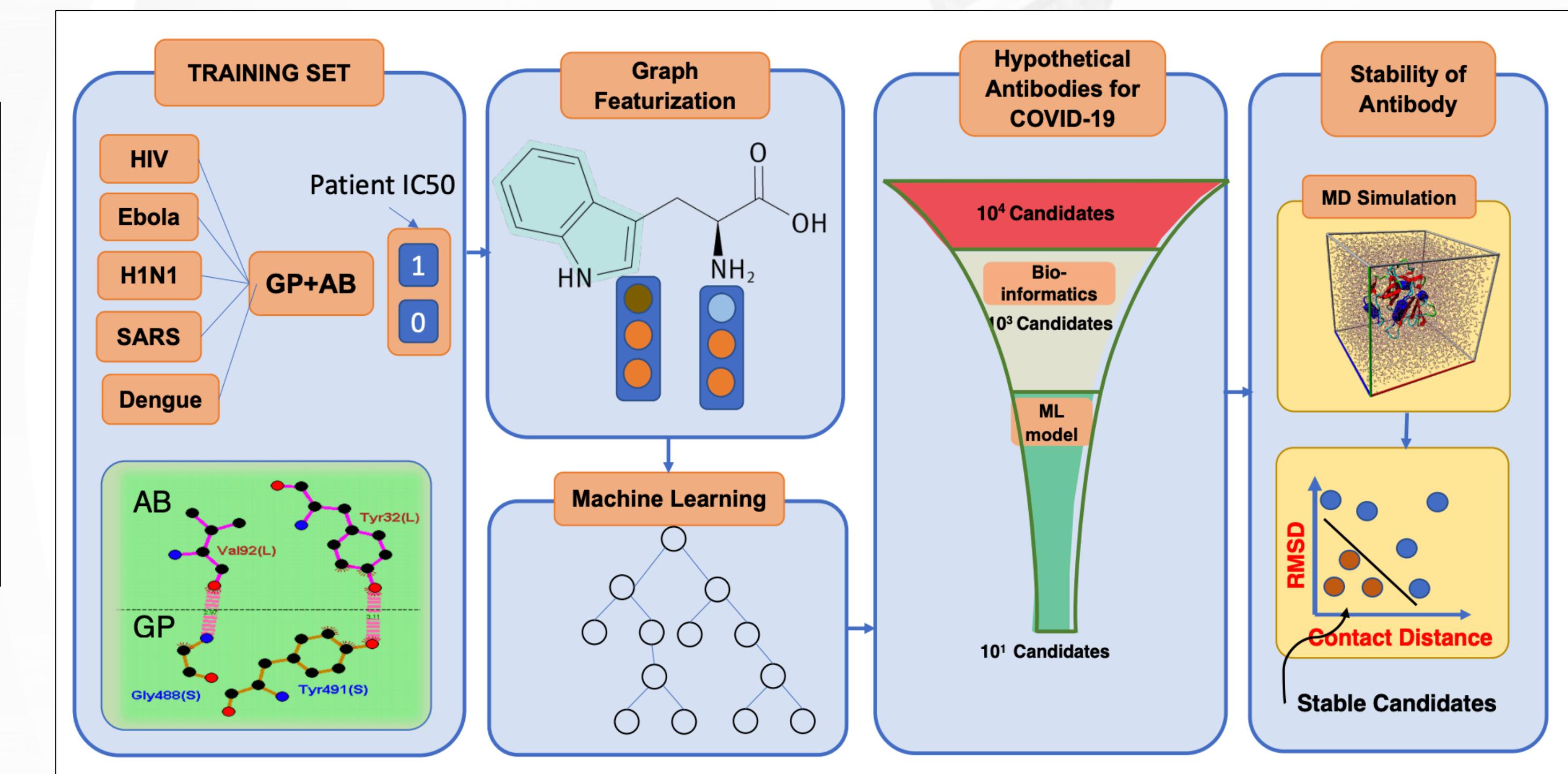
LITERATURE BASED DR



bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.14.992156>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Potential Neutralizing Antibodies Discovered for Novel Corona Virus Using Machine Learning

Rishikesh Magar ¹, Prakarsh Yadav ² and Amir Barati Farimani ^{1*}



Fast Drug Listing

bioRxiv preprint doi: <https://doi.org/10.1101/2020.02.05.936013>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Potentially highly potent drugs for 2019-nCoV

Duc Duy Nguyen¹, Kaifu Gao¹, Jiahui Chen¹, Rui Wang¹, and Guo-Wei Wei^{1,2,3 *}

¹ Department of Mathematics, Michigan State University, MI 48824, USA

² Department of Biochemistry and Molecular Biology
Michigan State University, MI 48824, USA

³ Department of Electrical and Computer Engineering
Michigan State University, MI 48824, USA

February 5, 2020

Abstract

The World Health Organization (WHO) has declared the 2019 novel coronavirus (2019-nCoV) infection outbreak a global health emergency. Currently, there is no effective anti-2019-nCoV medication. The sequence identity of the 3CL proteases of 2019-nCoV and SARS is 96%, which provides a sound foundation for structural-based drug repositioning (SBDR). Based on a SARS 3CL protease X-ray crystal structure, we construct a 3D homology structure of 2019-nCoV 3CL protease. Based on this structure and existing experimental datasets for SARS 3CL protease inhibitors, we develop an SBDR model based on machine learning and mathematics to screen 1465 drugs in the DrugBank that have been approved by the U.S. Food and Drug Administration (FDA). We found that many FDA approved drugs are potentially highly potent to 2019-nCoV.

Key words: 2019-nCoV, Drug repositioning, DrugBank, deep learning, algebraic topology.

* Address correspondences to Guo-Wei Wei. E-mail:wei@math.msu.edu

Predicting commercially available antiviral drugs that may act on the novel coronavirus (2019-nCoV), Wuhan, China through a drug-target interaction deep learning model

Bo Ram Beck¹, Bonggun Shin^{1,2}, Yoonjung Choi¹, Sungsoo Park^{1,*}, Keunsoo Kang^{3,*}

¹Deargen, Inc., Daejeon, Republic of Korea

²Department of Computer Science, Emory University, Atlanta, GA, United States

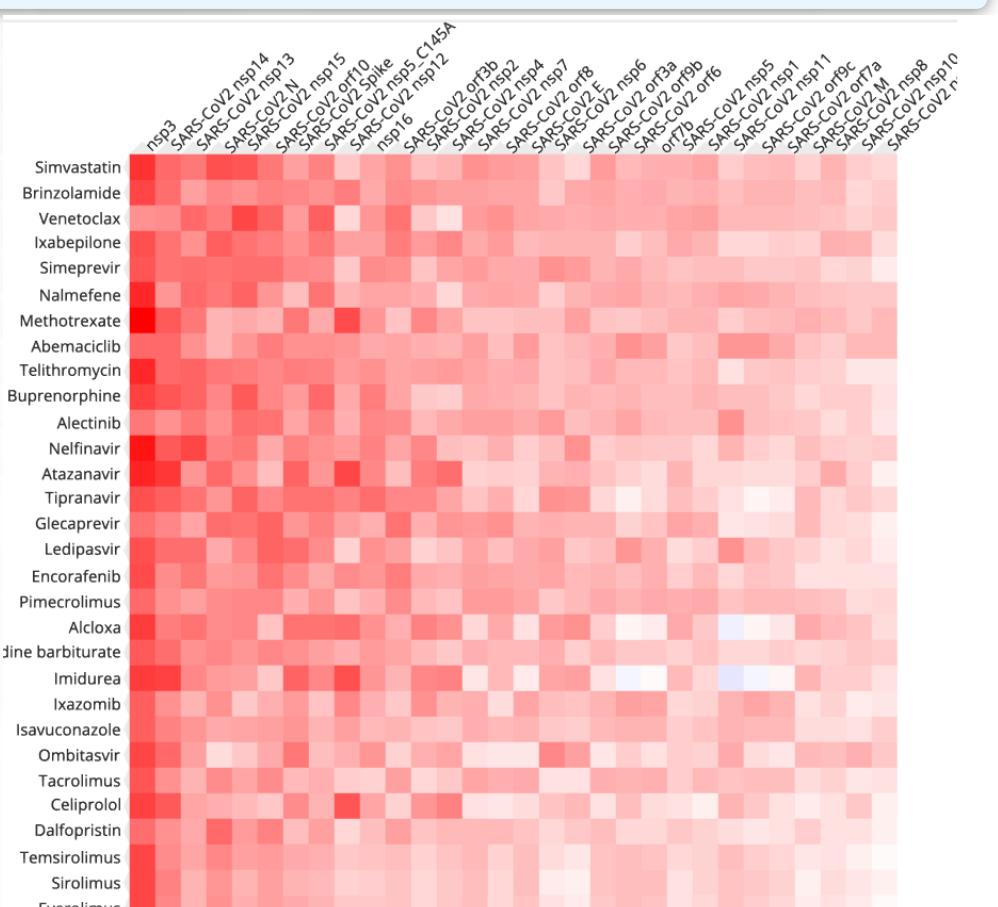
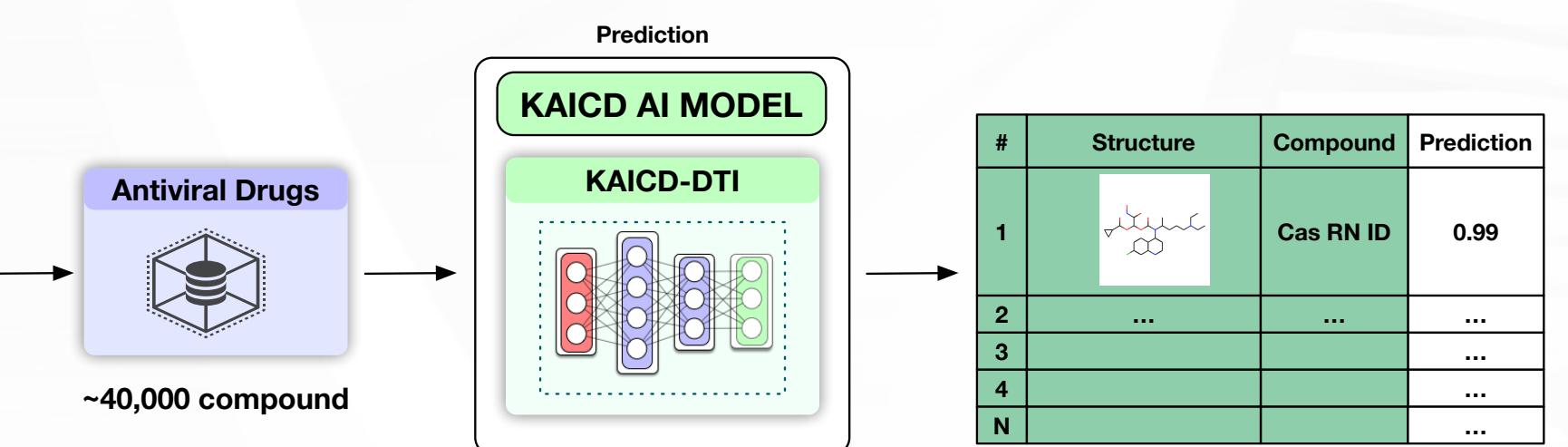
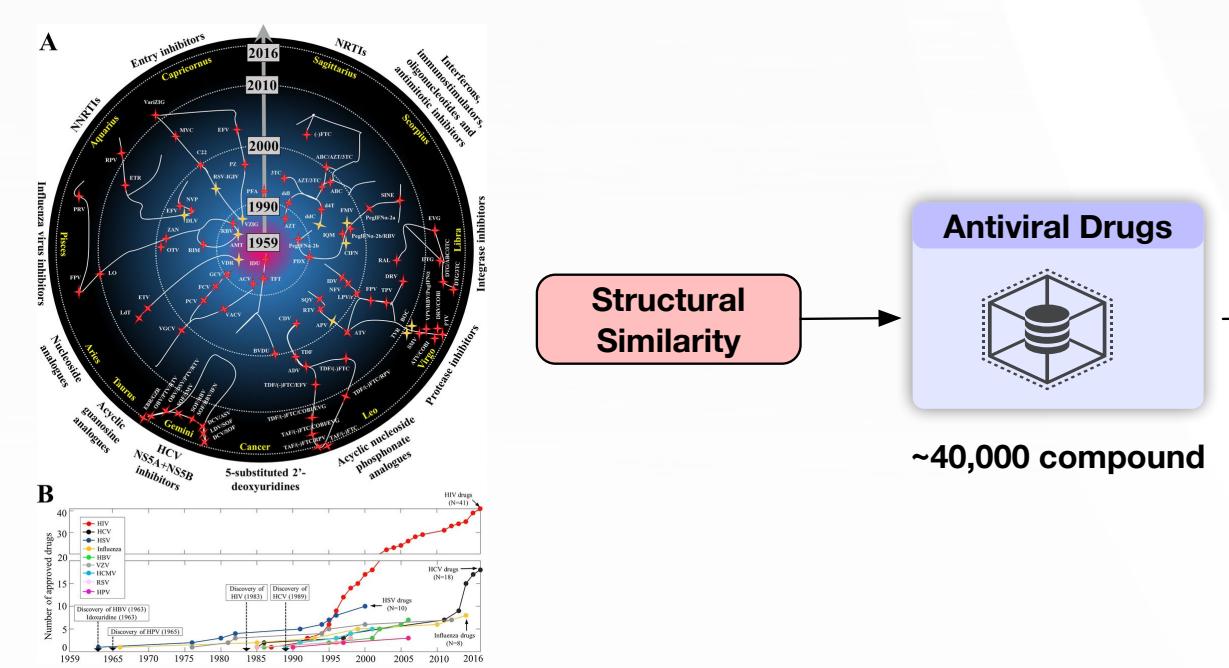
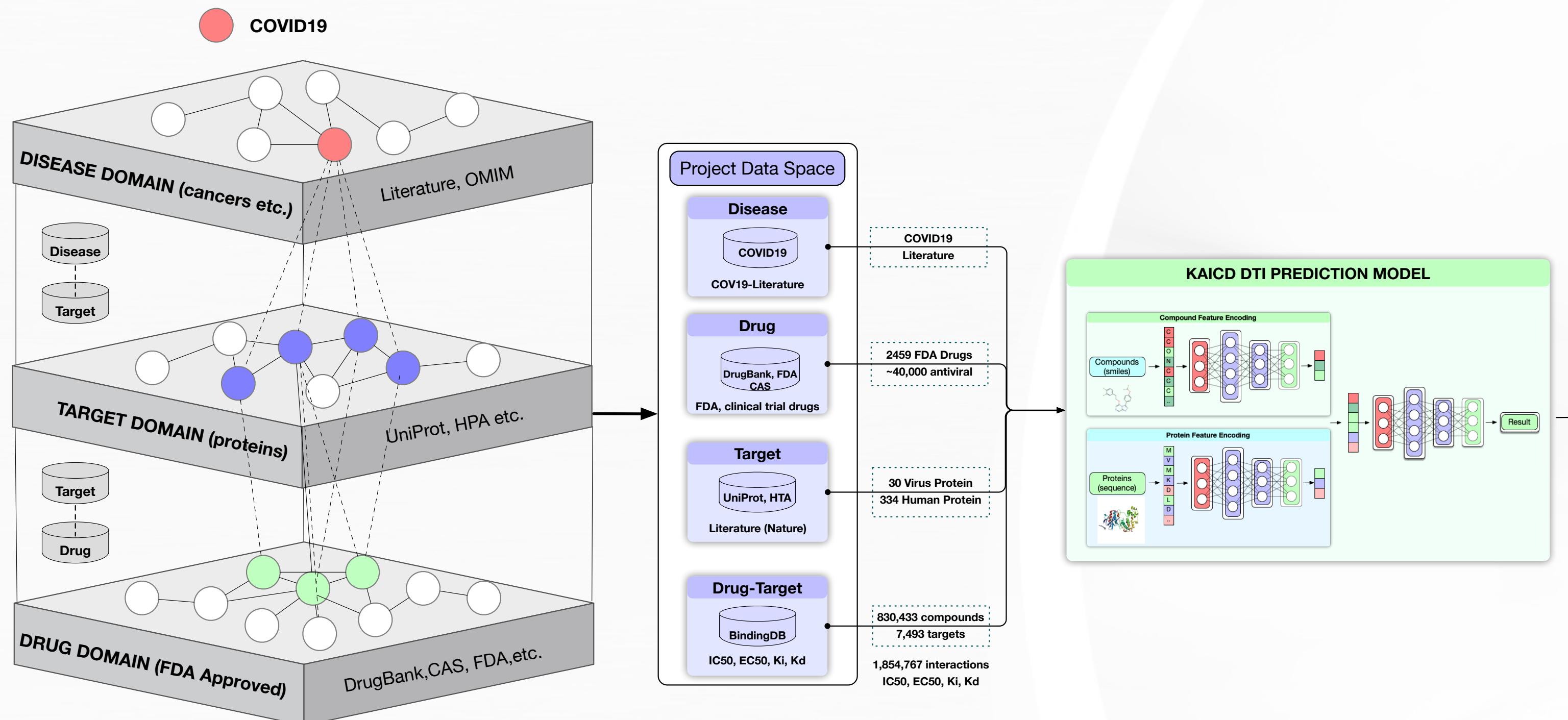
³Department of Microbiology, College of Natural Sciences, Dankook University, Cheonan, Republic of Korea

COVID19 SEMI REPOSITIONING

DRUG REPOSITIONING



KAICD
한국인공지능신약개발지원센터
Korea AI Center for Drug Discovery and Development



Data Space

■ Interaction DB (BindingDB)

- 830,433 small molecules
- 7,493 protein targets
- 1,854,767 interactions (IC₅₀, EC₅₀, Ki, Kd)

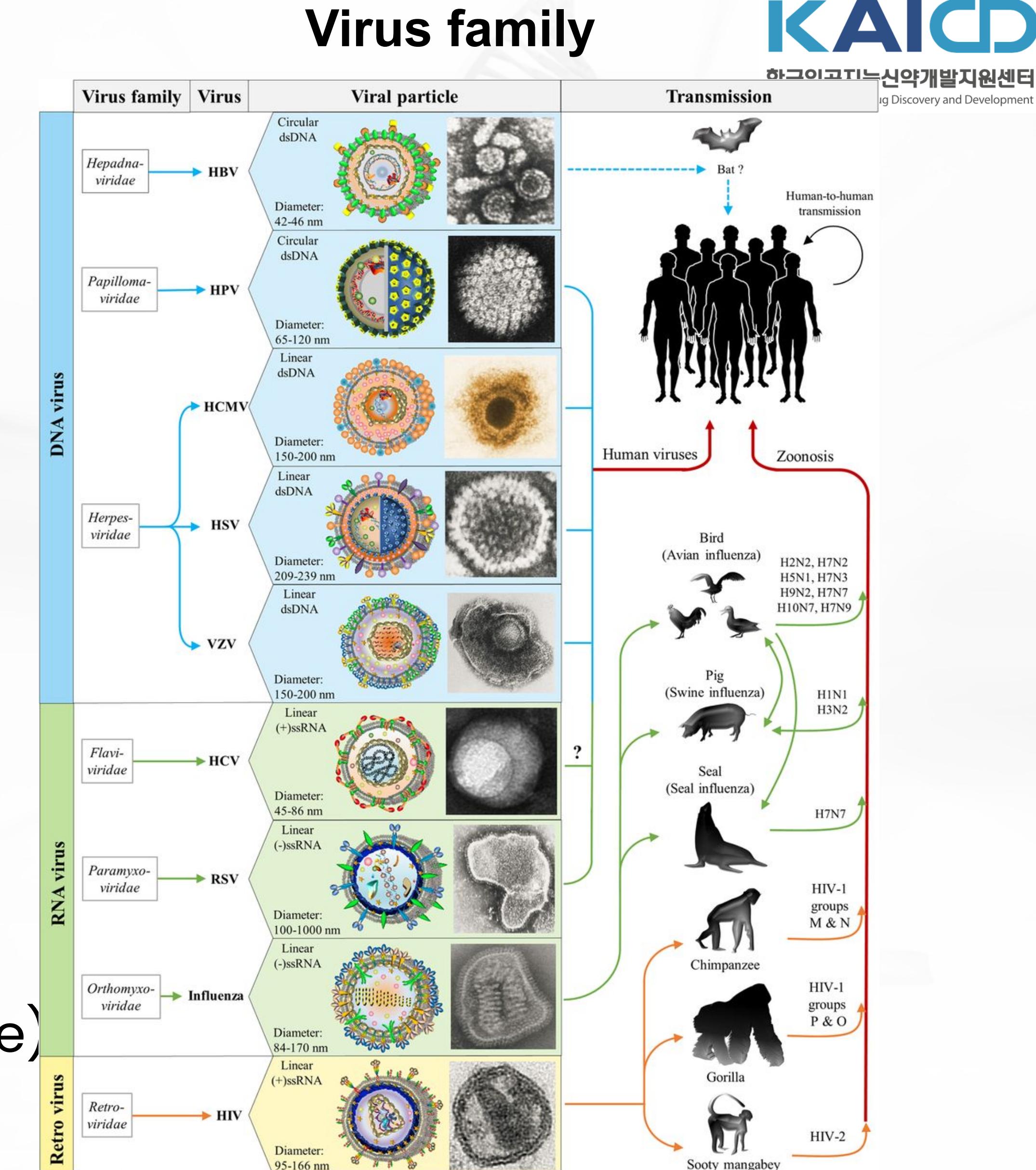
■ Drug DBs

▪ DrugBank

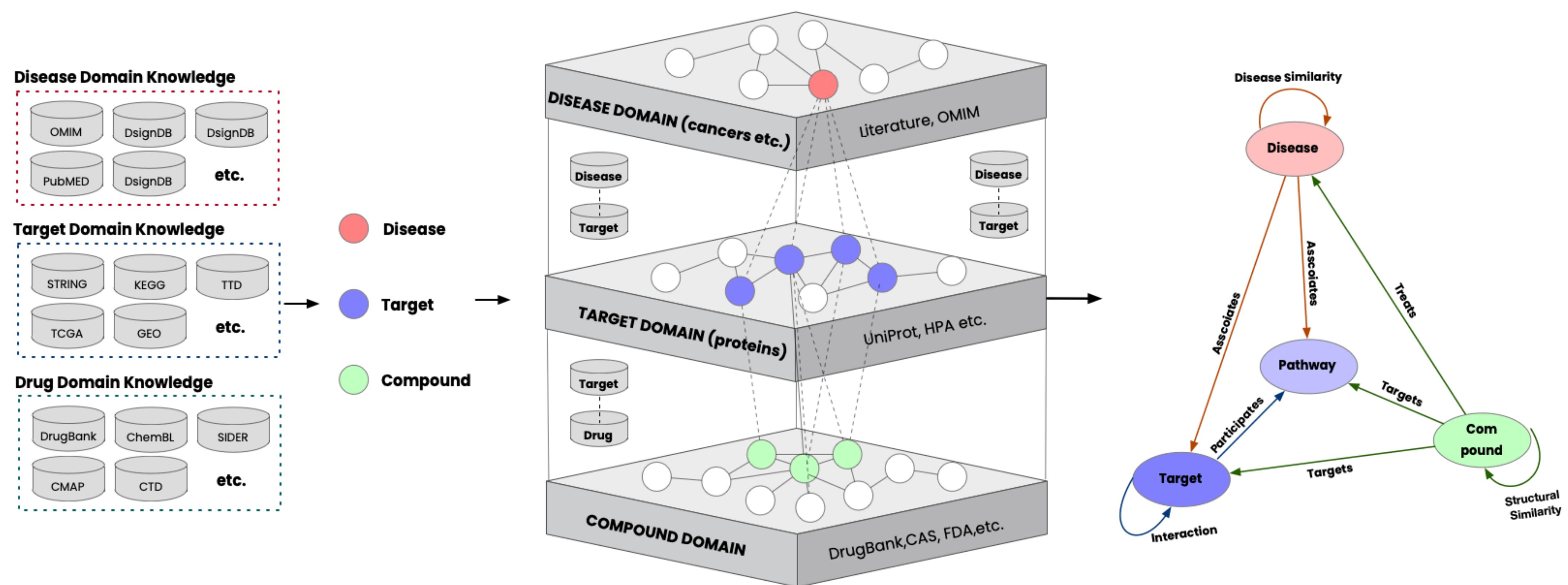
- 2459 FDA approved drugs
- 124 Clinical Trial Drugs
- ~40,000 antiviral compounds (CAS)

■ Target DBs

- 30 proteins: Virus: Sars-Cov2 Viral Proteins (Nature)
- 334 proteins: Human: (Nature)

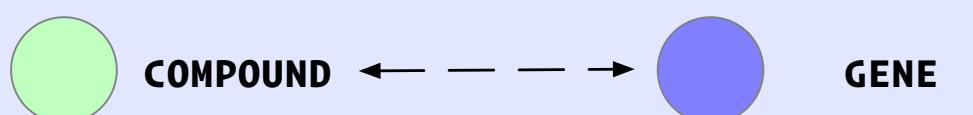
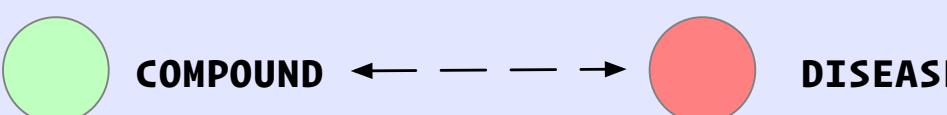


Erik De Clercq, and Guangdi Li Clin. Microbiol. Rev. 2016;
 doi:10.1128/CMR.00102-15



HEAD	RELATION	TAIL	HEAD	RELATION	TAIL	RELATION TYPES	RELATION TYPES
Target	↔	Target	Target	↔	Disease	PPI, literature , coexpressed etc.	target , association
Disease	↔	Disease	Target	↔	Drug	similarity, shares same target etc.	bioactivity, binding, association
Drug	↔	Drug	Drug	↔	Disease	structural (Tanimoto), functional similarity etc.	clinical, literatural, experimental

READING RELATIONSHIP TYPES (ONTOLOGY)



COMPOUND-DISEASE

Compound-To-Disease →

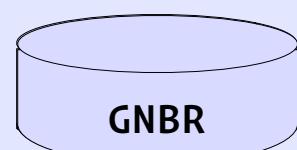
- (T) treatment/therapy (including investigatory)
- (C) inhibits cell growth (esp. cancers)
- (Sa) side effect/adverse event
- (Pr) prevents, suppresses
- (Pa) alleviates, reduces
- (J) role in disease pathogenesis

Disease-To-Compound ←

- (Mp) biomarkers (of disease progression)

33.3GB

LITERATURE (TEXT)



Biomedical Relationships (PUBCHEM)



Compound:1234 - GNBR::(T) Compound:Disease - Disease:MESH:D01234



GENE-DISEASE

Gene-To-Disease →

- (U) causal mutations
- (Ud) mutations affecting disease course
- (D) drug targets
- (J) role in pathogenesis
- (Te) possible therapeutic effect
- (Y) polymorphisms alter risk
- (G) promotes progression

Disease-To-Gene ←

- (Md) biomarkers (diagnostic)
- (X) overexpression in disease
- (L) improper regulation linked to disease

~ 24 million research articles

COMPOUND-GENE

Compound-To-Gene →

- (A+) agonism, activation
- (A-) antagonism, blocking
- (B) binding, ligand (esp. receptors)
- (E+) increases expression/production
- (E-) decreases expression/production
- (E) affects expression/production (neutral)
- (N) inhibits

Gene-To-Compound ←

- (O) transport, channels
- (K) metabolism, pharmacokinetics
- (Z) enzyme activity

Compound:1234 - GNBR::(B) Compound:Gene - Gene:4321



GENE-GENE

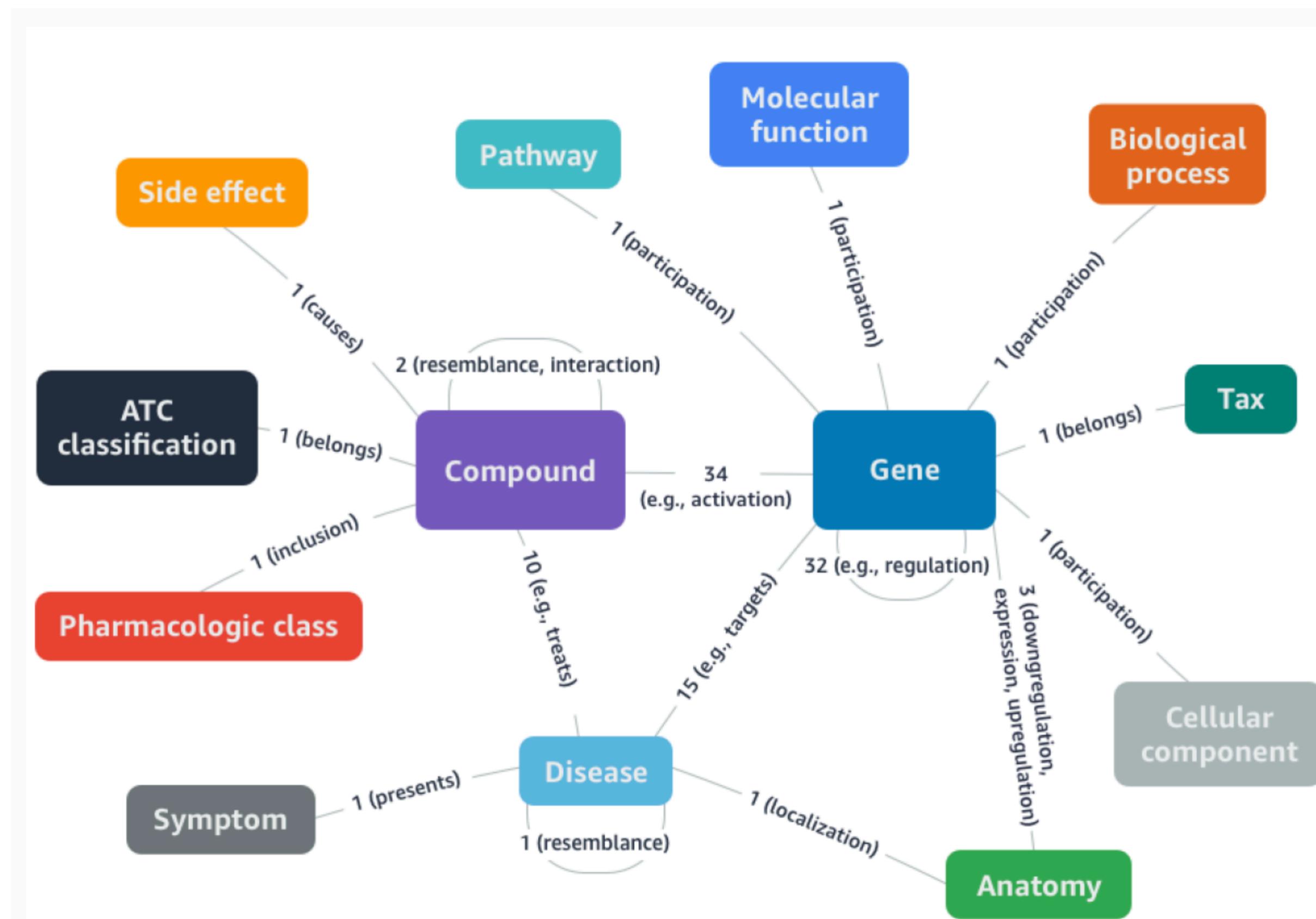
- (B) binding, ligand (esp. receptors)
- (W) enhances response
- (V+) activates, stimulates
- (E+) increases expression/production
- (E) affects expression/production (neutral)
- (I) signaling pathway
- (H) same protein or complex
- (Rg) regulation
- (Q) production by cell population

Gene:1234 - GNBR::(D) Gene:Disease - Disease:MESH:D01234

Gene:1234 - GNBR::(Rg) Gene:Gene - Gene:4321

Knowledge Graph Database

- Amazon Deep Engine Science Team
 - University of Minnesota, The Ohio State University and
 - **Drug Repurposing Knowledge Graph (DRKG)**



- 6 DBs
 - DrugBank, GNBR, Hetionet (Knowledge Database), STRING (PPI), IntAct (PPI), Bibliography
- 97,328 entities
 - 13 types of entities
- 5,874,461 triplets
 - 107 types of relations



Thank You