IDENTIFYING DRUG-REPURPOSING OPPORTUNITIES FOR COVID-19 USING AI AND DIFFUSION MODELS

Phong Lai Bao Minh[1,2], Nguyen Thi Nguyet[1,2], Bui Nguyen Phuong Linh[1,2], Ho Thanh Duy Khanh[1,2], and Do Trong Hop[1,2]

[1] VNUHCM - University of Information Technology, Viet Nam
[2] 20522217, 20521689, 20521527, 20521445@gm.uit.edu.vn, hopdt@uit.edu.vn

Abstract. December 2019 - A milestone to assess the origin of the COVID-19 pandemic in Wuhan, China. Since then, although three years have passed, humans have still been unable to completely control the variation of SARS-Cov-2 virus. The sudden outbreak of the virus has brought challenges to the research and finding treatment methods suitable for clinical use. However, the traditional way to develop new drugs is time-consuming and expensive. Therefore, drug repurposing is highly valued by scientists, who focus on studying new indications of existing drugs and reducing the challenges faced in drug development. In this project, we implemented two algorithms, including algorithms based on artificial intelligence and network diffusion. The initial results of the AI algorithm and the diffusion network ranked 1610 and 6254 drugs, respectively. From here, the data can be used to further develop clinical trials and screening effective drugs against coronavirus in order to quickly control the global epidemic.

Keywords: Drug-repurposing · Artificial intelligence · Network diffusion · COVID-19 pandemic.

1 Introduction the problems

SAR-CoV-2 virus was first discovered during the outbreak in China. Later, as the epidemic gradually spread to other parts of Asia and eventually spread to all parts of the world, efforts to contain the virus seemed to have failed. By January 30, 2023, there had been more than 670 million cases of infection and 6.83 million deaths, making COVID-19 become one of the most deadly epidemics in human history. In the face of this emergency, the main requirement of researchers and scientists is to quickly find drugs that inhibit the virus to limit infection. However, the usual way to develop new drugs may take 10 to 15 years. In view of the high consumption rate, considerable cost and slow discovery rate of new drugs, drug reuse is one of the scientific research directions of developing safe and effective COVID-19 treatment methods. Drug repositioning (also known as drug repurposing) is the repurposing of an approved drug for the treatment of a different disease or medical condition than that for which it was originally

developed. Drug repurposing usually requires three steps before taking the drug across the development pipeline: recognition of the right drug; systematic evaluation of the drug effect in clinical models; and estimation of usefulness in phase II clinical trials. Drug-repurposing algorithms rank drugs based on one or multiple streams of information, such as molecular profiles, chemical structures, adverse profiles, molecular docking, electronic health records, pathway analysis, genome wide association studies, and network perturbations. To quantify and compare their true predictive power, all algorithms must make predictions for the same set of candidates, and the experimental validation must focus not only on the top candidates, as it does now, but on a wider list of drugs chosen independently of their predicted rank. Based on practical necessity, we proceeded to implement algorithms based on artificial intelligence and network diffusion. Each algorithm ranked 1610 and 6254 drugs, respectively, based on their expected effectiveness against the SAR-CoV-2 virus.

2 Description of the dataset

The dataset includes 12 csv files, but to serve the research problem, we only use 7 files, including:

• **Dataset S1. SARS-COV2-Human Interactome** Protein-protein interactions between 26 SARS-COV2 proteins and 332 human proteins detected by affinity purification followed by mass spectrometry (dataset retrieved from Gordon et al (2020)).

• **Dataset S2. Protein-Protein Human Interactome:** 327,924 pairwise binding interactions between 18,508 human proteins.

• **Dataset S3. List of drugs and their respective targets:** Retrieved from the DrugBank database.

• **Dataset S4. List of 17,222 differentially expressed genes identified by exposure of 793 drugs in different cell lines:** Data obtained from the DrugBank database. Differential Expressed Genes used for P3.

• **Dataset S5. Network Overlap Between 299 Diseases and SARS-COV2 Targets:** The S_{vb} measure captures the network-based overlap between SARS-COV2 targets v and the gene pool associated with disease b.

• **Dataset S6. Embedding vectors:** Representations of diseases as learned by the GNN model. Each row in the file contains the embedding vector for a particular disease.

• **Dataset S7. Embedding vectors:** Representations of drugs as learned by

the GNN model. Each row in the file contains the embedding vector for a particular drug.

3 Definitions and Research methods

3.1 Protein- Protein Interaction (PPI)

Protein interactions are interactions between proteins or between proteins and other molecules in the cell. This affects cellular activities and affects the life processes of the body. This can be said to be the centre of biological processes. Proteins rarely work in isolation, they perform their functions through interactions with other biological units so the examination of protein-protein interaction (PPI) is essential for understanding the mechanism of biological processes. PPI is usually determined by biochemical and experimental methods, but the cost is high and the success rate cannot be guaranteed. Therefore, the development of error-generated calculation models improves the above shortcomings. This improves the ability to detect high probability interactions, thus providing a priority for practical experiments.

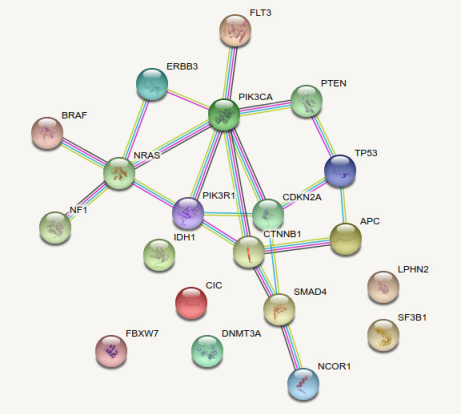

Fig. 1. Protein-Protein Interactions Network

Drawing PPI interaction diagrams can not only deeply understand the function of protein, but also help clarify the molecular mechanism in cells. According to Phizicky and Fields, PPI can change the properties of enzymes that produce new binding sites, inactivate or destroy proteins, and thus change the properties of proteins.

PPI network is composed of top and edge. Protein is the top of the graph and edge is the interaction between proteins. From the graph, we will have a computational model to design the organisation of the PPI network. Observing the

graph pattern can provide many insights, so we can predict the function of protein by observing the function of protein interaction and which protein network the protein belongs to. In addition, subgraphs are also interspersed as proteins with similar functions to individual components.

3.2 Human Interactome and SARS-CoV-2 and Drug Targets

The human interactome was assembled from 21 public databases that compile experimentally derived protein–protein interactions (PPI) data:

• Binary PPIs, derived from high-throughput yeast two-hybrid experiments, three-dimensional protein structures.

• PPIs identified by affinity purification followed by mass spectrometry.

• Kinase substrate interactions.

• Signalling interactions.

• Regulatory interactions.

The final interactome used in this study contains 18,505 proteins and 327,924 interactions (Dataset S2). We retrieved interactions between 26 SARS-CoV-2 proteins and 332 human proteins reported by Gordon, et. al. (2020) (Dataset S1) and drug target information from the DrugBank database, which contains 24,648 interactions between 6,253 drugs and their 3,909 targets, and drug target interaction data curated from the literature for 25 drugs (Dataset S3).

3.3 Disease Comorbidities

Pre-existing conditions worsen prognosis and recovery of COVID-19 patients. Previous work showed that the disease relevance of human proteins targeted by a virus can predict the signs, symptoms, and diseases caused by that pathogen. This prompted us to identify diseases whose molecular mechanisms overlap with cellular processes targeted by SARS-CoV-2, allowing us to predict potential comorbidity patterns. We retrieved 3,173 disease-associated genes for 299 diseases, finding that 110 of the 332 proteins targeted by SARS-CoV-2 are implicated in other human diseases; however, the overlap between SARS-CoV-2 targets and the pool of the disease-associated genes was not statistically significant. We evaluated the network-based overlap between the proteins associated with each of the 299 diseases and the host protein targets of SARS-CoV-2 using the S_{vb} metric, where $S_{vb} < 0$ signals a network-based overlap between the SARS-CoV-2 viral targets v and the gene pool associated with disease b . We found that $S_{vb} > 0$ for each disease, indicating that COVID-19 disease module does not directly overlap with any major disease module. The diseases closest to the COVID-19

disease module (smallest S_{vb}) included several cardiovascular diseases and cancers, whose comorbidity in COVID-19 patients is well documented. The same metric predicted comorbidity with neurological diseases, in line with our observation that the host protein targets are expressed in the brain (Dataset S5).

In summary, we found that the SARS-CoV-2 host protein targets do not overlap with proteins associated with any major diseases, indicating that a potential COVID-19 treatment cannot be derived from the arsenal of therapies approved for a specific disease. These findings argue for a strategy that maps drug targets without regard to their localization within a particular disease module. However, the disease modules closest to the SARS-CoV-2 viral targets are those with noted comorbidity for COVID-19 infection, such as pulmonary and cardiovascular diseases, and obesity.

3.4 Graph Neural Networks (GNN)

Graph Neural Networks are a special class of neural networks that are capable of working with data that is represented in graph form. These networks are heavily motivated by Convolutional Neural Networks (CNNs) and graph embedding. The model could process graphs that are acyclic, cyclic, directed, and undirected. The objective of GNN is to learn a state embedding that encapsulates the information of the neighbourhood for each node. These networks have recently been applied in multiple areas including; combinatorial optimization, recommender systems, computer vision, etc. These networks can also be used to model large systems such as social networks, protein-protein interaction networks, knowledge graphs among other research areas.

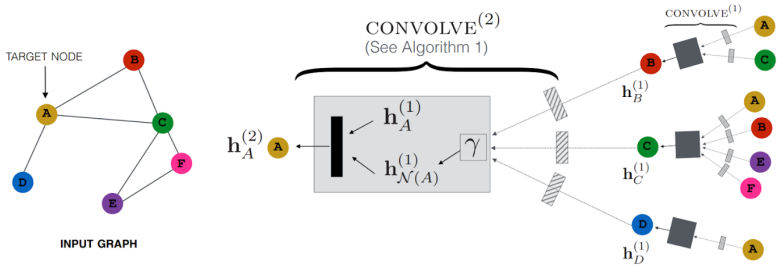

Fig. 2. Graph Neural Network.

3.5 Diffusion State Distances (DSD)

Diffusion state distance (DSD) is a metric on the vertices of a graph, motivated by bioinformatic modelling. Previous results on the convergence of DSD to a limiting metric relied on the definition being based on symmetric or reversible

random walk on the graph. Consider the undirected graph G(V,E) on the vertex set $V = v_1, v_2, v_3, ..., v_n$ and $|V|$ = n. Recall that He^k (A,B) is defined as the expected number of times that a random walk starting at node A and proceeding for k steps, will visit node B. In what follows, assume k is fixed, and when there is no ambiguity, in the value of k, we will denote He^k (A,B) by He(A,B). We further define a n - dimensional vector He(v_i) $\forall$v$_i$ $\in$V where:

$$\text{He}(v_i) = (\text{He}(v_i, v_1), \text{He}(v_i, v_2), ..., \text{He}(v_i, v_n))$$

Then, the Diffusion State Distance (DSD) between two vertices u and v, $\forall u, v \in$V is defined as:

$$\text{DSD(u,v)} = ||He(u) - He(v)||_1$$

Where $||He(u) - He(v)||_1$ denotes the L_1 norm of the He vectors of u and v.

4 Drug Repurposing Prediction Algorithms

The COVID-19 pandemic presents both the societal imperative and the rationale to test drugs at a previously unseen scale. It offers a unique opportunity to quantify and improve the efficacy of the available predictive algorithms, while also identifying potential treatments for COVID-19. In this study, we implemented two drug reuse algorithms in the network:

4.1 Artificial Intelligence Based Algorithm (A1-A4)

The AI algorithm is a graph neural network (GNN) architecture that takes as input a multimodal graph with three types of nodes (representing drugs, proteins, and diseases) and edges capturing different types of interactions between these nodes. The algorithm generates embedding vectors of drug and disease nodes. A multimodal graph is a G = (V,R) heterogeneous graph. The N micro-nodes in V represent three different types of biomedical entities (drugs, proteins and diseases). The marker edges (v_i, r, v_j) belonging to R represent four entity relationships (protein-protein, drug target association, disease-protein association, and drug disease index).

$\Rightarrow$ We define the problem as Link prediction on a multimodal graph, the task of the problem is to predict new edges based on the graph between drug (v_i) and disease (v_j) nodes which means drug vi specified for disease v_j.

In this research, we use graph neural network that is an end-to-end trainable model for link prediction on the multimodal graph and has two main components: (1) an encoder: a graph convolutional network operating on G and producing embeddings for nodes in G ; and (2) a decoder: a model optimizing embeddings such that they are predictive of known drug-disease indications. The neural message-passing encoder took as input a graph G and produced a

node d-dimensional embedding $z_i \in R^d$ for every drug and disease node in the graph. We used the encoder that learned a message-passing algorithm and aggregation procedure to compute a function of the entire graph that transformed and propagated information across graph G. The graph convolutional operator took into account the first-order neighbourhood of a node and applied the same transformation across all locations in the graph. Successive application of these operations then effectively convolved information across the K^{th} order neighbourhood, where K is the number of successive operations of convolutional layers in the neural network model. The graph convolutional operator takes the form

$$h_i^{k+1} = \phi(\sum_r \sum_{j \in N_T^i} a_r^{ij} W_r^k h_j^k + a_r^i h_i^k)$$

Where $h_i^k \in R^{d(k)}$ is the hidden state of node v_i vi in the $K^t h$ layer of the neural network with d(k) being the dimensionality of this layer's representation, r is an edge type, matrix W_r^k is an edge-type specific parameter matrix, ϕ denotes a non-linear element-wise activation function, and a_r denote attention coefficients. To arrive at the final embedding $z_i \in R^d$ of node vi, we compute its representation as $z_i = h_i^k$. Next, the decoder takes node embeddings and combines them to reconstruct labeled edges in G. In particular, the decoder scores a (v_i, r, v_j) triplet through a function g whose goal is to assign a score $g(v_i, r, v_j)$ representing how likely it is that drugs v_i will treat disease v_j.

We generated four lists of candidate drugs for COVID-19. To generate the lists, we used embeddings returned by the graph neural network, in particular, embeddings learned for nodes representing either COVID-19 or drugs in multimodal graph G. The embedding vectors for diseases and drugs are provided in Dataset S6 and Dataset S7, respectively. The pipeline A1 searches for drugs that are in the vicinity of the COVID-19 disease by calculating the cosine distance between COVID-19 and all drugs in the decoded embedding space. The decoding is based on the N = 10 nearest neighboring nodes in the embedding space, with a minimum distance between nodes of D = 0.25. The pipeline A2 prevents that nodes in the decoding embedding space from packing together too closely, by using D = 0.8 and keeping N unchanged. These constraints push the structures apart into softer, more general features, offering a better overarching view of the embedding space at the loss of the more detailed structure. Pipeline A3 forces the decoding to concentrate on the very local structure by using N = 5, to explore a smaller neighborhood, while setting the minimum distance at a mid range point of D = 0.5. Pipeline A4 focuses on a broader view of the embedding space by setting N=10 and D = 1. Finally, to obtain lists of candidate drugs, each pipeline ranked drugs based on the pipeline-defined distances of drugs to COVID-19. Intuitively, parameter N constrained the size of the local

neighborhood each pipeline looked at in the embedding space when calculating the distances, and parameter D controlled how tightly the pipeline was allowed to pack the embeddings together.

4.2 Diffusion Based Algorithm

Diffusion-Based Algorithm is inspired by diffusion state distance (DSD). It uses the diffusion property to define a similarity metric for pairs of nodes to calculate the similarity between nodes and its effect on the entire network. When we obtain the similarity score between nodes, we can calculate the similarity between the target drug and SARS-VoV-2 proteome. The proximity measurement is based on the shortest average distance from the drug target to the SARS-CoV-2 target. Inspired by DSD, we developed five new metrics to calculate the impact of drug-targets on SARS-CoV-2 targets. Pipeline D1 is defined as:

$$J_{DSD}^{min} = \frac{1}{||T||}\sum\nolimits_{t\in T} \min DSD\ (t, v)$$

Where DSD (t, v) represents the diffusion state distance between nodes t and v. We continue to use the next formula (pipeline D2):

$$J_{KL}^{min} = \frac{1}{||T||}\sum\nolimits_{t\in T} \min KL\ (t, v)$$

And pipeline D3:

$$J_{KL}^{med} = \frac{1}{||T||}\sum\nolimits_{t\in T} median\ KL\ (t, v)$$

Where KL is the relative entropy between the vectors representing nodes t and v. Finally, to provide symmetric measures, we tested the measurements (pipeline D4):

$$I_{JS}^{min} = \frac{1}{||T||}\sum\nolimits_{t\in T} \min JS\ (t, v)\ |$$

And pipeline D5:

$$I_{JS}^{med} = \frac{1}{||T||}\sum_{t\in T} median\ JS\ (t, v)$$

JS is the Jensen-Shannon difference between vectors representing nodes t and v. All five formulas must meet different v.

4.3 Explanatory Subgraph

For each pipeline, we identified "explanatory subgraphs" to help understand the predictions made by the respective pipeline. The key idea was to summarize where in the data the pipeline seeks evidence for their predictions. Given a particular prediction, an explanatory subgraph is a small sub-network of the entire network considered by the pipeline that is most influential for the prediction and contributes most to the predictive power. For the artificial intelligence-based methods (AI), the subgraphs were extracted using a GNN Explainer algorithm. GNN Explainer specifies an explanation as a subgraph of the entire network the GNN was trained on, such that the subgraph maximizes the mutual information with the GNN's prediction. This is achieved by formulating a mean field variational approximation and learning a real-valued graph mask, which selects the important subgraph using counterfactual reasoning. For the diffusion method, we first identified the SARS-CoV-2 targets (seeds) that have the maximum (or median, depending on the pipeline) similarity with the drug targets under consideration. Once the seeds are identified for each drug target, we extract the vector representation of the target and the corresponding seeds. Each element of these vectors corresponds to a node in the network:

$$\text{t:}[r_1, r_2, r_3, ..., r_n]$$
$$\text{s:}[\text{w}_1, w_2, w_3, ..., w_n]$$

Each pipeline performs an element-wise comparison of these two vectors to calculate similarity values, defined as similarity terms, using:

$$term_i^{DSD}(t,s) = |\ r_i - w_i|$$

$$term_i^{KL}(t,s) = r_i \log\left(\frac{r_i}{w_i}\right)$$

$$term_i^{KS}(t,s) = \frac{1}{2}\left[r_i \log\left(\frac{r_i}{m_i}\right) + w_i \log\left(\frac{w_i}{m_i}\right)\right], m_i = \frac{r_i + w\_i}{2}$$

These distance similarity terms collectively contribute to each drug's ranking score. Among all 18,446 nodes, we are only interested in those whose variations lead to the current ranking (drug prediction scores). Therefore, we applied a feature selection algorithm to eliminate the network nodes (features) that do not contribute to the predicted scores (outcomes). This task is done by training a regression tree model where feature values are the similarity terms (as defined above) between drug targets and the corresponding seeds. This resulted in 2,507 important features for pipeline D1 (DSD-min), 2198 for D2 (KL-min), 2,263 for D3 (KL-med), 1,655 for D4 (JS-min), and 1,817 for D5 (JS-med). Important features are those with non-zero importance value as characterized by the Regressor model.

Once the important features/nodes are extracted, we search this space to identify the explanatory network of each set of drug targets. To do so, we rank the similarity terms of each target and the corresponding. seeds on the space of important features and identify the nodes with the highest contribution to the similarity measure such that they satisfy the following equation:

$$log_{10}\left(\frac{l}{term_i}\right) \leq 1, l = \left(term_i\right),\ i \in V$$

If a drug has multiple targets or if each target has multiple corresponding seeds (seeds with the same similarity to a target), the results are aggregated. The explanatory network of a target that happens to be a seed is that seed itself.

5 Achievements

The main way to re-use drugs is to find similarities in molecular expression between the disease being studied and some other diseases. Therefore, to search for diseases with a molecular mechanism that overlaps with that of COVID-19, we first mapped experimentally determined host protein targets of respiratory syndrome-causing proteins. Severe acute illness caused by the SAR-CoV-2 virus (Dataset S1) with the human interaction dataset (Dataset S2)- a set of 332,749 pairwise binding interactions between 18,508 human proteins.

The obtained results show that the proteins targeted by the SAR-CoV-2 virus are not randomly distributed in the human interaction system, but form a large connected part (LCC) consisting of 208 proteins and many subgraphs. Most of the SAR-CoV-2 LCC virus proteins are expressed in lung tissue, potentially explaining the effectiveness of the virus in inducing pulmonary manifestations.

Fig. 3. Interactions between proteins causing severe acute respiratory syndrome coronavirus and human proteins.

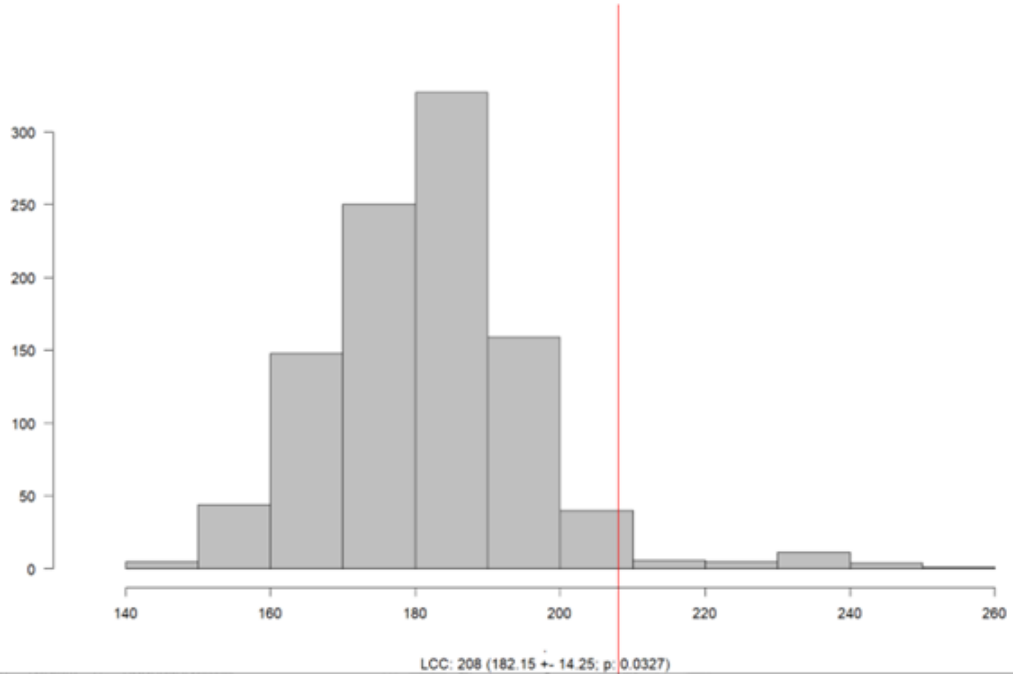

Fig. 4. Chart showing random expectation of LCC size.

The red line on the graph represents the LCC size of COVID-19 that is randomly larger than expected.

We next assessed the network-based overlap between the 299 disease-associated proteins (d) and the corona host protein targets (v) using the S_{vd} index. With $S_{vd} > 0$ for all diseases, it means that the COVID-19 disease module does not directly overlap with disease proteins associated with any single disease.

⇒ Effective COVID-19 treatment cannot be provided from the approved specific disease treatment library. Therefore, a network-based strategy is needed to identify reusable drugs without considering the identified disease indications.

We have implemented two competitive network reuse algorithms:

• **AI-based algorithm:** maps drug and disease-associated protein targets to points in low-dimensional vector space. Up to 4 predicted paths (A1 to A4). Corresponding to 4 csv files with the ranking values of each model and the drug score. The file contains the following columns.

→ Drug ID: the keys of the drug2targets dictionary.

→ Distance score: distance between drug to COVID-19 disease,the lower the better.

• **Diffusion Based Algorithm:** Inspired by diffusion distance and ranking drugs based on capturing the network similarity of drug protein targets with SAR-CoV-2 host protein targets. The algorithm gives 5 ranking pipelines from D1 to D5. File output is a csv file with the ranking values of each model and the drug score. There is also a ranking value (integer) that corresponds to each ranking score. The file contains the following columns.

→ Drug: the keys of the drug2targets dictionary.

→ APP-Drugs: a binary value indicating whether or not the drug is approved. If app-drugs given by the user is empty, all binary values will be marked 'no'.

→ Degree: the number of targets.

→ Targets: human gene targets.

→ DSD-min-Rank: an integer showing drug rank by DSD-min; the lower, the better.

→ KL-med-Rank: an integer showing drug rank by KL-med; the lower, the better.

→ KL-min-Rank: an integer showing drug rank by KL-min; the lower, the better.

→ JS-med-Rank: an integer showing drug rank by JS-med; the lower, the better.

→ JS-min-Rank: an integer showing drug rank by JS-min; the lower, the better.

→ DSD-min: the score given by DSD-min to each drug; the lower, the better.

→ KL-med: the score given by KL-med to each drug; the lower, the better.

→ KL-min: the score given by KL-min to each drug; the lower, the better.

→ JS-med: the score given by JS-med to each drug; the lower, the better.

→ JS-min: the score given by JS-min to each drug; the lower, the better.

The methodological advances presented here not only suggest potential drug candidates for COVID-19, but offer a principled algorithmic toolset to identify future treatments for diseases underserved by the cost and the timelines of conventional de novo drug discovery processes. Moreover, drug combinations could increase the potency of some drugs, and given a synergistic effect, could also improve outcomes.

6 Conclusion and Development direction

The project provides an initial development direction to help create a network medicine framework to identify drug reuse opportunities for COVID-19. COVID disease is the product of damage by the virus itself and damage by immune overreaction (cytokine storm). As the assay used for the experimental screening only detects the inhibition of the viral replication cycle, an immunomodulatory drug that reduces the cytokine storm without interfering with virus replication would not show up as a hit in our screen. We have identified drugs that reduce the viral load enough such that the immune system is not overstimulated, potentially lowering the chance of a cytokine storm. However, the results obtained from the new algorithm are only predictive. Due to limitations in biomedical knowledge, research methods, instruments, and study locations, we have not been able to conduct clinical trials of drug rating predictors. Therefore, in the near future, we want to further improve the quality of drug ratings, and expect to experimentally screen the efficacy against the SAR-CoV-2 virus of the reusable pipelines. drug against VeroE6 cells- renal epithelial cells derived from African green monkeys. In addition, the project can apply some other research directions as well as refer to some algorithms such as Network Proximity to improve the quality of the results.

References

1. S. Dakshanamurthy et al., Predicting new indications for approved drugs using a proteochemometric method. J. Med. Chem. 55, 6832–6848 (2012).

2. A. I. Casas et al., From single drug targets to synergistic network pharmacology in ischemic stroke. Proc. Natl. Acad. Sci. U.S.A. 116, 7129–7136 (2019).

3. F. Cheng et al., Network-based approach to prediction and population-based validation of in silico drug repurposing. Nat. Commun. 9, 2691 (2018).

4. S. Sadegh et al., Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. Nat. Commun. 11, 3518 (2020).

5. E. Guney, J. Menche, M. Vidal, A.-L. Barábasi, Network-based in silico drug efficacy screening. Nat. Commun. 7, 10331 (2016).

6. N. Gulbahce et al., Viral perturbations of host networks reflect disease etiology. PLoS Comput. Biol. 8, e1002531 (2012).

7. D. E. Gordon et al., A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature 583, 459–468 (2020).

8. K. Luck, G. M. Sheynkman, I. Zhang, M. Vidal, Proteome-scale human interactomics. Trends Biochem. Sci. 42, 342–354 (2017).

9. J. H. Beigel et al., Remdesivir for the treatment of Covid-19—Preliminary report. N.Engl. J. Med. 383, 1813–1826 (2020).

10. K. Lin, P. Gallay, Curing a viral infection by targeting the host: The example of cyclophilin inhibitors. Antiviral Res. 99, 68–77 (2013).

11. A. Sharma et al., A disease module in the interactome explains disease heterogeneity, drug response and captures novel pathways and genes in asthma. Hum. Mol. Genet. 24, 3005–3020 (2015).