



Published anti-SARS-CoV-2 *in vitro* hits share common mechanisms of action that synergize with antivirals

Jing Xing , Shreya Paithankar, Ke Liu , Katie Uhl, Xiaopeng Li, Meehyun Ko, Seungtaek Kim, Jeremy Haskins and Bin Chen

Corresponding author: Bin Chen, Department of Pediatrics and Human Development, College of Human Medicine, Michigan State University, Grand Rapids, Michigan, 49503, USA. Tel.: +01-6162342819; E-mail: chenbi12@msu.edu

Abstract

The global efforts in the past year have led to the discovery of nearly 200 drug repurposing candidates for COVID-19. Gaining more insights into their mechanisms of action could facilitate a better understanding of infection and the development of therapeutics. Leveraging large-scale drug-induced gene expression profiles, we found 36% of the active compounds regulate genes related to cholesterol homeostasis and microtubule cytoskeleton organization. Following bioinformatics analyses revealed that the expression of these genes is associated with COVID-19 patient severity and has predictive power on anti-SARS-CoV-2 efficacy *in vitro*. Monensin, a top new compound that regulates these genes, was further confirmed as an inhibitor of SARS-CoV-2 replication in Vero-E6 cells. Interestingly, drugs co-targeting cholesterol homeostasis and microtubule cytoskeleton organization processes more likely present a synergistic effect with antivirals. Therefore, potential therapeutics could be centered around combinations of targeting these processes and viral proteins.

Key words: SARS-CoV-2; drug mechanisms; transcriptomics; drug combination

Jing Xing received a PhD in drug design from the University of the Chinese Academy of Sciences in 2018 and then pursued the postdoctoral training at Michigan State University, focusing on drug discovery from the deep-learning and system biology perspective.

Shreya Paithankar received a master's degree in medical and bioinformatics from Grand Valley State University in 2018. She is now a technical assistant at Dr. Bin Chen's lab, leading the efforts in RNA-Seq processing and integration.

Ke Liu received a BA in computer science from Shandong University in 2008 and a PhD in biology from Tsinghua University in 2014. He is now an assistant professor at Michigan State University.

Katie Uhl received a master's degree in cellular and molecular biology from Grand Valley State University in 2017. She is now a lab manager at Michigan State University College of Human Medicine.

Xiaopeng Li received a PhD in Physiology from Michigan State University, followed by postdoctoral training at the University of California San Francisco and the University of Iowa. He is now an associate professor at the Department of Pediatrics and Human Development at Michigan State University.

Meehyun Ko is a scientist at Dr. Seungtaek Kim's Lab at Institut Pasteur Korea. Her expertise is antiviral effect evaluation in cell models.

Seungtaek Kim received a PhD from Iowa State University in 2004 and then pursued the postdoctoral training at the University of Wisconsin-Madison and University of North Carolina at Chapel Hill. He is now the head of Zoonotic Virus Lab, Institut Pasteur Korea. His research interest is in the virology of new emerging viruses and therapeutics development.

Jeremy Haskins received a master's degree in radiation physics and radiation cancer biology from Colorado State University in 2017. He is now a research assistant at Michigan State University.

Bin Chen received a PhD in informatics from Indiana University, Bloomington and pursued the postdoctoral training at Stanford University. He then joined the University of California San Francisco as a faculty member in 2015. He is now an assistant professor at Michigan State University.

Introduction

As of 15 February 2021, SARS-CoV-2 has infected 108 million people and claimed 2 million lives. In addition, a growing number of researchers even speculate the society may have to endure the current circumstances implicated with SARS-CoV-2 for an extended period. Vaccines are promising for a cure, yet the emerging mutations of SARS-CoV-2 impose challenges to target the viral proteins; thus, targeting host cells remains a viable therapeutic approach. The global efforts in the trailing months have led to the discovery of at least 184 drug repurposing candidates *in vitro* (Supplementary Data 1). For example, Jeon *et al.* [1] screened a panel of 48 FDA-approved drugs against SARS-CoV-2, which were preselected by an assay of SARS-CoV-1, and found several candidates with very low IC_{50} values, including niclosamide and ciclesonide. Riva *et al.* [2] screened the ReFRAME library with approximately 12 000 bioactives and identified 21 candidates. Among all the repurposing hits, a few were verified to target host processes, such as the eIF4A inhibitor zotatifin [3] and the eEF1A inhibitor plitidepsin [4] that disrupt the cap-dependent mRNA translation of coronaviruses. However, the majority of the repurposing hits only exhibited moderate antiviral potential, and their mechanisms of action (MoA) remain largely unelucidated because of the haste to test candidates for clinical trials. The following disappointing results of chloroquine, ivermectin and others suggest that additional insights into their mechanisms may facilitate an enhanced understanding of infection and the development of better therapeutics for COVID-19.

Compared to the sparse drug-target matrix, transcriptomic profiles induced by individual perturbagens provide a full-matrix of drug-gene relations that could be explored to understand drug mechanisms. Furthermore, these profiles are widely used to evaluate the mimicking or reversal effect between drugs and diseases, under the assumption that drugs with similar MoA present similar effects on gene expression profiles, or effective drugs likely restore the gene expression dysregulated in the disease state [5–8]. Through integrative bioinformatics analyses of transcriptomic profiles and *in vitro* experiments, we found that those positive compounds regulate cholesterol homeostasis and microtubule cytoskeleton organization pathways, which also align with COVID-19 patients' severity, compound efficacy and synergism with antiviral drugs (Figure 1A). Therefore, we propose that potential therapeutics could be centered around combinations of targeting these processes and viral proteins.

Results

Identify anti-SARS-CoV-2 compound signature genes

Leveraging large-scale drug-induced gene expression profiles from the Library of Integrated Network-Based Cellular Signatures (LINCS) project [9], we computed the average expression change of 978 landmark genes after treatment of 7879 compounds individually across multiple cell lines. This collection contains 97 anti-SARS-CoV-2 positive compounds. Then we identified the differential gene expression induced by these positive compounds compared with the remaining 7782 samples. This resulted in 63 genes specifically dysregulated, including 30 downregulated and 33 upregulated (Supplementary Data 2). For example, NPC1, an intracellular cholesterol transporter, showed a much higher expression induced by anti-SARS-CoV-2 compounds than other compounds ($P = 2.2E-09$, Wilcox rank sums test, Figure 1B). Particularly, it is highly upregulated ($z\text{-score} \geq 2$) by 17 positive compounds, such as pevonedistat and cepharanthine. In addition, the expression change of NPC1 is

positively correlated with drug EC_{50}/IC_{50} (Spearman Correlation $Rho = -0.44$, $P = 8.2E-05$, Supplementary Data 2). An opposite pattern exists in CCNA2, a G1/S and G2/M transition regulator ($Rho = 0.50$, $P = 6.69E-11$, Figure 1B and Supplementary Data 2). Among all the 97 positive compounds, 80 of them regulate at least one of the 63 shared signature genes toward the proposed direction (Supplementary Data 3). To quantify the regulation effects of each compound on the 63 anti-SARS-CoV-2 signature genes, we added up the expression Z scores of the 33 upregulated genes, then subtracted the Z scores of the 30 downregulated genes. As shown in Figure 1C and Supplementary Data 3, for compounds with greatly changed (GC) profiles (more than 1% of the genes in the transcriptomic profile were dysregulated upon a specific compound treatment), the 63 signature genes were more affected by the positive compounds than other LINCS compounds. The separation of these two groups was clear and significant (median values are 39.54 and 4.21, respectively; $P = 9.35E-18$, Wilcox rank sums test). For the weakly changed (WC) profiles, although their effects on the 63 genes are generally minor, the difference is still significant ($P = 3.11E-07$, Wilcox rank sums test). Non-human-targeting drugs like saquinavir and amodiaquine might cause the weak change. These results suggest the 63 genes' expression change represents the host-targeting anti-SARS-CoV-2 compound activity.

The differential expression of the 63 genes might serve as a cellular response signature for anti-SARS-CoV-2 candidates. We then validated this signature with five independent genome-wide CRISPR screening datasets from [10–14]. These studies identified 152 host factor genes critical in SARS-CoV2-infection (Supplementary Data 4), termed pro-viral genes, among which 49 were mapped to the LINCS shRNA perturbagen collection. Assuming that knock-down of individual pro-viral genes might benefit the host cells to respond to the SARS-CoV-2 challenge, we computed the mimic effect between the anti-SARS-CoV-2 compound signature and the expression profiles perturbed by shRNA of each pro-viral gene through a gene-set enrichment analysis (adopted from RGENES) [7]. The RGENES values of 49 pro-viral genes knock-down are significantly higher than the remaining 4321 genes ($P = 7.38E-03$, Wilcox rank sums test, one-tailed, Figure 1D), suggesting inhibition of pro-viral genes has similar effects with the active compounds, and the anti-SARS-CoV-2 signature captures essential biological processes involved in viral infection.

Key pathways targeted by anti-SARS-CoV-2 compounds

By incorporating gene co-expression knowledge, we performed gene ontology (GO) enrichment analysis with a background beyond LINCS 978 genes. For the 33 upregulated genes, cholesterol biosynthetic and metabolic processes ('cholesterol homeostasis' [15] hereafter) were enriched, with genes NPC1, INSIG1 and HMGCS1 involved (Supplementary Figure S1A). For the 30 downregulated genes, the 'microtubule cytoskeleton organization' process was enriched, with genes DAG1, CCNB1, AURKA, PSRC1, STMN1, KIF20A, TUBB6 and MYBL2 annotated to this term (Supplementary Figure S1B). Although pathways related to the mitotic cell cycle were also significant, we did not further investigate them because of the biased distribution of the cancer-enriched genes in the LINCS data. We also investigated whether there were a few 'druggable' proteins that the anti-SARS-CoV-2 compounds commonly target on. The active compounds were annotated and mapped to the drug repurposing list including 1824 bioactives from the Broad Institute Drug Repurposing Hub (Supplementary Data 5). Then we calculated the enrichment

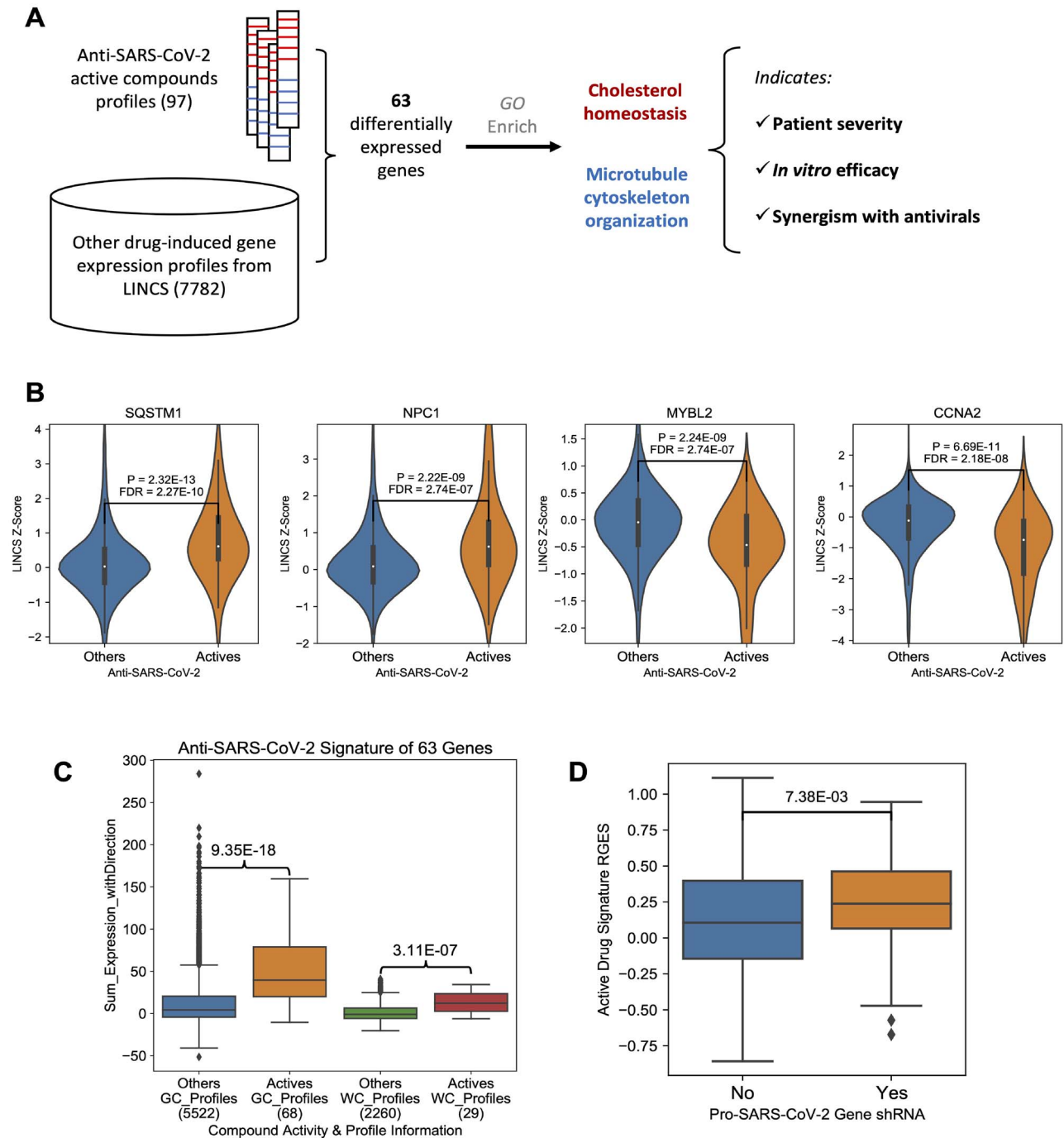


Figure 1. Anti-SARS-CoV-2 compound signature gene expression change. **A**, The workflow of this research. **B**, Example genes were induced (SQSTM1 and NPC1) or suppressed (MYBL2 and CCNA2) by anti-SARS-CoV-2 compounds. The y-axis indicates the LINCIS z-score of a specific compound and a higher score means higher expression change. P-values were derived from Wilcoxon rank sum tests, and further corrected across all LINCIS 978 genes. **C**, A boxplot showing the comparison of the effects on the 63 signature genes between active compounds and other compounds. The y-axis indicates an overall effect of a specific compound on the 63 genes, and a higher score means a better alignment with the up/downregulation pattern. On the x-axis, the number of gene expression profiles in each group is labeled under the group name. P-values were derived from Wilcoxon rank sum tests. **D**, A boxplot comparison between anti-SARS-CoV-2 CRISPR screening gene hits and non-hits. A higher RGES score on y-axis indicates a query host cell shRNA knockdown-induced gene expression profile more closely resembles the summarized gene expression signature of anti-SARS-CoV-2 active compounds. x-axis denotes whether a query host gene knock-out makes the cells resistant to SARS-CoV-2 infection. In the box plot, the central line represents the median value, and the bounds represent the 25th and 75th percentiles. The whiskers are 25th/75th quartiles plus 1.5 times the interquartile range.

of each MoA term pertaining to the active compounds using the Fisher exact test (Supplementary Data 5). However, none of the MoA terms were significantly enriched (false discovery rate (FDR) < 0.05), which emphasized the necessity of decoding

drug mechanisms from transcriptomic data. Overall, 36% (35 out of 97) of the active compounds upregulate the expression of the three genes mapped to 'cholesterol homeostasis' and/or downregulate the expression of the eight genes mapped to

'microtubule cytoskeleton organization' (average z-score ≥ 1.5), whereas three randomly chosen drug profiles did not match this pattern (Figure 2A). Strikingly, these active compounds are not antivirals, and their primary mechanism of action varies. Common mechanisms include NF- κ B inhibitors, selective estrogen receptor modulators, and histamine receptor antagonists. Therefore, their antiviral activity might be an off-target or indirect effect on cholesterol homeostasis and/or microtubule cytoskeleton organization.

Cholesterol homeostasis and Microtubule Cytoskeleton organization gene expression indicates patient severity and compound efficacy *in vitro*

Next, we examined the expression of the genes involved in the two pathways using two independent COVID-19 patient cohorts, including 33 samples from PBMC and 65 samples from leukocytes. The gene expression differences were compared between healthy (or non-COVID-19 donors) and COVID-19 patient groups with different severity levels. As shown in Figure 2B, Supplementary Table S1, and Supplementary Data 6 and 7, the comparison between the patient and healthy group showed a reversal pattern between disease gene expression and the summarized gene expression of the active compounds, meaning the three genes upregulated by anti-SARS-CoV-2 compounds were strongly suppressed in patient samples, whereas the eight genes downregulated by active compounds were mostly induced in the infection status. We then designed the Cholesterol homeostasis and Microtubule Cytoskeleton organization (ChoM-Cyto) score to quantify the reversal (negative) or mimicking (positive) pattern. This score associates with patients' severity. The ChoM-Cyto score successively dropped from -0.72 for the moderate patient group to -1.05 for the severe group, and then to -1.16 for the intensive care unit (ICU) group. In another cohort, the ChoM-Cyto score significantly correlates with patients' sequential organ failure assessment (SOFA) (Spearman correlation -0.52 , $P=0.019$, Figure 2C, Supplementary Table S2 and Supplementary Data 7). In addition to the severity, we also investigated the effect of age or gender on the ChoM-Cyto gene expression. The male patient group showed a more negative ChoM-Cyto score than the female group in the first cohort but not in the second one (Supplementary Figure S2A). We did not find an association between ChoM-Cyto and patient ages in both cohorts (Supplementary Figure S2A and B). Together, cholesterol homeostasis and microtubule cytoskeleton organization pathways are associated with COVID-19 severity, suggesting that co-targeting these two pathways may improve the outcome.

We then investigated if the ChoM-Cyto score has predictive power on the anti-SARS-CoV-2 activity. Cortese *et al.* [16] found SARS-CoV-2 caused cytoskeleton remodeling imperative for viral replication. After testing a few drugs altering cytoskeleton integrity and dynamics, the authors only observed withaferin A had a robust antiviral effect. Gene expression profiles of these drugs suggested they acted differently on ChoM-Cyto genes, although all targeted cytoskeleton-related processes. Withaferin A showed a similar pattern with the active drugs, whereas paclitaxel showed a significantly opposite pattern (Figure 2D). In addition, withaferin A also regulated cholesterol homeostasis, whereas the other three only exhibited minor effects (Figure 2D and Supplementary Table S3, P -values were one-tailed and derived from randomly shuffled signatures permutation); thus, withaferin A achieved a higher ChoM-Cyto score, explaining

its better antiviral effect. A stronger pattern was observed for the positive control, cepharanthine ($EC_{50}=4.47 \mu\text{M}$ [1]). It was reported to block cholesterol trafficking via targeting NPC1 [17]. Also, the qPCR experiment showed the actin expression was undetectable after the human lung primary small airway cells were treated with cepharanthine (Supplementary Table S4), confirming its role on cytoskeleton organization.

To further evaluate its predictive power, we applied the ChoM-Cyto score to all LINCS compounds (Supplementary Data 8). Among the FDA-approved drugs, two top candidates, lomitapide (Microsomal triglyceride transfer protein large subunit (MTTP) inhibitor for hypercholesterolemia treatment) and monensin (an ionophore with reported anti-MERS activity), were not evaluated by published studies, thus selected to test anti-SARS-CoV-2 activity *in vitro*. We found that monensin inhibited SARS-CoV-2 replication in Vero-E6 cells [multiplicity of infection (MOI) of 0.01] with IC_{50} of $11 \mu\text{M}$, and its CC_{50} was $>50 \mu\text{M}$ (Figure 2E). Although lomitapide was inactive under this experimental setting (Supplementary Figure S3), Mirabelli *et al.* [18] recently reported its IC_{50} as 765 nM in Huh7 cells. In addition, among the top 20 candidates, bazedoxifene, dronedarone and osimertinib were already reported active [1]. To further confirm the expression change upon drug treatment, the mRNA level of NPC1, INSIG1 and HMGCS1 was measured in human lung primary small airway cells after treated with cepharanthine, lomitapide, and monensin, respectively. As shown in Figure 2F and Supplementary Data 9, each compound induced a 3~8-fold change of NPC1 and HMGCS1 at $0.5 \mu\text{M}$. In addition to cepharanthine, lomitapide and monensin also inhibited actin expression at 5 and $10 \mu\text{M}$ in human lung primary small airway cells (Supplementary Table S4). This suggests that regulating cholesterol homeostasis and microtubule cytoskeleton organization might contribute to antiviral efficacy. In addition, a comparison between cytotoxic ($CC_{50} < 50 \mu\text{M}$) and non-toxic hits suggests that ChoM-Cyto genes expression change does not significantly contribute to the cytotoxicity (Supplementary Data 10).

ChoM-Cyto gene expression indicates synergistic effects with antiviral drugs

Because the SARS-CoV-2 entry and infection of cells comprises multiple critical biological processes inside infected cells, we further evaluated the potential of the combination of targeting ChoM-Cyto genes and other processes such as viral replication in COVID-19 treatment. To do so, we elicited a recent combination screening study from The National Center for Advancing Translational Sciences (NCATS) [19, 20], where 15 host-targeting compounds were combined with at least one of 11 antivirals [nine known antivirals, two potent SARS-CoV-2 candidates nitazoxanide and hydroxychloroquine (HCQ)] (Figure 3A, Supplementary Data 11). For each host-targeting compound in the columns of Figure 3A, we summarized its synergistic effect with the 11 antivirals (see Methods, Row 'Synergism' in Figure 3A) as well as the ChoM-Cyto score (the middle row in Figure 3A). The drugs with higher ChoM-Cyto scores are more likely to present a synergistic effect with antivirals (Spearman correlation of -0.70 , $P=1.96\text{E-}03$, Figure 3A and B). For instance, the mefloquine's effect on ChoM-Cyto genes shares a similar pattern with that from these reported anti-SARS-CoV-2 active compounds, i.e., upregulation of NPC1, INSIG1 and HMGCS1, whereas downregulation of the other eight genes, and it strongly synergizes with arbidol, a viral envelope fusion

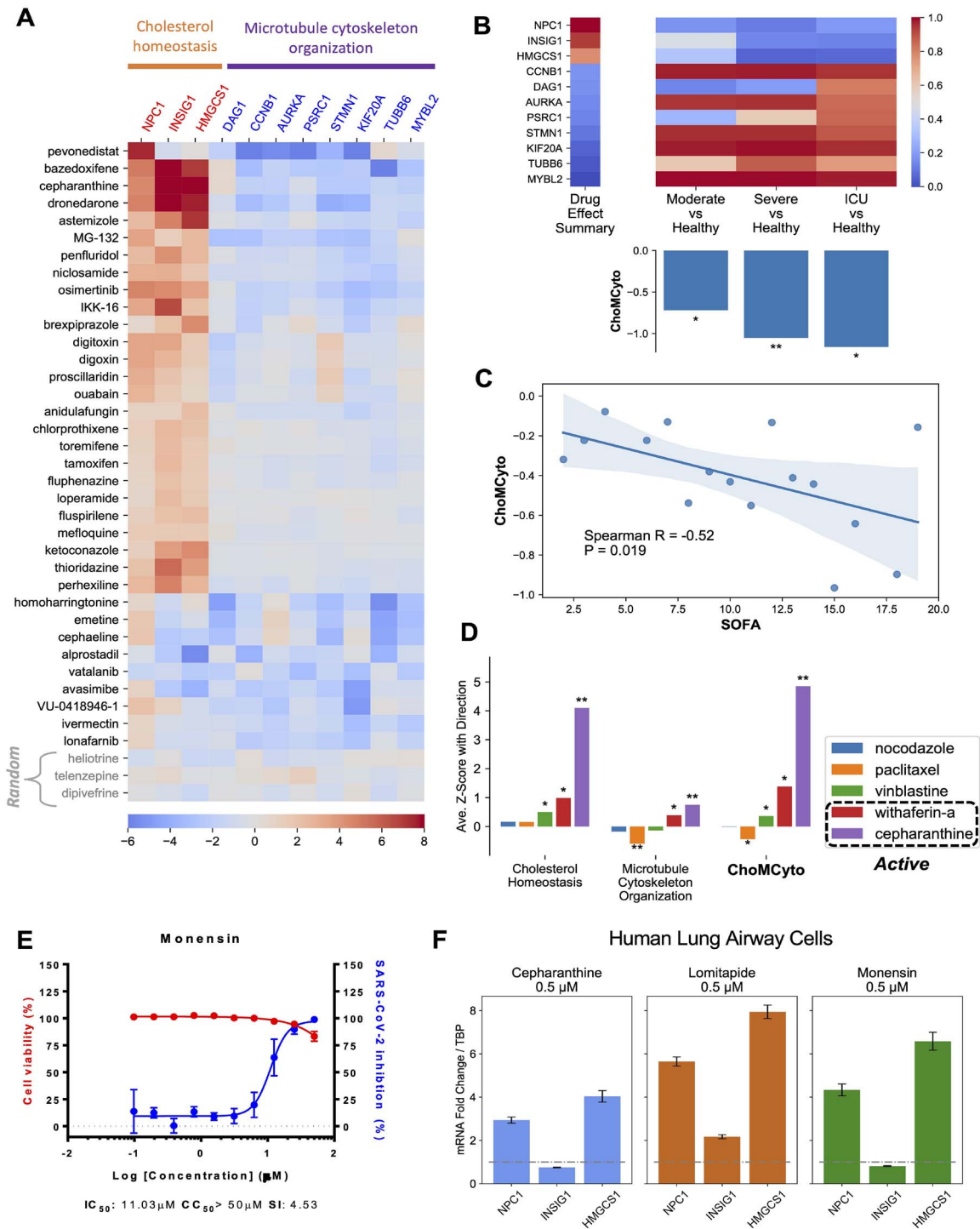


Figure 2. Key pathways regulated by anti-SARS-CoV-2 active compounds indicate patients severity and compound efficacy in vitro. **A**, Expression change of genes involved in cholesterol homeostasis or microtubule cytoskeleton organization induced by anti-SARS-CoV-2 compounds or three randomly selected compounds namely heliotrine, telenzepine, and dipivefrine. **B**, The expression change of the selected ChoMcyto genes in different COVID-19 patient groups (Sequence Read Archive ID: SRP267176). The bar plot shows the overall ChoMcyto scores. **C**, ChoMcyto scores correlate with patients' SOFA (Sequence Read Archive ID: SRP279280). The P-value is one-tailed. **D**, Drugs targeting cytoskeleton, and their effects on cholesterol homeostasis, microtubule cytoskeleton organization and ChoMcyto genes. Dashed rectangle incorporates anti-SARS-CoV-2 active compounds. **E**, The dose-response (blue) and dose-viability (red) curves of monensin, with IC₅₀, CC₅₀ and selectivity index labeled. **F**, mRNA fold changes (compared with TATA-box-binding protein (TBP)) of NPC1, INSG1 and HMGCs1 induced by cepharanthine, lomitapide and monensin at 0.5 µM in human lung primary small airway cells. Error bars denote standard deviations. For reference, the fold change of 1.0 is shown with grey dashed lines. In both heatmaps and their labels, red indicates upregulation, and blue means downregulation. *P < 0.05; **P < 0.001.

inhibitor. On the contrary, leflunomide has a negative ChoMcyto score and antagonizes with lopinavir or nelfinavir. Remdesivir showed a synergistic effect with amodiaquine, nitazoxanide

and emetine; it might retain a potential synergistic effect with triflupromazine, which was not tested yet, but with the highest ChoMcyto score. Among the top of all repurposing candidates,

cepharanthine (ChoMCyto=9.44, [Supplementary Data 8](#)) was also reported to present synergism with remdesivir. The treatment of both compounds under the concentration of 2.5 μM could inhibit >95% cytopathic effect *in vitro*, whereas the treatment of each under the same concentration could only inhibit 35 and 10% cytopathic effect, respectively. [19]. Although either agent alone only presented weak antiviral activity (EC_{50} at micromolar level), their combination exerted a marked effect. This suggests the therapeutic potential of the anti-inflammatory drug cepharanthine combined with antiviral treatment such as remdesivir. Co-targeting the two pathways, especially the ChoMCyto genes, might boost the efficacy of known antiviral drugs for COVID-19 treatment.

Potential biochemical mechanisms of co-targeting

The aforementioned studies demonstrated that ChoMCyto gene expression was closely associated with COVID-19 patient severity and anti-SARS-CoV-2 activity *in vitro*. Next, we employed compound-target and protein-protein interaction (PPI) data to explore the biochemical mechanisms of co-targeting cholesterol homeostasis and microtubule cytoskeleton organization. Thus, the 35 positive compounds listed in [Figure 2A](#) were mapped to their binding or functional target proteins, which were then connected with the ChoMCyto gene products in a PPI network ([Supplementary Figure S4](#) and [Supplementary Data 12](#)). Interestingly, 15 out of the 35 compounds directly target neurotransmitter receptors, including dopamine receptors, serotonin receptors, adrenergic receptors, histamine receptors, opioid receptors and muscarinic acetylcholine receptors, which were densely clustered at one side of the PPI network. The ChoMCyto genes were located on the other side. Between them were a few hub protein targets which interact with many other proteins, implying the potential PPI connections from the compound primary targets to the expression of genes involved in cholesterol homeostasis and microtubule cytoskeleton organization. Among all the target proteins, TP53 interacts with the most ChoMCyto genes, including *INSIG1*, *MYBL2*, *AURKA*, *CCNB1* and *STMN1*. It was also reported regulating the cholesterol level and cytoskeleton remodeling [21, 22]. Thus, TP53 might be a pivotal factor in co-targeting cholesterol homeostasis and microtubule cytoskeleton organization. The ChoMCyto genes could be regulated by prelamin-A/C and microtubule-associated protein tau, which are the mutual targets of 10 and 6 positive compounds, respectively, through TP53. The other two important hub proteins, STAT3 and EGFR, intermediate the interactions between ChoMCyto genes and the neurotransmitter receptors. This PPI network suggests that there are multiple protein targets that directly or indirectly regulate cholesterol homeostasis and microtubule cytoskeleton organization against SARS-CoV-2 infection.

Discussion

Our survey of the positive hits from SARS-CoV-2 drug screenings reveals 36% of them share common mechanistic effects through the regulation of cholesterol homeostasis and/or microtubule cytoskeleton organization. We designed a ChoMCyto score to quantify this summarized drug effect pattern, which is associated with COVID-19 patient severity. By applying the ChoMCyto pattern to predict anti-SARS-CoV-2 efficacy, we discovered monensin with EC_{50} of 11 μM in Vero-E6 cells.

Literature survey suggested the antiviral mechanism of monensin might be via blocking viral transport within the Golgi complex [23]. This indirect antiviral mechanism inspired us to investigate the synergism between established antiviral drugs and repurposed drugs targeting host cellular ChoMCyto genes.

Our findings corroborate the emerging evidence from other studies. For instance, the analysis of clinical data reports hypolipidemia is associated with the severity of COVID-19 [24]. Zang et al. [25] found that cholesterol 25-hydroxylase suppresses SARS-CoV-2 replication by blocking membrane fusion. In addition, its product, 25-hydroxycholesterol, was elevated in a fatal COVID-19 patient and infected mice [26]. The large-scale profiling of SARS-CoV-2 virus-host interactions found that viral proteins Orf8 and Orf9c directly bind to NPC2 and SCAP [3], important components for cholesterol transport and monitoring. In addition, the ‘microtubule-based process’ is enriched in the host protein targets of viral NSP10 and NSP13 [3]. In our recent work [27], 10 out of 11 ChoMCyto genes showed a contrasting expression pattern with that dysregulated by coronavirus infection. Together, we conclude that cholesterol homeostasis and microtubule cytoskeleton organization pathways are disrupted by the viral infection, and one-third of the published anti-SARS-CoV-2 compounds tend to restore these biological processes inside the host cells. Co-targeting the two pathways might boost the efficacy of known antiviral drugs for COVID-19 treatment.

Methods

Collect anti-SARS-CoV-2 positive compounds

By August 2020, 184 compounds with reported *in vitro* anti-SARS-CoV-2 activity were manually collected from 14 studies ([Supplementary Data 1](#)). These compounds could inhibit the SARS-CoV-2 induced-cytopathic effect and/or SARS-CoV-2 replication in the host cells with $\text{EC}_{50}/\text{IC}_{50}$ less than 50 μM . The compound name, efficacy, cytotoxicity, chemical structure, MoA and source links were compiled into this positive compound collection.

Summarize compound- or shRNA- induced gene expression changes

The drug-induced gene expression profiles produced in the LINCS L1000 project [9] are accessible in Gene Expression Omnibus (GEO) with IDs GSE92742 (Phase 1) and GSE70138 (Phase 2). The level-5 profile derived from the comparison of gene expressions between the perturbagen- or vehicle control-treated samples represents gene expression changes upon compound or shRNA treatment. Only 978 landmark genes and high-quality profiles (*is_gold* = 1, annotated in the meta-information) were included in the analysis. As one compound could be profiled under different concentrations, treatment durations, and cellular contexts, given a specific compound, we took the median LINCS z-scores of its profiles measured at 10 μM of drug treatment, regardless of the time and cellular context, resulting in a matrix of 978 genes by 7879 compounds at the end. Similarly, we took the median z-scores of each shRNA across different treatment conditions. Of note, only consensus gene (CGS) knockdown signatures were used unless no ‘CGS’ profiles were measured for a given shRNA perturbagen. This resulted in 4370 different shRNA profiles after merging. A z-

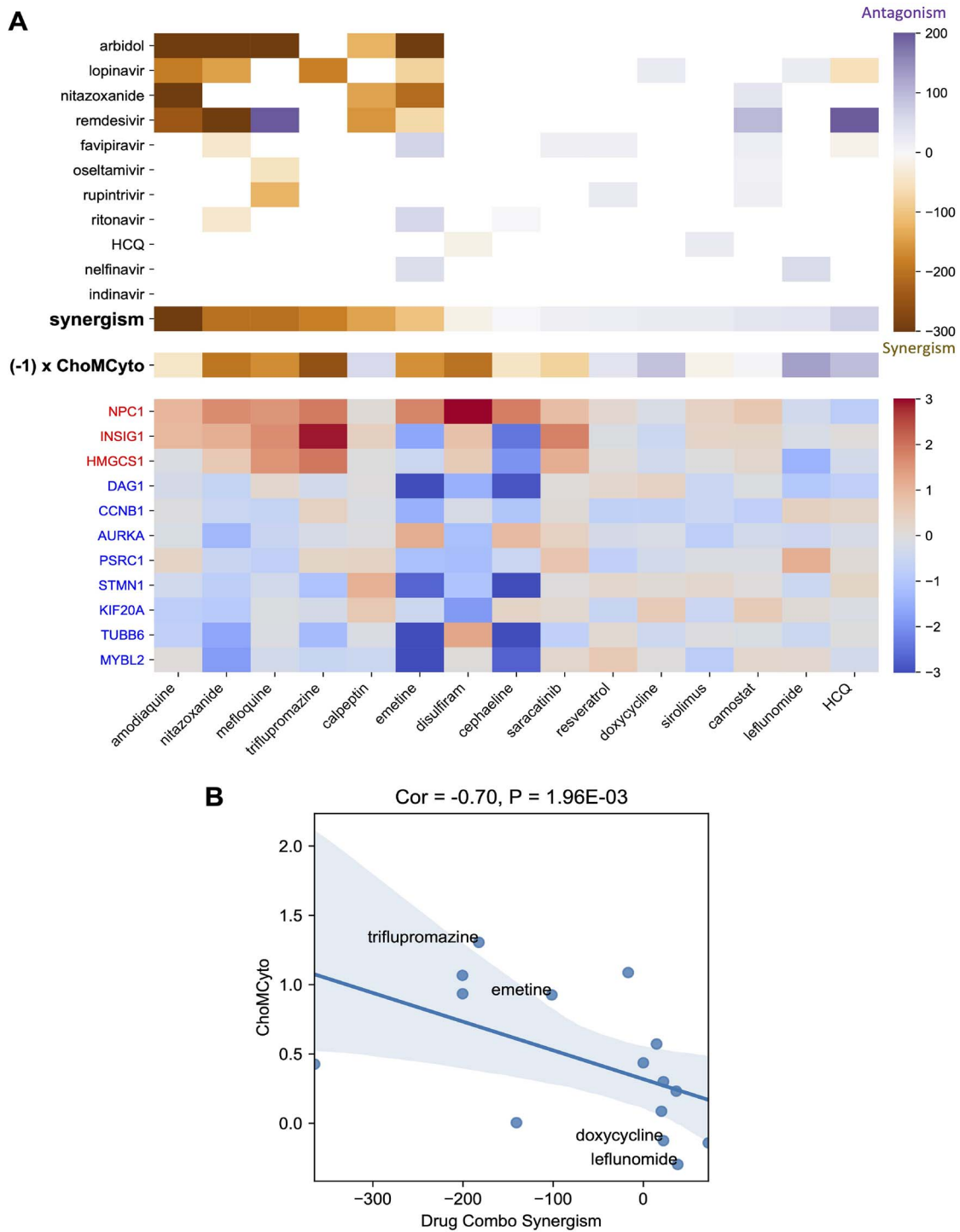


Figure 3. The drug effect on ChoMcyto genes correlates with antiviral synergism. **A**, heatmaps of synergism effects (top), ChoMcyto score (middle) and ChoMcyto genes expression change (bottom). Heatmaps share drug columns. In the top panel, the row 'Synergism' summarizes the average synergistic effect of antiviral drugs with each host-targeting compound in the x-axis. White: missing values, purple: synergism, orange: antagonism. The middle row illustrates the ChoMcyto scores multiplied by -1 , for a better color agreement with 'Synergism'. In the bottom heatmap and gene labels, red indicates upregulation, and blue means downregulation. **B**, ChoMcyto scores and average synergism effects of repurposed anti-SARS-CoV-2 candidates in a scatter plot. Each dot represents a compound. Spearman R correlation and the P-value (one-tailed) are labeled.

score larger than 1.5 indicates that the drug can upregulate the gene expression, whereas a z-score less than -1.5 means downregulation. A larger absolute z-score means a higher

magnitude of expression change. The LINCS profiles have been extensively explored for therapeutic discovery in our previous works [7, 8, 28, 29].

Evaluate genes expression affected by anti-SARS-CoV-2 active compounds

The compounds profiled in LINCS were divided into two groups: anti-SARS-CoV-2 actives and others. The *in vitro* anti-SARS-CoV-2 efficacy data (Supplementary Data 1) were collected from the literature and the NCATS OpenData Portal (<https://opendata.ncats.nih.gov/covid19/assay?aid=14>, accessed by 2 July 2020). Given the concern about data quality from high-throughput screenings, for the efficacy data from NCATS, only compounds with AC_{50} less than 5 μ M were included as actives. The expression z-scores of each landmark gene were compared between actives and other compounds using both the Wilcoxon rank sums test and Fisher exact test. For the Fisher exact test, the following 2 by 2 table was calculated to derive the one-tailed P-values of up/downregulation, respectively.

[(Number of active compounds up/downregulating a gene, Total number of actives),

(Number of other compounds up/downregulating a gene, Total number of other compounds)]

The smaller P-value between up- and down- regulation was selected as the final P-value for this gene. Then we used the Benjamini-Hochberg false discovery rate to correct P-values derived from the Wilcoxon rank sums and Fisher exact test, respectively. The genes with both P-values less than 0.05 and both FDRs less than 0.05 were chosen as significant genes affected by anti-SARS-CoV-2 drugs.

Validate active compounds' signature using CRISPR screening datasets

We first compiled a pro-viral genes list based on five recent studies on the CRISPR screening of anti-SARS-CoV-2 host factor genes [10–14]. Based on the description from these papers, the following criteria were applied to define hits: Daniloski et al.: $FDR < 0.1$ and lfc (\log_2 fold change) > 1 for either MOI 1 or MOI 3; Hoffmann et al.: $FDR < 0.1$ with a positive z-score; Schneider et al.: $FDR < 0.05$ and z-score > 4 ; Wang et al.: score $< 10E-04$ and Wei et al.: Cas9-v2 average > 3 . These hits were mapped to the LINCS shRNA collection. For each LINCS shRNA perturbation, its average profile was summarized, followed by the calculation of RGES to evaluate whether it could mimic (a positive RGES score) the effect of the anti-SARS-CoV-2 active compounds. Finally, the Wilcoxon rank sums test was performed to compare RGES scores between the knock-down of pro-viral genes and other genes in the LINCS shRNA perturbation collection.

Interpret gene expression signatures induced by active compounds

To gain biological insights from the significant genes commonly regulated by active anti-SARS-CoV-2 compounds, we investigated their GO enrichment in STRING-db [30] (<https://string-db.org/>). Enrichment analyses for up- or down- regulated genes were performed separately. As the LINCS profiles only measured the expression change of 978 landmark genes, which cover a small part of the whole transcriptome, we incorporated the gene co-expression knowledge to overcome this limitation. For each part of the dysregulated gene set, a co-expression network was retrieved from STRING-db with the highest confidence interaction scores and no more than 50 interactors of the second shell included. Then the GO enrichment analysis was performed on the interaction network with the whole transcriptome as background.

Evaluate genes expression affected by anti-SARS-CoV-2 active compounds in human lung primary small airway cells

Human small airway epithelia at the air-liquid-interface were cultured using the same protocol for pig small airway culture as published before [31]. Human donor lungs were collected from Spectrum Health Accelerator of Research Excellence protocol with institutional review board approval. All human samples were de-identified and without personal health identifiers.

Briefly, small airway tissue is dissected from distal lungs from within 2 cm from the edge of lung parenchyma. Small airway tissue was digested with pronase and epithelial cells were co-cultured with irradiated feeder cells. Expanded small airway cells were seeded on transwells for at least 2 weeks before it is ready for usage. Anti-SARS-CoV-2 active compounds were added to small airway cells with different doses for 24 h, DMSO was used as vehicle control. RNA was isolated based on the TRIzol™ Reagent (Invitrogen, 15 596 026) protocol. To the 1 ml of Trizol homogenized sample, 200 μ l of chloroform were added and thoroughly mixed by shaking before incubating on ice for 2–3 min. Samples were centrifuged for 15 min at $12\,000 \times g$ at 4°C. The aqueous phase was removed and transferred into a new sample tube. One volume of 70% ethanol was added to each sample, and the entire contents were used for the RNA extraction process. RNA was extracted using the PureLink™ RNA Mini Kit (Invitrogen, 12 183 020) and eluted in 30 μ l of RNase-free water. Complementary DNA (cDNA) from the extracted RNA was synthesized using the SuperScript™ IV VIL0™ Master Mix (Invitrogen, 11 756 050) with the recommended ezDNase enzyme treatment. The resulting cDNA was quantified and diluted to a concentration of 1 μ g/ μ l. A master mix was prepared for each set of primers by utilizing the TB Green® Premix Ex Taq II (TLi RNaseH Plus) reagent kit (Takara Bio, RR820A). Primers were designed and ordered through Sigma-Aldrich (Supplementary Table S5). To prepare the reaction master mix, the recommended protocol for the StepOnePlus Real-Time PCR System in the product manual was followed. Briefly, the TB Green® Premix Ex Taq II (TLi RNaseH Plus) was diluted to a final concentration of 1X and the final concentration for each primer was 0.4 μ M. Except for the template, 18 μ l of the master mix was added to the appropriate well of a 0.2 ml Non-skirted Black-Lettered 96-well PCR plate (Thermo Scientific, AB-0600-L). Of the prepared 1 μ g/ μ l template cDNA, 2 μ l was added to each well in the plate and mixed gently. The plate was covered and centrifuged briefly to remove air bubbles from the wells. The reaction was run using a QuantStudio™ 3 Real-Time PCR System (Applied Biosystems™, A28567). Raw data were uploaded to the Thermo Fisher Cloud Dashboard for data analysis. Gene expression relative to the TBP housekeeping gene was calculated using the $\Delta\Delta$ CT method.

Process patient blood transcriptome samples and compare between patient groups

Raw FASTQ files for SRP267176 (<https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP267176>) and SRP279280 (<https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP279280>) were downloaded from public database NCBI SRA (<https://www.ncbi.nlm.nih.gov/sra>). Data were processed using RNA-seq by expectation maximization (RSEM) [32, 33] 1.3.1+ STAR [34] 2.6.1 pipeline. \log_2 transformed (addition of pseudocount 1) TPM values as gene expression measures were used for analysis. The RNA-Seq processing code is available at GitHub (<https://github.com/>

Bin-Chen-Lab/chenlab_toil). Sample metadata were obtained from GEO (GSE152418 and GSE157103). For the first dataset, gene expression differences were compared between healthy donors (control samples) and moderate/severe/ICU patients (case samples). For the second, gene expression differences were compared between non-COVID-19 non-severe patients with SOFA as 'NA' (control samples) and patient groups with different SOFA scores (case samples). When evaluating the relationship between ChoMcyto and patients' age or gender, samples in each dataset were grouped by age or gender instead of their severity level. Then the gene expression differences were compared between COVID-19 patients and healthy or non-COVID-19 donors within the same age or gender group. For example, fold change values of all genes were derived by comparing COVID-19 patients aged 50–59 versus healthy donors aged 50–59. The \log_2 fold change values were calculated using the `diffExp` function in the open Cancer Therapeutic discovery (OCTAD) [28] R package. For better visualization in a heatmap, the \log_2 fold change values from different comparisons were converted into gene rankings within a group.

Quantify the enrichment between a gene set expression pattern and a disease/drug signature

Given the expression-changing pattern of a set of genes, i.e., significantly dysregulated genes together with their directions of changing (up/downregulated), such as the ChoMcyto pattern, we calculated how strongly it matched with a disease signature (e.g. moderate patients versus healthy donors) or a drug profile (e.g. the summarized LINCS signature of withaferin-A). In detail, ES_{up} was derived by averaging the \log_2 fold change or LINCS z-score values of the genes expected to be upregulated in the gene set of interest. Similarly, ES_{down} was derived by averaging the downregulated genes of interest. Then the final score was defined as $ES_{up} - ES_{down}$. A positive score indicates the disease signature or drug profile exhibits a similar pattern with the gene set of our interest (e.g. ChoMcyto genes direction summarized from active anti-SARS-CoV-2 compounds). In contrast, a negative score indicates the reversal of this pattern. A P-value was derived by scoring 5000 randomly shuffled signatures.

Evaluate inhibition of SARS-CoV-2 replication in vitro

We evaluated whether the proposed compounds could inhibit SARS-CoV-2 replication in an immunofluorescence assay, as previously reported [1]. In brief, Vero cells were seeded at 1.2×10^4 cells per well in the 384-well-plate 24 h prior to infection. Cells were treated with different concentrations (0.1 to 50 μ M, 10 points) of compounds approximately 30 min prior to infection. Then the cells were infected with SARS-CoV-2 (strain: β CoV/Korea/KCDC03/2020, NCCP43326) at a MOI of 0.01, followed with incubation at 37°C for 24 h. After fixation at room temperature for 30 min and permeabilization with 0.25% tritonX-100 for 10 min, the primary antibody, anti-SARS-CoV-2 Nucleoprotein (Sino Biological, 40143-T62), was attached at 37°C for 1.5 h. Then the secondary antibody was attached using goat-anti-rabbit-IgG-alexa-488 + Hoechst 33342, and nucleus staining was conducted at 37°C for 1 h. Then the assay plate was submitted to the Operetta microscope (Perkin Elmer) for imaging. The Columbus software (Perkin Elmer) was used to calculate the infectivity and cell number. Cell viability was measured by comparing the cell number to mock infection. Finally, the dose-response curve was generated with Prism 7 (GraphPad) to calculate IC_{50} with the four-parametric nonlinear fitting algorithm.

Associate drug combination synergism and drug-induced gene expression

Bobrowski et al. [20] measured the anti-SARS-CoV-2 synergism of 73 drug combinations in vitro. As they calculated synergistic (HSA.Neg) and antagonistic (HSA.Pos) effects separately, we took the one with a larger absolute value as the synergism score of each drug combination. We investigated those pairs comprising one known antiviral drug (e.g. remdesivir and arbidol) and one non-typical-antiviral drug (e.g. amodiaquine and mefloquine). For each of the latter drugs, its general synergistic ability was summarized as an average synergism value across all the known antiviral drugs tested. Accordingly, the ChoMcyto scores of these non-typical-antiviral drugs were calculated based on their LINCS profiles. Then we calculated the Spearman correlation between their average synergism values and ChoMcyto scores.

Compare genes expression change induced by cytotoxic and non-toxic compounds

The positive anti-SARS-CoV-2 hits from screenings were annotated as cytotoxic if reported CC_{50} less than 50 μ M. For each landmark gene, its expression z-scores were compared between cytotoxic and non-toxic compounds using the Wilcoxon rank sums test. Then we used the FDR to correct P-values.

Build a PPI network containing positive compounds' primary targets and ChoMcyto genes

First, from the ChEMBL [35] database, we retrieved the activity records of the 35 positive compounds, which regulate cholesterol homeostasis and/or microtubule cytoskeleton organization. The activity records were filtered based on the following criteria: (i) single protein targets, (ii) pChEMBL values ≥ 5 , (iii) neither CYP450 nor HERG proteins and (iv) active to ≥ 3 query compounds according to the second requirement. Of note, the pChEMBL value is a number of roughly comparable measures of half-maximal response concentration, potency or affinity to be compared on a negative logarithmic scale. For example, an IC_{50} measurement of 1 nM would have a pChEMBL value of 9. Then the gene symbols of these targets, together with ChoMcyto gene symbols, were searched in the STRING-db [30] for their PPI connections with the default setting.

Software tools and statistical methods

All analyses were conducted in Python v3.7.6 programming language. Data matrices merging and basic statistics were performed with the Pandas package (v1.0.1). LINCS compound profile extraction was conducted using the `cmapPy` package (v4.0.1). Wilcoxon rank sums test, Fisher exact test and Spearman correlation were calculated with the Scipy package (v1.4.1). All P-values were two-sided unless specified otherwise. FDR values were calculated with the Statsmodels package (v0.11.0). Data visualization was implemented with the Matplotlib (v3.1.3) and Seaborn (v0.10.0) packages.

Key Points

- Published SARS-CoV-2 inhibitors share common mechanisms of action, which are related to cholesterol homeostasis and microtubule cytoskeleton organization.

- The transcriptomic pattern of these pathways correlates with COVID-19 patients' severity and indicates anti-SARS-CoV-2 activity *in vitro*.
- The transcriptomic pattern led to the discovery of monensin as an inhibitor of SARS-CoV-2 replication in Vero-E6 cells.
- Compounds co-targeting cholesterol homeostasis and microtubule cytoskeleton organization processes more likely present a synergistic effect with antivirals.

Supplementary Data

Supplementary data are available online at *Briefings in Bioinformatics*.

Data Availability

The processed LINGS dataset is available upon request. Other datasets are available in the supplementary data.

Contributions

J.X. and B.C. conceived the study. J.X. led and performed all the analyses with the input from S.P., K.L., J.H., and B.C., K.U. and X.L. performed the experiments in human lung primary small airway cells. M.K. and S.K. performed *in vitro* efficacy studies. J.X. and B.C. wrote the manuscript with the input from all co-authors. B.C. supervised the study.

Acknowledgments

The research is supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number R01GM134307, the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number K01ES028047, National Science Foundation (Award ID: 2028717), Spectrum Health-MSU Alliance Corporation, and the MSU Global Impact Initiative. The pathogen resource (NCCP43326) for this study was provided by the National Culture Collection for Pathogens. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2017M3A9G6068245 and NRF-2020M3E9A1041756). The content is solely the responsibility of the authors and does not necessarily represent the official views of sponsors.

References

1. Jeon S, Ko M, Lee J, et al. Identification of antiviral drug candidates against SARS-CoV-2 from FDA-approved drugs. *Antimicrob Agents Chemother* 2020; **64**: e00819–20.
2. Riva L, Yuan S, Yin X, et al. Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature* 2020; **586**: 113–9.
3. Gordon DE, Jang GM, Bouhaddou M, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 2020; **583**: 459–68.
4. White KM, Rosales R, Yildiz S, et al. Plitidepsin has potent preclinical efficacy against SARS-CoV-2 by targeting the host protein eEF1A. *Science* 2021; **371**: 926.
5. Lee BKB, Tiong KH, Chang JK, et al. DeSigN: connecting gene expression with therapeutics for drug repurposing and development. *BMC Genomics* 2017; **18**: 934.
6. van Noort V, Schölch S, Iskar M, et al. Novel drug candidates for the treatment of metastatic colorectal cancer through global inverse gene-expression profiling. *Cancer Res* 2014; **74**: 5690–9.
7. Chen B, Ma L, Paik H, et al. Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets. *Nat Commun* 2017; **8**: 16022.
8. Chen B, Wei W, Ma L, et al. Computational discovery of niclosamide ethanolamine, a repurposed drug candidate that reduces growth of hepatocellular carcinoma cells *in vitro* and *in mice* by inhibiting cell division cycle 37 signaling. *Gastroenterology* 2017; **152**: 2022–36.
9. Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 2017; **171**: 1437–1452.e17.
10. Hoffmann H-H, Schneider WM, Rozen-Gagnon K, et al. TMEM41B is a pan-flavivirus host factor. *Cell* 2021; **184**: 133–148.e20.
11. Schneider WM, Luna JM, Hoffmann H-H, et al. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* 2021; **184**: 120–132.e14.
12. Wang R, Simoneau CR, Kulsuptrakul J, et al. Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell* 2021; **184**: 106–119.e14.
13. Daniloski Z, Jordan TX, Wessels H-H, et al. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell* 2021; **184**: 92–105.e16.
14. Wei J, Alfajaro MM, DeWeirdt PC, et al. Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cell* 2021; **184**: 76–91.e13.
15. Luo J, Yang H, Song B-L. Mechanisms and regulation of cholesterol homeostasis. *Nat Rev Mol Cell Biol* 2020; **21**: 225–45.
16. Cortese M, Lee J-Y, Cerikan B, et al. Integrative imaging reveals SARS-CoV-2-induced reshaping of subcellular morphologies. *Cell Host Microbe* 2020; **28**: 853–866.e5.
17. Lyu J, Yang EJ, Head SA, et al. Pharmacological blockade of cholesterol trafficking by cepharanthine in endothelial cells suppresses angiogenesis and tumor growth. *Cancer Lett* 2017; **409**: 91–103.
18. Mirabelli C, Wotring JW, Zhang CJ, et al. Morphological cell profiling of SARS-CoV-2 infection identifies drug repurposing candidates for COVID-19. *bioRxiv* 2020; 2020.05.27.117184. <https://doi.org/10.1101/2020.05.27.117184>
19. Brimacombe KR, Zhao T, Eastman RT, et al. An OpenData portal to share COVID-19 drug repurposing data in real time. *bioRxiv* 2020. [10.1101/2020.06.04.135046](https://doi.org/10.1101/2020.06.04.135046) preprint: not peer reviewed.
20. Bobrowski T, Chen L, Eastman RT, et al. Synergistic and antagonistic drug combinations against SARS-CoV-2. *Mol Ther* 2021; **29**: 873–85.
21. Aylon Y, Oren M. The hippo pathway, p53 and cholesterol. *Cell Cycle* 2016; **15**: 2248–55.
22. Araki K, Ebata T, Guo AK, et al. p53 regulates cytoskeleton remodeling to suppress tumor progression. *Cell Mol Life Sci* 2015; **72**: 4077–94.
23. Griffiths G, Quinn P, Warren G. Dissection of the Golgi complex. I. Monensin inhibits the transport of viral membrane proteins from medial to trans Golgi cisternae in baby hamster kidney cells infected with Semliki Forest virus. *J Cell Biol* 1983; **96**: 835–50.

24. Wei X, Zeng W, Su J, et al. Hypolipidemia is associated with the severity of COVID-19. *J Clin Lipidol* 2020; **14**: 297–304.
25. Zang R, Case JB, Yutuc E, et al. Cholesterol 25-hydroxylase suppresses SARS-CoV-2 replication by blocking membrane fusion. *Proc Natl Acad Sci U S A* 2020; **117**: 32105.
26. Zu S, Deng Y-Q, Zhou C, et al. 25-Hydroxycholesterol is a potent SARS-CoV-2 inhibitor. *Cell Res* 2020; **30**, **30**(11, 11):1043, 1043–5, 5.
27. Xing J, Shankar R, Drelich A, et al. Analysis of infected host gene expression reveals repurposed drug candidates and time-dependent host response dynamics for COVID-19. *bioRxiv* 2020. [10.1101/2020.04.07.030734](https://doi.org/10.1101/2020.04.07.030734) preprint: not peer reviewed.
28. Zeng B, Glicksberg BS, Newbury P, et al. OCTAD: an open workspace for virtually screening therapeutics targeting precise cancer patient groups using gene expression features. *Nat Protoc* 2021; **16**: 728–53.
29. Chen B, Greenside P, Paik H, et al. Relating chemical structure to cellular response: an integrative analysis of gene expression, bioactivity, and structural data across 11,000 compounds. *CPT Pharmacometrics Syst Pharmacol* 2015; **4**: 576–84.
30. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; **47**: D607–13.
31. Li X, Tang XX, Vargas Buonfiglio LG, et al. Electrolyte transport properties in distal small airways from cystic fibrosis pigs with implications for host defense. *Am J Physiol Lung Cell Mol Physiol* 2016; **310**: L670–9.
32. Li B, Ruotti V, Stewart RM, et al. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* 2010; **26**: 493–500.
33. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; **12**: 323.
34. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**: 15–21.
35. Mendez D, Gaulton A, Bento AP, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res* 2019; **47**: D930–40.