scientific reports

OPEN

Multiscale interactome analysis coupled with off-target drug predictions reveals drug repurposing candidates for human coronavirus disease

Michael G. Sugiyama[1,21], Haotian Cui[2,3,21], Dar'ya S. Redka[4,21], Mehran Karimzadeh[3], Edurne Rujas[5,6,7], Hassaan Maan[3,8,9], Sikander Hayat[10,20], Kyle Cheung[1,11], Rahul Misra[1], Joseph B. McPhee[1,11], Russell D. Viirre[1,11], Andrew Haller[12,13], Roberto J. Botelho[1,11], Raffi Karshafian[11,14,15,16], Sarah A. Sabatinos[1,11], Gregory D. Fairn[6,14,17], Seyed Ali Madani Tonekaboni[4], Andreas Windemuth[4], Jean-Philippe Julien[5,6,18], Vijay Shahani[4], Stephen S. MacKinnon[4✉], Bo Wang[2,3,9,19✉] & Costin N. Antonescu[1,11,14✉]

The COVID-19 pandemic has highlighted the urgent need for the identification of new antiviral drug therapies for a variety of diseases. COVID-19 is caused by infection with the human coronavirus SARS-CoV-2, while other related human coronaviruses cause diseases ranging from severe respiratory infections to the common cold. We developed a computational approach to identify new antiviral drug targets and repurpose clinically-relevant drug compounds for the treatment of a range of human coronavirus diseases. Our approach is based on graph convolutional networks (GCN) and involves multiscale host-virus interactome analysis coupled to off-target drug predictions. Cell-based experimental assessment reveals several clinically-relevant drug repurposing candidates predicted by the in silico analyses to have antiviral activity against human coronavirus infection. In particular, we identify the MET inhibitor capmatinib as having potent and broad antiviral activity against several coronaviruses in a MET-independent manner, as well as novel roles for host cell proteins such as IRAK1/4 in supporting human coronavirus infection, which can inform further drug discovery studies.

Human coronaviruses have emerged as significant agents in serious human disease, highlighting the need for rapid development of therapeutics. Recently, the severe acute respiratory syndrome (SARS) coronavirus-2

[1]Department of Chemistry and Biology, Ryerson University, Toronto, ON, Canada. [2]Department of Computer Science, University of Toronto, Toronto, ON, Canada. [3]Vector Institute, Toronto, ON, Canada. [4]Cyclica Inc., Toronto, ON, Canada. [5]Program in Molecular Medicine, The Hospital for Sick Children Research Institute, Toronto, ON, Canada. [6]Department of Biochemistry, University of Toronto, Toronto, ON, Canada. [7]Biofisika Institute (CSIC, UPV/EHU) and Department of Biochemistry and Molecular Biology, University of the Basque Country (UPV/EHU), Bilbao, Spain. [8]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada. [9]Peter Munk Cardiac Centre, University Health Centre, Toronto, ON, Canada. [10]Precision Cardiology Laboratory, Bayer US LLC, Cambridge, MA, USA. [11]Graduate Program in Molecular Science, Ryerson University, Toronto, ON, Canada. [12]Phoenox Pharma, Toronto, ON, Canada. [13]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada. [14]Keenan Research Centre for Biomedical Science, St. Michael's Hospital, Toronto, ON, Canada. [15]Institute for Biomedical Engineering, Science and Technology (iBEST), a partnership between Ryerson University and St. Michael's Hospital, Toronto, ON, Canada. [16]Department of Physics, Ryerson University, Toronto, ON, Canada. [17]Department of Surgery, University of Toronto, Toronto, ON, Canada. [18]Department of Immunology, Toronto, ON, Canada. [19]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada. [20]Present address: Institute of Experimental Medicine and Systems Biology, RWTH Aachen University, Aachen, Germany. [21]These authors contributed equally: Michael G. Sugiyama, Haotian Cui, and Dar'ya S. Redka. ✉email: stephen.mackinnon@cyclicarx.com; bowang@vectorinstitute.ai; cantonescu@ryerson.ca

(SARS-CoV-2) has caused a pandemic with > 125 M individuals infected and over 2.5 M deaths globally[1]. Prior to this, the related human coronaviruses SARS-CoV-1 and Middle East Respiratory Syndrome coronavirus (MERS-CoV) were responsible for outbreaks of severe human coronavirus disease with significant mortality[2,3]. Additional human coronaviruses including NL63, OC43, and 229E elicit milder disease such as the common cold[4], although more severe human diseases related to these viruses have been reported[5]. Human coronaviruses thus represent an important family of related viruses that impact human health worldwide with new therapies for these agents urgently needed. The recent emergence of SARS-CoV-2 variants of concern[6,7] indicates that therapeutic strategies with broad antiviral activity to human coronaviruses are a high priority.

SARS-CoV-2, SARS-CoV-1 and NL63 enter cells by using their Spike (S) outer protein, which interacts with angiotensin converting enzyme 2 (ACE2) on host cells[8–11], while 229E and OC43 depend on host cell aminopeptidase N and glycosaminoglycans, respectively[12,13]. In addition, other host cell receptors contribute to SARS-CoV-2 entry, such as neuropilin-1[14,15]. The coronavirus-bound receptor protein(s) enter cells via clathrin-mediated endocytosis[16,17] or other membrane traffic mechanisms[18], after which the viral RNA genome undergoes replication and expression of viral proteins, leading to assembly of viral progeny at a specialized ER-derived coronavirus replication organelle[19]. This is followed by coronavirus release to the extracellular milieu by a mechanism involving lysosomal exocytosis[20], allowing spread of the viral particles to nearby cells[21,22].

Human coronavirus cellular entry, replication, assembly and egress depend on a wide range of host cell proteins and functions. For SARS-CoV-2, 26 viral proteins were found to interact with 332 host cell proteins, spanning a range of functions including membrane trafficking, centrosome structure and function, membrane transport, DNA replication and stress granule formation[23]. This viral protein interactome provides a rich source of information from which to identify host proteins that are functionally important for viral entry and replication and thus may serve as antiviral drug targets. Infection studies have also demonstrated conserved host-protein interaction patterns across different coronaviruses, suggesting that some host proteins can be targeted for broader antiviral activities[24–26]. The virus/host protein interactome and the identification of proteins functionally required for viral entry and replication of common cold coronaviruses may reveal important novel pan-human coronavirus antiviral drug targets.

Antivirals targeting SARS-CoV-2 such as remdesivir have modest clinical benefits[27], indicating a need for more effective antivirals to complement anti-inflammatory therapy approaches for the treatment of COVID-19. As the timeline for development of new drugs is ~ 10 years[28], and given the current dire need for effective antiviral agents, the identification of new drugs with antiviral activity cannot provide antiviral therapies in time to address the current pandemic. Drug repurposing is a preferred method for developing rapid response therapies[28,29], as it prioritizes the identification of additional drugs that can be repurposed as antivirals.

Repositioning of approved drugs can be achieved either by repurposing for a related disease, repurposing of a target with a known drug for another indication, or repurposing of a drug for a novel target by taking advantage of off-target interactions. The relative prevalence of the three approaches reflects the balance of their relative feasibility and broadness of their applicability. The disease-centric approach, while the most direct and the most common, representing 59% of repositioning successes[30], suffers from a more narrow generalizability in the disease space, as exemplified by the modest utility of repurposing of antivirals described above. Target-centric repurposing, while having the potential for broader application as a less direct approach, is less common since discovery of a novel link between a target and a disease is a rare finding. Even if such a link is discovered, the generalizability of the target-centric approach is limited by the drug-target coverage in the human proteome: as of 2019, only 667 of the roughly 20,000 human proteins (~ 3%) are directly targeted by FDA- approved drugs[31]. For that reason, drug-centric repurposing, as the least direct methodology, offers access to a much broader range of repurposing opportunities. However, since it is aimed at finding novel, off-target drug-target interactions, and therefore relies on the fundamental understanding of structural information about both the drug and the target, it has been the least prevalent repositioning approach, representing only 6% of the successes[30].

Successful drug repurposing discoveries to date with either of the three approaches have been largely accidental or hypothesis-driven[32]. Computational approaches, however, provide an alternative opportunity to identify repurposing candidates that may have been overlooked. These hypothesis-free strategies reference broad data sources to identify new protein targets, identify new compounds for a pre-selected target, or pair phenotypic signatures of a disease state with drug actions[33]. Notably, multiscale interactome approaches combine known relationships between disease, biological pathways, genes, proteins, and other -omic data to predict new potential indirect relationships[34]. Furthermore, when those interactomes are built into Graph Convolutional Networks (GCNs), they offer a systematic, Machine-Learning (ML) based approach to model the value of each relationship in the network empirically based on their intrinsic properties[35,36]. Such models can be applied to the target-centric repurposing approach. ML models that provide predictions for drug-targets interaction, on the other hand, and are capable of cross-screening libraries of clinically relevant compounds with large sets of proteins, can assist with the drug-centric repurposing programs[37].

In this study, we aimed to exploit the therapeutic potential of approved drugs by taking advantage of both the directness of the target-centric and the broader potential of the drug-centric repurposing approaches by combining GCN-based multiscale host-viral interactome approaches for target discovery with off-target interaction predictions from the PolypharmDB database[37] to shortlist clinically-relevant repurposing candidates for screening in coronavirus infectivity assays. This specific combination of approaches was selected to simultaneously explore a variety of mechanistically-informed host-based targets in a focused phenotypic driven experiment. In contrast, a single-target virtual screen would require a prior target validation stage or risk having no observed activities upon experimental testing. The proposed approach maximizes the diversity of targets and compounds explored to increase the likelihood of a bioactive discovery in an experimentally efficient manner.

We then examine the effectiveness of these predictions using a combination of human coronavirus entry and infection assays. We reveal several compounds that have antiviral activity, in particular capmatinib, a drug
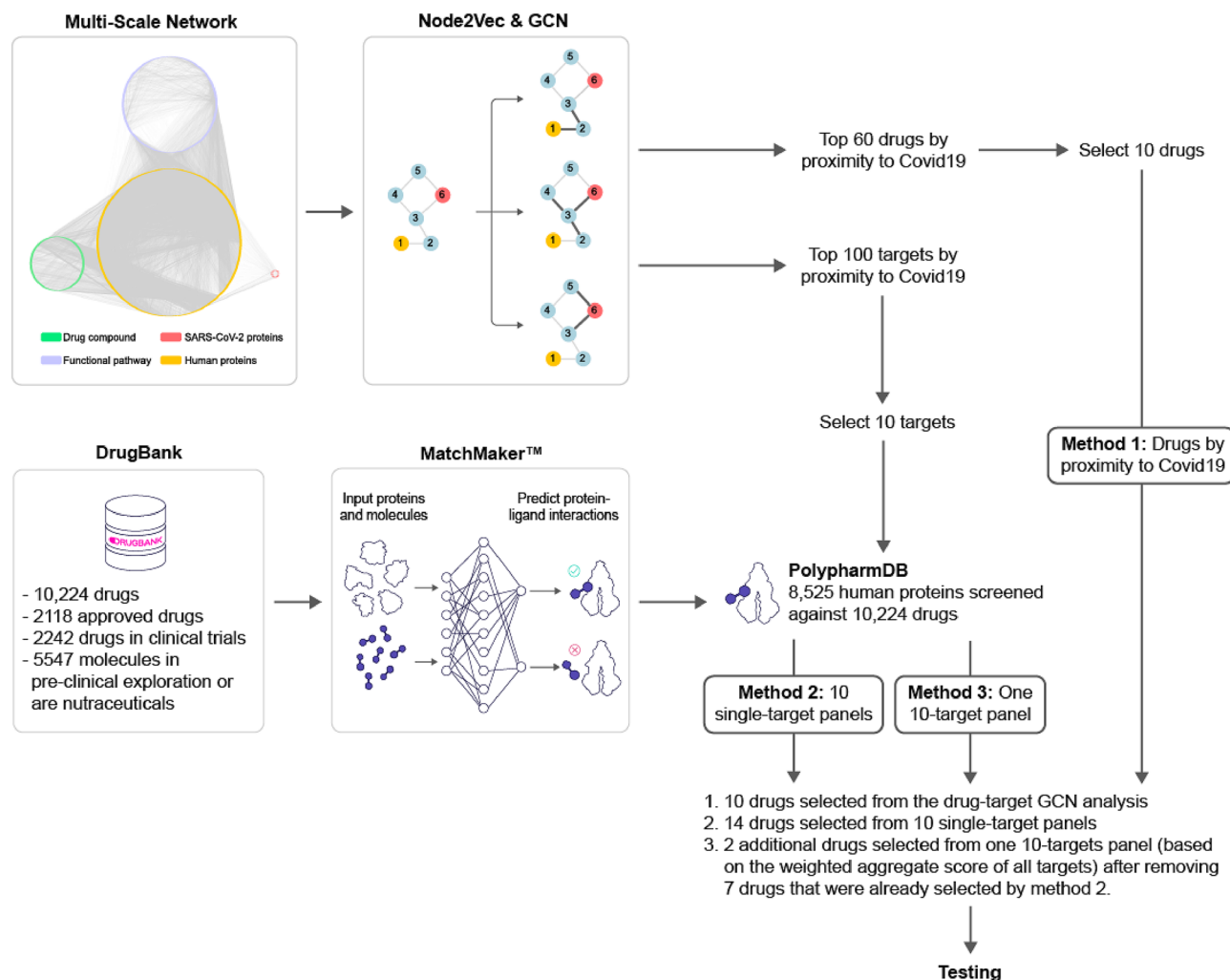
Figure 1. Three methods of selecting top drug candidates based on the GCN and MatchMaker prediction. Twenty-six drugs were selected for experimental testing by applying the predictions made by Node2Vec/GCN and MatchMaker models using three different methods. For *Method 1*, ten drugs were selected directly from top 60 drugs that were predicted to be most proximal to COVID-19 based on the GCN alone. For *Methods 2* and *3*, first, ten human protein targets were selected from the list of top 100 proteins that are most proximal to COVID-19 based on the GCN prediction. Following that, for *Method 2*, 14 drug candidates were selected from top ten top ranking PolypharmDB (i.e., 10,224 drugs from DrugBank[95] screened against 8525 human proteins; see *Materials and Methods*) candidates for each protein (i.e. ten single-target panels resulting in 100 candidates in total). For *Method 3*, nine out of the top 25 ranking PolypharmDB candidates for the ten-targets panel were selected as candidates, seven of which were already present in the list of candidates selected with *Method 2*. Please see Table 1 and Materials and Methods for further details.

known to inhibit the receptor tyrosine kinase MET. Notably, we find that capmatinib has potent anti-human coronavirus activity in a MET-independent manner. Further, we find novel roles for human proteins such as IRAK1/4 in supporting human coronavirus infection.

Results

A total of 26 drug repurposing candidates for SARS-CoV-2 were identified using three separate approaches involving a multiscale interactome GCN and the PolypharmDB drug repurposing database[37], as described in Fig. 1 and in "*Materials and Methods*". Drugs assert their function by binding to proteins and regulating biological pathways. Therefore, we first constructed a multiscale interactome network to represent the known relation of 1661 drug molecules, 17,660 human proteins, 9,798 functional pathways and finally the 26 expressed proteins from SARS-CoV-2. This representation enables fine grained analysis on the probable drugs and protein targets for COVID-19, based on the assumption that the targets interacting with the viral protein nodes on the network are more likely to play a role in the potential treatment. Thus, we proposed a combined approach of node2vec and GCN to learn and generate the node embeddings in the multi-scale network. The embedding is optimized in an unsupervised manner to the objective that related nodes on the network should have higher embedding similarity (see *Materials and Methods*). All embeddings are then sorted by the proximity to the COVID-19

Method 1: Drug-target GCN analysis		Method 2: PolypharmDB query for 10 single GCN-target panels		Method 3: PolypharmDB query for one 10 GCN-targets panel	
Drug	Target	Drug	Target	Drug	Panel of Targets
Bucladesine	PRKACA	Bortezomib	IRAK4	Anidulafungin	UGGT2, SDF2, NLRX1, MOGS, HEPACAM, IRAK4, ADAM15, CD46, LILRA3, and CHPF2
Cinnarizine	CACNA1C	Cefotiam	TARS2	Capmatinib	
Doxycycline	PADI4	Dapagliflozin	UGGT2	*Bortezomib*	
Eflornithine	SLC25A21	Degarelix	CD46	*Dapagliflozin*	
Flucytosine	DNMT1	Fosamprenavir	ADAM15	*Fosamprenavir*	
Glyburide	ABCA1	Gentamicin	HEPACAM	*Glipizide*	
Nelarabine	POLA1	Glipizide	ADAM15	*Palbociclib*	
Opicapone	COMT	Palbociclib	IRAK4	*Streptozocin*	
Otilonium	CACNA1C	Polidocanol	MOGS	*Tofacitinib*	
Simvastatin	ITGB2	Saquinavir	CD46		
		Streptozocin	HEPACAM		
		Sugammadex	MOGS		
		Telithromycin	CD46		
		Tofacitinib	IRAK4		

Table 1. Drugs that were selected for testing using three methods based on GCN analysis and MatchMaker predictions. Top 26 drug candidates were selected based on GCN and MatchMaker predictions as described in Fig. 1 and in Materials and Methods. Cefotiam, under Method 2, was selected from the list of 10 top ranking candidates for TARS2, which was among the top 100 human proteins predicted to be associated with COVID-19 through target-only GCN analysis, but was not prioritized in the initial selection of top 10 targets described above. Drug candidates that were selected using Method 3 and were also predicted by Method 2 are shown in grey font and italics.

node. When first reviewing proximity distances between COVID-19 and human targets, very few had known small molecule modulators, which is an expected limitation of the target-centric drug repurposing strategy. To overcome the sparsity of drug-target interaction in the network, additionally, we further applied the drug-centric approach: off-target predictions for relevant low-data targets were retrieved from the PolypharmDB database. PolypharmDB is a database of precompiled all-by-all Drug Target Interaction (DTI) predictions performed by the MatchMaker deep-learning engine, for 8535 human proteins with 10,244 clinically-tested small molecules (see *Materials and Methods*).

For *Method 1*, a target-centric approach, the top 10 drug candidates were selected among the 60 drugs with the highest proximity scores to COVID-19, as determined by the GCN approach. For *Method 2*, a drug-centric approach, 14 candidates were selected among single-target PolypharmDB hits for protein targets with high proximity scores to COVID-19 as determined by the GCN approach. For *Method 3*, an extension of the drug-centric approach, the final two candidates were selected with a variation on *Method 2*, which prioritizes compounds with multiple predicted interactions to GCN-identified SARS-CoV-2 targets. The compounds selected using all three methods are summarized in Table 1.

Since this approach generates numerous potential hits for further validation, we chose to use a system that could be widely used and scalable without restricted access to BSL3 laboratories, as required for monitoring of SARS-CoV-2 infection. This method may bias the identification of drug compounds with pan-human coronavirus antiviral activity, rather than compounds with selective antiviral activity towards SARS-CoV-2. Thus, to examine whether the 26 selected compounds exhibit antiviral activity, we established a cell culture-based immunofluorescence (IF) screening system using the 229E human alphacoronavirus, based on detection of the 229E S protein. This assay was designed to allow ~ 100% of control cells to express the S protein following 48 h of infection, allowing robust detection of antiviral activity as reduction of S protein abundance. We validated this IF assay by treating 229E-infected cells with the nucleoside analogue prodrug remdesivir[27] (Fig. S1), showing that this BSL2-based screening system allows for the safe and straightforward identification of novel antiviral compounds.

Using this assay, we identified four putative antivirals within the 26 predicted hits (4/26, ~ 15%) that reduced S protein abundance following 229E infection by at least 50% (Fig. 2) with no apparent cytotoxicity. Treatment with palbociclib and anidulafungin caused partial attenuation of 229E infection, while treatment with capmatinib or polidocanol resulted in nearly 100% inhibition of 229E infection. Based on our in silico framework, these four compounds were predicted to target several host proteins, each with potential novel and noncanonical roles in supporting coronavirus infection (Table 1). Interestingly, palbociclib, capmatinib and anidulafungin were predicted to target IRAK4, either as a primary target or by a polypharmacology panel. Another compound, bortezomib, was also predicted to target IRAK4, however cytotoxicity prevented further analysis of this compound.

We focused further investigation into the role of capmatinib as a potential antiviral therapeutic against human coronavirus disease, given its robust impairment of 229E infection in the IF assay. Capmatinib is an
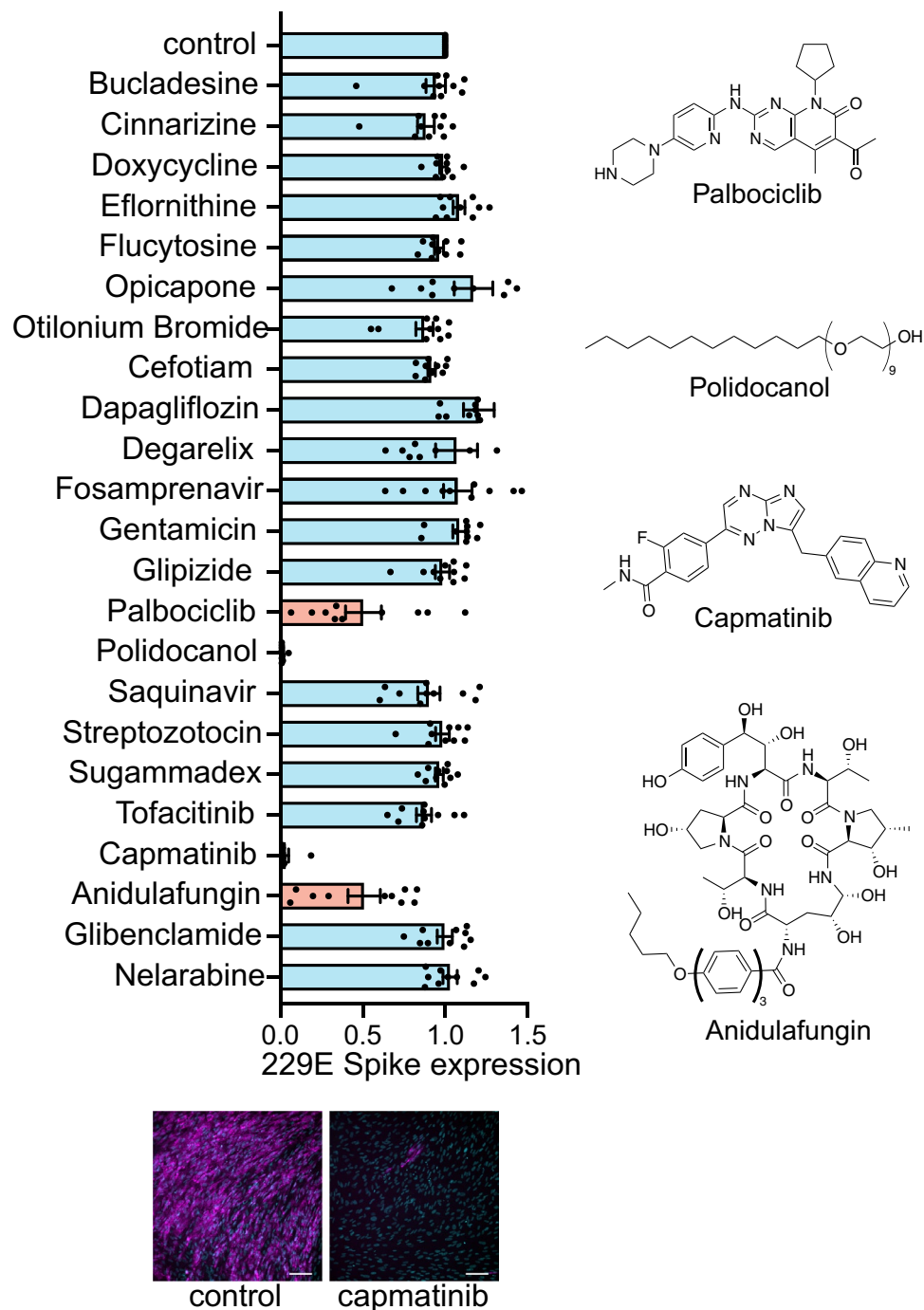
Figure 2. Screening of predicted compounds identifies capmatinib and other drugs as host-targeted compounds with antiviral activity against human coronaviruses. Graph depicting mean ± SE and individual measurements of 229E Spike protein expression as measured by IF assay upon incubation with drugs as per Table S4. Results are expressed as mean 229E Spike expression relative to the DMSO vehicle (control) condition. (bottom) Representative images showing S protein expression (magenta) or DAPI (cyan) of the DMSO vehicle (control) or capmatinib (10 µM) treated conditions. Scale, 100 µm. Also shown are structures of palbociclib, polidocanol, capmatinib and anidulafungin, compounds that showed antiviral activity.

orally-available inhibitor of the receptor tyrosine kinase (RTK) MET, and is used in the clinic for the treatment of MET-amplified non-small cell lung cancer[38]; however, our analyses predict that capmatinib may have antiviral activity due to inhibition of targets other than MET.

To further characterize the antiviral activity of capmatinib on human coronaviruses, we developed several complementary cell-based assays to probe the different aspects of infection by several different human coronaviruses, including 229E, OC43, and NL63. Noteworthy, capmatinib treatment impaired 229E viral replication
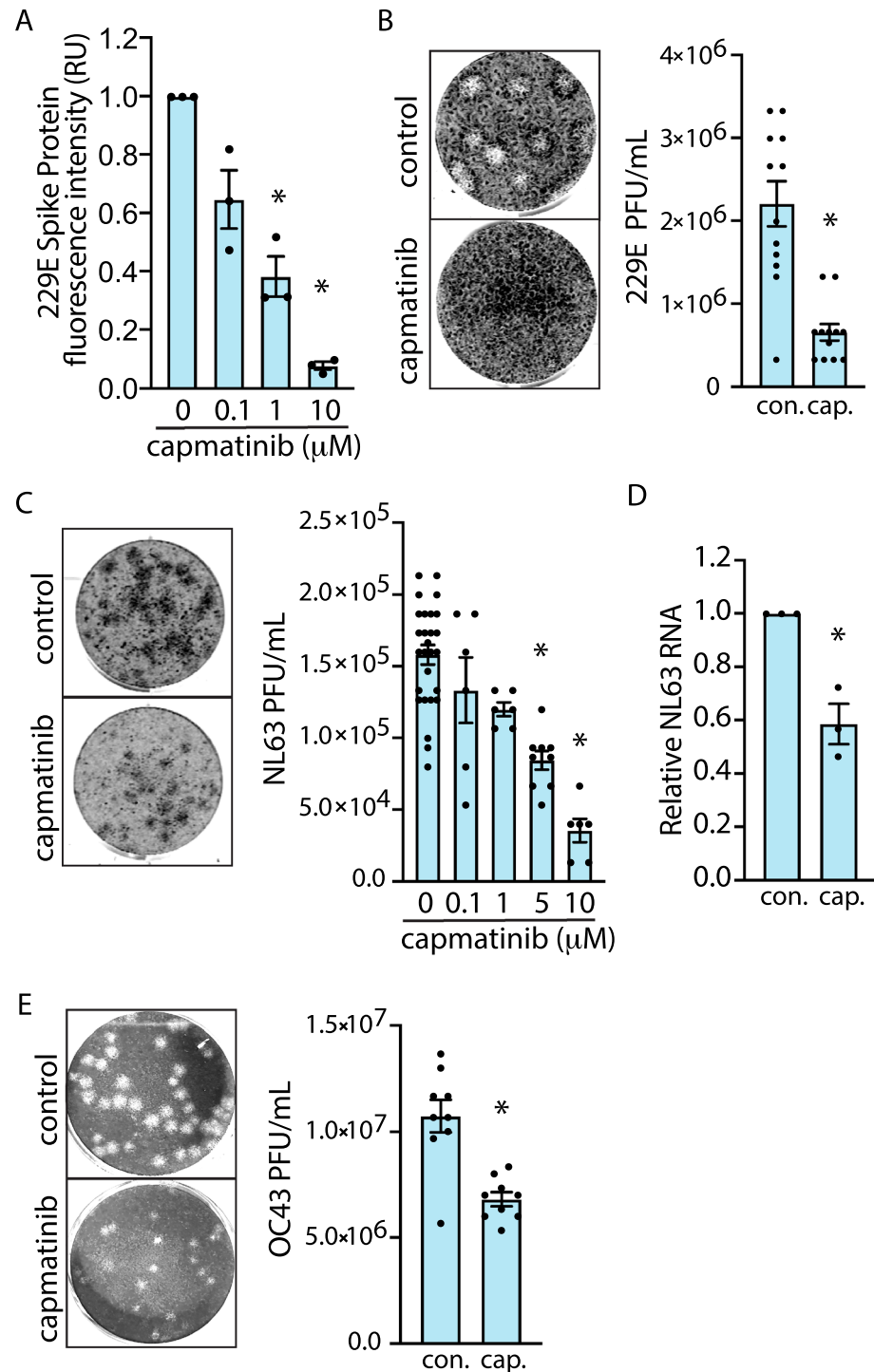
Figure 3. Capmatinib has a broad range of antiviral activity against human coronaviruses. (**A**) Quantification of 229E Spike protein abundance in MRC-5 cells treated with increasing doses of capmatinib in the IF assay (48 h infection), as mean ± SE (n = 3) expressed relative to DMSO (vehicle) control. *$P < 0.05$ relative to control (**B**) (left) Representative images of 229E plaques observed in MRC-5 cells treated with 10 μM capmatinib or DMSO (vehicle) for 6 days. (right) Quantification of viral titer from the DMSO (vehicle) control or capmatinib plaque assays, expressed as PFU/mL. *$P < 0.05$ relative to control (**C**). (left) Representative images of NL63 plaques observed in LLC-MK2 treated with 10 μM capmatinib or DMSO (vehicle) control for 5 days. (right) Quantification of NL63 PFU/mL in LLC-MK2 cells treated with the indicated doses of capmatinib. *$P < 0.05$ relative to control (**D**). Relative NL63 N protein RNA abundance 3 days post-infection in LLC-MK2 cells treated with 10 μM capmatinib relative to DMSO (vehicle) control. *$P < 0.05$ relative to control, n = 3 experimental replicates. E (left) Representative images of OC43 plaques observed in LLC-MK2 cells treated with 10 μM capmatinib or DMSO (vehicle) for 5 days. (right) Quantification of OC43 PFU/mL in LLC-MK2 cells treated with capmatinib or DMSO (vehicle) control. *$P < 0.05$ relative to control.

in a dose-dependent manner, with concentrations as low as 1.0 µM resulting in significant attenuation of 229E S protein abundance following infection (Fig. 3A). The ability of capmatinib to impair viral replication was further confirmed using a plaque-forming unit (PFU) assay in MRC-5 cells. 229E infection of control cells resulted in the formation of distinct, large circular plaques indicating cytopathic effect (CPE) (Fig. 3B). In contrast, capmatinib treatment resulted in reduced total number and size of individual plaques; viral quantification revealed that capmatinib treatment resulted in a > 50% decrease in viral production compared to control. Importantly, this method for viral quantification is based entirely on total plaque number and does not consider plaque size, thus, these results likely represent an underestimate of the antiviral effects of capmatinib, given the differences in plaque morphology between control and treatment groups (Fig. 3B).

To determine the breadth of the antiviral effects of capmatinib, we measured whether capmatinib also exhibited antiviral activity for other BSL2 human coronavirus infections, including OC43 and NL63. We adapted our MRC-5/229E plaque assay protocol for infection of LLC-MK2 cells with the alphacoronavirus NL63, which like SARS-CoV-2 depends on ACE2 for infection[8]. Treatment of LLC-MK2 cells with increasing concentrations of capmatinib resulted in a dose-dependent decrease in NL63 PFU with minimal cytotoxicity observed at 5 days post-infection (dpi), with > 70% reduction in PFU at 10 µM capmatinib (Fig. 3C). Consistent with results obtained with the PFU assay, qRT-PCR analysis showed that capmatinib attenuated NL63 N RNA abundance by approximately 50%, measured at 3 dpi (Fig. 3D). We also established a PFU assay based on the infection of LLC-MK2 cells with human betacoronavirus OC43. Treatment of OC43-infected LLC-MK2 cells with capmatinib resulted in a nearly 50% reduction in viral particles (Fig. 3E). As with our results obtained in the MRC-5/229E model, capmatinib treatment reduced both plaque number and the overall size of the plaques, indicating that our quantification of PFU/mL likely underestimated the actual antiviral effect of the drug. Taken together, our results from 3 different human coronaviruses and cell-based assays of coronavirus infection demonstrate that capmatinib has a broad range of antiviral activity against human coronaviruses.

We next explored the mechanism of the antiviral action of capmatinib. To determine whether the potent antiviral activity of capmatinib was due to its canonical role as a MET inhibitor, we compared the effects of capmatinib with another distinct MET inhibitor, AMG-337 that has similar in vitro $IC_{50} \leq 1$ nM for MET as capmatinib[39,40]. We first treated NL63-infected LLC-MK2 cells with capmatinib or an equimolar amount of AMG-337. As expected, capmatinib treatment greatly attenuated CPE as measured by PFU (Fig. 4A). In contrast, AMG-337 treatment, while similarly tolerated by the cells, did not result in an appreciable reduction in PFU. Similar results were observed in OC43-infected LLC-MK2 cells and 229E-infected MRC-5 cells; AMG-337 treatment had no effect on PFU in either case (Fig. 4B and Fig. S2). Capmatinib, but not AMG-337 was also effective at reducing 229E S protein abundance in the IF assay (Fig. 4C).

Spike-pseudotyped viral inhibition assays also demonstrate the ability of capmatinib to interfere with certain aspects of viral infection. SARS-CoV-1 and SARS-CoV-2 pseudoviruses (PsV) require human ACE2 for cell entry[41]. Incubation of HeLa cells stably expressing ACE2 (HeLa-ACE2 cells) with capmatinib showed a potent inhibition of infection of both SARS-CoV-1 or SARS-CoV-2 PsV with half-inhibitory concentrations (IC_{50}) of 14.1 ± 0.2 and 26.0 ± 0.1 µM, respectively (Fig. 4D). In contrast, treatment with concentrations of AMG-337 up to the mM range had no appreciable effect on PsV neutralization (Fig. 4D). Both compounds displayed minimal cytotoxicity in HeLa cells (Fig. S3). Together, these results indicate that capmatinib exhibits human coronavirus antiviral activity by inhibiting a target other than MET. Moreover, the pseudotyped virus assay selectively monitors virus uptake and reporter gene delivery, suggesting that capmatinib inhibits early stages of virus entry.

We next sought to gain insight into the possible mechanism by which capmatinib exhibits human coronavirus antiviral activity. Based on MatchMaker, capmatinib was predicted to exhibit off-target interactions with IRAK4 and other protein targets (Table 1), which are part of pathways (e.g. interleukin-1 receptor) that have been broadly implicated in mediating SARS-CoV-2 infection and COVID-19 disease[42,43]. In addition, our in silico analyses also predicted palbociclib and anidulafungin as binders of IRAK4, and as these drugs also demonstrated antiviral activity (Fig. 2), this suggests that IRAK4 may be an important novel drug target for human coronavirus infection and that capmatinib and other drugs may exhibit antiviral activity by novel action on IRAK4.

IRAK4 (Interleukin-1 receptor-associated kinase 4) is a serine/threonine kinase that forms part of the myddosome complex (consisting of MyD88 and other IRAK kinases) to transduce signals from Toll-like receptors (TLRs) or interleukin receptors[44–46]. In canonical myddosome signal transduction, activation of membrane TLRs results in recruitment of the adaptor protein MyD88 and its associated IRAKs. MyD88-bound IRAK4 recruits and phosphorylates additional IRAK kinases, such as IRAK1 and IRAK2, which can then dissociate from the myddosome and trigger a signaling cascade involving subsequent TRAF6 and TAK-1 activation[44–46]. These signaling events ultimately result in activation of the transcription factors NF-κB and MAPKs, which together contribute to the innate immune response to pathogens. Importantly, IRAK4 exhibits some functional redundancy with the related kinase IRAK1[47], which led us to consider inhibition of IRAK1 and IRAK4 as an antiviral mechanism for human coronavirus infection.

We examined the effect of JH-I-25, a dual specific IRAK1 and IRAK4 inhibitor with no clinical relevance with IC_{50} values of 9.3 nM and 17.0 nM[48] respectively, on human coronavirus infection. Treatment of LLC-MK2 cells with JH-I-25 (10 µM) greatly reduced CPE in cells infected with OC43 (Fig. 4F). Similar results were obtained with the IF assay in MRC-5 cells infected with 229E (Fig. 4G). To verify that our results were specific to the combined inhibition of IRAK1 and IRAK4, we silenced both using siRNA gene silencing. In concordance with results obtained from the JH-I-25 experiments, combined silencing of IRAK1 and IRAK4 attenuated 229E S protein abundance (Fig. 4H). Consistent with a role for IRAK1/4 in coronavirus infection, inhibition of p38 MAPK, known to be activated downstream of IRAK1/4[44–46], also impaired coronavirus infection (Fig. S4). Taken together, these data demonstrate that the effects of combined IRAK1/4 inhibition recapitulate the effects of capmatinib, as predicted by our in silico analyses.

Figure 4. The antiviral activity of capmatinib is not attributed to its canonical role as an inhibitor of MET. (**A**) (left) Representative images of plaques in LLC-MK2 cells treated with 10 μM capmatinib, 10 μM AMG-337, or DMSO (vehicle) control and infected with NL63 for 5 days and (right) quantification of NL63 viral titer, shown as mean PFU ± SE (n = 3 with 3 technical replicates per experiment). (**B**) (left) Representative images of plaques in LLC-MK2 cells treated with DMSO (vehicle) control, 10 μM capmatinib, or 10 μM AMG-337, and infected with OC43 and (right) quantification of OC43 viral titer shown as mean PFU ± SE (n = 3 with 3 technical replicates per experiment). (**C**) (left) Representative images from IF assay of MRC-5 cells treated with 10 μM capmatinib or AMG-337 and infected with 229E. (right) Quantification of 229E S protein expression (20 images per condition, n = 3). (**D**) Representative neutralization curves from n = 5 independent experiments showing the relative antiviral activity of capmatinib vs. AMG-337 in pseudovirus assays performed with the SARS-CoV-1 and SARS-CoV-2 Spike protein. (**E**) Structures of capmatinib and AMG-337. (**F**) (left) Representative images of plaques in LLC-MK2 cells treated with 10 μM JH-I-25 and infected with OC43 for 5 days and (right) quantification of the OC43 viral titer shown as mean PFU ± SE (n = 3 with 3 technical replicates per experiment). (**G**) (left) Representative images from MRC-5 cells treated with 10 μM JH-I-25 and infected with 229E for 2 days and (right) quantification of 229E Spike protein expression (10 images per condition, n = 3). H (left) Representative images of MRC-5 cells transfected with IRAK1/4 siRNA or control and infected with 229E for 2 days. (right) Quantification of 229E Spike protein expression shown as mean ± SE, expressed relative to DMSO (vehicle) control (n = 3). *$P < 0.05$.

Discussion

Computational analysis and antiviral protein target predictions.

In silico drug repurposing approaches have the innate advantage of accessing very large collections of information to uncover indirect associations that may have been otherwise overlooked. While the target-centric methodologies provide a more direct opportunity for drug repurposing, they are limited due to the overwhelming majority of human proteins not being targeted by existing small molecule drugs. The drug-, or structure-centric methodologies, on the other hand, while limited by the physicochemical properties of available approved drugs, provide access to a broader range of human targets. In this study, we addressed these challenges simultaneously by both exploring the targets with known approved drugs and searching for previously unappreciated modulators of undrugged targets among approved drugs. To that end, we paired two ML approaches designed for target-identification and drug-target interaction predictions, demonstrating the viability of an in silico first repurposing workflow, coupled with robust bioactivity assays.

We integrated drug-target interactions, proteins and functional pathways in a multiscale interactome network. Based on the assumption that drugs take effect by binding to proteins and regulating pathways, the multiscale interactome traces the biological processes of available treatments via the interactions across proteins, functional pathways, drugs and the target disease, COVID-19. To capture this process, our approach used biased random walks and a Graph Convolutional Network (GCN) to model the correlations of nodes of multiple types and build embeddings for each drug, protein and pathway. The GCN module refines the protein and drug embeddings by further aggregating the relations in the network. GCN has been reported to be capable of encoding both graph structure and node features very efficiently. The model has a sequence of non-linear filter layers which aggregates the information of every node's vicinity, making it effective to learn both local and global relations from lower to higher layers. The GCN embeddings were optimized in an unsupervised manner to encode the direct and indirect relations in the multiscale interactome. Then by ranking the proximity between the embeddings of candidate proteins to the target disease, the GCN model determined the potential efficacy of drugs or protein targets for this disease.

Matchmaker, PolypharmDB, and drug repurposing.

The multiscale interactome GCN unveiled multiple plausible viral-host targets for repurposing leads. Rather than subjecting a single protein to a target-centric computational screen, the subsequent screening stage maximized diversity in both targets and their putative binders. Repurposing targets prioritized by the GCN were cross-referenced to PolypharmDB, a precompiled database of off-target interaction predictions between 8525 human proteins and 10,244 small molecule compounds with prior clinical evaluation. This approach contrasts sharply with disease-centric or target-centric approaches, which make up the overwhelming majority (~ 90%) of drug repurposing studies[49]. Evidence that capmatinib's antiviral activity is driven by interactions other than its known target MET is provided by the lack of antiviral activity of AMG-337, another potent inhibitor of MET RTK (Fig. 4A–E). AMG-337 s is chemically distinct from capmatinib, which likely drives its differential polypharmacology.

Criteria were set to maximize the diversity of compound and target selection for downstream evaluation. The recognition that polypharmacology may play a role in effective therapeutics motivated the evaluation of multiple targeted agents, leading to the nomination of capmatinib and other compounds for testing. Underappreciated polypharmacology may play a larger role in the activity of small molecule drugs; a study investigating oncology medications found for several drugs investigated that on-target interactions were inconsequential for drug activity, with off-target effects driving efficacy[50]. Alternatively, the use of multiple targets as separate or as combined objectives may have contributed to the success of this investigation simply by providing more possibilities for favorable outcomes. However, since few compounds are evaluated per presumed target and since these are mostly based on off-target interaction predictions, negative results obtained through this process are unable to inform on the involvement of their respective proteins on cell infectivity. Nonetheless, there may be broader opportunities within a drug-centric approach for drug repurposing, where validated therapeutic targets can be linked to small molecule drugs through approaches like PolypharmDB.
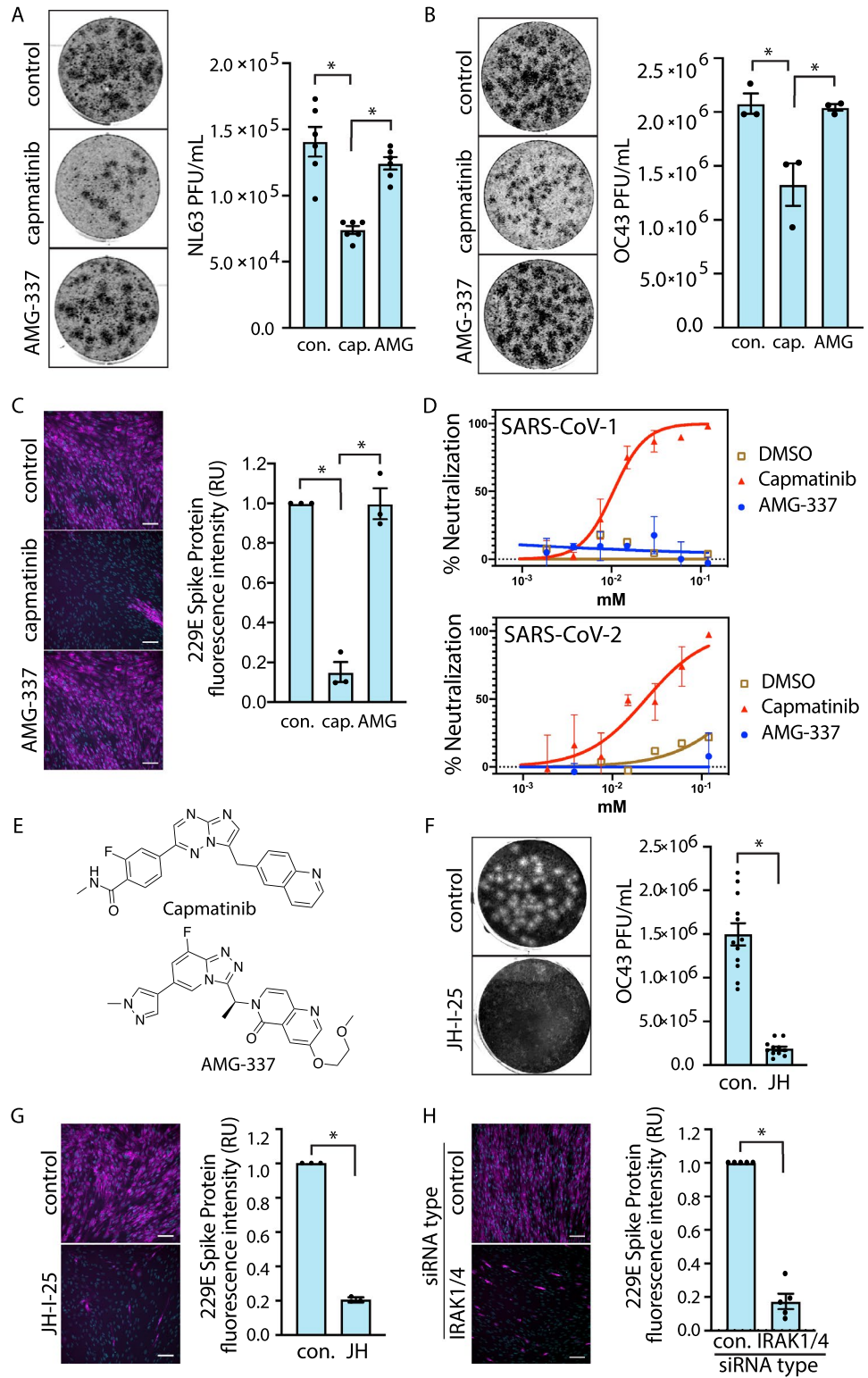
To that end, several other groups have used cell-based drug screening or computational approaches to identify drug repurposing candidates for COVID-19 (reviewed by[51–53]). Our approach, which uses a unique computational pipeline for identification of clinically-relevant drugs followed by experimental testing using a broad range of coronaviruses, is complimentary to these approaches. We are encouraged that our approach identified several compounds/families of drugs that have also been identified and tested for their ability to inhibit coronavirus infection in vitro, including polidocanol[54] and anidulafungin[55,56]. We also identified additional drugs including capmatinib and palbociclib, which to our knowledge have not been previously reported as having antiviral activity for human coronavirus infection. This may reflect certain advantages to our method, which considered > 10,000 clinically-relevant drugs, which may allow identification of molecules not considered in smaller-scale screens.

Capmatinib as an antiviral drug for COVID-19 and other human coronavirus diseases. Capmatinib was developed as a MET RTK inhibitor for the treatment of various MET-amplified tumors[57]. The MET proto-oncogene encodes an RTK that serves as the receptor for hepatocyte growth factor (HGF). A number of small-molecule MET inhibitors have recently been developed and are currently undergoing clinical trials to determine their efficacy in reducing cancer morbidity and mortality[58,59]. The availability of data on the safety of MET inhibitors in the clinic, and in particular capmatinib, provides support for this strategy of drug repurposing via polypharmacology.

We observed that capmatinib exhibited antiviral activity in the low micromolar range. Indeed, using the SARS-CoV-2 and SARS-CoV-1 pseudotype virus infection assay, we determined that the IC_{50} for capmatinib for neutralization of infection was 14.1 ± 0.2 and 26.0 ± 0.1 μM, which is in the range reported for action of other drugs targeting SARS-CoV-2 infection when tested in Vero and Calu-3 cells in culture, including remdesivir[60]. Hence, capmatinib should be considered as a candidate for therapeutic testing in further preclinical and clinical trials for the treatment of human coronavirus diseases, including COVID-19. In addition, while we have not tested the action of capmatinib against SARS-CoV-2 variants that have emerged in 2021[6,7], the broad antiviral activity that capmatinib exhibits against 5 genetically different human coronaviruses (229E, OC43 and NL63 live virus infection and SARS-CoV-1 and SARS-CoV-2 pseudotyped virus assays) suggests that capmatinib may hold promise in broadly treating SARS-CoV-2 variants of concern, or other variants that may arise from further antigenic drift as well as other emerging viruses. To our knowledge, this is the first demonstration of antiviral activity of capmatinib in cell-based coronavirus infection assays. While this manuscript was in preparation, other computational approaches predicted binding of capmatinib to the SARS-CoV-2 S protein, viral proteases, RNA-dependent RNA polymerase and/or viral endoribonuclease[61–63]. While our analysis indicates that capmatinib may act by targeting IRAK signaling, and we find a novel requirement for IRAK1/4 in supporting human coronavirus infection, the mechanism of antiviral action of capmatinib warrants further investigation in future studies.

Novel role for IRAK signaling in supporting human coronavirus infection. Myddosome signaling is considered an essential component of the innate immune response to bacterial PAMPs, and loss of function mutations in IRAK4 or MyD88 results in a primary immunodeficiency syndrome associated with a dramatic increase in susceptibility to specific pyogenic bacterial infections[64–67]. In contrast, MyD88 and IRAK4 may have distinct roles in response to other pathogens, as their perturbation often does not impact susceptibility to some viral, fungal, protozoal, or other infectious agents[64–68]. Notably, MyD88 perturbation contributes to coronavirus infection and symptom severity in animal models[69]. However, whether and how human coronaviruses may engage TLR and IRAK signals during infection remains poorly understood. While IRAK1 and 4 are broadly expressed, in single-cell datasets deposited in the covid19atlas, IRAK1 and IRAK4 appear highly expressed in Secretory3 cells in the bronchial epithelial dataset[70]. Interestingly, Secretory3 cells have the highest expression of ACE2 as compared to other cell types in the primary human bronchial epithelial cell dataset, consistent with a role for IRAK1/4 in supporting coronavirus infection.

Our results from two different human coronaviruses and two different host cell lines suggest that human coronavirus infection requires IRAK1 and/or IRAK4. We focused our studies on concomitant perturbation of IRAK1 and IRAK4, given the redundancy that has been reported between these kinases in some contexts[71]. Consistent with our results, another study performed an analysis of phosphorylated sequences within host cells and predicted activation of IRAK4 within 15 min of infection of Vero cells with SARS-CoV-2, and also revealed that a different inhibitor of IRAK1/4 impaired SARS-CoV-2 infection[72]. Together with the results presented here, this suggests that IRAK1/4 may contribute a non-canonical function to support human coronavirus infection. Recent work has also established that modulation of IRAK4 dependent immune responses is crucial for mounting an appropriate immune response during SARS-CoV-2 infection, supporting this observation[73].

As such, therapeutic modulation of IRAK signaling, and perhaps also that of TLRs and MyD88, may be a useful strategy for treatment of patients with COVID-19. In fact, a phase II clinical trial is ongoing to probe the use of the IRAK4 selective inhibitor PF-06650833 to treat COVID-19 patients with acute respiratory distress syndrome[74]. Supporting the role of this signaling pathway in the progression of the disease, obese individuals have increased TLR/MyD88 signaling that may predispose them to severe COVID-19 symptoms[75].

In this study, we use multiple methods involving a multiscale interactome GCN and the PolypharmDB drug repurposing database to identify new drug targets and drug repurposing candidates for the treatment of human coronavirus disease. We also provide evidence that several drug molecules predicted by this method have previously unknown antiviral activity against human coronavirus infection, in particular capmatinib. Further, we identify IRAK1/4 as new and unexpected coronavirus drug targets, required for coronavirus infection. This work highlights the potential of this computational approach, but it is important to note that additional pre-clinical and clinical testing is required before conclusions can be made about efficacy of treatment coronaviruses diseases in humans. Nonetheless, this indicates that the methods described herein are a novel and powerful approach for

the rapid identification of new therapeutic strategies to identify antiviral drugs and could also be applied more widely for novel therapeutic intervention to other classes of disease.

Methods

Materials. *Viruses.* 229E and OC43 coronaviruses were obtained from the American Type Culture Collection (ATCC) (ATCC VR-740™ and ATCC VR-1558™). NL63 coronavirus was kindly provided by Dr. Scott Gray-Owen (University of Toronto). Original viral stocks were stored at -80ºC until use and all subsequent viral stocks were produced from the original parental stocks.

Cell lines. MRC-5 (lung fibroblast) cells were obtained from ATCC (ATCC CCL-171™). CoronaGrow LLC-MK2 (kidney epithelial) cells, a subclonal line of parental LLC-MK2 cells were obtained from VectorBuilder Inc. (Chicago, IL, Cat. No. CL0004). HEK293T cells expressing the full-length SARS-CoV-2 spike (BEI NR52310) were obtained from BEI Resources (Manassas, VA). HEK-293 T cells expressing the SARS-CoV-2 spike were kindly provided by S. Pöhlmann (Leibniz Institute for Primate Research, Göttingen, Germany). HeLa-ACE2 cells were kindly provided by D.R. Burton (The Scripps Research Institute).

Multiscale Interactome. A combined host–pathogen multiscale interactome was assembled by augmenting a pre-constructed, base human network with viral-host protein–protein interaction data. The base multiscale interactome network consisting of drug-protein interactions (8,568), human protein–protein interactions (387,626) and protein-pathway interactions (22,545) was retrieved from Ruiz et al.[34], which aggregates data from multiple primary sources[76–86]. An additional 332 experimentally-derived, viral-host protein–protein interactions were added to adapt the network for SARS-CoV2 repurposing[23]. Then for a given viral protein that interacts with human proteins, we investigated the pathways in which these human proteins are involved. We added direct connections between viral proteins and the retrieved human functional pathways. In our experiments, adding direct connections between viral proteins and human pathways shows significant improvement in performance. Lastly, COVID-19 was introduced as a final entity to the multiscale interactome network and linked to all SARS-CoV-2 proteins. The multiscale interactome was represented as a graph $G = (V, E)$, where individual proteins, pathways and drugs form vertices (V) and interactions form the edges (E).

The goal of our approach is to learn meaningful embeddings of the nodes in the multiscale interactome network so that we can predict drug candidates or protein targets using these embeddings. To propose a proper embedding function, naively, we can introduce the bias of graph homophily, i.e. drugs/targets that connect to the viral protein via the shortest paths are the most likely to be effective. This is reflective of the assumption that drug effects propagate along the biological network to treat diseases. However, non-homophily relations can be equally important—for example, a protein can be a promising target if it is involved in several important related pathways, although it may not be directly connected to viral proteins. To address this challenge, we use graph convolutional networks to learn this hierarchy of complex relations from the interactome data.

Graph convolutional network. To prepare the Graph Convolutional Network (GCN), initial node embeddings were generated using Node2Vec[87]. The return parameter p and "in–out" parameter q, in Node2Vec were set to 0.25 in order to balance global and local views of the random walk process[88], which helped capture the aforementioned homophily and non-homophily relations. The embedded dimensions size D was set to 64 and node2vec was applied to the multiscale interactome graph G to convergence. The Node2vec output was an embedded matrix $H_n \in \mathbb{R}^{K,D}$, where $K = |V|$ is the number of nodes, and each row in H_n is the learned representation for the corresponding node.

These embeddings were used as the initial input feature layer, $H^{(0)}$, in our GCN model. The model consists of stacked multiple Graph Convolution and each one of them is defined by Eq. 1[35]. In this model, α is the activation function ReLU, $\tilde{A} = A + I_N$ is a transformation of the multiscale interactome adjacency matrix (A) with the additional self-associating nodes (identity matrix I_N), $H^{(l)}$ and $W^{(l)}$ representing layer-specific feature vectors and trainable weights, while $\tilde{D}$ is a diagonal degree matrix defined by $\tilde{D}_{ij} = \sum_j \tilde{A}_{ij}$.

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \qquad (1)$$

The adjacency matrix $A \in \mathbb{R}^{K,K}$ in Eq. 1 is generated by the edges E in the multiscale interactome. Each element in A is either the connection weight or zero if there is no connection.

The output features of the last layer $H^{(L)} = \left[h_1^{(L)}, h_2^{(L)}, \ldots, h_K^{(L)}\right]^T$ are then normalized to produce the final embedding for each node.

$$h_i = \frac{h_i^{(L)}}{\left\|h_i^{(L)}\right\|_2} \qquad (2)$$

To train the GCN model, we first compute cosine similarity between the embeddings of all nodes, $s_{ij} = h_i^T h_j$ and use a diffusion loss inspired by Liu et al.[89] to train the model parameters with gradient backpropagation,

$$L(s_{ij}) = -\frac{\alpha}{2}(s_{ij} - \beta)^2,$$

where $\alpha > 0$, and β is a predefined threshold and set to 0.25 in this work.

The idea of the loss function is to cluster embeddings of relevant nodes in the multiscale network and separate those irrelevant embeddings conversely. This effect can be observed from the derivative of L:

$$\frac{\partial L(s_{ij})}{\partial s_{ij}} = -\alpha(s_{ij} - \beta).$$

Therefore similarity values larger than threshold β are encouraged and embeddings of relevant nodes will cluster. Conversely, the loss function also diverges the embeddings of irrelevant nodes in the network even further away. Finally, we train the model for fixed 20 epochs and obtain the updated embeddings h_i for each node. The final embeddings were then saved for further computation.

Lastly, the proximity values (i.e. cosine similarities) from COVID-19 to each drug and each human protein present in the multiscale interactome were calculated and ranked to identify direct repurposing candidates and plausible targets (see section "Compound Selection"). Source code for generating the proximity values is made available at https://github.com/bowang-lab/gcn-drug-repurposing.

Drug-target interaction predictions. To identify small molecule drug candidates for host-based targets proposed by the GCN with few or no known binders, we looked up Drug Target Interaction (DTI) predictions in PolypharmDB[37]. PolypharmDB is a drug-repurposing database of pre-computed drug-target interaction predictions, evaluated by using the 2020Q2 beta release of MatchMaker[90]. Matchmaker is a deep learning model that predicts interaction of a small molecule drug/binding site pair using paired structural features of the drug with the 3D structural features of the protein binding sites. MatchMaker models with positive training examples complexes obtained by threading drug-target interaction (DTI) data[91] onto 3D structures of protein–ligand binding sites obtained from the Protein DataBank[92] and SwissModel[93], on the basis of chemical similarity[90]. Negative training examples are obtained by random shuffling of positive DTI pairs, and models are trained on progressive thresholds of increasing stringency in accordance to the confidence in the source DTI and in the selection of representative 3D structures[94].

The compound screening library consists of 10,244 small molecule drugs, retrieved from DrugBank[95] on February 19th, 2020. The library excluded compounds with fewer than four carbon atoms, or whose SMILES chemical structure was unable to be parsed by RD-KIT (Release_2018_09_1) RDKit: Open-source cheminformatics; http://www.rdkit.org)[95]. The screening set includes 2118 approved drugs, 2242 drugs in clinical trials and 5547 molecules in preclinical development or nutraceuticals.

The protein screening library represents 29,290 pockets from 8525 human proteins obtained by the PDB[92] (retrieved January 2018) and SwissModel[93] database (Retrieved June 2018). Specifically, pockets correspond to binding sites of drug-like molecules observed as cocrystalized ligands in the PDB source files, or superimposed ligands from template structures from SwissModel source files. Drug-target interaction pairs were evaluated using all combinations of small molecule drugs and available binding site structures. Individual human proteins were ranked on the basis of their top-scoring pockets. Additional details related to the construction of PolypharmDB and screening libraries are available from Redka et al. 2020[37].

Compound selection. Small molecule repurposing candidates were selected from three methods combining the GCN network and PolypharmDB (Fig. 1) approaches. For each method, a systematic selection criterion was applied, as described below, to maximize the diversity of assayed compounds and to ensure the relevance of the hit to the infectivity-based assays, its availability, and repurposing appropriateness. Selected drugs for all three methods are provided in Table 1.

For *Method 1*, a target-centric drug repurposing approach, compounds were selected from the top 60 drug candidates suggested directly by the GCN, on the basis of their network distance to COVID-19. Since the GCN input multiscale interactome contained a broad variety of drugs, the following criteria were applied to reduce the number of candidates from 60 to 10 compounds:

1. Non-small molecule drugs, such as recombinant proteins were eliminated (e.g., peginterferon alfa-2b);
2. Only one candidate was chosen if multiple drugs shared the same protein in the final node of the path to COVID-19 (e.g., only top scoring flucytosine was chosen from a total of 4 candidates connected to DNMT1);
3. Candidates that were a part of networks involved with the cytokine storm, adaptive immunity, or lymphocyte response were not selected (e.g., chloramphenicol connected to CD55, CD4-positive, alpha–beta T cell cytokine production node was not selected);
4. Natural amino acids (e.g., L-asparagine) and non-drug like compounds (e.g., urea and calcium chloride) were not selected;
5. Only FDA approved candidates were selected.

Once 10 compounds were selected in the order of their network distance to COVID-19, the selection process was completed corresponding to the last compound having a rank of 40.

Methods 2 and *3*, drug-centric repurposing approaches, combined the GCN for viral-host target selection and PolypharmDB to predict small molecule binders for the proposed targets. Target selection was performed by ranking human proteins in accordance with their GCN proximity scores to COVID-19. The top 100 human proteins of 17,660 represented in the network were considered in the selection process. From those 100 proteins, 66 were eliminated because they were missing from the proteome included in PolypharmDB. Out of the remaining 34 proteins, 11 proteins were eliminated because the descriptions of their functions on Uniprot or

GeneCards contained terms associated with the stages of the viral infection that were outside the scope of the infectivity-based assay in MRC-5 human lung fibroblast cells used in this study: lymphocyte response, cytokine storm, natural killer cell cytotoxicity, immune synapse formation, T-cell or B-cell response. The remaining 23 proteins were manually curated to prioritize 10 proteins that were most likely to be associated with the different stages of the viral cycle, such as attachment and entry, translation, replication, assembly, or release, based either on the description of their function or previous reports. Those selected ten proteins were as follows: UGGT2, SDF2, NLRX1, MOGS, HEPACAM, IRAK4, ADAM15, CD46, LILRA3, and CHPF2. Another set of 5 proteins out of 23 proteins had functions with a potential of being involved in the viral infection of lung fibroblasts, and therefore those proteins were reserved as a supplemental list of targets if insufficient predicted binders were found for the primary list of 10 targets. That supplemental list of targets was as follows: TARS2, GOLGA3, MDN1, THUMPD2, ZBTB37. The remaining 8 of 23 proteins were removed from consideration.

Following the selection of the primary list of 10 targets, *Methods 2* and *3* diverge in their compound selection process. For *Method 2*, the top ten ranking FDA-approved small-molecule repurposing candidates for each protein (i.e. 10 single-target panels; a total of 100 candidates) were filtered to 13 compounds, as described below. For Method 3, top 25 ranking FDA-approved small-molecule candidates for a single multi-target polypharmacological panel that included all 10 proteins (i.e., a single 10-targets panel) were analysed to yield 2 additional compounds that were distinct from those that were already selected by Method 2. The analysis of the latter was based on the weighted aggregate score of all targets. Since both lists of 100 and 25 top ranking compounds were already filtered to contain only small molecules and FDA-approved drugs, for both methods, the compounds were further selected by following these criteria:

1. The compound is predicted to be interacting with the target by MatchMaker, in addition to having a high rank relative to the proteome;
2. The compound is drug-like (i.e., natural compounds and endogenous ligands, natural amino acids, urea, metal chelators were excluded);
3. Tracer compounds or topical medications were not selected;
4. Duplication of the class of drug and its indication was minimized (e.g., only two antivirals out of 9 candidates, and only one sodium glucose co-transporter-2 inhibitor out of 3 candidates in the list of top 100 by Method 1 were selected);
5. The number of targets from the primary list of 10 represented in the final selection was maximized (i.e., select one candidate per target screen first, but if some targets did not result in any compounds with a significant probability of interaction predicted by MatchMaker, then select additional compounds from the supplemental list of targets; as a result, some targets are represented 2–3 times in Table 1).

Following the criteria described above, 13 compounds out of 100 were selected by Method 2, and 9 compounds out of 25 were selected by Method 3. However, 7 of the repurposing candidates were common to compounds from Method 2, resulting in only 2 unique compounds (see Table 1). An additional compound scored highly for the supplemental targets of interest and was nominated to be included in the testing panel given its predicted performance and the distinct target mechanism (Table 1).

Cell culture and human coronavirus propagation. For propagation of cell lines for use in coronavirus infection experiments, cells were maintained in a standard tissue culture incubator maintained at 37 °C with 5% CO_2. For infection of cells and propagation of coronaviruses, cells were maintained in a standard cell culture incubator maintained at 33 °C with 5% CO_2. This temperature was determined to support optimal propagation of human coronaviruses and yielded higher viral titers than preparations of the virus grown at 37 °C.

MRC-5 and LLC-MK2 cell lines were maintained in *growth medium*, consisting of Minimum Essential Medium Eagle (MEM), with Earle's salts, L-glutamine and sodium bicarbonate (Sigma Aldrich Cat. No. M4655) and supplemented with 10% heat-inactivated fetal bovine serum (FBS, ThermoFisher Scientific Cat. No. 10082147) and 1X penicillin–streptomycin (P/S, ThermoFisher Scientific Cat. No. 15070063). For all experiments involving coronavirus infection, cells were maintained in *infection medium*, which is the same as *growth medium*, except the concentration of heat-inactivated FBS was 2%. For plaque assays, cells monolayers were overlaid with *plaque media*, consisting of 2 × MEM Temin's modification, no phenol red (ThermoFisher Scientific Cat. No. 11-935-046), 2% heat-inactivated FBS, 1 × P/S, and mixed 1:1 with pre-warmed 0.6% agarose (ThermoFisher Scientific Cat. No. 16500-500). Alternatively, see Table S1 for specific formulations of culture media.

Propagation of human coronaviruses was performed based on suppliers' instructions. Briefly, MRC-5 or LLC-MK2 cells were grown to 90% confluence in a T75 tissue culture flask in standard growth medium. Cell monolayers were washed 2 × with infection medium prior to infection. Cell monolayers were infected with 229E (MRC-5 cells), OC43 (LLC-MK2), or NL63 (LLC-MK2) at a MOI of 0.01 in a total volume of 4 mL infection medium for 2 h adsorption at 33 °C with 5% CO_2. Following viral adsorption, unbound virus was aspirated, and cell monolayers were washed 2 × with infection medium, and then 12 mL fresh infection medium was added. Flasks were then placed back in the 33 °C incubator for a period of 2–4 days depending on the coronavirus strain to achieve maximal viral titer. To harvest the virus, supernatant was collected in a 15 mL conical tube and centrifuged at 1000 × g for 10 min to pellet cell debris. Viral stocks were stored as single use aliquots at − 80 °C. Viral concentration of new preparations of viral stocks were measured by PFU assay.

Immunofluorescence detection of 229E infection. Briefly, MRC-5 human lung fibroblast cells (seeded on glass coverslips) were infected with 229E at a MOI of 0.01 in the presence of drug (1–10 μM) or vehicle control for 1 h at 33 °C and 5% CO_2. Following incubation, excess unbound virus was removed, and

cells were incubated in fresh infection medium with drugs for an additional 48 h. This time point was chosen because it was initially determined to result in infection of ~ 100% of control cells and thus provided a baseline for examination of antiviral activity. This time point was also chosen because the infected cells remained in a virtually intact monolayer with minimal CPE—extensive CPE could lead to sampling errors and could be confounded with potential cytotoxic effects of the drugs. Drugs that caused signs of cytotoxicity (identified by DAPI staining) were excluded from further analysis as they confounded interpretation of potential antiviral effects in the IF assay.

For the IF protocol, cells infected with coronaviruses were washed 2 × with PBS (with Mg^{2+} and Ca^{2+}) and then immediately fixed for 1 h with 4% paraformaldehyde. This was followed by 15 min treatments with: 0.15% glycine, 0.1% triton X-100, and 3% bovine serum albumin with PBS washes in between each step. 229E S protein was detected with treating the cells with 100 µL (1:50 dilution) of the mouse anti-229E S protein antibody 9.8E12[97] by the inverted drop technique for 1 h at room temperature. Following primary antibody incubation, cells were washed and treated with AlexaFluor488-conjugated anti-mouse secondary antibody (1:1000 dilution) (Cedarlane Labs Cat. No. 115–545-003) and DAPI (1 µg/mL) for 1 h at room temperature. Following another wash, coverslips were mounted on glass slides with DAKO mounting media (Agilent Technologies Cat. No. S3023) and then incubated at room temperature overnight to solidify. Slides were visualized on an inverted microscope by widefield epifluorescence (Leica DMi8 microscope, Andor Zyla 4.2-megapixel camera, run by Quorum WaveFX by Metamorph software). For each IF experiment, a total of 10–20 randomly chosen fields were selected in the DAPI channel for acquisition of the 229E S protein with a 10 × objective lens. Images were quantified by measuring the total fluorescence signal using ImageJ software (National Institutes of Health, Bethesda, MD)[98].

Human coronavirus plaque assays. All coronavirus plaque assay protocols were adapted from previously described methods for measuring viral concentration using PFU assays[99–101]. Briefly, cells were grown to confluency on 6-well tissue culture plates and then infected with serially diluted virus in a volume of 300 µL for 1 h adsorption at 33 °C with 5% CO_2, with gentle agitation every 15 min. Following adsorption, unbound virus was removed, and cells were washed 2 × with infection media and then overlaid with plaque media. In experiments involving the testing of potential antiviral compounds, drugs were added to the infection media during adsorption and to the plaque media. Following an incubation period of several days to establish CPE (see below for virus-specific information), cells were fixed with 10% neutral buffered formalin overnight, followed by removal of the agarose plug and counterstaining with 1% crystal violet solution to visualize the plaques. For all plaque assays, each condition was performed in technical triplicates.

229E: Confluent monolayers of MRC-5 cells were grown on 6-well tissue culture plates and infected with serially diluted 229E virus. After washing away unbound virus, the cells were overlaid with a semi-solid agarose medium to restrict the spread of virus to adjacent cells. After 5–7 days incubation, cells were fixed and stained to quantify the number of plaques in each well.

NL63: NL63 viruses were propagated in the monkey kidney epithelial cell line LLC-MK2, which has been reported to support NL63 production[102]. Confluent monolayers of LLC-MK2 cells were grown on 6-well plates and infected with serially diluted NL63 virus. After washing away unbound virus, the cells were overlaid with a semi-solid agarose medium to restrict the spread of virus to adjacent cells. After 4 days incubation, cells were fixed and stained to quantify the number of plaques in each well.

OC43: We determined that OC43 viruses could be readily propagated in LLC-MK2 cells, similar to NL63 viruses. Confluent monolayers of LLC-MK2 cells were grown on 6-well plates and infected with serially diluted OC43 virus. After washing away unbound virus, the cells were overlaid with a semi-solid agarose medium to restrict the spread of virus to adjacent cells. After 4 days incubation, cells were fixed and stained to quantify the number of plaques in each well.

qRT-PCR detection of NL63. LLC-MK2 cells seeded on 6-well tissue culture plates were infected with NL63 at a MOI of 0.01 for a 1 h adsorption period in the presence of 10 µM capmatinib in a volume of 300 µL. Following viral adsorption, cells were washed 2 × with infection media and then 500 µL fresh media/drug was added to the cells for a 3-day incubation period. To extract total NL63 RNA, 1 mL TRIzol™ Reagent (ThermoFisher Scientific) was added directly to the samples (cells + supernatant) and cells were lysed by scraping. Total RNA was purified using the Direct-zol™ RNA Miniprep Kits (Zymo Research, Irvine, CA) according to manufacturer's instructions. Reverse transcription and qPCR were performed in a one-step reaction using Luna Universal One-Step RT-qPCR (NEB, Ipswich MA) according to manufacturer's instructions. RT-qPCR reactions were performed on a CFX96 thermal cycler (Bio-Rad, Mississauga, ON). Experiments were performed with at least 2 technical replicates to monitor variation between wells, no template/no RT controls, and melt curves. Reactions were performed at a final volume of 20 µL with 20 ng input RNA and a primer concentration of 500 nM. All results were normalized to GAPDH RNA levels and to the infected, non-drug treated condition. Relative change in NL63 RNA was calculated using the $2^{-\Delta\Delta Ct}$ method[103]. A minimum of 3 experimental replicates were used to assess NL63 Nucleocapsid RNA using previous established primers[104]. For a complete list of qPCR primers see Table S2. For thermal cycler conditions see Table S3.

siRNA transfection. siRNA transfection was performed by two individual transfections at 72 h and 48 h prior to infection or other experimental manipulation, respectively. Briefly, MRC-5 cells were grown to 40% confluency before the first round of transfection. On the day of transfection, cells were placed in fresh growth medium. A transfection master mix was made in Opti-MEM (ThermoFisher Cat. No. 31985070) by diluting stock siRNA solutions to a final concentration of 50 nM and 6.3 µL Lipofectamine RNAiMAX transfection

reagent (ThermoFisher Cat. No. 13778030) per well of a 6-well tissue culture dish. Cells were transfected for 4 h before switching to fresh growth medium. For a list of siRNA sequences see Table S2.

Pseudotyped virus particle assay. SARS-CoV-2 pseudotyped viruses (PsV) were prepared using an HIV-based lentiviral system as previously described[11] with few modifications. Briefly, PsVs were produced by transfection of human kidney HEK293T cells with the full-length SARS-CoV-2 spike (BEI NR52310) or SARS-CoV-1 spike (kindly provided by S. Pöhlmann, Leibniz Institute for Primate Research, Göttingen, Germany). Cells were co-transfected with a lentiviral backbone encoding the luciferase reporter gene (BEI NR52516), a plasmid expressing the Spike (BEI NR52310) and plasmids encoding the HIV structural and regulatory proteins Tat (BEI NR52518), Gag-pol (BEI NR52517) and Rev (BEI NR52519). After 24 h at 37 °C, 5 mM sodium butyrate was added to the media and the cells were incubated for an additional 24–30 h at 30 °C. Next, the PsV particles were harvested, passed through 0.45 µm pore sterile filters and finally concentrated using a 100 K Amicon (Merck Millipore Amicon-Ultra 2.0 Centrifugal Filter Units).

Neutralization was determined in a single-cycle neutralization assay using HeLa-ACE2 cells (kindly provided by D.R. Burton; The Scripps Research Institute). To that end, 50 µL of twofold serial dilutions of the small molecules were incubated with 10,000 cells/well seeded the day before (100 µL/well) for 1 h at 37 °C. After 1 h incubation, 50 µL of PsVs was added to each well and incubated for 48 h-60 h in the presence of 10 µg/mL of polybrene (Sigma Aldrich, TR-1003-G). Infection levels were inferred from the amount of luminescence in relative light units (RLUs) after adding 50 µL Britelite plus reagent (PerkinElmer) to 50 µL of media containing cell (i.e. after removing 130 µL/well to account for evaporation). After 2 min incubation, the volume was transferred to a 96-well white plate (Sigma-Aldrich) and the luciferase intensity was read using a Synergy Neo2 Multi-Mode Assay Microplate Reader (Biotek Instruments). Two to three biological replicates with two technical replicates each were performed. Culture media was prepared by supplementing DMEM media with 2% inactivated FBS and 50 µg/ml of gentamicin. IC$_{50}$ values were calculated using Prism.

In order to confirm that the reduced infection was not related to cell toxicity, HeLa-ACE2 cell viability upon incubation with serial dilutions of the small molecules was assessed. 10,000 cells/ well of pre-seeded HeLa-ACE2 cells were co-cultured with twofold serial dilutions of the small molecules at 37 °C for 48 h-60 h under the same conditions as in the neutralization assay. Cell viability was monitored by adding 50 µL of CellTiter-Glo 2.0 reagent (Promega) to 200 µL of media containing cells. After 10 min incubation, 100 µL volume was transferred to a 96-well black plate (Sigma-Aldrich) to measure luminescence in relative light units (RLUs) using a Synergy Neo2 Multi-Mode Assay Microplate Reader (Biotek Instruments).

Drug library. Unless indicated otherwise, all compounds were purchased from MedChemExpress (Monmouth Junction, NJ) and were reconstituted according to manufacturer's specifications. A complete list of compounds is provided in the Table S4. All compounds were reconstituted to stock concentrations of 1–10 mM and frozen in individual aliquots at − 80 °C until use.

Statistical analysis. All statistical analysis for biological experiments were performed with GraphPad Prism 9 software using student t-tests when comparing two conditions (Figs. 3B,D,E, 4B,E,F,G, S2A) or one-way ANOVA with Tukey post-hoc test when comparing multiple conditions (Figs. 3A,C, 4A,C, S4).

Data availability
Datasets generated during and/or analysed during the current study are available from the corresponding authors on reasonable request, including the ranks and scores of compounds screened using the commercial PolyPharmDB for all targets of interest identified within the manuscript, licensed from Cyclica Inc.

References
1. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet. Infect. Dis.* **20**, 533–534 (2020).
2. Yin, Y. & Wunderink, R. G. MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology* **23**, 130–137 (2018).
3. De Wit, E., Van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016).
4. Gagneur, A. *et al.* Coronavirus-related nosocomial viral respiratory infections in a neonatal and paediatric intensive care unit: a prospective study. *J. Hosp. Infect.* **51**, 59–64 (2002).
5. da Veiga, A. B. G. *et al.* More than just a common cold: endemic coronaviruses OC43, HKU1, NL63, and 229E associated with severe acute respiratory infection and fatality cases among healthy adults. *J. Med. Virol.* **93**, 1002–1007 (2021).
6. Fontanet, A. *et al.* SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet* **397**, 952–954 (2021).
7. Abdool, K. S. S. & de Oliveira, T. New SARS-CoV-2 variants-clinical, public health, and vaccine implications. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMc2100362 (2021).
8. Wu, K., Li, W., Peng, G. & Li, F. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19970–19974 (2009).
9. Zhang, H., Penninger, J. M., Li, Y., Zhong, N. & Slutsky, A. S. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Med.* https://doi.org/10.1007/s00134-020-05985-9 (2020).
10. Yang, J. *et al.* Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nat. Commun.* **11**, 1–10 (2020).
11. Hoffmann, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280. https://doi.org/10.1016/j.cell.2020.02.052 (2020).
12. Yeager, C. L. *et al.* Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357**, 420–422 (1992).

13. Hulswit, R. J. G. *et al.* Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 2681–2690 (2019).

14. Daly, J. L. *et al.* Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science (80-.).* **370**, 861–865 (2020).

15. Cantuti-Castelvetri, L. *et al.* Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science (80-.).* **370**, 856–860 (2020).

16. Bayati, A., Kumar, R., Francis, V. & McPherson, P. S. SARS-CoV-2 infects cells following viral entry via clathrin-mediated endocytosis. *J. Biol. Chem.* **296**, 100306 (2021).

17. Inoue, Y. *et al.* Clathrin-dependent entry of severe acute respiratory syndrome coronavirus into target cells expressing ACE2 with the cytoplasmic tail deleted. *J. Virol.* **81**, 8722–8729 (2007).

18. Wang, H. *et al.* SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Res.* **18**, 290–301 (2008).

19. Wolff, G. *et al.* A molecular pore spans the double membrane of the coronavirus replication organelle. *Science (80-.).* **369**, 1395–1398 (2020).

20. Ghosh, S. *et al.* β-coronaviruses use lysosomes for egress instead of the biosynthetic secretory pathway. *Cell* **183**, 1520-1535.e14 (2020).

21. Ziebuhr, J. The coronavirus replicase. *Curr. Top. Microbiol. Immunol.* **287**, 57–94 (2005).

22. Sawicki, S. G., Sawicki, D. L. & Siddell, S. G. A contemporary view of coronavirus transcription. *J. Virol.* **81**, 20–29 (2007).

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

24. Gordon, D. E. *et al.* Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science (80-.).* **370**, 10 (2020).

25. Wang, R. *et al.* Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell* **184**, 106-119.e14 (2021).

26. Schneider, W. M. *et al.* Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* **184**, 120-132.e14 (2021).

27. Lukowski, S. W. *et al.* Integrated analysis of mRNA and miRNA expression in response to interleukin-6 in hepatocytes. *Genomics* **106**, 107–115 (2015).

28. Chong, C. R. & Sullivan, D. J. New uses for old drugs. *Nature* **448**, 645–646 (2007).

29. Hernandez, J. J. *et al.* Giving drugs a second chance: overcoming regulatory and financial hurdles in repurposing approved drugs as cancer therapeutics. *Front. Oncol.* **7**, 273 (2017).

30. Parisi, D. *et al.* Drug repositioning or target repositioning: a structural perspective of drug-target-indication relationship for available repurposed drugs. *Comput. Struct. Biotechnol. J.* **18**, 1043–1055 (2020).

31. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2016).

32. Edwards, A. What are the odds of finding a COVID-19 drug from a lab repurposing screen?. *J. Chem. Inf. Model.* **60**, 5727–5729 (2020).

33. Park, K. A review of computational drug repurposing. *Transl. Clin. Pharmacol.* **27**, 59 (2019).

34. Ruiz, C., Zitnik, M. & Leskovec, J. Identification of disease treatment mechanisms through the multiscale interactome. *Nat. Commun.* **12**, 1–15 (2021).

35. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. In: *5th International Conference on Learning Representations ICLR 2017-Conference Track Proceedings* (2016).

36. Schlichtkrull, M. *et al.* Modeling Relational Data with Graph Convolutional Networks. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **10843 LNCS**, 593–607 (2017).

37. Redka, D. S. *et al.* PolypharmDB, a deep learning-based resource, quickly identifies repurposed drug candidates for COVID-19. *ChemRxiv* https://doi.org/10.26434/chemrxiv.12071271.v1 (2020).

38. Baltschukat, S. *et al.* Capmatinib (INC280) is active against models of non–small cell lung cancer and other cancer types with defined mechanisms of MET activation. *Clin. Cancer Res.* **25**, 3164–3175 (2019).

39. Hughes, P. E. *et al.* In vitro and in vivo activity of AMG 337, a potent and selective MET kinase inhibitor, in MET-dependent cancer models. *Mol. Cancer Ther.* **15**, 1568–1579 (2016).

40. Liu, X. *et al.* A novel kinase inhibitor, INCB28060, blocks c-MET-dependent signaling, neoplastic activities, and cross-talk with EGFR and HER-3. *Clin. Cancer Res.* **17**, 7127–7138 (2011).

41. Kucharska, I. *et al.* Title: Multivalency transforms SARS-CoV-2 antibodies into broad and ultrapotent 1 neutralizers 2 3 4 Edurne Rujas. *bioRxiv* https://doi.org/10.1101/2020.10.15.341636 (2020).

42. Del Valle, D. M. *et al.* An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat. Med.* **26**, 1636–1643 (2020).

43. Solun, B. & Shoenfeld, Y. Inhibition of metalloproteinases in therapy for severe lung injury due to COVID-19. *Med. Drug Discov.* **7**, 100052 (2020).

44. Balka, K. R. & Nardo, D. Understanding early TLR signaling through the Myddosome. *J. Leukoc. Biol.* **105**, 339–351 (2019).

45. Leifer, C. A. & Medvedev, A. E. Molecular mechanisms of regulation of Toll-like receptor signaling. *J. Leukoc. Biol.* **100**, 927–941 (2016).

46. Suzuki, N., Suzuki, S. & Yeh, W. C. IRAK-4 as the central TIR signaling mediator in innate immunity. *Trends Immunol.* **23**, 503–506 (2002).

47. Song, K. W. *et al.* The kinase activities of interleukin-1 receptor associated kinase (IRAK)-1 and 4 are redundant in the control of inflammatory cytokine expression in human cells. *Mol. Immunol.* **46**, 1458–1466 (2009).

48. Wang, L. *et al.* Crystal structure of human IRAK1. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 13507–13512 (2017).

49. Parisi, D. *et al.* Drug repositioning or target repositioning: a structural perspective of drug-target-indication relationship for available repurposed drugs. *Comput. Struct. Biotechnol. J.* **18**, 1043–1055 (2020).

50. Lin, A. *et al.* Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci. Transl. Med.* **11**, eaaw8412 (2019).

51. Galindez, G. *et al.* Lessons from the COVID-19 pandemic for advancing computational drug repurposing strategies. *Nat. Comput. Sci.* **1**, 33–41 (2021).

52. Singh, N. & Villoutreix, B. O. Resources and computational strategies to advance small molecule SARS-CoV-2 discovery: lessons from the pandemic and preparing for future health crises. *Comput. Struct. Biotechnol. J.* **19**, 2537–2548 (2021).

53. Kuleshov, M. V. *et al.* The COVID-19 gene and drug set library. *SSRN* https://doi.org/10.21203/RS.3.RS-28582/V1 (2020).

54. Ellinger, B. *et al.* A SARS-CoV-2 cytopathicity dataset generated by high-content screening of a large drug repurposing collection. *Sci. data* **8**, 70 (2021).

55. Tsegay, K. *et al.* A repurposed drug screen identifies compounds that inhibit the binding of the COVID-19 spike protein to ACE2. *Front. Pharmacol.* **12**, 103 (2021).

56. Jeon, S. *et al.* Identification of antiviral drug candidates against SARS-CoV-2 from FDA-approved drugs. *Antimicrob. Agents Chemother.* **64**, 1039 (2020).

57. Liu, X. *et al.* A novel kinase inhibitor, INCB28060, blocks c-MET-dependent signaling, neoplastic activities, and cross-talk with EGFR and HER-3. *Clin. Cancer Res.* **17**, 7127–7138 (2011).

58. De Mello, R. A. *et al.* The role of MET inhibitor therapies in the treatment of advanced non-small cell lung cancer. *J. Clin. Med.* **9**, 1918 (2020).

59. Puccini, A. *et al.* Safety and tolerability of c-MET inhibitors in cancer. *Drug Saf.* **42**, 211–233 (2019).
60. Ko, M., Jeon, S., Ryu, W. S. & Kim, S. Comparative analysis of antiviral efficacy of FDA-approved drugs against SARS-CoV-2 in human lung cells. *J. Med. Virol.* **93**, 1403–1408 (2021).
61. Liang, H., Zhao, L., Gong, X., Hu, M. & Wang, H. Virtual screening FDA approved drugs against multiple targets of SARS-CoV-2. *Clin. Transl. Sci.* **14**, 1123–1132. https://doi.org/10.1111/cts.13007 (2021).
62. Parveen, S. & Alnoman, R. B. Potential exploration of recent FDA-approved anticancer drugs against models of SARS-CoV-2's main protease and spike glycoprotein: a computational study. *Biointerface Res. Appl. Chem.* **11**, 10059–10073 (2021).
63. Al-Rashedi, N. A. M., Munahi, M. G. & Alobaidi, L. A. H. Prediction of potential inhibitors against SARS-CoV-2 endoribonuclease: RNA immunity sensing. *J. Biomol. Struct. Dyn.* https://doi.org/10.1080/07391102.2020.1863265 (2020).
64. Gobin, K. *et al.* IRAK4 deficiency in a patient with recurrent pneumococcal infections: case report and review of the literature. *Front. Pediatr.* **5**, 83 (2017).
65. Von Bernuth, H. *et al.* Pyogenic bacterial infections in humans with MyD88 deficiency. *Science (80-.).* **321**, 691–696 (2008).
66. Jia, A. *et al.* Clinical IRAK4 deficiency caused by homozygosity for the novel IRAK4 (c.1049delG, p.Gly350Glufs*15) variant. *Cold Spring Harb. Mol. Case Stud.* **6**, 1036 (2020).
67. Picard, C. *et al.* Clinical features and outcome of patients with IRAK-4 and MyD88 deficiency. *Medicine* **89**, 403–425 (2010).
68. Reynolds, L. A. *et al.* MyD88 signaling inhibits protective immunity to the gastrointestinal helminth parasite heligmosomoides polygyrus. *J. Immunol.* **193**, 2984–2993 (2014).
69. Sheahan, T. *et al.* MyD88 is required for protection from lethal infection with a mouse-adapted SARS-CoV. *PLoS Pathog.* **4**, e1000240 (2008).
70. Lukassen, S. *et al.* SARS -CoV-2 receptor ACE 2 and TMPRSS 2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* **39**, e105114 (2020).
71. Song, K. W. *et al.* The kinase activities of interleukin-1 receptor associated kinase (IRAK)-1 and 4 are redundant in the control of inflammatory cytokine expression in human cells. *Mol. Immunol.* **46**, 1458–1466 (2009).
72. Liu, S. *et al.* Potential antiviral target for SARS-CoV-2: a key early responsive kinase during viral entry. *CCS Chem.* **3**, 559–568 (2021).
73. Onodi, F. *et al.* SARS-CoV-2 induces human plasmacytoid predendritic cell diversification via UNC93B and IRAK4. *J. Exp. Med.* **218**, 1002 (2021).
74. IRAK4 Inhibition in Treatment of COVID-19 With ARDS (I-RAMIC)-full text view-ClinicalTrials.gov. Available at: https://clinicaltrials.gov/ct2/show/NCT04575610. Accessed 5th April 2021.
75. Cuevas, A. M., Clark, J. M. & Potter, J. J. Increased TLR/MyD88 signaling in patients with obesity: is there a link to COVID-19 disease severity?. *Int. J. Obes.* https://doi.org/10.1038/s41366-021-00768-8 (2021).
76. Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).
77. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
78. Corsello, S. M. *et al.* The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.* **23**, 405–408 (2017).
79. Carbon, S. *et al.* The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
80. Oughtred, R. *et al.* The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, D529–D541 (2019).
81. Menche, J. *et al.* Uncovering disease-disease relationships through the incomplete interactome. *Science (80-.).* **347**, 1257601–1257601 (2015).
82. Venkatesan, K. *et al.* An empirical framework for binary interactome mapping. *Nat. Methods* **6**, 83–90 (2009).
83. Rual, J. F. *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173–1178 (2005).
84. Yu, H. *et al.* Next-generation sequencing to generate interactome datasets. *Nat. Methods* **8**, 478–480 (2011).
85. Rolland, T. *et al.* A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226 (2014).
86. Salwinski, L. *et al.* The database of interacting proteins: 2004 update. *Nucleic Acids Res.* **32**, D449–D451 (2004).
87. Grover, A. & Leskovec, J. Node2vec: Scalable feature learning for networks. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 13–17-August-2016, 855–864 (Association for Computing Machinery, 2016).
88. Yue, X. *et al.* Graph embedding on biomedical networks: methods, applications and evaluations. *Bioinformatics* **36**, 1241–1251 (2019).
89. Liu, C. *et al.* Guided Similarity Separation for Image Retrieval. (2019).
90. WO2020140156A1-Method and system for predicting drug binding using synthetic data - Google Patents. https://patents.google.com/patent/WO2020140156A1/en. Accessed 7th April 2021.
91. Szklarczyk, D. *et al.* STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* **44**, D380–D384 (2016).
92. Berman, H. M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
93. Bienert, S. *et al.* The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res.* **45**, D313–D319 (2017).
94. Tonekaboni, S. A. M. *et al.* Learning across label confidence distributions using Filtered Transfer Learning. In *Proceedings of 19th IEEE International Conference Machine Learning with Applications ICMLA 2020* 1117–1123 (2020).
95. Wishart, D. S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **34**, 2096 (2006).
96. Landrum, G. *et al.* rdkit/rdkit: 2018_09_1 (Q3 2018) Release. (2018). https://doi.org/10.5281/ZENODO.1468109
97. Wong, A. H. M. *et al.* Receptor-binding loops in alphacoronavirus adaptation and evolution. *Nat. Commun.* **8**, 1–10 (2017).
98. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
99. Poppe, M. *et al.* The NF-κB-dependent and -independent transcriptome and chromatin landscapes of human coronavirus 229E-infected cells. *PLoS Pathog.* **13**, 1069 (2017).
100. Herzog, P., Drosten, C. & Müller, M. A. Plaque assay for human coronavirus NL63 using human colon carcinoma cells. *Virol. J.* **5**, 569 (2008).
101. Baer, A. & Kehn-Hall, K. Viral concentration determination through plaque assays: Using traditional and novel overlay systems. *J. Vis. Exp.* https://doi.org/10.3791/52065 (2014).
102. Milewska, A. *et al.* Entry of Human Coronavirus NL63 into the Cell. *J. Virol.* **92**, 10239 (2017).
103. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, 1039 (2001).
104. Wan, Z. *et al.* A melting curve-based multiplex RT-qPCR assay for simultaneous detection of four human coronaviruses. *Int. J. Mol. Sci.* **17**, 10000 (2016).

Acknowledgements

Author contributions

The study was conceived or designed by M.G.S., H.C., D.S.R., M.K., H.M., S.K., J.B.M., R.D.V., A.H., R.J.B., R.K., S.A.S., G.D.F., S.A.M.T., A.W., J.P.J., V.S., S.S.M., B.W. and C.N.A. The data was acquired and/or analyzed by M.G.S., H.C., D.S.R., M.K., E.R., H.M., S.H., K.C., R.M., S.A.S., V.S., S.S.M., B.W., and C.N.A. The new software used in this work was created by H.C., D.S.R., M.K., H.M., S.K., S.A.M.T., A.W., V.S., S.S.M. and B.W. The manuscript was prepared by M.G.S., H.C., D.S.R., E.R., R.D.V., J.P.J., V.S., S.S.M., B.W. and C.N.A.

Competing interests

DSR, SAMT, AW, VS, SSM are employees of Cyclica Inc, and may own stock in Cyclica Inc. SH is an employee of Bayer US LLC (a subsidiary of Bayer AG) and may own stock in Bayer AG. AH is founder and CEO of Phoenox Pharma and may own stocks in Phoenox Pharma. CNA, RK and RDV serve on the scientific advisory board of Phoenox Pharma. All other authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-02432-7.

Correspondence and requests for materials should be addressed to S.S.M., B.W. or C.N.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.