JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

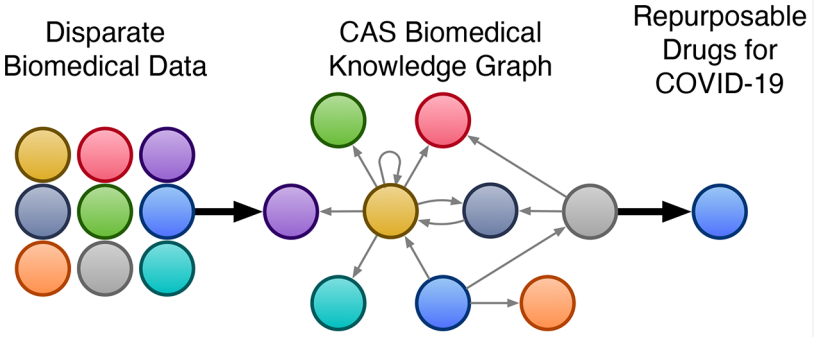
pubs.acs.org/jcim Article

Knowledge Graph-Based Approaches to Drug Repurposing for COVID-19

Jacob Al-Saleem,* Roger Granet, Srinivasan Ramakrishnan, Natalie A. Ciancetta, Catherine Saveson, Chris Gessner, and Qiongqiong Zhou

Cite This: https://doi.org/10.1021/acs.jcim.1c00642 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: The COVID-19 pandemic has motivated researchers all over the world in trying to find effective drugs and therapeutics for treating this disease. To save time, much effort has focused on repurposing drugs known for treating other diseases than COVID-19. To support these drug repurposing efforts, we built the CAS Biomedical Knowledge Graph and identified 1350 small molecules as potentially repurposable drugs that target host proteins and disease processes involved in COVID-19. A computer algorithm-driven drug-ranking method was developed to prioritize those identified small molecules. The top 50 molecules were analyzed according to their molecular functions and included 11 drugs in clinical trials for treating COVID-19 and new candidates that may be of interest for clinical investigation. The CAS Biomedical Knowledge Graph provides researchers an opportunity to accelerate innovation and streamline the investigative process not just for COVID-19 but also in many other diseases.

INTRODUCTION

To date, very few treatments have received FDA approval as therapeutics for COVID-19, while the need for such drugs remains high. To reduce development time and costs, much research has focused on repurposing small molecules that have either already been approved as drugs or have been clinically studied.[1] Because COVID-19 is characterized by the impact of multiple, interlinked physiological systems, including pulmonary hyperinflammation, severe lung injury, blood coagulopathy, renal and neurological problems, and the cellular pathways that underlie these systems,[2,3] it is proposed that a knowledge graph approach would be of value in identifying the connections between these systems as well as potential therapeutics.[4−6]

Knowledge graphs are a type of database that allow users to organize and connect pieces of data based on the relationships that exist between them. Each unit of data can be thought of as a dot (or node) connected to other units by lines (or edges) that represent the relationships between the nodes. This type of database places as much importance on the relationships that connect data as on the data itself. Knowledge graphs can also combine data from multiple sources. These features allow insights that would not be possible using the individual data sources and traditional databases. One small, highly simplified example is:

- Alpelisib (unit of information, or node, and a small molecule) inhibits (relationship, or edge) tumor protein p53 (unit of information, or node)
- Tumor protein p53 upregulates transcription factor Fos
- Transcription factor Fos upregulates transcription factor STAT3
- Transcription factor STAT3 is associated with vascular inflammation

This simple knowledge graph is depicted visually in Figure 1. If a user were to query the knowledge graph to predict what drugs might inhibit vascular inflammation, the graph could

Received: June 4, 2021

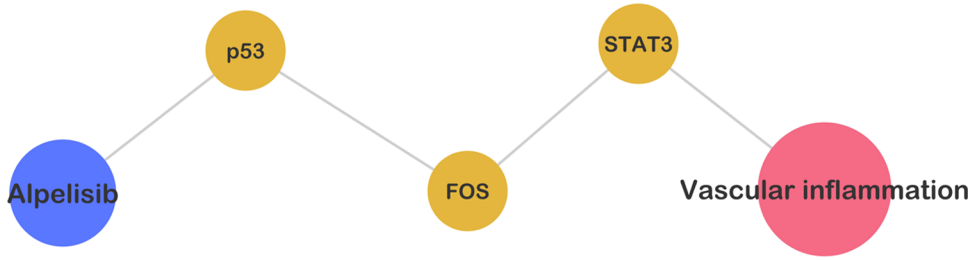


Figure 1. Visual depiction of simplified knowledge graph relating alpelisib to vascular inflammation.

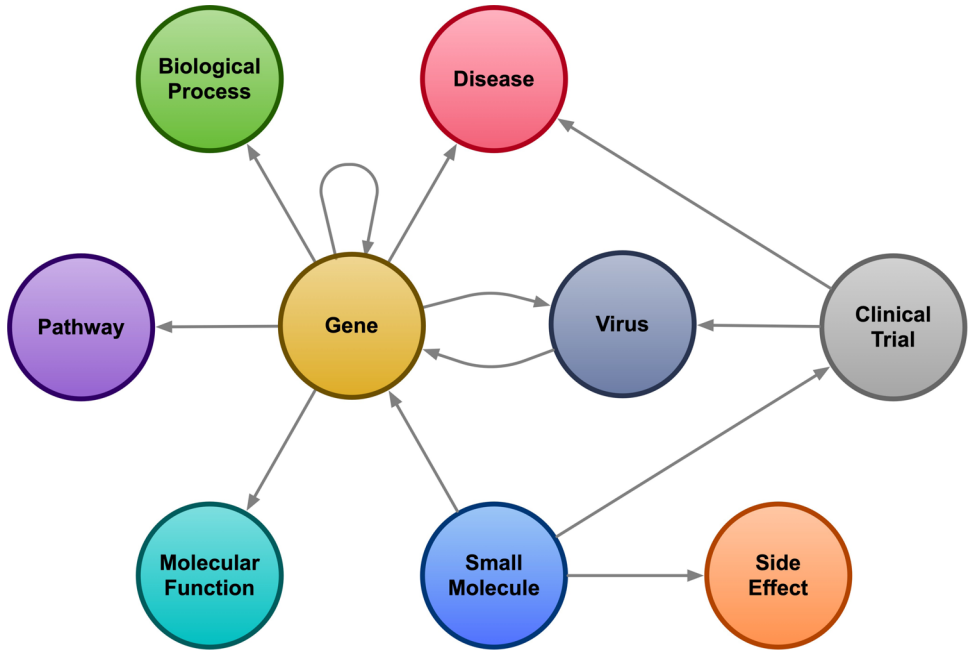


Figure 2. Simple schematic diagram of the CAS Biomedical Knowledge Graph.

provide the answer of alpelisib. This may not be obvious from traditional databases, which might show only direct inhibitors of transcription factor STAT3 or of vascular inflammation itself, but because a knowledge graph links multiple nodes via relationships, the second-level inhibitor alpelisib can also be found. This example illustrates how knowledge graphs can be used to manage, explore, and navigate through the interactions and connections between disparate pieces of information to gain insights and make predictions. As a result of their value, knowledge graphs have grown in importance in the last 10 years in both industry and academia.[7]

It is important to note that knowledge graphs are both scalable and modular, so they can be used in many different areas of research or other activities. For example, a pharmaceutical researcher could use a knowledge graph to identify potential drug candidates or drug targets for diseases. In other fields, material scientists could use a knowledge graph to identify the best compounds for inclusion in designing a new type of material. A graph could also power the hunt for new light-absorbing compounds to use in creating more efficient solar cells. Combined with nutritional data, a knowledge graph could assist food scientists in identifying ingredients that could promote health or improve a recipe. These are just a few of the many possible uses that knowledge graphs could provide.

In this CAS Biomedical Knowledge Graph, we incorporated human diseases, proteins, small-molecule inhibitors, virus, and COVID-19-specific data for identifying small molecules that show potential for repurposing as COVID-19 therapeutics. This CAS Biomedical Knowledge Graph features the human-curated substance data in the CAS Content Collection linked to biomedical data from both CAS and external databases. The information units, or nodes, in this graph are human proteins (denoted by their gene names), biological processes, diseases, and small molecules, including drugs and drug candidates. The links, or edges, between them are relationships such as drug X targets protein Y and protein Y is involved in biological process Z. A novel algorithmic method was also developed for ranking the most promising drug candidates. It prioritized those identified molecules that target unique proteins involved in COVID-19 disease pathways, while minimizing side effects. The most highly ranked substances are discussed in terms of their possible relevance to COVID-19.

RESULTS

The CAS Biomedical Knowledge Graph was constructed using data from the CAS Content Collection and public repositories. In total, the graph contains over 6 million nodes and 18 million relationships. A simple visual schema is shown in Figure 2. Genes and gene products (i.e., proteins or microRNAs) are all referenced by their respective gene node in the graph. A detailed description of the knowledge graph and its

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

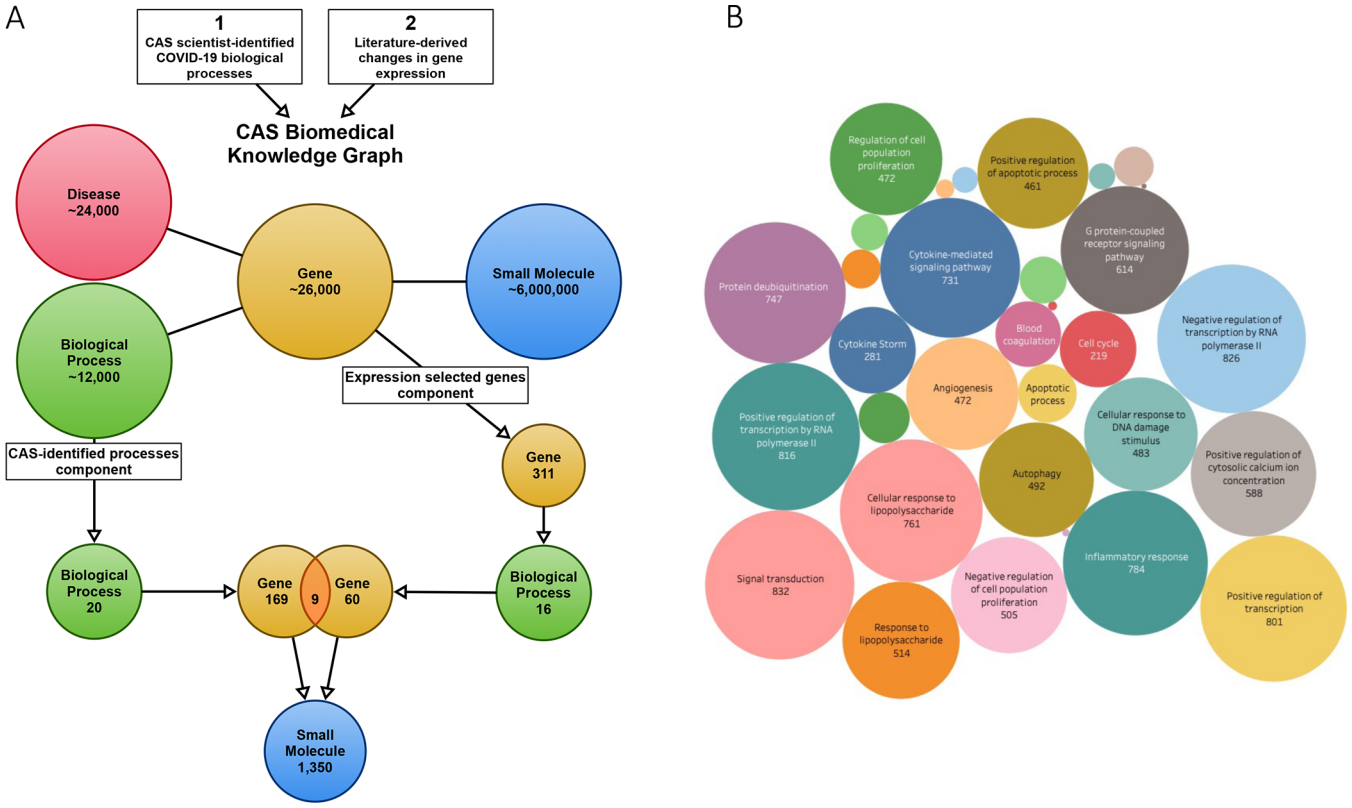


Figure 3. Identification of small molecules targeting biological processes involved in COVID-19. (A) Flowchart of two-component approach to identify potential COVID-19 therapeutics. (B) Diagram displaying the number of small molecules that target the biological process/disease nodes selected from the two-component approach. The larger the circle, the larger the number of small molecules that connect to that node. The total number of small molecules that connect to the node is also shown below the node description. Note that there is an extensive overlap of small molecules between the nodes.

construction can be found in the Materials and Methods section.

A two-component approach was designed to identify COVID-19 drug repurposing candidates; a flowchart of this approach is shown in Figure 3A. The first step of both components was the collection of biological processes deemed important in the SARS-CoV-2 infection process and COVID-19. For the first component, CAS scientists identified a selection of CAS controlled vocabulary headings and associated synonyms to search the CAS Content Collection for SARS-CoV-2 infection-related documents and collected those containing potential drug targets. Intellectual analysis of the resulting documents along with author terminology was then used to gather a list of 20 biological processes deemed important in the SARS-CoV-2 infection process and COVID-19, as summarized in Zhou et al.[8] Some of the biological processes identified include viral entry, endocytosis, autophagy, cytokine storm, and blood coagulation (full list in Table S1). For the second component, genes that were significantly upregulated (>2-fold) by SARS-CoV-2 infection as described in ref 9 were extracted, and the biological processes associated with four or more of these genes were identified, of which there were 16 in total. These included, for example, inflammatory response, angiogenesis, and negative regulation of RNA transcription (full list in Table S2). The 36 processes collected in total from both components were then matched against Gene Ontology (GO), and the corresponding GO terms were used from this point on. Any SARS-CoV-2-specific processes were mapped to the corresponding general viral processes in GO to gather a larger set of potential targets. The graph was then queried for small molecules that modulate the genes associated with these biological processes. The resulting small molecules from both components were then combined, resulting in a set of 1350 small molecules. The number of compounds connected to each biological process/disease node is shown in Figure 3B. The graph queries used for both components are provided in the Supporting Information.

To rank the identified small molecules, the following equation was developed and used to score each of the 1350 molecules individually

$$\text{score} = \frac{(\sum \text{GR}_{\text{CAS}} + \sum \text{BPR}_{\text{CAS}}) + \left(\frac{\sum \text{GR}_{\text{EXP}} + \sum \text{BPR}_{\text{EXP}}}{5}\right)}{\text{LOG(SEP)}} + 2 \times \text{CS} + 2 \times \text{AG}$$

Gene rarity (GR) is a measure of the number of small molecules that directly connect to a given gene, defined as

$$\text{GR} = \frac{1}{\text{LOG(count of small molecules connected to the given gene)}}$$

Biological process rarity (BPR) is a measure of the number of small molecules that connect to a given biological process separated by one gene, defined as

$$\text{BPR} = \frac{1}{\text{LOG(count of small molecules connected to the given bio process)}}$$

The queries used to calculate GR and BPR are provided in the Supporting Information.

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. Top 50 Drug Repurposing Candidates with CAS Registry Number, Drug Name, Drug Class, and Clinical Trial Status[a]

Rank	CAS Registry Number	Drug Name	Drug Class	Clinical Trial
1	149647-78-9	vorinostat	HDAC inhibitors	
2	179324-69-7	bortezomib	protease inhibitors	
3	23214-92-8	doxorubicin	DNA metabolism-related	
4	284461-73-0	sorafenib	kinase inhibitors	
5	183321-74-6	erlotinib	kinase inhibitors	
6	231277-92-2	lapatinib	kinase inhibitors	
7	114977-28-5	docetaxel	microtubule-regulating agents	
8	667463-62-9	MLS 2052	kinase inhibitors	
9	404950-80-7	panobinostat	HDAC inhibitors	
10	152459-95-5	imatinib	kinase inhibitors	yes
11	56-65-5	adenosine 5′ triphosphate	other	
12	872511-34-7	BGJ 398	kinase inhibitors	
13	2447-54-3	sanguinarine	other	
14	1339928-25-4	fimepinostat	other	
15	183506-66-3	apicidin	HDAC inhibitors	
16	58880-19-6	trichostatin A	HDAC inhibitors	
17	943319-70-8	ponatinib	kinase inhibitors	
18	112953-11-4	7-hydroxystaurosporine	kinase inhibitors	
19	1256448-47-1	nanatinostat	HDAC inhibitors	
20	287383-59-9	scriptaid	HDAC inhibitors	
21	1210608-43-7	PIM 447	kinase inhibitors	
22	477600-75-2	tofacitinib	kinase inhibitors	yes
23	868540-17-4	carfilzomib	protease inhibitors	
24	989-51-5	epigallocatechin gallate	DNA metabolism-related inhibitors	yes
25	23541-50-6	daunorubicin hydrochloride	DNA metabolism-related inhibitors	
26	870262-90-1	letaxaban	coagulation factor Xa inhibitors	
27	1195765-45-7	dabrafenib	kinase inhibitors	
28	25316-40-9	doxorubicin hydrochloride	DNA metabolism-related inhibitors	
29	491-80-5	biochanin	other	
30	405169-16-6	dovitinib	kinase inhibitors	
31	50-65-7	niclosamide	other	yes
32	957054-30-7	pictilisib	kinase inhibitors	
33	1108743-60-7	entrectinib	kinase inhibitors	
34	97-77-8	tetraethylthiuram disulfide	other	yes
35	75706-12-6	leflunomide	other	yes
36	726169-73-9	mocetinostat	HDAC inhibitors	
37	637-03-6	phenylarsine oxide	other	
38	1951-25-3	amiodarone	other	yes
39	630-60-4	ouabain	other	
40	58-00-4	(−)-apomorphine	other	
41	64-86-8	colchicine	microtubule-regulating agents	yes
42	90-34-6	primaquine	other	yes
43	936563-96-1	ibrutinib	kinase inhibitors	yes
44	31431-39-7	mebendazole	microtubule-regulating agents	
45	361442-04-8	saxagliptin	protease inhibitors	
46	1032900-25-6	ceritinib	kinase inhibitors	
47	446-72-0	genistein	kinase inhibitors	yes
48	20830-81-3	daunorubicin	DNA metabolism-related	
49	480449-70-5	edoxaban	coagulation factor Xa inhibitors	
50	153436-53-4	tyrphostin AG 1478	kinase inhibitors	

[a]Drugs that were difficult to classify are listed as "other". The numbers of drugs in each class in the top 50 are: 18 kinase inhibitors, 7 HDAC inhibitors, 5 DNA metabolism-related, 3 microtubule-regulating agents, 2 coagulation factor Xa inhibitors, and 12 in other classes.

Side effect proxy (SEP) is defined as the number of biological processes the small molecule is connected to in the graph. Cytokine storm (CS) is assigned the value of 1 when the small molecule connects to the cytokine storm node or 0 if it does not. Likewise, activated gene (AG) is given the value of 1 if the small molecule activates a gene or 0 if it does not.

This scoring equation measures all of the interactions identified in our two-component approach (GR/BPR_{CAS} represents the results of component one identified by CAS scientists, and GR/BPR_{EXP} represents the upregulated-expression results of component two). Importance is given to genes and biological processes that are not connected to large numbers of small molecules in the graph (GR + BPR).

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

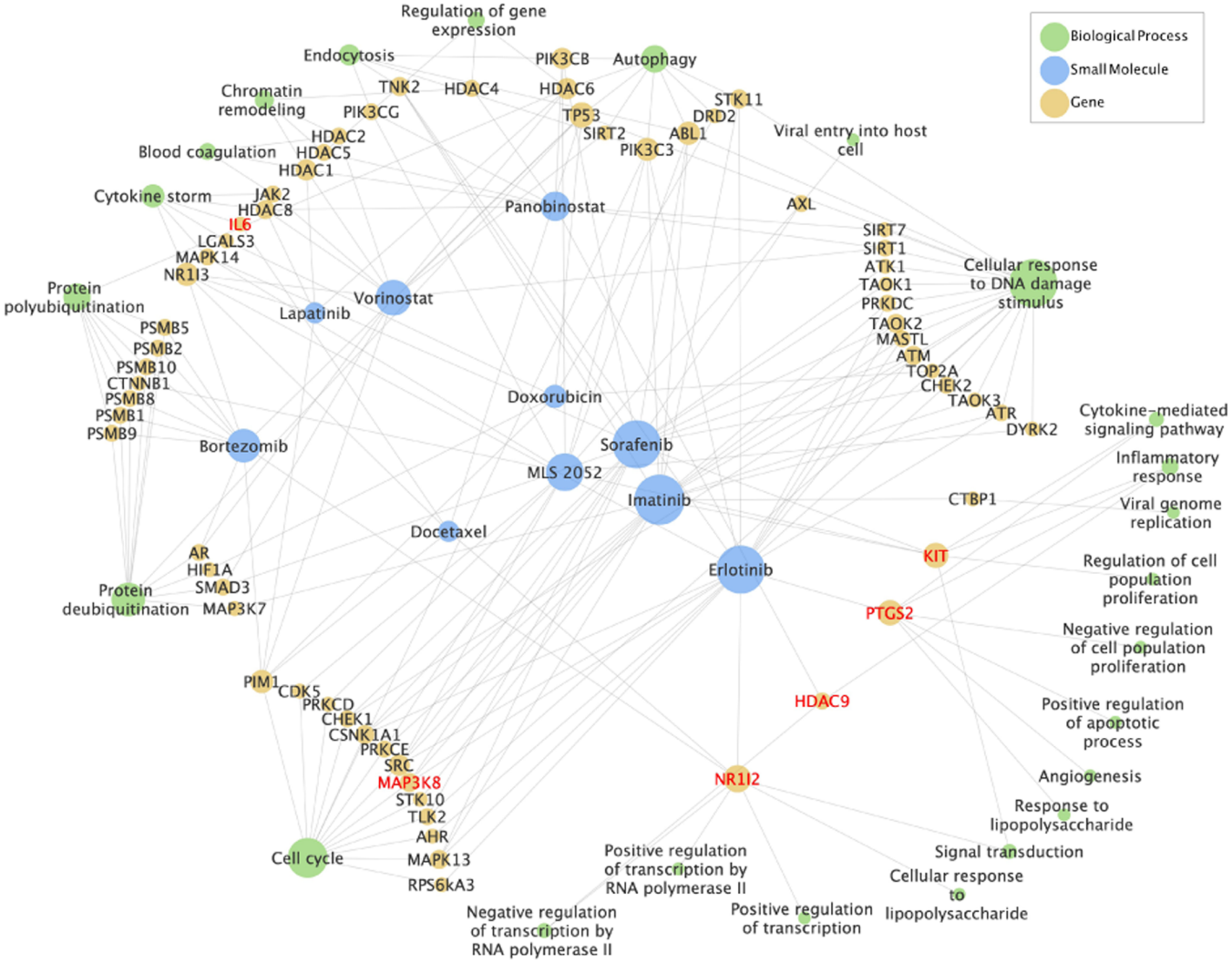


Figure 4. Network diagram showing the connections of the top 10 scoring drugs from the results. Gene names in red represent genes that have a greater than 2-fold change in expression in response to SARS-CoV-2 infection. The size of the node corresponds to the number of connections to other nodes.

These genes and biological processes were postulated as being of higher interest because they are targeted by fewer small molecules. To normalize for the promiscuity of small molecules, a penalty is applied to all small molecules that scales with the number of biological processes it connects to in the graph (SEP). Due to the inherent importance of the cytokine storm module, a score boost was given to small molecules that connect to that node (CS). A score boost was also applied to small molecules that have an activating relationship with genes as this is a rare relationship (AG). The values of this equation can be fine-tuned based on the experimental objectives. In our final equation, the importance of the upregulated-expression results was lowered by dividing the score by 5; this number was derived empirically to increase the presence of clinical trial drugs in the top results. This adjustment placed more emphasis on the CAS scientist-defined biological process scores while still allowing the upregulated-expression scores to influence the final ranking. The 50 top-ranked potential drug repurposing candidates along with their drug class and clinical trial status are shown in Table 1. The individual score components for the top 50 small molecules are provided in Table S3, and the complete result set is provided in Table S4. The top 10 small molecules are visualized in the network diagram in Figure 4.

Among the top 50 drug repurposing candidates, 11 have been or are in clinical trials for treating COVID-19, thus supporting the validity of our results.[10] An error analysis was performed by varying constants of the equation to determine their effect on the number of clinically investigated small molecules in the top 50. We found that the values chosen for our constants supplied the highest number of clinically investigated small molecules within the top 50. Interestingly, AG had no effect on the number of clinical trial drugs present in the top 50 results. However, the impact of AG is still apparent as 46 out of the top 50 feature this rare connection. The error analysis results are described in the Supporting Information.

The largest class of drugs found in our results was kinase inhibitors, which accounted for 36% of the top 50 drug repurposing candidates in Table 1. The high prevalence of this drug class can be explained by the fact that kinases are involved in almost all biological processes, and their activities are dysregulated in many diseases. As such, kinase inhibitors are one of the most studied drug classes in pharmacology.[11]

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Indeed, it has been estimated that 20−33% of all drug discovery research involves protein kinases alone.[12] Furthermore, kinases have long been shown to be involved in the viral infection process, including in coronavirus infections.[13] For instance, receptor tyrosine kinases are involved in the cell entry of many different viruses.[14] Bekerman et al. have shown that kinase inhibitors impair intracellular viral trafficking and exert broad-spectrum antiviral effects.[15] Inhibitors of kinases PKC, IRAK4, p38,[16] and GSK-3[17] suppress SARS-CoV-2 replication. Given this, the large number of kinase inhibitors in our top 50 results is within expectations, and the enrichment of this class is likely due to their high prevalence in drug discovery research.

In this study, the kinase inhibitors we identified include those affecting receptor tyrosine kinases (RTKs) such as the EGF, FGF, PDGF, and ALK receptors as well as nonreceptor tyrosine kinases such as Bruton tyrosine kinase (BTK). Also included were serine/threonine kinases such as B-RAF, PKC, PIM, and GSK-3beta and lipid kinases such as phosphatidylinositol 3-kinase (PI3K). Four of these, the tyrosine kinase inhibitors imatinib, tofacitinib, ibrutinib, and genistein, have been or are in clinical trials for COVID-19. Additionally, Treon et al. found that ibrutinib (BTK inhibitor) may offer protection against the severe form of COVID-19 and may mitigate lung injury due to SARS-CoV-2.[18]

Another of the larger drug classes from our top 50 results was histone deacetylase inhibitors (HDIs). This makes sense in relation to COVID-19 in that (1) HDACs regulate gene expression by reducing histone deacetylation, and HDIs have been shown to reduce the expression of both angiotensin-converting enzyme 2 (ACE2), the main cell surface receptor for SARS-CoV-2, and the ABO glycosyltransferase, an enzyme-regulating blood type, a known COVID-19 risk factor;[19] (2) HDACs regulate several of the chemokines and cytokines involved in the immune response in COVID-19;[20] and (3) the SARS-CoV-2 proteinase MPro directly binds to HDAC2.[21] Additionally, Liu et al.[22] showed that HDAC inhibitors such as romidepsin can block SARS-CoV-2 entry in a pseudotyped SARS-CoV-2 virus model. Further investigation is warranted, however, because HDACs have also been shown to be required for the transcription of interferon-stimulated genes and antiviral responses.[23]

Microtubules are filaments composed of tubulin subunits. They are constantly going through the process of assembly and disassembly at their ends, giving them a dynamic and unstable quality.[24] Many studies have shown that SARS-CoV-2 proteins interact with microtubules or microtubule-associated proteins. For example, NSP13 interacts with many proteins in the centrosome, where microtubule minus ends are organized. The microtubule-regulating agents, such as docetaxel, colchicine, and mebendazole, in Table 1, may therefore be of use in disrupting SARS-CoV-2 infection. In fact, colchicine (ranked 41 in Table 1), a microtubule polymerization blocker, and VERU-111, an α- and β-tubulin inhibitor/cytoskeleton disruptor, are currently in clinical trials for the treatment of COVID-19 patients.

Another drug class shown in our results are protease inhibitors, most of which are proteasome inhibitors. It has been previously shown that the ubiquitin-proteasome system (UPS) is involved in viral replication and the cytokine storm[25] including in coronavirus-associated diseases,[26] so it seems rational that proteasome inhibitors would be of value in treating COVID-19. Several such inhibitors are already being investigated as COVID-19 therapeutics, and several were found in our results (bortezomib, carfilzomib, and saxagliptin).[27]

The category labeled "other" from Table 1 includes drugs that are difficult to classify. While we will not discuss most of these in detail, two, (−)-apomorphine and ouabain, were of interest. The dopamine agonist and aporphine-type alkaloid (−)-apomorphine is linked in the knowledge graph to the well-studied COVID-19 drug target called the sigma-1 receptor (Sigma1R, gene SIGMAR1). Sigma1R is a ligand-regulated membrane chaperone usually localized to endoplasmic reticulum (ER)-mitochondrial membrane junctions. It regulates many processes including protein folding, ER and oxidative stress, autophagy, and ion transport. Viruses often use cell stress pathways to aid replication, and accordingly, Sigma1R ligands have been studied as general antiviral agents for many years and, more recently, as anti-SARS-CoV-2 agents.[28] Further, the SARS-CoV-2 NSP6 protein directly binds to the sigma-1 receptor[21] and a sigma-1 receptor ligand, fluvoxamine, has been shown to reduce the chances of deterioration in patients with symptomatic COVID-19.[29] This suggests that (−)-apomorphine may be a worthwhile drug repurposing candidate for COVID-19.

Another drug of interest in the other category, ouabain, is a Na+/K+-ATPase inhibitor. Na+/K+-ATPase is a membrane transporter that exports cellular sodium in exchange for importing potassium. It regulates cell−ion concentrations, cell volume, membrane potential, and reactive oxygen species. While the mechanisms are not fully understood, many viruses, including coronaviruses, are known to be inhibited by Na+/K+-ATPase-inhibiting cardiac glycosides. Indeed, coronavirus cell entry is inhibited when the Na,K-ATPase alpha1 subunit is silenced or inhibited.[30] Further, a peptide (NaKtide) derived from the alpha1 subunit of Na+/K+-ATPase reduces the inflammatory cytokines present in chronic obesity and therefore may be of value in treating the cytokine storm often seen in severe COVID-19.[31] Indeed, others have recently provided supporting in vitro[32] and in vivo[33] evidence for the effectiveness of Na+/K+-ATPase inhibitors in inhibiting SARS-CoV-2.

DISCUSSION

In this paper, we describe the construction of the CAS Biomedical Knowledge Graph and its application, along with a novel results-ranking method, in predicting potential drug repurposing candidates for COVID-19. The graph contains 6 million nodes and 18 million relationships and is built on data from the CAS Content Collection and external databases. Overall, we identified and ranked 1350 small-molecule repurposing candidates and analyzed the top 50 in greater detail. The validity of the knowledge graph and ranking method is supported by the fact that 11 of the top 50 results have been or are currently in clinical trials for treating COVID-19 and that many of the drug classes for these small molecules are well known to play important roles in viral infections. While we focused on COVID-19 in this study, the CAS Biomedical Knowledge Graph described here can also be used to analyze other diseases such as Alzheimer's, Parkinson's, cancer, and even rare, or orphan, diseases. Beyond the life sciences,knowledge graphs building on our vast collection of scientific information can be appliedin many areas of science, including other areas of chemistry, materials science, food science, energy technology, and environmental research.

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

The advantages of knowledge graphs have become more widely known in the last 10 years. Most importantly, and beyond their use just as an information management system, knowledge graphs allow users to grasp information in a visual and intuitive way. Users can easily zoom in on specific modules, or subsets, of a large data set and then zoom back out to see how that subset fits in with the whole of the data. They can visually navigate through the pathways connecting data to see how modules, including nonadjacent ones, affect each other. This allows users a different perspective on a research problem. Another important advantage of knowledge graphs is that because they are both modular and flexible, data sources can be substituted in and out, such as was done here by adding COVID-19 clinical trial data to the CAS graph.

An example in pharmaceutical research of the nonadjacency benefit mentioned above is that by linking disease-associated pathways, the proteins in those pathways, and the small molecules that regulate them, a knowledge graph can enable a researcher who has identified a novel interaction between two proteins to quickly identify which pathways, biological processes, or diseases this interaction could alter. Use of the graph in this manner has the potential to greatly increase the speed of basic biomedical research. This same kind of approach allows life sciences researchers to identify a wider variety of drug targets "upstream" or "downstream" of those already known. Many of these targets may have been previously overlooked or were not considered to participate in other disease processes. The present results illustrate this nicely. If the goal is to reduce the blood coagulation associated with COVID-19, traditional methods may suggest only blood coagulation factors and the blood vessel wall proteins they directly interact with as targets. But our results also suggest histone deacetylases (HDACs) as possible targets because two HDACs have been linked to blood coagulation within the graph. By the same token, using a knowledge graph allows the prediction of potential upstream and downstream drug candidates. The widely known HDI vorinostat affects three proteins/genes involved in blood coagulation and can therefore be considered a potential repurposing candidate for treating COVID-19-related coagulopathy. That an HDI may reduce blood clotting is supported by the evidence that another HDI, valproic acid, upregulates expression of tissue-type plasminogen activator and reduces thrombus size after vascular injury.[34] In drug discovery research, therefore, the wider, more comprehensive view provided by knowledge graphs can lead to cost- and time-savings in initial drug screening. Of course, all identified drug candidates would still have to be validated by experimental and clinical testing.

In addition to their strengths, knowledge graphs have some of the same limitations common to all data management systems. For instance, by linking modules in a complex network, they may give the false visual impression that they contain a complete picture of what is known about a subject. However, knowledge is always incomplete, so like all databases, they must be maintained and periodically updated. Furthermore, the power of a graph depends on the quality and comprehensiveness of the data sources used to build it.

The equation developed for ranking small molecules in this study focuses on identifying relevant and significant small molecule-to-protein relationships. We hypothesized that small molecules with extensive connections to target proteins would be more likely to have significant side effects. An inherent drawback to this approach is that small molecules that could be effective in COVID-19 therapies may be down-ranked if they are highly connected. Our two-component query combined human-designed and data-driven approaches, which we hypothesize may allow our results to capture potentially unknown molecular mechanisms of COVID-19. The flexibility of our equation in combining this two-component approach allows us to independently adjust the importance of each component. For our final ranking, the CAS-designed component was given higher importance as we felt the chosen biological processes covered a more specific and important spectrum of COVID-19 pathology. This scoring could be altered as other applications require. We also explored the use of machine learning with our results. We employed a decision tree to determine which of the variables identified in our equation was most important for the prediction of a small molecule's use in a COVID-19 clinical trial. We found that the side effect proxy was of highest importance followed by gene rarity and biological process rarity. The machine learning results also demonstrated the effectiveness of our equation in ranking the small molecules. The use of machine learning to improve the returned results could be of great value in a different application.

While the CAS Biomedical Knowledge Graph contains a massive wealth of entities and relationships, there is also potential for improvement. One improvement that could be made is the addition of a new layer of information within the protein−protein interaction relationships. If these relationships included the effect one protein had on another, such as phosphorylation, activation, and/or physical inhibition, the graph would allow for more powerful queries. Searching for an inhibitor of an upstream activator of a protein involved in a disease is one such line of questioning this data would allow. Another addition planned for the graph is the expression level of genes in different tissues and cellular compartments. The accuracy of our queries for a given disease would increase if we could limit our search to genes expressed only in the tissue of interest. An example of this for COVID-19 would be restraining our results based on gene expression in the lungs. These are just two of many possible improvements, which will require additional data analytics work to integrate into the CAS Biomedical Knowledge Graph. We plan to leverage the data and expertise of CAS to add these capabilities to the CAS Biomedical Knowledge Graph in the future.

In conclusion, we have leveraged a century's worth of CAS scientific information curation expertise to create the CAS Biomedical Knowledge Graph, which combines an extensive list of small molecules, present in the CAS Content Collection, with external databases of human genes, molecular processes, pathways, and diseases. We used the CAS Biomedical Knowledge Graph to identify 1350 small molecules with potential to be repurposed as COVID-19 therapeutics. Because knowledge graphs are both scalable and modular, this application to COVID-19 is only one example of the vast array of possible uses for CAS-powered knowledge graphs.

MATERIALS AND METHODS

CAS Biomedical Knowledge Graph. The CAS Biomedical Knowledge Graph was constructed in Neo4j (DBMS Version 4.1.0). In total, the graph contains over 6 million nodes and 18 million relationships. During data ingestion, all references to proteins and genes were normalized to their NCBI/HUGO gene abbreviation and all small-molecule references were normalized to their CAS Registry Number.

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Small Molecules and Bioactivity Data. Over 6 million small molecules were added to the graph from the CAS Content Collection. All small molecules were cross-referenced with data from PubChem and ChEMBL. Data scientists at CAS then connected small-molecule nodes to gene nodes where experimental assays have been performed linking the two, with details from these assays stored in the small molecule-to-gene relationships. These details include information about the activity being measured, raw values from the assay, and the source of the experimental data. Over 10 million relationships between small molecules and genes are present in the graph. Side effects associated with the small molecules were obtained from SIDER (version 4.1) and ingested into the graph.

Human Genes, Viral Genes, and Viruses. Gene nodes serve as a representation of a gene's DNA, RNA, and protein forms. Over 26 000 human and viral genes were obtained from the UniProt database and stored in the graph using their NCBI/HUGO gene abbreviations. Protein−protein interactions between human proteins were obtained from STRING-DB (version 11.0) resulting in over 5 million protein−protein interactions (PPIs) with STRING-DB confidence values stored in each relationship. Over 1000 virus nodes were added from UniProt, which were linked to the viral genes they express.

Diseases, Pathways, Molecular Functions, and Biological Processes. Over 24 000 human disease designations, obtained from NCBI's MedGen, were added to the graph. A hierarchy of disease inheritance was established using MedGen's parent−child relationships of diseases (over 14 000 links). Disease−gene associations were also obtained from MedGen, resulting in over 5 million connections. Pathways and pathway−gene associations were obtained from NCBI, resulting in over 8000 pathways and over 121 000 links between genes and pathways. Molecular functions (over 4000) and biological processes (over 12 000) were obtained from the Gene Ontology knowledgebase along with their gene associations (over 59 000 and over 138 000, respectively).

SARS-CoV-2-Specific Data. We identified several data sources that were used to establish connections between SARS-CoV-2 and the CAS Biomedical Knowledge Graph. One such data source measured human gene expression-level changes in response to SARS-CoV-2 infection.[9] This data was added as a relationship between the SARS-CoV-2 virus and the human genes where each relationship contains the expression fold change and p values (over 18 000 links). We also added SARS-CoV-2- and COVID-19-related clinical trial information obtained from clinicaltrials.gov. Relationships in the graph were generated between the clinical trial, the diseases/viruses being investigated, and the small molecules used in the trial. In addition, biological processes related to COVID-19 that were identified in our previous work[8] were included.

Graph Queries and Image Preparation. Graph queries were performed in Neo4j using the Cypher query language. Small-molecule results were filtered to ensure they fit the following three criteria: (1) the small molecule has been identified by CAS scientists as having pharmacological activity; (2) the assay that generated the small molecule-to-gene relationship measured IC_{50}, EC_{50}, K_d, or potency values; and (3) the raw value results from the assay were 10 μM or lower. Query code and returned results can be found in the Supporting Information. All network graph images were generated using Cytoscape (version 3.8.2).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00642.

Code excerpts from project (TXT)
Error analysis of ranking equation (PDF)
CAS Scientist-identified COVID-19 target processes (Table S1) and SARS-CoV-2 expression change-identified COVID-19 target processes (Table S2) (PDF)
Top 50 small molecules with score components (Table S3) (XLSX)
Complete list of 1350 small molecules for two-component approach (Table S4) (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Jacob Al-Saleem − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States;* orcid.org/0000-0001-8929-7098; Email: JAl-Saleem@cas.org

Authors

Roger Granet − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States*
Srinivasan Ramakrishnan − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States*
Natalie A. Ciancetta − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States*
Catherine Saveson − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States*
Chris Gessner − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States*
Qiongqiong Zhou − *CAS, A division of the American Chemical Society, Columbus, Ohio 43202, United States;* orcid.org/0000-0001-6711-369X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00642

Notes

The authors declare no competing financial interest.

Neo4J Community Edition 4.1.4 and Neo4j Desktop 1.3.4 are free software programs used to create, edit, and interact with the graph database. These software packages can be downloaded from Neo4j directly at https://neo4j.com/download-center/. Neo4j Python Driver 4.0 is a free software toolkit used to access the Neo4j-based graph database using the python scripting language. The Neo4j Python driver can be downloaded at https://pypi.org/project/neo4j/. All python code was executed using Visual Studio Code, which can be downloaded at https://code.visualstudio.com/download. Cytoscape 3.8.2 is a free software program that allows for graphical representations of graph database query results. Cytoscape can be downloaded at https://cytoscape.org/download.html. Cytoscape Neo4j plugin (version 0.4) is a free software plugin used to access the Neo4j-based graph database within Cytoscape. The plugin can be downloaded at http://apps.cytoscape.org/apps/cytoscapeneo4jplugin. All small molecule and small molecule relationship data is proprietary to CAS. Gene and virus information was obtained

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

from UniProt. Diseases and links to genes were obtained from NCBI MedGen. Protein−protein interactions were obtained from STRING-DB. Molecular functions and biological process were obtained from the Gene Ontology knowledgebase.

■ ACKNOWLEDGMENTS

The authors would like to acknowledge Rumiana Tenchov for her assistance in reference curation, Laura Czuba for project coordination, and Peter Jap and Cristina Tomeo for insightful discussions. We thank the CAS Data, Analytics & Insights team for extracting chemical names. We are also grateful to Manuel Guzman, Gilles Georges, Michael Dennis, Carmin Gade, Dawn George, Cynthia Casebolt, and Hong Xie for executive sponsorship.

■ REFERENCES

(1) Ng, Y. L.; Salim, C. K.; Chu, J. J. H. Drug repurposing for COVID-19: Approaches, challenges and promising candidates. *Pharmacol. Ther.* **2021**, *228*, No. 107930.

(2) Yao, X.-H.; Luo, T.; Shi, Y.; He, Z.-C.; Tang, R.; Zhang, P.-P.; Cai, J.; Zhou, X.-D.; Jiang, D.-P.; Fei, X.-C.; Huang, X.-Q.; Zhao, L.; Zhang, H.; Wu, H.-B.; Ren, Y.; Liu, Z.-H.; Zhang, H.-R.; Chen, C.; Fu, W.-J.; Li, H.; Xia, X.-Y.; Chen, R.; Wang, Y.; Liu, X.-D.; Yin, C.-L.; Yan, Z.-X.; Wang, J.; Jing, R.; Li, T.-S.; Li, W.-Q.; Wang, C.-F.; Ding, Y.-Q.; Mao, Q.; Zhang, D.-Y.; Zhang, S.-Y.; Ping, Y.-F.; Bian, X.-W. A cohort autopsy study defines COVID-19 systemic pathogenesis. *Cell Res.* **2021**, DOI: 10.1038/s41422-021-00523-8.

(3) Delorey, T. M.; Ziegler, C. G. K.; Heimberg, G.; Normand, R.; Yang, Y.; Segerstolpe, Å.; Abbondanza, D.; Fleming, S. J.; Subramanian, A.; Montoro, D. T.; Jagadeesh, K. A.; Dey, K. K.; Sen, P.; Slyper, M.; Pita-Juárez, Y. H.; Phillips, D.; Biermann, J.; Bloom-Ackermann, Z.; Barkas, N.; Ganna, A.; Gomez, J.; Melms, J. C.; Katsyv, I.; Normandin, E.; Naderi, P.; Popov, Y. V.; Raju, S. S.; Niezen, S.; Tsai, L. T. Y.; Siddle, K. J.; Sud, M.; Tran, V. M.; Vellarikkal, S. K.; Wang, Y.; Amir-Zilberstein, L.; Atri, D. S.; Beechem, J.; Brook, O. R.; Chen, J.; Divakar, P.; Dorceus, P.; Engreitz, J. M.; Essene, A.; Fitzgerald, D. M.; Fropf, R.; Gazal, S.; Gould, J.; Grzyb, J.; Harvey, T.; Hecht, J.; Hether, T.; Jané-Valbuena, J.; Leney-Greene, M.; Ma, H.; McCabe, C.; McLoughlin, D. E.; Miller, E. M.; Muus, C.; Niemi, M.; Padera, R.; Pan, L.; Pant, D.; Pèer, C.; Pfiffner-Borges, J.; Pinto, C. J.; Plaisted, J.; Reeves, J.; Ross, M.; Rudy, M.; Rueckert, E. H.; Siciliano, M.; Sturm, A.; Todres, E.; Waghray, A.; Warren, S.; Zhang, S.; Zollinger, D. R.; Cosimi, L.; Gupta, R. M.; Hacohen, N.; Hibshoosh, H.; Hide, W.; Price, A. L.; Rajagopal, J.; Tata, P. R.; Riedel, S.; Szabo, G.; Tickle, T. L.; Ellinor, P. T.; Hung, D.; Sabeti, P. C.; Novak, R.; Rogers, R.; Ingber, D. E.; Jiang, Z. G.; Juric, D.; Babadi, M.; Farhi, S. L.; Izar, B.; Stone, J. R.; Vlachos, I. S.; Solomon, I. H.; Ashenberg, O.; Porter, C. B. M.; Li, B.; Shalek, A. K.; Villani, A.-C.; Rozenblatt-Rosen, O.; Regev, A. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature* **2021**, *595*, 107−113.

(4) Domingo-Fernández, D.; Baksi, S.; Schultz, B.; Gadiya, Y.; Karki, R.; Raschka, T.; Ebeling, C.; Hofmann-Apitius, M.; Kodamullil, A. T. COVID-19 Knowledge Graph: a computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology. *Bioinformatics* **2021**, *37*, 1332−1334.

(5) Zhang, R.; Hristovski, D.; Schutte, D.; Kastrin, A.; Fiszman, M.; Kilicoglu, H. Drug Repurposing for COVID-19 via Knowledge Graph Completion. *J. Biomed. Inf.* **2021**, *115*, No. 103696.

(6) Hsieh, K.; Wang, Y.; Chen, L.; Zhao, Z.; Savitz, S.; Jiang, X.; Tang, J.; Kim, Y. Drug Repurposing for COVID-19 using Graph Neural Network with Genetic, Mechanistic, and Epidemiological Validation, 2020. arXiv:2009.10931v1. https://arxiv.org/abs/2009.10931v1.

(7) Hogan, A.; Blomqvist, E.; Cochez, M.; D'Amato, C.; De Melo, G.; Gutierrez, C.; Labra Gayo, J. E. Knowledge Graphs, 2021. arXiv:2003.02320. https://arxiv.org/abs/2003.02320v5.

(8) Zhou, Q. A.; Kato-Weinstein, J.; Li, Y.; Deng, Y.; Granet, R.; Garner, L.; Liu, C.; Polshakov, D.; Gessner, C.; Watkins, S. Potential Therapeutic Agents and Associated Bioassay Data for COVID-19 and Related Human Coronavirus Infections. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 813−834.

(9) Blanco-Melo, D.; Nilsson-Payant, B. E.; Liu, W. C.; Uhl, S.; Hoagland, D.; Møller, R.; Jordan, T. X.; Oishi, K.; Panis, M.; Sachs, D.; Wang, T. T.; Schwartz, R. E.; Lim, J. K.; Albrecht, R. A.; ten Oever, B. R. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **2020**, *181*, 1036−1045.

(10) Clinical Trials. https://clinicaltrials.gov/ (accessed April 29, 2021).

(11) Cohen, P.; Tcherpakov, M. Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell* **2010**, *143*, 686−693.

(12) Roskoski, R., Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. *Pharmacol. Res.* **2020**, *152*, No. 104609.

(13) Pillaiyar, T.; Laufer, S. Kinases as Potential Therapeutic Targets for Anti-coronaviral Therapy. *J. Med. Chem.* **2021**, DOI: 10.1021/acs.jmedchem.1c00335.

(14) Haqshenas, G.; Doerig, C. Targeting of host cell receptor tyrosine kinases by intracellular pathogens. *Sci. Signaling* **2019**, *12*, No. eaau9894.

(15) Bekerman, E.; Neveu, G.; Shulla, A.; Brannan, J.; Pu, S. Y.; Wang, S.; Xiao, F.; Barouch-Bentov, R.; Bakken, R. R.; Mateo, R.; Govero, J.; Nagamine, C. M.; Diamond, M. S.; De Jonghe, S.; Herdewijn, P.; Dye, J. M.; Randall, G.; Einav, S. Anticancer kinase inhibitors impair intracellular viral trafficking and exert broad-spectrum antiviral effects. *J. Clin. Invest.* **2017**, *127*, 1338−1352.

(16) Liu, S.; Zhu, L.; Xie, G.; Mok, B. W.-Y.; Yang, Z.; Deng, S.; Lau, S.-Y.; Chen, P.; Wang, P.; Chen, H.; Cai, Z. Potential Antiviral Target for SARS-CoV-2: A Key Early Responsive Kinase during Viral Entry. *CCS Chem.* **2021**, *3*, 559−568.

(17) Liu, X.; Verma, A.; Ramage, H.; Garcia, G.; Myers, R. L.; Lucas, A.; Michaelson, J. J.; Coryell, W.; Kumar, A.; Charney, A.; Kazanietz, M. G.; Rader, D. J.; Ritchie, M. D.; Berrettini, W. H.; Damoiseaux, R.; Arumugaswami, V.; Schultz, D.; Cherry, S.; Klein, P. S. Targeting the Coronavirus Nucleocapsid Protein through GSK-3 Inhibition, 2021. DOI: 10.1101/2021.02.17.21251933.

(18) Treon, S. P.; Castillo, J. J.; Skarbnik, A. P.; Soumerai, J. D.; Ghobrial, I. M.; Guerrera, M. L.; Meid, K.; Yang, G. The BTK inhibitor ibrutinib may protect against pulmonary injury in COVID-19-infected patients. *Blood* **2020**, *135*, 1912−1915.

(19) Takahashi, Y.; Hayakawa, A.; Sano, R.; Fukuda, H.; Harada, M.; Kubo, R.; Okawa, T.; Kominato, Y. Histone deacetylase inhibitors suppress ACE2 and ABO simultaneously, suggesting a preventive potential against COVID-19. *Sci. Rep.* **2021**, *11*, No. 3379.

(20) Gatla, H. R.; Muniraj, N.; Thevkar, P.; Yavvari, S.; Sukhavasi, S.; Makena, M. R. Regulation of Chemokines and Cytokines by Histone Deacetylases and an Update on Histone Decetylase Inhibitors in Human Diseases. *Int. J. Mol. Sci.* **2019**, *20*, No. 1110.

(21) Gordon, D. E.; Jang, G. M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K. M.; O'Meara, M. J.; Rezelj, V. V.; Guo, J. Z.; Swaney, D. L.; Tummino, T. A.; Hüttenhain, R.; Kaake, R. M.; Richards, A. L.; Tutuncuoglu, B.; Foussard, H.; Batra, J.; Haas, K.; Modak, M.; Kim, M.; Haas, P.; Polacco, B. J.; Braberg, H.; Fabius, J. M.; Eckhardt, M.; Soucheray, M.; Bennett, M. J.; Cakir, M.; McGregor, M. J.; Li, Q.; Meyer, B.; Roesch, F.; Vallet, T.; Mac Kain, A.; Miorin, L.; Moreno, E.; Naing, Z. Z. C.; Zhou, Y.; Peng, S.; Shi, Y.; Zhang, Z.; Shen, W.; Kirby, I. T.; Melnyk, J. E.; Chorba, J. S.; Lou, K.; Dai, S. A.; Barrio-Hernandez, I.; Memon, D.; Hernandez-Armenta, C.; Lyu, J.; Mathy, C. J. P.; Perica, T.; Pilla, K. B.; Ganesan, S. J.; Saltzberg, D. J.; Rakesh, R.; Liu, X.; Rosenthal, S. B.; Calviello, L.; Venkataramanan, S.; Liboy-Lugo, J.; Lin, Y.; Huang, X.-P.; Liu, Y.; Wankowicz, S. A.; Bohn, M.; Safari, M.; Ugur, F. S.; Koh, C.; Savar, N. S.; Tran, Q. D.; Shengjuler, D.; Fletcher, S. J.; O'Neal, M. C.; Cai, Y.; Chang, J. C. J.; Broadhurst, D. J.; Klippsten, S.; Sharp, P. P.; Wenzell, N. A.; Kuzuoglu-Ozturk, D.; Wang, H.-Y.; Trenker, R.; Young, J. M.; Cavero, D. A.; Hiatt, J.; Roth, T. L.; Rathore, U.; Subramanian, A.; Noack, J.; Hubert, M.; Stroud, R.

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

M.; Frankel, A. D.; Rosenberg, O. S.; Verba, K. A.; Agard, D. A.; Ott, M.; Emerman, M.; Jura, N.; von Zastrow, M.; Verdin, E.; Ashworth, A.; Schwartz, O.; d'Enfert, C.; Mukherjee, S.; Jacobson, M.; Malik, H. S.; Fujimori, D. G.; Ideker, T.; Craik, C. S.; Floor, S. N.; Fraser, J. S.; Gross, J. D.; Sali, A.; Roth, B. L.; Ruggero, D.; Taunton, J.; Kortemme, T.; Beltrao, P.; Vignuzzi, M.; García-Sastre, A.; Shokat, K. M.; Shoichet, B. K.; Krogan, N. J. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020**, *583*, 459−468.

(22) Liu, K.; Zou, R.; Cui, W.; Li, M.; Wang, X.; Dong, J.; Li, H.; Li, H.; Wang, P.; Shao, X.; Su, W.; Chan, H. C. S.; Li, H.; Yuan, S. Clinical HDAC Inhibitors Are Effective Drugs to Prevent the Entry of SARS-CoV2. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 1361−1370.

(23) Chang, H.-M.; Paulson, M.; Holko, M.; Rice, C. M.; Williams, B. R. G.; Marié, I.; Levy, D. E. Induction of interferon-stimulated gene expression and antiviral responses require protein deacetylase activity. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9578−9583.

(24) Wen, Z.; Zhang, Y.; Lin, Z.; Shi, K.; Jiu, Y. Cytoskeleton-a crucial key in host cell for coronavirus infection. *J. Mol. Cell Biol.* **2021**, *12*, 968−979.

(25) Longhitano, L.; Tibullo, D.; Giallongo, C.; Lazzarino, G.; Tartaglia, N.; Galimberti, S.; Li Volti, G.; Palumbo, G. A.; Liso, A. Proteasome Inhibitors as a Possible Therapy for SARS-CoV-2. *Int. J. Mol. Sci.* **2020**, *21*, No. 3622.

(26) Kircheis, R.; Haasbach, E.; Lueftenegger, D.; Heyken, W. T.; Ocker, M.; Planz, O. Potential of proteasome inhibitors to inhibit cytokine storm in critical stage COVID-19 patients, 2020. arXiv:2008.10404. https://arxiv.org/abs/2008.10404v1.

(27) Limanaqi, F.; Busceti, C. L.; Biagioni, F.; Lazzeri, G.; Forte, M.; Schiavon, S.; Sciarretta, S.; Frati, G.; Fornai, F. Cell Clearing Systems as Targets of Polyphenols in Viral Infections: Potential Implications for COVID-19 Pathogenesis. *Antioxidants* **2020**, *9*, No. 1105.

(28) Vela, J. M. Repurposing Sigma-1 Receptor Ligands for COVID-19 Therapy? *Front. Pharmacol.* **2020**, *11*, No. 582310.

(29) Lenze, E. J.; Mattar, C.; Zorumski, C. F.; Stevens, A.; Schweiger, J.; Nicol, G. E.; Miller, J. P.; Yang, L.; Yingling, M.; Avidan, M. S.; Reiersen, A. M. Fluvoxamine vs Placebo and Clinical Deterioration in Outpatients With Symptomatic COVID-19: A Randomized Clinical Trial. *JAMA* **2020**, *324*, 2292−2300.

(30) Amarelle, L.; Lecuona, E. The Antiviral Effects of Na,K-ATPase Inhibition: A Minireview. *Int. J. Mol. Sci.* **2018**, *19*, No. 2154.

(31) Xie, Z. J.; Novograd, J.; Itzkowitz, Y.; Sher, A.; Buchen, Y. D.; Sodhi, K.; Abraham, N. G.; Shapiro, J. I. The Pivotal Role of Adipocyte-Na K peptide in Reversing Systemic Inflammation in Obesity and COVID-19 in the Development of Heart Failure. *Antioxidants* **2020**, *9*, No. 1129.

(32) Cho, J.; Lee, Y. J.; Kim, J. H.; Kim, Si.; Kim, S. S.; Choi, B.-S.; Choi, J.-H. Antiviral activity of digoxin and ouabain against SARS-CoV-2 infection and its implication for COVID-19. *Sci. Rep.* **2020**, *10*, No. 16200.

(33) Plante, K. S.; Dwivedi, V.; Plante, J. A.; Fernandez, D.; Mirchandani, D.; Bopp, N.; Aguilar, P. V.; Park, J. G.; Tamayo, P. P.; Delgado, J.; Shivanna, V.; Torrelles, J. B.; Martinez-Sobrido, L.; Matos, R.; Weaver, S. C.; Sastry, K. J.; Newman, R. A. Antiviral activity of oleandrin and a defined extract of Nerium oleander against SARS-CoV-2. *Biomed. Pharmacother.* **2021**, *138*, No. 111457.

(34) Larsson, P.; Alwis, I.; Niego, B.; Sashindranath, M.; Fogelstrand, P.; Wu, M. C.; Glise, L.; Magnusson, M.; Daglas, M.; Bergh, N.; Jackson, S. P.; Medcalf, R. L.; Jern, S. Valproic acid selectively increases vascular endothelial tissue-type plasminogen activator production and reduces thrombus formation in the mouse. *J. Thromb. Haemostasis* **2016**, *14*, 2496−2508.

https://doi.org/10.1021/acs.jcim.1c00642
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX