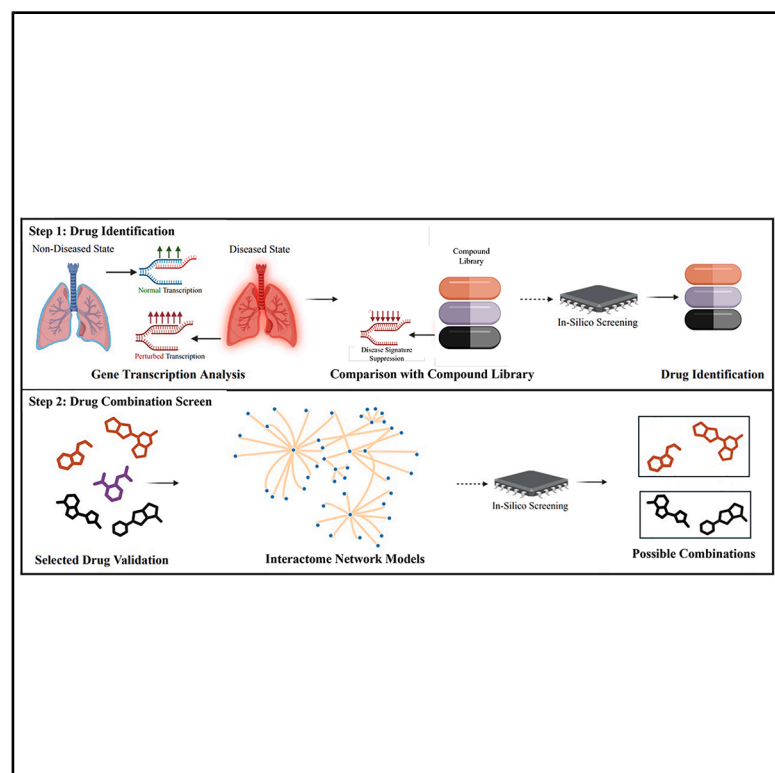


# RepurNet: A network medicine framework for systematic drug repurposing and disease vulnerability discovery

## Graphical abstract



## Authors

Sairahul R. Pentaparthi, Raymond Zou, Prabal Chhibbar, ..., Nidhi Sahni, Erxi Wu, S. Stephen Yi

## Correspondence

nidhi.sahni.2025@gmail.com (N.S.),  
erxi.wu@bswhealth.org (E.W.),  
song.yi@bcm.edu (S.S.Y.)

## In brief

Medicine; Biological sciences; Computer science

## Highlights

- *In silico* screening identifies hundreds of possible drug classes for repurposing
- Several predicted drugs could interfere with host signal transduction networks
- Network-based approach to assess the compatibility of repurposed drug combinations
- Comparative analysis reveals common dysregulated pathways across diseases



## Article

# RepurNet: A network medicine framework for systematic drug repurposing and disease vulnerability discovery

Sairahul R. Pentaparthi,<sup>1,16,17</sup> Raymond Zou,<sup>2,17</sup> Prabal Chhibbar,<sup>3</sup> Aditya Shrawat,<sup>1</sup> Matheus B. Belchior,<sup>4</sup> Weiyi Peng,<sup>5</sup> Werner J. Geldenhuys,<sup>4</sup> Muhammad Shahzeb Khan,<sup>6</sup> Jason H. Huang,<sup>7,8</sup> Juan I. Fuxman Bass,<sup>9,10</sup> Jishnu Das,<sup>3,11</sup> Nidhi Sahni,<sup>7,12,13,\*</sup> Erxi Wu,<sup>1,7,8,14,\*</sup> and S. Stephen Yi<sup>1,7,15,18,\*</sup>

<sup>1</sup>Dell Medical School, The University of Texas at Austin, Austin, TX 78712, USA

<sup>2</sup>Department of Microbiology and Immunology, Columbia University, New York, NY 10032, USA

<sup>3</sup>Center for Systems Immunology, Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA

<sup>4</sup>Department of Microbiology, Immunology and Cell Biology, West Virginia University, Morgantown, WV 26506, USA

<sup>5</sup>Department of Biology and Biochemistry, University of Houston, Houston, TX 77004, USA

<sup>6</sup>Baylor Research Institute, and Department of Medicine, Baylor College of Medicine, Temple, TX 76508, USA

<sup>7</sup>Department of Neurosurgery, Baylor Research Institute and Baylor College of Medicine, Temple, TX 76508, USA

<sup>8</sup>Department of Neurosurgery, Neuroscience Institute, Baylor Research Institute, Baylor Scott & White Health, Temple, TX 76508, USA

<sup>9</sup>Department of Biology, Boston University, Boston, MA 02215, USA

<sup>10</sup>Center for Cancer Systems Biology (CCSB), and Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

<sup>11</sup>Department of Computational and Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA

<sup>12</sup>Quantitative and Computational Biosciences Program, Baylor College of Medicine, Houston, TX 77030, USA

<sup>13</sup>Department of Epigenetics and Molecular Carcinogenesis, and Department of Bioinformatics and Computational Biology, MD Anderson Cancer Center, Houston, TX 77054, USA

<sup>14</sup>Naresh K. Vashisht College of Medicine and Irma Lerma Rangel College of Pharmacy, Texas A & M University, College Station, TX 77843, USA

<sup>15</sup>Dan L Duncan Comprehensive Cancer Center, and Dan L Duncan Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX 77030, USA

<sup>16</sup>Present address: The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

<sup>17</sup>These authors contributed equally

<sup>18</sup>Lead contact

\*Correspondence: [nidhi.sahni.2025@gmail.com](mailto:nidhi.sahni.2025@gmail.com) (N.S.), [erxi.wu@bswhealth.org](mailto:erxi.wu@bswhealth.org) (E.W.), [song.yi@bcm.edu](mailto:song.yi@bcm.edu) (S.S.Y.)

<https://doi.org/10.1016/j.isci.2026.115896>

## SUMMARY

Drug repurposing has become a versatile strategy in rapidly identifying drug candidates to treat disease. Using the early COVID-19 pandemic as an example, we perform a systems-scale *in silico* screen to identify compounds that induce a gene regulatory signature reverse of those found in the disease state. We design a disease module integrating protein interactions to suggest rational combinations of drug candidates to counteract functional changes during disease. We present this strategy as RepurNet, a high-throughput *in silico* framework to rapidly prioritize compounds in early drug discovery. We determine network similarities between distinct diseases that could share common morbidities to illustrate the extension and application of this approach to other diseases, such as cancer. Together, we present a generalizable integrative systems biology and network medicine approach to unveil signaling perturbation and specificity in disease, providing optimized repurposed drug candidates, with broad implications for precision medicine in various disease contexts.

## INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) caused the COVID pandemic. The release of the Pfizer-BioNTech, Moderna, Janssen, Oxford/AstraZeneca, and Novavax SARS-CoV-2 vaccines has enabled the control and prevention of COVID worldwide.<sup>1,2</sup> However, current vaccines do not appear to induce durable cellular or humoral responses, and the level of protection offered by current vaccines against

newly emerging SARS-CoV-2 variants is much lower than the level of protection offered by these vaccines against the original strain.<sup>3</sup> Merck and Pfizer have developed Molnupiravir and Paxlovid, respectively, which are targeted antivirals to treat mild to moderate COVID.<sup>4,5</sup> While antivirals such as Remdesivir can treat severe, hospitalized cases of COVID, early administration is critical.<sup>6,7</sup> Therefore, identifying drug candidates to treat COVID and other infectious diseases is still of global importance.<sup>8</sup>



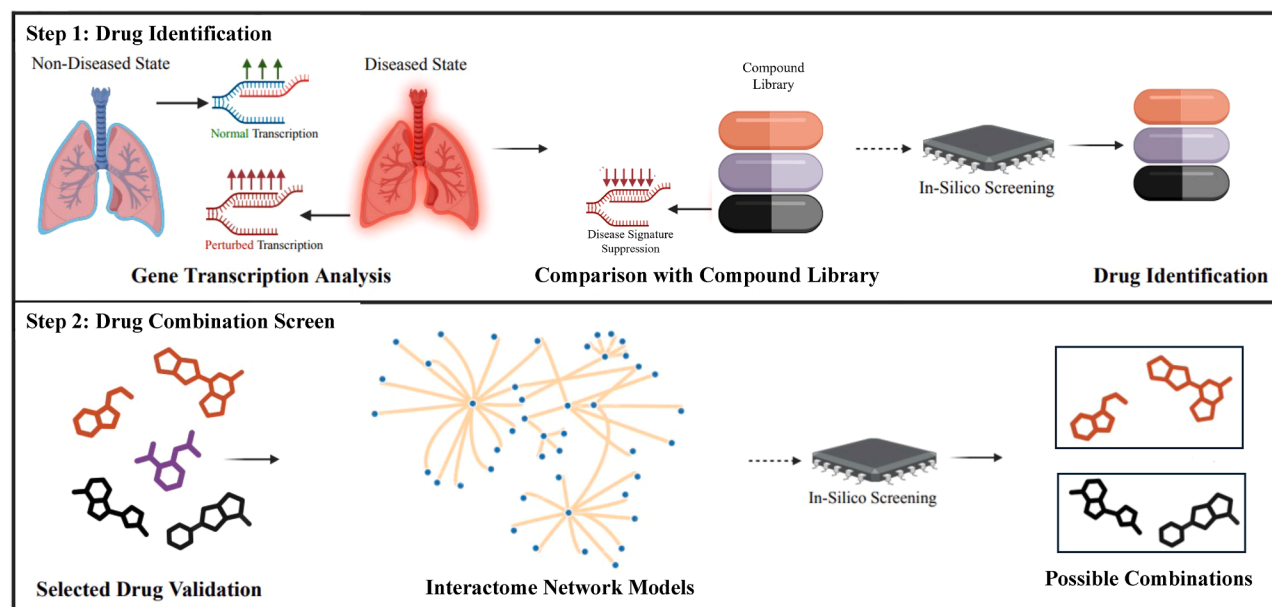


**Figure 1. An overview of the RepurNet framework for drug repurposing**

Firstly, transcriptomic data are analyzed to identify disease-associated gene expression signatures. Then, a library of compound-associated changes in gene expression is compiled. These two datasets were then modeled *in silico* to identify candidate drugs that could putatively suppress disease-associated gene expression signatures. Secondly, candidate drugs are validated by a network-based approach. Briefly, a protein-protein interaction (PPI) interactome that connects disease proteins with their target proteins is compiled, and candidate drugs are then analyzed for which nodes they target in the interactome. Further *in silico* screening can suggest rational drug combinations where candidate drugs simultaneously suppress disease-induced perturbations in independent nodes of the interactome.

Drug repurposing has become a popular strategy in rapidly identifying drug candidates to treat COVID, as it bypasses the time and resource commitments required to develop new drugs.<sup>9</sup> Drugs considered for repurposing can either potentially inhibit SARS-CoV-2 replication or counteract the effects of COVID.<sup>9</sup> For example, Remdesivir, which was developed originally as an RNA polymerase inhibitor of broad spectrum to target Ebola, was repurposed to treat COVID after promising results from clinical trials, and it is one of the few drugs approved by the United States Food and Drug Administration for treating COVID.<sup>10,11</sup> The identification of drug candidates can either follow biological approaches through experimentation or clinical trials, or computational approaches through various large-scale data analysis.<sup>12</sup> While biological approaches are valuable in providing experimental evidence,<sup>12</sup> they are often time-consuming, labor-intensive, and expensive. In contrast, computational approaches are advantageous in integrating large quantities of data and simultaneously investigating the potential of many drugs as therapeutic candidates for drug repurposing.<sup>13</sup> In the case of drug repurposing against SARS-CoV-2, several studies have leveraged different strategies to identify hundreds of putative therapeutic candidates. For example, Egieyeh et al. (2021) conducted a structure-based strategy involving molecular docking analysis and molecular dynamics simulation to screen drugs for their potential to bind and inhibit the SARS-CoV-2 spike protein, which mediates viral entry into host cells through binding with the human angiotensin receptor 2 (hACE2).<sup>14</sup> Saha et al. (2022) applied a network-based approach by screening for

drugs that could target key interactions in an interaction network of human and SARS-CoV-2 proteins, and they identified Fostatinib as a potential candidate for drug repurposing.<sup>15</sup>

We therefore developed a series of computational approaches to identify and validate promising drug candidates to potentially treat COVID. We first characterized transcriptomic perturbations during COVID to generate disease-associated gene expression signatures and identify perturbed gene regulatory networks. Then, we conducted a systems-scale *in silico* screen to identify drugs that induce a gene signature that is opposite to that found in patients with COVID, as that would theoretically reverse the gene signature caused by COVID (Figure 1). We validated this approach by further examining clustering patterns among the identified compounds, accounting for their complete transcriptomic profiles beyond the disease signature. Additionally, we employed a cheminformatics approach to assess possible additional interactions between identified compounds and both host and viral proteins.

After validating the initial compounds, we developed and incorporated a COVID disease module integrating prior knowledge of protein-protein interactions between SARS-CoV-2 viral proteins and human host proteins, and used the disease module to suggest rational combinations of the identified drug candidates to counteract the cellular phenotypic changes associated with infection by SARS-CoV-2. In combination with the prior *in silico* screen, we present this combined approach as RepurNet, an extensible framework to identify early drug candidates in new or understudied diseases.

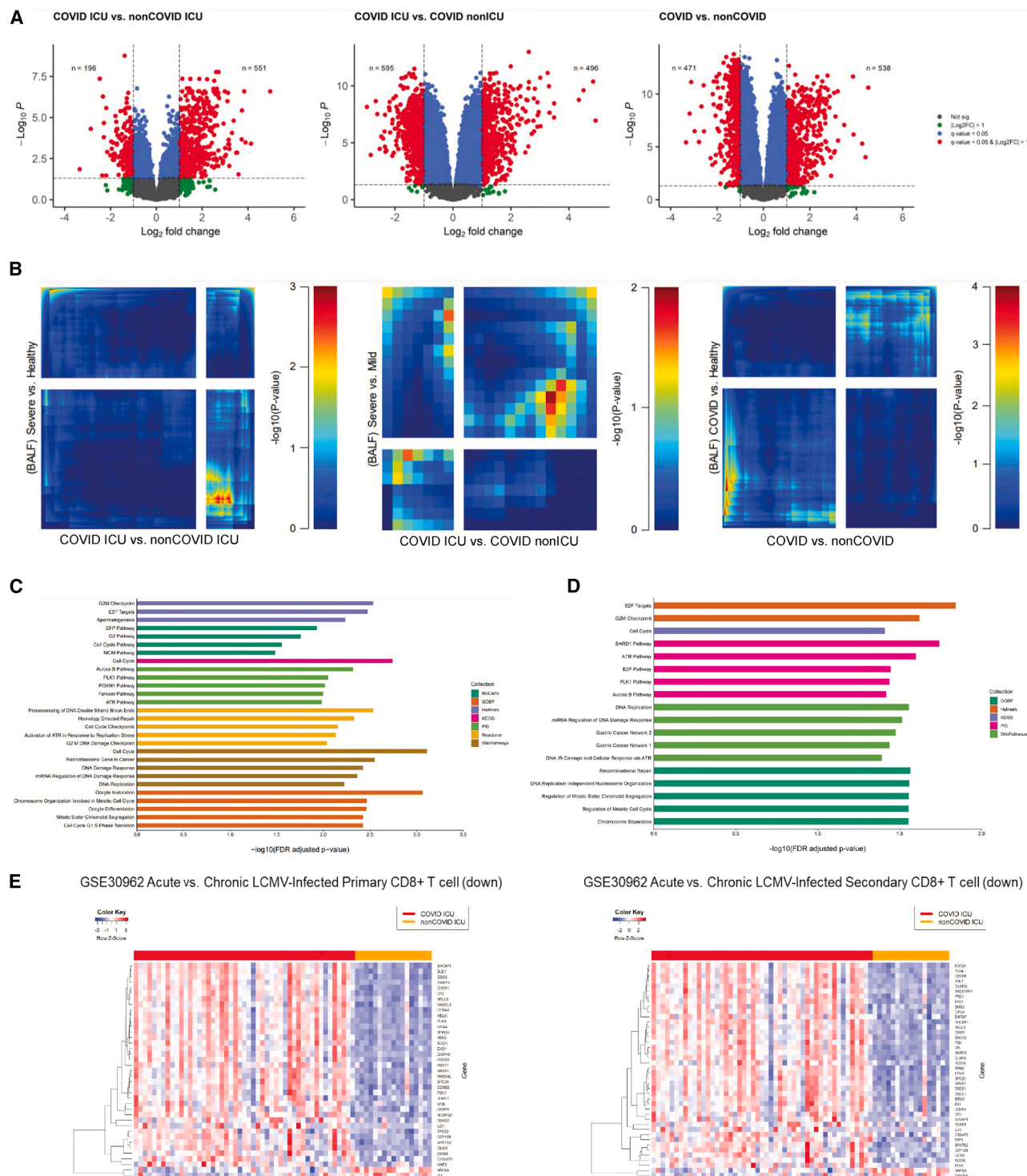


**Figure 2. Gene regulatory network signatures and widespread rewiring associated with COVID disease severity**

(A) Volcano plots summarize the DEGs in the COVID ICU vs. nonCOVID ICU, COVID ICU vs. COVID nonICU, and COVID vs. nonCOVID DEG profiles. The cutoffs for differential expression are FDR-adjusted  $p$ -value  $< 0.05$  and  $|\log_2(\text{fold-change})| \geq 1$ , which are represented by the dotted lines. The number of significantly upregulated or downregulated genes in each DEG profile is annotated on the respective sides of their x axes.

(B) RRHO heatmaps visualize concordance or discordance between blood-derived DEG profiles and BALF-derived DEG profiles. The statistical significance of concordantly or discordantly expressed genes is expressed and visualized in terms of  $-\log_{10}(p\text{-value})$ , where values above 1.3 (which correspond to  $p$ -values  $< 0.05$ ) indicate significantly concordant or discordant expression patterns. Heatmaps are interpreted as demonstrated in Cahill et al. (2018), where discordant

(legend continued on next page)

Lastly, we explored network similarities between COVID disease and lung cancers (TCGA-LUAD and TCGA-lung squamous cell carcinoma (LUSC)). Together, our integrative systems biology and network medicine approach represents a generalizable model to unveil signaling perturbation and specificity in COVID and other human diseases, providing a set of optimized repurposed drug candidates, with major and broad clinical implications for precision medicine in various disease contexts.

## RESULTS

### Unique gene regulatory network signatures associated with COVID severity derived from blood samples

We obtained patient samples of whole blood leukocyte mRNA expression data from the Gene Expression Omnibus (GEO) dataset GSE157103 that classified samples based on whether patients were diagnosed with COVID and whether patients were admitted to an intensive care unit (ICU), which modeled disease status and severity, respectively.<sup>16</sup> Patients who tested positive with COVID were not disease-free as all samples were obtained from hospitalized patients, and specific comorbidities of each patient were unavailable.<sup>16</sup> We generated differentially expressed gene (DEG) profiles comparing samples from ICU patients with COVID against ICU patients without COVID (COVID ICU vs. nonCOVID ICU); samples from ICU patients with COVID against nonICU patients with COVID (COVID ICU vs. COVID nonICU); and samples from patients with COVID against patients without COVID (COVID vs. nonCOVID) (Figure 2A; Table S1). Perturbed genes with a  $|\log_2FC| > 1$  and a false discovery rate (FDR) adjusted  $p$ -value  $< 0.05$  were considered differentially expressed. Notably, our COVID ICU vs. COVID nonICU profile revealed significant upregulation of *S100A12* ( $\log_2FC = 1.45$ , FDR adjusted  $p$ -value =  $3.47 \times 10^{-7}$ ) and *OSM* ( $\log_2FC = 1.03$ , FDR adjusted  $p$ -value =  $6.94 \times 10^{-5}$ ) in COVID ICU patients compared to COVID nonICU patients, and both of these genes' high expression was previously identified to be correlated with clinical severity.<sup>17</sup> After deconvoluting immune cell type proportions from transcript per million (TPM) normalized gene expression data, which was also provided in the study GSE157103, we observed significantly decreased proportions of myeloid dendritic cells ( $p$ -value =  $8.81 \times 10^{-9}$ ), NK cells ( $p$ -value =  $2.01 \times 10^{-4}$ ), CD8<sup>+</sup> T cells ( $p$ -value =  $1.30 \times 10^{-7}$ ), and non-regulatory CD4<sup>+</sup> T cells ( $p$ -value =  $4.63 \times 10^{-3}$ ) in whole blood of COVID ICU patients compared to COVID nonICU patients (Figure S1). Moreover, from the genes included in the Kyoto Encyclopedia of Genes and Genomes (KEGG)<sup>18</sup> p53 signaling pathway, we observed significant downregulation of *IGFBP3*

( $\log_2FC = -1.54$ , FDR adjusted  $p$ -value =  $5.00 \times 10^{-8}$ ), *ADGRB1* ( $\log_2FC = -1.05$ , FDR adjusted  $p$ -value =  $1.78 \times 10^{-5}$ ), and *BCL2* ( $\log_2FC = -1.17$ , FDR adjusted  $p$ -value =  $1.00 \times 10^{-8}$ ) from the COVID ICU vs. COVID nonICU profile. We also observed significant upregulation of *CCNB1* ( $\log_2FC = 1.87$ , FDR adjusted  $p$ -value =  $2.56 \times 10^{-7}$ ), *RRM2* ( $\log_2FC = 2.49$ , FDR adjusted  $p$ -value =  $2.78 \times 10^{-7}$ ), *CCNB2* ( $\log_2FC = 2.13$ , FDR adjusted  $p$ -value =  $1.44 \times 10^{-6}$ ), *GTSE1* ( $\log_2FC = 1.98$ , FDR adjusted  $p$ -value =  $7.63 \times 10^{-7}$ ), *CDK1* ( $\log_2FC = 2.50$ , FDR adjusted  $p$ -value =  $2.56 \times 10^{-7}$ ), and *CHEK1* ( $\log_2FC = 1.87$ , FDR adjusted  $p$ -value =  $4.43 \times 10^{-8}$ ) and significant downregulation of *SIVA1* ( $\log_2FC = -1.20$ , FDR adjusted  $p$ -value =  $1.08 \times 10^{-3}$ ) from both COVID ICU vs. nonCOVID ICU and COVID vs. nonCOVID profiles (reported  $\log_2FC$ s and FDR adjusted  $p$ -values are based on the COVID ICU vs. nonCOVID ICU profile).

Additionally, Xu et al. (2020) identified differences between the peripheral immune response and the local immune response in cases of severe COVID, and we sought to examine whether results from our blood-derived DEG profiles were consistent with these observations when compared to gene expression in the infected lung.<sup>20</sup> As such, we analyzed the differences between our three blood-derived DEG profiles against bronchoalveolar lavage fluid (BALF)-derived DEG profiles using rank-rank hypergeometric overlap (RRHO) analysis, a method developed by Cahill et al. (2018) designed to compare significant similarities or differences between two sets of expression data.<sup>19,21</sup> First, we obtained BALF immune cell mRNA expression data from the GEO dataset GSE145926 COVID and generated DEG profiles comparing patients with severe COVID against healthy controls (severe vs. healthy); patients with severe COVID against patients with mild COVID (severe vs. mild); and patients with COVID against healthy controls (COVID vs. Healthy).<sup>22</sup> The BALF-derived DEG profiles are summarized in Table S2. Then, we performed RRHO analysis to compare blood-derived and BALF-derived DEG profiles in the pairwise fashion, as summarized in Table 1, and our results are summarized in Figure 2B. The pairwise comparisons were designed for the BALF DEG profiles to best parallel their respective whole blood DEG profiles, first in terms of disease status and then in terms of severity. While some comparisons are not ideal, such as comparing BALF samples from healthy patients to whole blood samples of nonCOVID ICU patients, we could not make more accurate comparisons due to data availability constraints. As such, we prioritized whether patients were diagnosed with COVID in constructing comparisons. Concordance and/or discordance of profile pairs were identified by whether the max  $-\log_{10}(p\text{-value})$  was located

genes are enriched in the upper left and lower right quadrants of each plot, and concordant genes are enriched in the upper right and lower left quadrants of each plot.<sup>19</sup>

(C and D) Barplots summarize the top significantly enriched gene sets by FDR adjusted  $p$ -value. The top 5 most significant gene sets from the Hallmark, BioCarta, KEGG, PID, Reactome, WikiPathways, and GOBP collections that enrich the COVID ICU profile compared to the nonCOVID ICU profile (C) and those that enrich the COVID profile compared to the nonCOVID profile (D) are displayed. If fewer than 5 gene sets from any collection are enriched for a given profile, then all of those gene sets are displayed.

(E) Heatmaps summarizing expression patterns of genes in the Acute vs. Chronic LCMV-Infected Primary CD8<sup>+</sup> T cell (down) and the Acute vs. Chronic LCMV-Infected Secondary CD8<sup>+</sup> T cell (down) gene sets from ImmuneSigDB (study: GSE30962). Heatmaps are constructed using  $\log_2(TMM+1)$  mRNA expression data from GSE157103. Clustering is performed using Pearson's correlation distance and Ward's D2 clustering. Only differentially expressed genes from the COVID ICU vs. nonCOVID ICU DEG profile are shown. The gene sets have FDR adjusted  $p$ -values of 0.00567 and 0.00202, respectively. See also Figure S1 and Tables S1, S2, S3, S4, and S5.

**Table 1. Levels of concordance and discordance between blood-derived DEG profiles and BALF-derived DEG profiles and their corresponding maximum  $-\log_{10}(p\text{-value})$**

| Blood DEG Profile          | BALF DEG Profile   | Level of concordance or discordance | max $-\log(p\text{-value})$ |
|----------------------------|--------------------|-------------------------------------|-----------------------------|
| COVID ICU vs. nonCOVID ICU | Severe vs. Healthy | Weak Discordance                    | 2.89                        |
| COVID ICU vs. COVID nonICU | Severe vs. Mild    | Weak Discordance                    | 1.91                        |
| COVID vs. nonCOVID         | COVID vs. Healthy  | Weak Concordance                    | 4.03                        |

in quadrants corresponding to either concordance or discordance according to the stratified interpretation proposed by Cahill et al. (2018).<sup>19</sup> Lists of genes that RRHO analysis identified as having maximal statistical overlaps either concordantly or discordantly between each profile pair are provided in Table S3. We also sought to validate the COVID ICU vs. COVID nonICU DEG profile against proteomic expression data, and we examined whether the protein products of DEGs in the COVID ICU vs. COVID nonICU profile were also differentially dysregulated in sera of patients with severe COVID compared to patients with mild COVID described by Shen et al. (2020).<sup>23</sup> We observed that significantly higher expression of genes *ORM1* ( $\log_2FC = 2.20$ , FDR adjusted  $p\text{-value} = 4.79 \times 10^{-9}$ ), *ORM2* ( $\log_2FC = 2.04$ , FDR adjusted  $p\text{-value} = 1.87 \times 10^{-6}$ ), *LRG1* ( $\log_2FC = 1.27$ , FDR adjusted  $p\text{-value} = 8.21 \times 10^{-8}$ ), and *HP* ( $\log_2FC = 1.41$ , FDR adjusted  $p\text{-value} = 2.07 \times 10^{-4}$ ) but lower expression of *IGFBP3* ( $\log_2FC = -1.54$ , FDR adjusted  $p\text{-value} = 5.00 \times 10^{-8}$ ) corresponded to parallel profiles of their protein products. Taken together, our results revealed unique gene regulatory network signatures associated with COVID severity derived from blood samples in the patients.

### Widespread rewiring of signaling pathways in distinct COVID patient samples

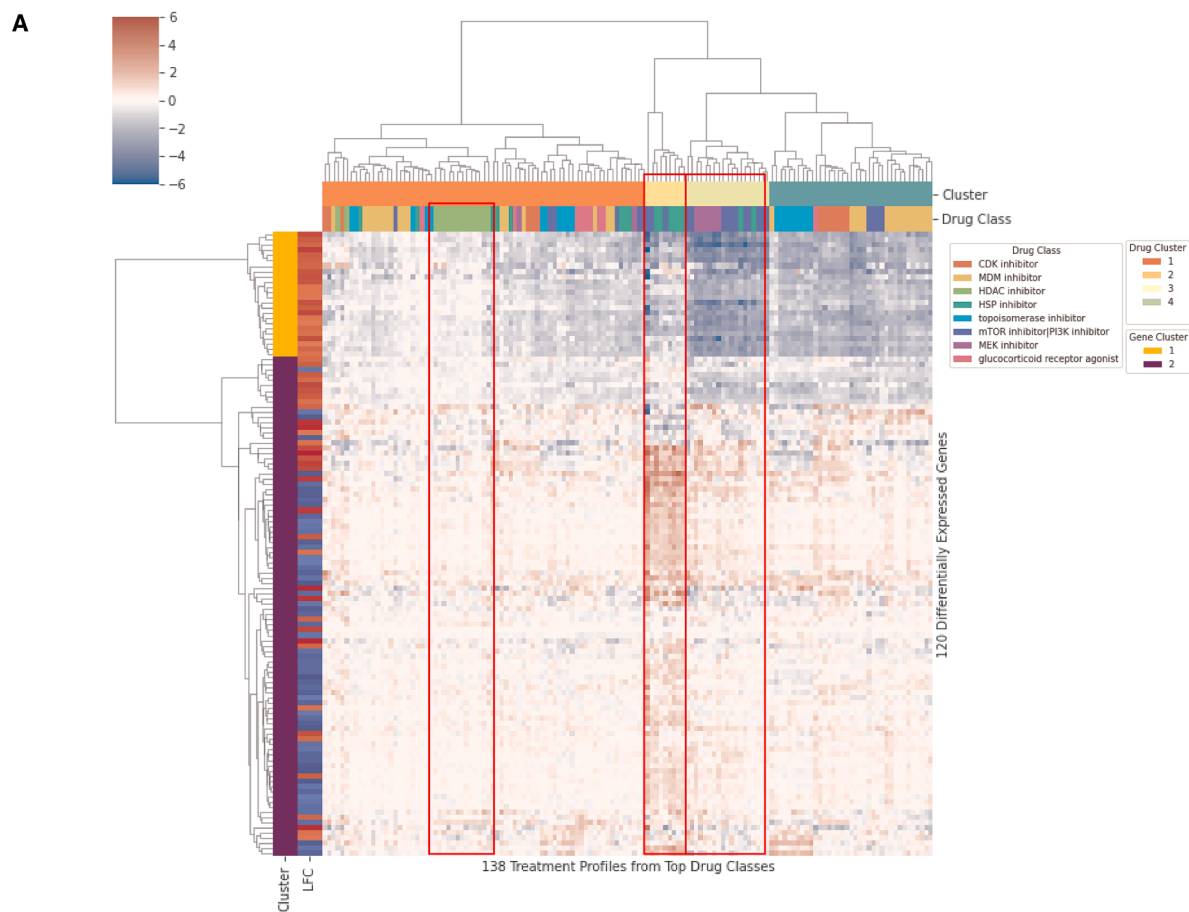
We also sought to identify perturbed signaling pathways during COVID. We conducted gene set enrichment analysis (GSEA) on our three blood-derived DEG profiles using an FDR-adjusted  $p\text{-value} < 0.05$  as a cutoff for statistical significance (Figures 2C and 2D). We tested for enriched gene sets in the Hallmark,<sup>24</sup> BioCarta,<sup>25</sup> KEGG,<sup>18</sup> Pathway Interaction Database (PID),<sup>26</sup> Reactome,<sup>27</sup> WikiPathways,<sup>28</sup> Gene Ontology Biological Pathways (GOBP),<sup>29,30</sup> Gene Ontology Molecular Function (GOMF),<sup>29,30</sup> oncogenic signature,<sup>25</sup> and ImmuneSigDB<sup>31</sup> collections available on MSigDB (v7.4).<sup>25</sup> The number of gene sets comprising these collections is summarized in Table S4. Testing for enriched gene sets within each gene set collection was done separately in the GSEA software developed by UC San Diego and the Broad Institute.<sup>25,32</sup> 300 gene sets were positively enriched in COVID ICU patients compared to nonCOVID ICU patients, and 124 gene sets were positively enriched in COVID patients compared to nonCOVID patients. All our GSEA results are summarized in Table S5. Many of the significantly enriched GOBP terms in patients with COVID were related to cell cycle processes. Addition-

ally, the PLK1 pathway (PID) and cell cycle (WikiPathways) pathways were enriched in COVID ICU patients with FDR-adjusted  $p\text{-values}$  of  $8.91 \times 10^{-3}$  and  $7.77 \times 10^{-4}$ , respectively.<sup>26,33</sup>

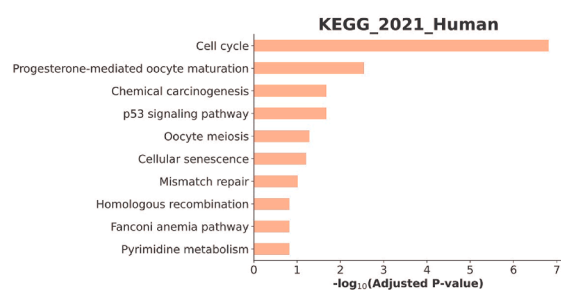
Interestingly, the acute vs. chronic LCMV-infected primary CD8<sup>+</sup> T cell (down) and acute vs. chronic LCMV-infected secondary CD8<sup>+</sup> T cell (down) gene sets from the ImmuneSigDB collection (study GSE30962) were enriched in the COVID ICU patients with FDR-adjusted  $p\text{-values}$  of 0.00567 and 0.00202, respectively.<sup>34</sup> These gene sets describe genes in primary and secondary murine CD8<sup>+</sup> T cells during lymphocytic choriomeningitis virus (LCMV) infection that are downregulated in an acute infection model compared to a chronic infection model.<sup>34</sup> Thus, some of these downregulated genes were upregulated in ICU patients with COVID compared to ICU patients without COVID. We constructed heatmaps of the  $\log_2(\text{TMM}+1)$  transformed mRNA expression data from GSE157103 to examine expression patterns of DEGs from the COVID ICU vs. nonCOVID ICU profile that were in these two gene sets (Figure 2E). Using Pearson's correlation distance and Ward's D2 hierarchical clustering, we identify distinct, separate clusters of COVID ICU and nonCOVID ICU patients, which suggest that these genes are characteristically dysregulated in ICU patients with COVID compared to ICU patients without COVID. No significantly enriched gene sets were shared between the blood-derived profiles and the BALF-derived profiles, indicating a highly differentiated response to SARS-CoV-2 infection in blood and in lung tissue from a pathway perspective.

### Discovery of repurposed compounds to combat COVID infection

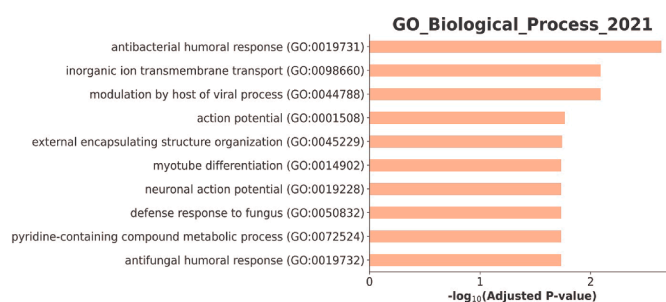
After identifying DEG profiles, we began the first part of the RepurNet analysis to identify candidate compounds. We used the Connectivity Map's 2021 LINCS\_COVID\_PUBLIC data release to screen for compounds associated with effective reversal of the observed gene signature.<sup>35</sup> The dataset contains  $\log_2FC$  values of gene expression changes after high-throughput treatment with drug repurposing candidate compounds. This release featured a wider range of compounds compared to previous LINCS data releases, with compounds from both the Broad Institute and Scripps Research represented in the dataset. Additionally, these assays were focused on cell lines from tissues strongly affected by COVID infection, such as lung (A375), hematologic (THP1, BICR6), and intestinal (OUMS23) cell lines. To focus our search for compounds that could reverse the severe COVID disease state, the top 60 upregulated and downregulated genes by magnitude of  $\log_2FC$  were selected from the COVID ICU vs. nonCOVID ICU DEG profile. The total pool of treatments was filtered down to 24-h treatments in the A375 lung cells to standardize comparisons between compounds and associated changes in gene expression. The filtered treatments were downloaded and ranked for signature reversal using a Pearson's correlation metric ( $-1 < R < 1$ ) comparing the  $\log_2FC$  values from our DEG profiles against the  $\log_2FC$  values for each treatment, with the assumption that a more negative Pearson's R value corresponds to a stronger reversal of the disease signature (see STAR Methods; Table S6). From an initial pool, 448 compounds were identified as strong candidates for reversal based on 2 selection criteria: First, if a compound had



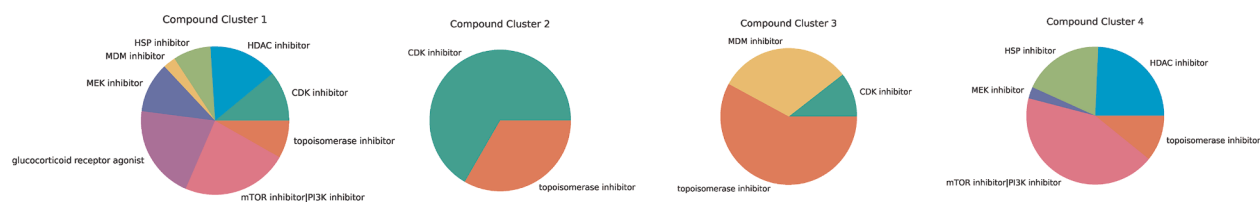
**B** Gene Cluster 1



**C** Gene Cluster 2



**D**



(legend on next page)

at least one replicate with a Pearson's R value of strong significance ( $p < 0.0001$ ), it was included as a prioritized candidate. Otherwise, if a compound was in at least 2 replicates with an R value of moderate significance ( $p < 0.05$ ), it was also included as a prioritized candidate. Among the identified 448 compounds, 8 drug classes were strongly represented (more than 7 candidates in the final compound list): mTOR/PI3K inhibitors, topoisomerase inhibitors, HDAC inhibitors, CDK inhibitors, glucocorticoid receptor agonists, HSP inhibitors, MEK inhibitors, and MDM inhibitors. Compounds in these eight classes were clustered to assess patterns in their transcriptomic activities (Figure 3A). 4 compound clusters were identified based on similar induced changes in gene expression patterns (Figures S2A–S2D). Similarly, 2 gene clusters were identified, highlighting groups of genes that respond similarly to chemical perturbation.

GO enrichment analysis of the gene clusters provided insight into the functions associated with each cluster (Figures 3B and 3C). Gene cluster 1, featuring genes that were upregulated in the COVID ICU signature, was most enriched in cell cycle activity. As such, many of the selected compounds showed primary activity in cell-cycle arrest. Gene cluster 2, mostly featuring genes that were downregulated in the COVID ICU signature, showed a less defined enrichment profile. Interestingly, many of the top activities were associated with transmembrane protein activity. Compounds that upregulated these genes could therefore help restore some of the extracellular communication that is suppressed during SARS-CoV-2 infection.

The compound clusters provided further insight into the activities of some of these compounds (Figure 3D). Compound cluster 1, visualized in red, showed compounds that had mild suppressive activity with regard to our transcriptomic signature. Among other compounds, this cluster included most of the corticosteroid receptor agonists. This class included dexamethasone, one of the few NIH-recommended treatments for patients hospitalized with COVID.<sup>36,37</sup> This suggests that corticosteroids, which were selected largely on their systemic anti-inflammatory properties, could be supported or replaced by compounds that could target the COVID signature with more focused antiviral activity. We found such candidates in clusters 2 and 3, visualized in tan and yellow, respectively. Compound cluster 2 was notable for activating expression in gene cluster 2, indicating they might help restore extracellular activities that are dysregulated during infection. The cluster mostly featured CDK inhibitors with a few topoisomerase inhibitors. Compound cluster 3, on the other hand, showed the strongest suppressive activity of the 4 compound clusters with regard to gene cluster 1. The cluster was enriched in topoisomerase inhibitors and MDM inhibitors, suggesting a strong role for compounds that arrest the cell cycle.

### Distinct global transcriptomic profiles associated with identified drug candidates

To visualize global transcriptomic profiles, we calculated the top 40 uniform manifold approximation and projection (UMAP) embeddings used to represent the original LINCS COVID-specific dataset in a reduced dimension space (Figure 4A). After performing a K-means nearest neighbor clustering analysis using FAISS to cluster all the compounds in the LINCS\_COVID\_PUBLIC dataset, we tested for the enrichment of our 448 candidate compounds in any of the clusters using the hypergeometric test with the FDR controlled at 1%. We observed that our candidate drugs were enriched ( $p < 0.0001$ ) in cluster 0 (C0) and cluster 1 (C1) (Figure 4B).

The clustering is highly interpretable and has functionality to generate the top “n” nearest neighbor matrices for every data point. We leveraged this functionality to examine the twenty nearest neighbors for every candidate drug treatment. A treatment-associated nearest neighbor score was computed using the number of candidate drug treatment neighbors. This score was significantly correlated with the coefficient values indicative of reversal of the observed signature of DEGs between COVID ICU patients vs. COVID nonICU patients. (Spearman correlation coefficient =  $-0.57$ ,  $p < 0.0001$ ) (Figure 4C).

We further found that the concordance was robust to the number of nearest neighbors. Performing the analysis with 5, 10, and 50 nearest neighbors resulted in Spearman correlation coefficient values between the two metrics of  $-0.46$ ,  $-0.5$ , and  $-0.56$ , respectively. All three of these values correspond to  $p$ -values  $< 0.0001$  (Figures S3A–S3C).

Interestingly, most compounds were rated consistently on both scoring metrics (the Pearson's correlation reversal score and the treatment-associated nearest neighbor score). Several of the high-scoring compounds had produced intriguing results in preclinical investigations. For example, panobinostat (Pearson's  $R = -0.54$ , treatment-associated nearest neighbor score = 7), an HDAC inhibitor currently released for use in multiple myeloma, was found to reduce SARS-CoV-2 replication in infected Vero cell lines as well as reduce ACE2 transcription rates in KATOIII and NUGC-4 cell lines.<sup>38,39</sup> Additionally, mitoxantrone (Pearson's  $R = -0.53$ , treatment-associated nearest neighbor score = 7), a topoisomerase inhibitor, was found to interfere with SARS-CoV-2 viral entry as well as replication. Mitoxantrone interferes with the interaction between the SARS-CoV-2 spike protein and heparin sulfate to limit entry, and it also intercalates in viral RNA to interfere with nucleocapsid assembly.<sup>40,41</sup> Together, these compounds, and potentially several others identified by our approach, may have efficacy in inhibiting the activity of SARS-CoV-2 on a molecular level.

### Figure 3. Network-guided identification of repurposed compounds for COVID

- (A) The top 138 treatments based on Pearson's score reversal values. Rows represent the 60 most upregulated and 60 most downregulated genes in the COVID ICU vs. nonCOVID ICU DEG profile, while columns represent compounds from the 8 most highly represented drug classes in our 448-compound master list. Rows and columns are clustered using Ward's D2 algorithm, producing 2 gene clusters and 4 compound clusters. Heatmap contains LFC2 values associated with 24 h treatment by the corresponding compound.
- (B) GO term enrichment results for gene cluster 1.
- (C) GO term enrichment results for gene cluster 2.
- (D) Pie charts illustrate compound representation in each of the 4 clusters. See also Figure S2 and Table S6.

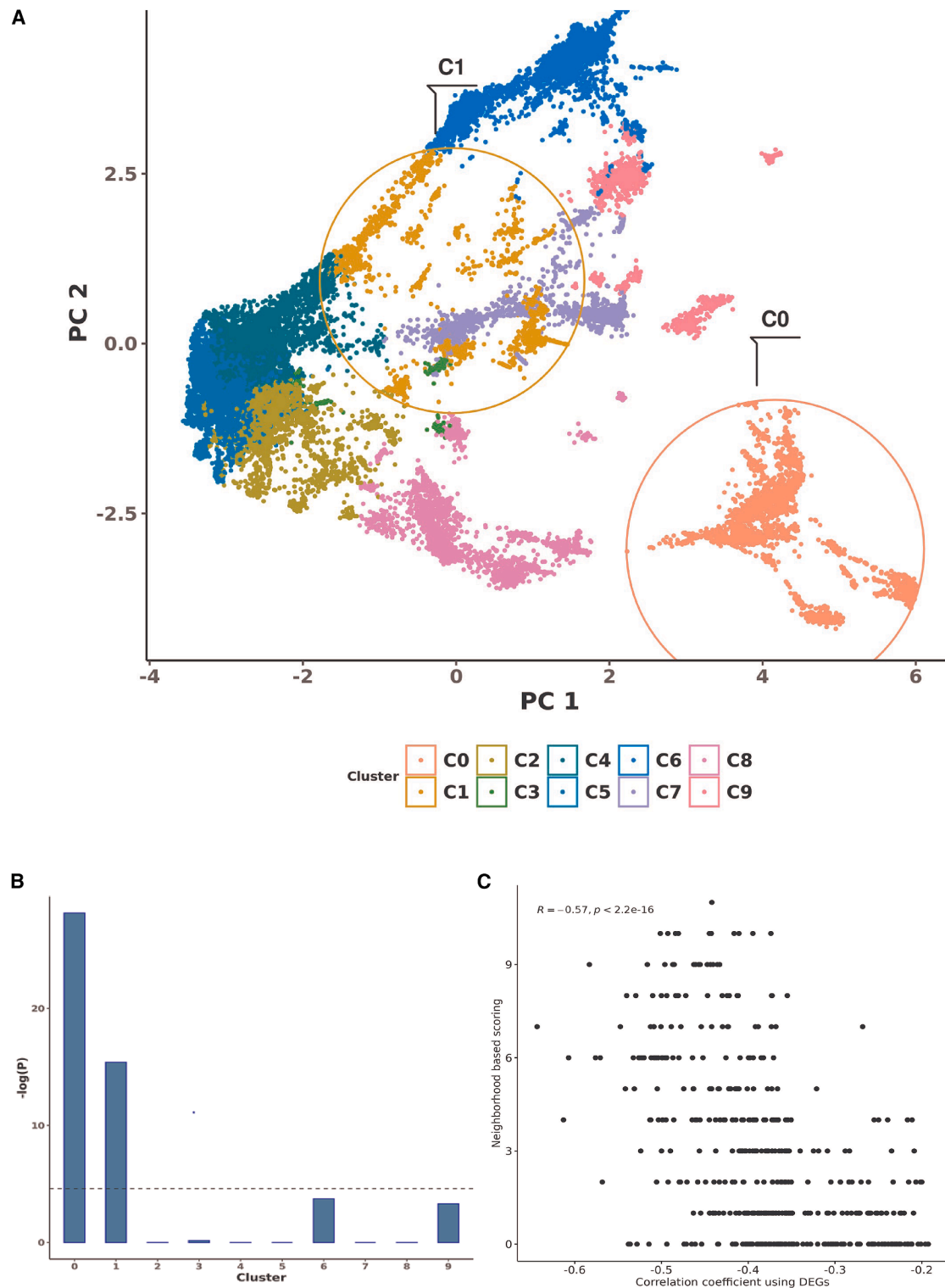


**Figure 4. Global transcriptomic network clustering further validates the identified drug candidates**

(A) Clustering of the top 40 UMAP embeddings calculated for the COVID LINCS dataset. The cluster membership of the different drug treatments is represented using the coordinates from the first two principal components. The encircled clusters show an enrichment for the candidate drug treatments (C0 and C1).

(legend continued on next page)

### Cheminformatic insights for the binding of repurposed compounds to key viral and host factors

To complement our findings that these screened compounds could potentially reverse transcriptomic perturbations during COVID, we conducted cheminformatic docking analysis using the molecular operating environment (MOE) suite on the top 40 candidate compounds by magnitude of their Pearson's correlation reversal scores to examine the potential docking of these top candidate compounds to the binding interface between the SARS-CoV-2 spike protein and hACE2. We hypothesize that such compounds could have dual activity in perturbing SARS-CoV-2 replication and, as such, could be prioritized for further validation as drug candidates. As seen in Figure 5A, the top 40 compounds clustered with MOE docking scores ranging from  $-9$  kcal/mol to  $-5.5$  kcal/mol. Given that there is a range of docking strengths associated with these compounds, these data allow for the prioritization of compounds when considering repurposing in the SARS-CoV-2 biological system, such that both a compound's potential to reverse transcriptomic perturbations and physically interrupt the spike/hACE2 binding interface may be taken into account. Further validation of our cheminformatic approach was to compare the list of compounds identified from the bioinformatic modeling and their docked scores to that of a decoy set of compounds. This decoy set was obtained from the directory of useful decoys (DUD), which is used to optimize docking and structure-based drug design workflows.<sup>42</sup> Comparing the ranking of the identified compounds to that of the DUD decoy set, we found that our identified compounds were ranked and identified as active as compared to the DUD decoy molecules.<sup>43</sup>

We next investigated whether any of the 448 candidates might directly interfere with infection by SARS-CoV-2 by perturbing host-viral protein interactions. To assess this possibility, we performed a cheminformatic analysis focused on the molecules' interaction capabilities with viable sites on host and viral proteins. Interestingly, 5 of the compounds showed direct ability to embed at the interaction site between hACE2 and the SARS-CoV-2 spike protein with low values for their respective free energies of binding ( $\Delta G_{\text{binding}}$ ). To exemplify this, we modeled the docking of Tyrphostin AG ( $\Delta G_{\text{binding}} = -13.34 \pm 3.64$  kcal/mol), Rapamycin ( $\Delta G_{\text{binding}} = -33.04 \pm 5.69$  kcal/mol), doxorubicin ( $\Delta G_{\text{binding}} = -17.53 \pm 4.22$  kcal/mol), and Copansilib ( $\Delta G_{\text{binding}} = -15.55 \pm 3.05$  kcal/mol) to the spike/hACE2 binding interface (Figures 5B–5E). The low values of  $\Delta G_{\text{binding}}$  for respective compounds to bind to the cavity between the SARS-CoV-2 spike protein and hACE2 suggest that these poses are favorable. Stabilizing hydrogen bonding sites were observed in each model. In certain docking models,  $\pi$ - $\pi$  interactions,  $\pi$ -cation interactions, salt bridges, and hydrophobic interactions were also observed to stabilize docking. Such compounds could therefore be prioritized higher for preclinical investigation due to potential dual activities.

Additionally, docking analyses focused on endogenous human proteins also provided some mechanistic insights into the identified compounds. First, doxorubicin was found to interact with SLC39A4 (Figure S4A), a zinc carrier protein. Zinc has been found to be protective in patients with COVID, so preclinical investigations may want to focus on whether the zinc restriction by this compound is driving the observed effects.<sup>44</sup> Additionally, copansilib, a PI3K inhibitor, was also found to interact with MELK (Figure S4B), a key kinase in the cell cycle, suggesting redundant stress in the cell cycle pathway. This may enable further functional analysis of cytotoxicity in future preclinical studies. Overall, the docking analysis provides important complementary insights into our compounds' utility, highlighting potential added anti-SARS-CoV-2 benefits in some compounds as well as underscoring some possible toxicity risk and need for careful dosage in others.

### Network-driven identification of candidate drug combinations

After validating the initial compound screen, we conducted the second part of the RepurNet analysis by building a COVID disease module to assess for possible drug combinations. Using a network degree-aware, diffusion-based strategy, we first built a COVID disease module connecting all 27 viral proteins using proteins in the human interactome.<sup>45</sup> The diffusion strategy used the 27 viral proteins as the inputs and used their projected interaction targets to assess their entry points into the human interactome. Subsequent simulated diffusion yielded a ranked list of genes based on log likelihood of deregulation at that node, and proteins were added to the list in rank order until all 27 viral proteins were connected in a large component of the disease module. This module represented a network-strategized subset of the interactome by which SARS-CoV-2 infection could perturb normal biological processes. The subnetwork was comprised of a large connected component featuring the 27 viral proteins, 349 human proteins, and 576 total edges (Figure 6A; Table S7). Enrichment analysis of the human proteins indicated the module was enriched in genes associated with protein processing in the ER and with endocytosis (Figure S5). These results are supported by prior findings indicating a role for ER stress in SARS-CoV-2 infection and coronavirus infections in general.<sup>46–48</sup>

While our identified compounds showed some impact on the relevant transcriptional profile, possible downstream mechanisms by which these compounds impact the transcriptome could further be elucidated by diffusing their impacts through the interactome, generating a target module containing the affected proteins. Therefore, we repeated the diffusion algorithm described above with the strict human interactome, removing the viral proteins from the network and setting the drug targets as the seed genes. Of the 448 compounds, 364 had drug target information, and of these, 117 were currently

(B) Enrichment analysis is performed to determine cluster membership of the candidate drugs. The hypergeometric test is performed to test for the enrichment of candidate drug treatments in individual clusters. The  $p$ -values are converted to the negative of the natural log, and the FDR is controlled at 1% (The dashed horizontal line).

(C) Treatment-associated nearest neighbor score has a negative correlation with the coefficient correlation calculated to quantify the perturbation caused by the candidate drug treatment (CDT) on the differentially expressed genes between COVID (ICU) v/s COVID (non-ICU) patients. The top 20 nearest neighbors are considered for calculating the score. Spearman's correlation coefficient is calculated for the two metrics. See also Figure S3.

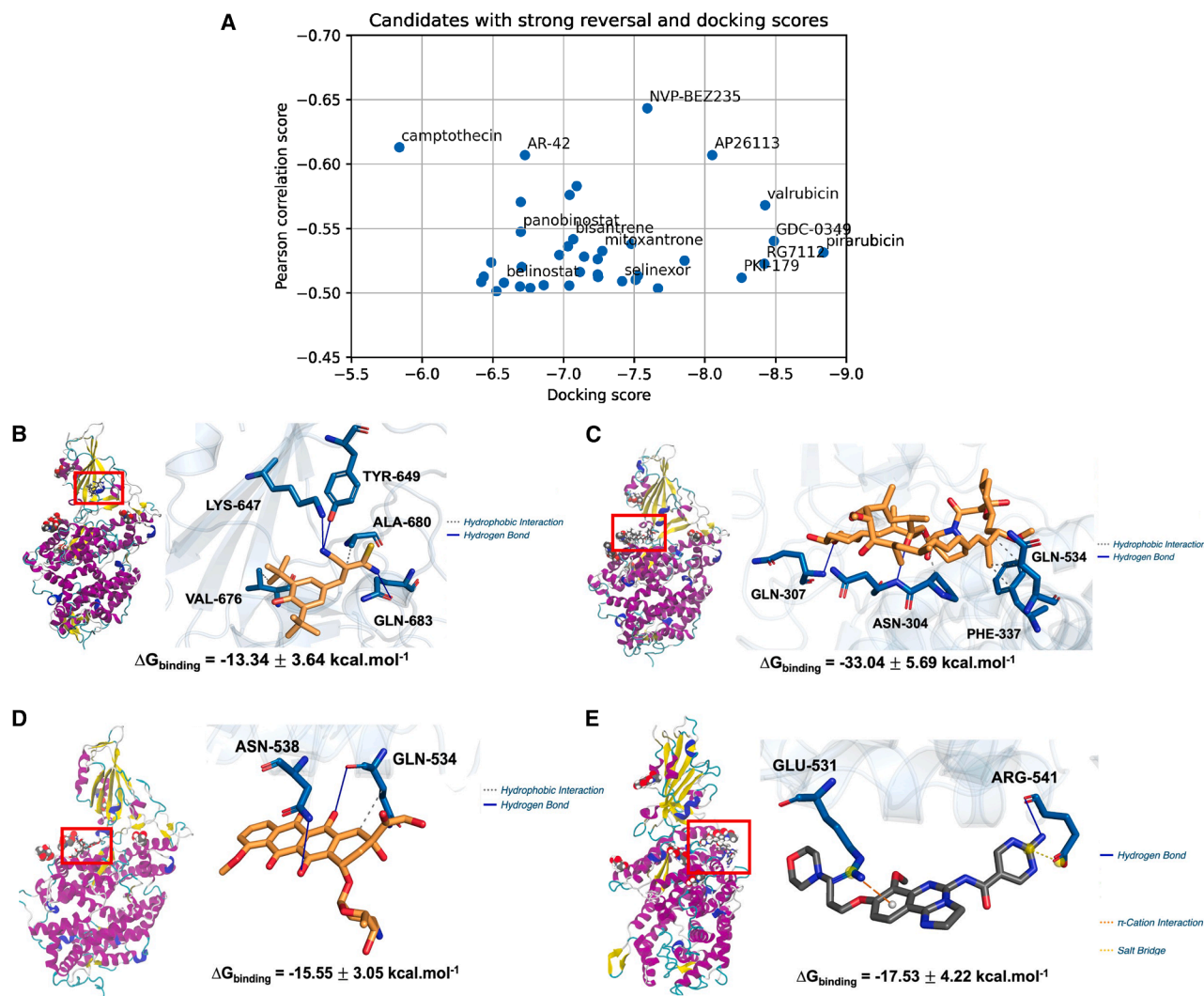


**Figure 5. Cheminformatic analysis on the docking of our identified compounds to key viral and host factors**

(A) Docking results for our top 40 identified compounds by Pearson's coefficient reversal score. As illustrated, our most highly prioritized compounds show variable tendency to interfere with the interaction site between the SARS-CoV-2 Spike protein and the host ACE2 receptor.

(B) Visualization of tyrphostin AG's ability to interfere with the SARS-CoV-2 spike protein and hACE2 interaction. The docking is stabilized with a calculated free energy of binding of  $-13.34 \pm 3.64 \text{ kcal mol}^{-1}$ .

(C) Visualization of rapamycin's ability to interfere with the SARS-CoV-2 spike protein and hACE2 interaction. The docking is stabilized with a calculated free energy of binding of  $-33.04 \pm 5.69 \text{ kcal mol}^{-1}$ .

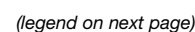
(D) Visualization of doxorubicin's ability to interfere with the SARS-CoV-2 spike protein and hACE2 interaction. The docking is stabilized with a calculated free energy of binding of  $-15.55 \pm 3.05 \text{ kcal mol}^{-1}$ .

(E) Visualization of copansilib's ability to interfere with the SARS-CoV-2 spike protein and hACE2 interaction. The docking is stabilized with a calculated free energy of binding of  $-17.53 \pm 4.22 \text{ kcal mol}^{-1}$ . (B–E) Stabilizing interactions for each docking model are indicated to the right of each model. See also Figure S4.

FDA-approved. We prioritized these 117 compounds to assess the safety of current co-treatments. Instead of attempting to build a target module for each of these compounds, we subset the diffusion results to the log likelihood values of the 349 human proteins in the disease module. In the search for candidate drug pairs, we sought to prioritize pairs that show good separation within the disease module. This property had been associated with a lower adverse effect profile, supporting safety for the proposed combinations.<sup>49</sup> To assess separation between a pair of compounds, we calculated an L1 norm value between

their log likelihood score vectors, with higher L1 norm values representing greater separation.

Clustering of L1 normalization values revealed two distinct groups of compounds in the examined set of candidates (Figure 6B; Table S8). The smaller cluster, occupying the left side of the heatmap, consisted of compounds that showed moderate to strong separation from the bulk of the evaluated compounds. By contrast, the larger cluster of compounds showed good separation from compounds in the smaller cluster but poor separation within the larger cluster. Interestingly, the larger



cluster contained many of the compounds identified in the top 8 drug classes in the drug reversal analysis, such as dexamethasone, mitoxantrone, and pirarubicin. By contrast, the smaller cluster contained compounds with strong R1 values that possessed other potential roles in treating COVID. We examined several intriguing compound combinations based on the L1 norm value in greater detail. In particular, we examined combinations of the 31 compounds in the small cluster with two subgroups of the larger cluster. One subgroup featured steroid medications such as dexamethasone and desonide (Figure 6C). The other subgroup featured topoisomerase inhibitors such as etoposide, mitoxantrone, and doxorubicin (Figure 6D). For the steroid group, the combinations of dexamethasone-sevoflurane and dexamethasone-homoharringtonine were selected for visualization in the disease module. In the pictured networks, node size and color intensity represent the differences in log likelihood of perturbation by each treatment. Blue nodes were more likely to be perturbed by the effects of dexamethasone, whereas red nodes were more likely to be affected by sevoflurane or homoharringtonine. In both combinations, dexamethasone was shown to target a similar core portion of the disease module. By contrast, homoharringtonine and sevoflurane targeted distinct parts of the module that were both outside of the area affected by dexamethasone. Similarly, two combinations involving mitoxantrone were visualized (mitoxantrone-ticagrelor and mitoxantrone-ouabain) in the disease module. The combination of mitoxantrone and ticagrelor had the highest L1 norm of the four combinations visualized, and this was reflected in the numerous red and blue nodes seen in the disease module that indicated strong separation of how the network was targeted. By contrast, the combination of mitoxantrone and ouabain showed less separability, indicated by the disappearance of several prominent blue nodes throughout the module. Altogether, the visualizations support that a higher L1 norm reflects stronger separation between drug pairs.

The compounds highlighted in the smaller cluster are intriguing for specific use cases in patients with COVID, and early clinical findings support their therapeutic utility.<sup>50–54</sup> Ticagrelor and ouabain, which registered effective separation with mitoxantrone and other topoisomerase inhibitors, had been associated with improved mortality rates and *in vitro* antiviral activity, respectively.<sup>50,51</sup> Homoharringtonine, which showed the potential for combination with dexamethasone, had been documented to interfere with SARS-CoV-2 replication and TMPRSS2 expression *in vitro*; the combination could therefore confer specific antiviral activity coupled with dexamethasone's general anti-inflammatory activity.<sup>52,53</sup> Finally, sevoflurane has been explored as an

anesthetic for ventilated patients with ARDS symptoms, so its compatibility with dexamethasone might support a combination in ventilated patients.<sup>54</sup>

### Shared molecular similarities and clinical implications between COVID and cancer

Previously, we noted that many positively enriched pathways in patients with COVID were related to cell cycle processes, suggesting significant upregulation in cell cycle activity in patients with COVID compared to other patients. Moreover, clinical evidence has suggested that patients with cancer infected with COVID display more severe symptoms and have a higher mortality rate compared to patients with COVID without a cancer diagnosis.<sup>55,56</sup> As such, we were curious as to whether patients with cancer could be predisposed to severe COVID due to already having a gene expression background characterized by upregulated cell cycle activity. Firstly, we investigated the genetic similarities between COVID and cancer by comparing our COVID vs. nonCOVID DEG profile and enriched gene sets in patients with COVID to those of lung adenocarcinoma (LUAD) and LUSC. We performed differential gene expression analysis on HTSeq gene expression count matrices provided by TCGA for each cancer type to obtain their corresponding DEG profiles (Table S9). We observed significant concordance (particularly in concordantly upregulated genes) between the COVID vs. nonCOVID DEG profile and the LUAD (max  $-\log_{10}(p\text{-value}) = 60.2$ ) and LUSC DEG profiles (max  $-\log_{10}(p\text{-value}) = 60.2$ ) after conducting RRHO analysis (Figure 7A). Given that patients with cancer have a higher susceptibility to COVID and experience more severe disease, these strong concordances in gene expression suggest possible molecular mechanisms for cancer patients' increased risk to COVID.<sup>57,58</sup> We further explored whether DEGs in the COVID vs. nonCOVID profile were also found to be highly relevant in LUAD or LUSC by cross-referencing those DEGs with genes found in the Broad Institute's MutSigCV, which documents genes mutated more often than by chance in various cancer types.<sup>59</sup> We noted that *MET*, which was highly mutated in LUAD, was also differentially overexpressed in COVID.

We also tested for enrichment of the previously mentioned gene set collections in TCGA-LUAD and TCGA-LUSC and subsequently examined which of their significantly enriched gene sets (FDR-adjusted  $p\text{-value} < 0.05$ ) were also enriched in COVID (Table S10). We observed that 15 pathways from the Hallmark, KEGG, PID, and WikiPathways collections were positively enriched in both COVID and either TCGA-LUAD or TCGA-LUSC (Figure 7B). Notably, each of these functional pathways corresponded to cell cycle or DNA repair pathways, suggesting that

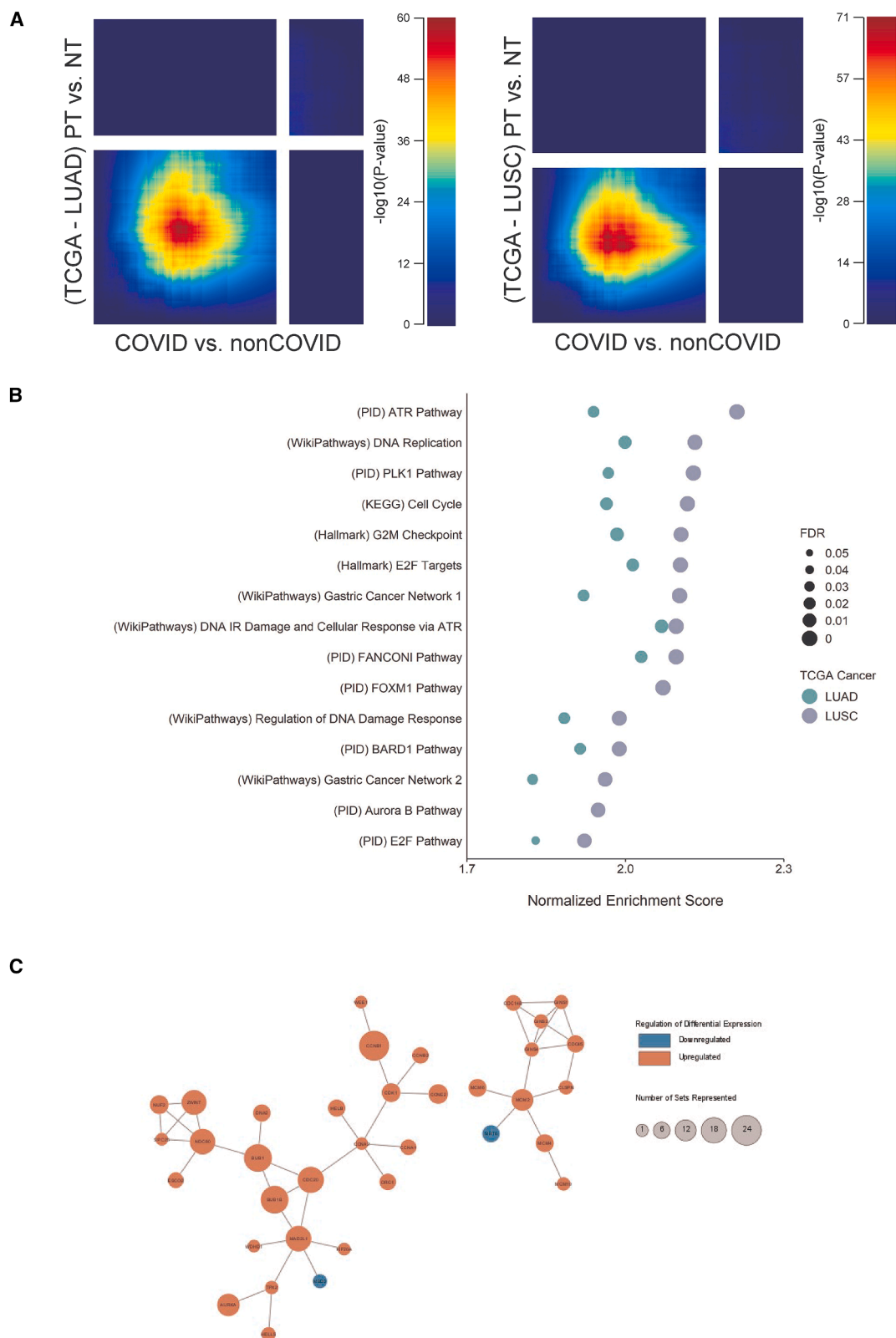
**Figure 6. Network medicine and disease module approach for identification of candidate drug combinations**

(A) The connected COVID disease module. Red nodes represent SARS-CoV-2 viral proteins while blue nodes represent host proteins. Nodes are sized by degree in the subnetwork.

(B) The L1 normalization heatmap. Red squares indicate compounds with increased separation in the network while blue squares indicate compounds with more similar activities.

(C) Subset of heatmap focusing on favorable partners for corticosteroid treatments. Network profiles for dexamethasone combinations with homoharringtonine and sevoflurane are visualized.

(D) Another subset of the heatmap focuses on favorable partners for topoisomerase treatments. Network profiles for mitoxantrone combinations with ticagrelor and ouabain are visualized. Of the highlighted networks, the ticagrelor-mitoxantrone combination was calculated to have the highest separation by L1 normalization. This is reflected in the increased quantity of large red or blue nodes in this network compared to the ouabain-mitoxantrone and dexamethasone combinations networks and suggests that a combination of ticagrelor and mitoxantrone may be better tolerated. See also Figure S5 and Tables S7 and S8.



(legend on next page)

SARS-CoV-2 induced similar changes to the cell cycle that were observed in tumor cells of LUAD or LUSC. To further investigate this idea, we used a network-based approach to examine DEGs from the COVID vs. nonCOVID DEG profile that were found in 54 GOBP terms (all of which also corresponded to cell cycle or DNA repair functions) enriched in COVID and either TCGA-LUAD or TCGA-LUSC. We subset the BioPlex 3.0 Interaction Network (293T cells) of the human interactome for DEGs in the COVID vs. nonCOVID profile that were also included in any of those 54 GOBP terms, and we identified two highly connected subnetworks (Figure 7C).<sup>60</sup> These two subnetworks could therefore highlight relevant genes that reveal how SARS-CoV-2 hijacks hosts' cellular machinery to replicate.<sup>61</sup>

To explore similarities with other cancers, we also expanded our analysis to include all 33 TCGA cancers, 21 of which had 3 or more normal tissue samples for differential gene expression analysis (DEG profiles in Table S9). After conducting RRHO analysis to compare concordance, the 21 TCGA cancers' DEG profiles to our COVID vs. nonCOVID DEG profile, we found that 19 TCGA cancers' DEG profiles showed strong concordance with our COVID vs. nonCOVID DEG profile ( $71.11 \geq \max -\log_{10}(p\text{-value}) \geq 8.43$  in concordant upregulated quadrants) (Figure S6). After using GSEA to examine whether these 19 TCGA cancers had any shared enriched pathways with COVID, we found that 123 unique gene sets were shared between any of the 19 TCGA cancers and COVID (Table S10). These data suggest that cancer treatments are a potential pool of drug targets for repurposing, as they may target similar genes and pathways common to COVID in the periphery and to cancer, and such drugs may be of interest for further investigation into whether they are beneficial for treating patients with cancer diagnosed with COVID.

We then obtained a pan-cancer network of putative oncogenic protein-protein interactions (onco-PPIs) from Cheng et al. (2020) to examine onco-PPIs that might be associated with gene dysregulation in COVID.<sup>62</sup> This putative onco-PPI network described interaction interfaces that were significantly enriched for mutations in cancer, so we selected this network for analysis to investigate overlaps specifically between mutated genes in cancer and dysregulated genes in COVID. After selecting for all DEGs in the COVID vs. nonCOVID DEG profile, we identified five subnetworks whose genes might be important for lung health and could be potential targets for drug development aimed at treating lung-related diseases (Figure S7).

## DISCUSSION

By analyzing whole blood leukocyte mRNA expression data from either ICU or nonICU patients, either with or without a COVID

diagnosis, we developed three gene regulatory signatures that characterize DEGs in patients with COVID, each from a different perspective. These DEG profiles enabled further analysis in examining significant changes in immune cell proportions and dysregulated gene sets relevant to COVID. We selected the COVID ICU vs. nonCOVID ICU DEG profile for further analysis, from which we identified potential drug clusters that could probably help treat severe COVID. These drug clusters may also serve as a potential guideline for screening drugs that could help with developing COVID therapies.

We recognize that the nonCOVID ICU patient group was not an ideal reference point for our analysis, as these patients are not healthy. As such, it is possible that our comparison could result in the search for drugs that convert the COVID disease state to another disease state. However, when we clustered samples in the COVID ICU vs. nonCOVID ICU profile based on the expression of the respective DEGs, the COVID ICU samples clustered separately from the nonCOVID ICU samples, suggesting that the perturbed gene expression was indeed a consequence of COVID (Figures S8A and S8B). Importantly, the nonCOVID ICU samples actually clustered closer together than did the COVID ICU samples. If these DEGs were associated with other disease states, the nonCOVID ICU samples comprising heterogeneous disease backgrounds would not cluster closely together, especially relative to the clustering of the COVID ICU samples. We observed a similar consistency in clustering of samples based on the gene expression of respective DEGs in the COVID vs. nonCOVID profile (Figure S7B). As such, we are confident that the DEGs derived from the COVID ICU vs. nonCOVID ICU profile are perturbed as a result of a severe COVID diagnosis, and the nonCOVID ICU patient group may therefore serve as a reference point to search for drugs that can reverse the COVID disease state.

Moreover, our gene regulatory network analysis results were consistent with those found in the literature. For example, we observed that myeloid dendritic cell and NK cell proportions were significantly decreased in the whole blood of COVID ICU patients compared to COVID nonICU patients, which is consistent with findings that conventional dendritic cells and NK cells are significantly decreased in severe COVID cases compared to healthy controls.<sup>63–66</sup> Additionally, our observation that CD8<sup>+</sup> T cells are drastically depleted in ICU patients with COVID is consistent with previous observations of lymphopenia in severe COVID cases.<sup>63,65,66</sup> We also identified multiple dysregulated p53 pathway genes in our COVID vs. nonCOVID and COVID ICU vs. nonCOVID ICU profiles, which supports the hypothesis that SARS-CoV-2 infection causes depletion of CD8<sup>+</sup> T cells through apoptosis.<sup>67</sup> Similar to our observations, other

### Figure 7. Molecular characteristics and comparative analyses of COVID and cancer

(A) RRHO heatmaps summarize the concordance of TCGA-LUAD and TCGA-LUSC's DEG profiles to the COVID vs. nonCOVID DEG profile. Concordance between the TCGA-LUAD and COVID vs. nonCOVID DEG profiles has a max  $-\log_{10}(p\text{-value}) = 60.2$ , and concordance between the TCGA-LUSC and COVID vs. nonCOVID DEG profiles has a max  $-\log_{10}(p\text{-value}) = 71.1$ . Heatmaps are interpreted as demonstrated in Figure 2B.

(B) Dotplot summarizes the 15 enriched gene sets/pathways that are overlapped between the COVID vs. nonCOVID profile and either the TCGA-LUAD or TCGA-LUSC profile.

(C) A subnetwork from the BioPlex 3.0 Interactions network that only includes genes that are differentially expressed in the COVID vs. nonCOVID DEG profile and that are included in any of the 54 shared enriched GOBP terms. Genes are annotated by whether they are significantly upregulated or downregulated in the COVID vs. nonCOVID DEG profile and by the number of GOBP terms they are part of. See also Figures S6–S8 and Tables S9 and S10.

transcriptomic analyses of peripheral blood samples have identified positive enrichment of cell cycle pathways and negative enrichment of cytokine-related pathways.<sup>67</sup> In regards to cytokine-related pathways, none of our profiles include significant upregulation of genes associated with the cytokine storm characterized in COVID (*IL-6*, *IL-2*, *IL-7*, *IL-10*, *TNF- $\alpha$* , *CSF2*, *CXCL10*, *CCL2*, or *CCL3*), and this observation is consistent with the hypothesis that elevated plasma cytokines in COVID originate from leukocytes in lung tissue rather than peripheral leukocytes in blood, so the lack of differential expression of these genes is not surprising.<sup>17,64,68</sup> Observed discordances between the blood-derived COVID ICU vs. nonCOVID ICU profile and the BALF-derived Severe vs. Healthy profile are also consistent with the observation that the immune response to COVID differs between peripheral blood and BALF.<sup>20</sup>

Interestingly, the positive enrichment of the Acute vs. Chronic LCMV-infected Primary and Secondary CD8<sup>+</sup> T cell (down) immunologic gene sets from GSE30962 suggests that the gene profiles of both peripheral CD8<sup>+</sup> naive and memory T cells during COVID resemble that of a chronic infection rather than that of an acute infection.<sup>34</sup> Investigating a chronic infection model could help study CD8<sup>+</sup> T cell lymphopenia in COVID pathology, as CD8<sup>+</sup> T cell exhaustion occurring during chronic antigen stimulation can lead to decreased numbers of CD8<sup>+</sup> T cells.<sup>34,69</sup> Taken together with our observations of significantly decreased CD8<sup>+</sup> T cell proportions in COVID ICU patients compared to nonCOVID ICU patients, the DEGs enriching these gene sets could help understand T cell lymphopenia in COVID, and manipulating the regulation of these genes (Figure 2E) could help improve COVID patient outcome, as lymphopenia is correlated with severe COVID cases.<sup>65,66</sup>

Using the RepurNet framework, we prioritized several intriguing drug candidates based on reversal of the gene expression signatures identified in SARS-CoV-2-infected samples. Three classes of drug classes – HDAC inhibitors, CDK inhibitors, and topoisomerase inhibitors – are enriched among the top-scoring drug candidates. These three drug classes target mechanisms associated with cell cycle progression, further highlighting a central role for cell cycle machinery in producing the COVID phenotype. Importantly, this phenotype is derived from patient samples rather than *in vitro* infection with SARS-CoV-2, and such patient samples produce cellular signals in response to both viral infection and the systemic physiological response to viral infection.<sup>70</sup> As such, our identified compounds aim to stabilize cell response to both viral infection and signs of physiological stress in COVID, such as hypoxia and inflammatory cytokines.<sup>70</sup> Several additional candidates are also found to show effective binding affinity to the interaction interface between SARS-CoV-2 viral proteins and host proteins, indicating an alternative strategy to combat the COVID phenotype. It is worth noting that the docking potential for drug candidates will differ between variant strains of SARS-CoV-2, especially if relevant surface residues of the spike protein are mutated in variant strains. As our drug candidates' gene reversal profiles were identified from clinical data derived from patients infected with canonical SARS-CoV-2, we used the canonical SARS-CoV-2 spike protein bound to hACE2 for our docking analyses to explore the potential for drug candidates to perturb spike protein binding to

hACE2. However, structures of mutant spike protein subunits, such as those from the Alpha and Beta variants of SARS-CoV-2, binding to hACE2 have been uploaded to the Protein DataBank, so future investigation of specific potential drug candidate(s) may be evaluated with different SARS-CoV-2 variants using this strategy.<sup>71</sup>

Integration of the SARS-CoV-2 disease interactome and the human interactome has further enabled us to propose a connected disease module connecting SARS-CoV-2 viral proteins to the drug candidates identified based on observed gene expression changes. This module was then used to conduct the second analysis of the RepurNet framework and create a preliminary prioritization of drug combinations based on previous work in network biology. A key limitation of this approach is that the use of pure protein-protein interaction network modules in assessing a drug's impact on an interactome misses possible contributions from off-target interactions between the drug and protein. However, numerous studies in network biology have supported that network expansion from a set of seed genes using a PPI network can help prioritize validated compounds and targets. As such, the proposed drug combinations still represent an appropriate prioritized set of drug combinations for further study. Given the extensibility of the RepurNet framework, we have made the scripts available online with representative input and output files, and the scripts can be flexibly repurposed for other disease states and networks.

Because our blood-derived DEG profiles and previous transcriptomic analyses of SARS-CoV-2 infection have revealed the upregulation of cell cycle pathways in blood leukocytes, we hypothesize that COVID could have similar alterations in gene expression as cancer, which is highly characterized by unrestricted cell division.<sup>20,63,72</sup> Our focused study into the similarities of altered gene expression between COVID and LUSC or LUAD also reveals many dysregulated genes in COVID that may provide insight into the observed increased susceptibility of patients with cancer to SARS-CoV-2 infection.<sup>73,74</sup> For example, *MET* is both highly mutated in LUAD and differentially overexpressed in COVID, and *MET* mutations cause aberrantly increased signaling that can increase cell proliferation, protect against apoptosis, and drive metastasis in cancer cells.<sup>75–77</sup> Our findings that COVID and various types of cancer share similar DEG profiles and upregulated pathways indicate the potential of using cancer disease models to understand COVID pathogenesis. Previous studies have already identified similar signaling pathways between sites of local SARS-CoV-2 infection and tumor microenvironments of cancer.<sup>73</sup> Previous experimental exploration of HDAC, CDK, and topoisomerase inhibitors, which are normally used to treat cancer, as rational therapeutic agents for COVID, especially highlights the potential of similar disease mechanisms between COVID in the periphery and cancer.<sup>78–80</sup>

### Limitations of the study

Though our strategy confidently predicts rational drug combinations of interest in drug repurposing, there are a few limitations to these predictions. Firstly, we did not experimentally validate these predictions. Ultimately, applications of our strategy in drug discovery would require experimental validation of

predicted drugs of interest. Secondly, our strategy intrinsically restricts the search space of drugs to those with documented gene expression profiles in one or more cell lines. However, this is a minor limitation given the sheer scope of this kind of data available on the Connectivity Map (<https://clue.io/data>) as their cumulative datasets include over 80,000 perturbagens associated with over 1 million gene profiles spanning various cell lines. Thus, our strategy remains highly generalizable despite some intrinsic limitations.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Stephen Yi ([song.yi@bcm.edu](mailto:song.yi@bcm.edu)).

### Materials availability

No unique materials were generated in this study.

### Data and code availability

- All associated data with this paper are listed in the [key resources table](#). No new data has been generated by this study. All data used are publicly available, as listed in the [key resources table](#).
- All original code has been deposited on GitHub (<https://github.com/Stephen-Yi-Laboratory/RepurNet>). The release version of the software (v1.0.0) is publicly available at Zenodo (<https://doi.org/10.5281/zenodo.18881178>) as of the date of publication, through this site: <https://zenodo.org/records/18881178>.
- Requests for further information may be directed to the [lead contact](#), Stephen Yi ([song.yi@bcm.edu](mailto:song.yi@bcm.edu)).

## ACKNOWLEDGMENTS

The authors are grateful for contributions from the TCGA Research Network Analysis Working Group, and also acknowledge the High-Performance Computing at Texas Advanced Computing Center (TACC) for assistance. Computational resources for cheminformatic docking analyses were provided by the WVU Research Computing Thorny Flat HPC cluster, which is funded in part by NSF OAC-1726534.

Grant Support: We would like to acknowledge the National Institutes of Health grant R35GM133658 (to S.Y.), R33CA281919 (to S.Y.), R35GM137836 (to N.S.), and R35GM128625 (to J.I.F.B.). S.Y. is an Affiliate Member of the NHGRI IGVF (Impact of Genomic Variation on Function) Consortium, a Partner Member of the NHGRI GREGoR (Genomics Research to Elucidate the Genetics of Rare diseases) Consortium, and an Associate Member of the NCI Cancer Systems Biology Consortium (CSBC). We also acknowledge the support from BioRender in figure generation.

## AUTHOR CONTRIBUTIONS

S.R.P., R.Z., N.S., and S.Y. conceived the study. S.R.P. and R.Z. conducted most of the computational modeling and analyses, with input from P.C., A.S., M.B.B., S.Y., and N.S. W.P., W.J.G., M.S.K., J.H.H., J.I.F.B., and J.D. provided intellectual input and constructive feedback. S.Y. and N.S. provided supervision throughout the course of the study. S.R.P., R.Z., N.S., and S.Y. wrote the manuscript with input from E.W., J.I.F.B., and J.D. All authors read and approved the manuscript.

## DECLARATION OF INTERESTS

S.S.Y., S.R.P., and N.S. disclose a patent filed under this technology (U.S. Provisional Application No. 63/627,196). All other authors declare that they have no conflicts of interest.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **METHOD DETAILS**
  - Regulatory network gene expression signature analysis
  - Immune cell type deconvolution
  - Gene set enrichment analysis (GSEA)
  - Comparison of blood-derived DEG profiles to BALF-Derived DEG profiles
  - Gene regulatory network comparison of COVID to TCGA cancer profiles
  - Drug reversal calculations with the LINCS1000 platform
  - Clustering on UMAP embeddings and nearest-neighbor based scoring
  - Cheminformatic structural docking analysis
  - COVID disease module construction
  - Network-guided rational design of drug combinations
  - Gene regulatory network analysis of similarities between cancers and COVID
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2026.115896>.

Received: March 22, 2023

Revised: July 27, 2025

Accepted: April 22, 2026

Published: April 27, 2026

## REFERENCES

1. Vosoughi, P., Naghib, S.M., Rahmadian, M., and Mozafari, M.R. (2025). A Comprehensive Review of mRNA-based Vaccines for COVID-19, A New Era in Pharmaceuticals: Unspecified and Unknown Aspects, Effects and Challenges. *Curr. Top. Med. Chem.* 25, 1467–1491. <https://doi.org/10.2174/0115680266325847241121034100>.
2. Saxena, S., Mandrah, V., Tariq, W., Das, P., Sambhav, K., and Devi, S.H. (2025). The Future of mRNA Vaccines: Potential Beyond COVID-19. *Cureus* 17, e84529. <https://doi.org/10.7759/cureus.84529>.
3. Mambelli, F., de Araujo, A.C.V.S.C., Farias, J.P., de Andrade, K.Q., Ferreira, L.C.S., Minoprio, P., Leite, L.C.C., and Oliveira, S.C. (2025). An Update on Anti-COVID-19 Vaccines and the Challenges to Protect Against New SARS-CoV-2 Variants. *Pathogens* 14, 23. <https://doi.org/10.3390/pathogens14010023>.
4. Kabinger, F., Stiller, C., Schmitzová, J., Dienemann, C., Kokic, G., Hillen, H.S., Höbartner, C., and Cramer, P. (2021). Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nat. Struct. Mol. Biol.* 28, 740–746. <https://doi.org/10.1038/s41594-021-00651-0>.
5. Niraj, N., Mahajan, S.S., Prakash, A., Sarma, P., and Medhi, B. (2022). Paxlovid: A promising drug for the challenging treatment of SARS-CoV-2 in the pandemic era. *Indian J. Pharmacol.* 54, 452–458. [https://doi.org/10.4103/ijp.ijp\\_291\\_22](https://doi.org/10.4103/ijp.ijp_291_22).
6. Gottlieb, R.L., Vaca, C.E., Paredes, R., Mera, J., Webb, B.J., Perez, G., Oguchi, G., Ryan, P., Nielsen, B.U., Brown, M., et al. (2022). Early Remdesivir to Prevent Progression to Severe Covid-19 in Outpatients. *N. Engl. J. Med.* 386, 305–315. <https://doi.org/10.1056/NEJMoa2116846>.
7. Mohammadi, D., Ghasemi, M., Manouchehrian, N., Zafarmand, M., Akbari, M., and Boroumand, A.B. (2024). COVID-19 vaccines: current and future challenges. *Front. Pharmacol.* 15, 1434181. <https://doi.org/10.3389/fphar.2024.1434181>.

8. Zhu, C., Pang, S., Liu, J., and Duan, Q. (2024). Current Progress, Challenges and Prospects in the Development of COVID-19 Vaccines. *Drugs* 84, 403–423. <https://doi.org/10.1007/s40265-024-02013-8>.
9. Sultana, J., Crisafulli, S., Gabbay, F., Lynn, E., Shakir, S., and Trifirò, G. (2020). Challenges for Drug Repurposing in the COVID-19 Pandemic Era. *Front. Pharmacol.* 11, 588654. <https://doi.org/10.3389/fphar.2020.588654>.
10. Beigel, J.H., Tomashek, K.M., Dodd, L.E., Mehta, A.K., Zingman, B.S., Kallil, A.C., Hohmann, E., Chu, H.Y., Luetkemeyer, A., Kline, S., et al. (2020). Remdesivir for the Treatment of Covid-19 — Final Report. *N. Engl. J. Med.* 383, 1813–1826. <https://doi.org/10.1056/NEJMoa2007764>.
11. Coronavirus (COVID-19) Drugs (2021). US Food and Drug Administration. 2021. [https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs#:~:text=Veklury%20\(Remdesivir\)%20is,-19%20requiring%20hospitalization](https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs#:~:text=Veklury%20(Remdesivir)%20is,-19%20requiring%20hospitalization).
12. Singh, T.U., Parida, S., Lingaraju, M.C., Kesavan, M., Kumar, D., and Singh, R.K. (2020). Drug repurposing approach to fight COVID-19. *Pharmacol. Rep.* 72, 1479–1508. <https://doi.org/10.1007/s43440-020-00155-6>.
13. Doto, S., Marabotti, A., Facchiano, A., and Tagliaferri, R. (2021). A review on drug repurposing applicable to COVID-19. *Brief. Bioinform.* 22, 726–741. <https://doi.org/10.1093/bib/bbaa288>.
14. Egieyeh, S., Egieyeh, E., Malan, S., Christofells, A., and Fielding, B. (2021). Computational drug repurposing strategy predicted peptide-based drugs that can potentially inhibit the interaction of SARS-CoV-2 spike protein with its target (human ACE2). *PLoS One* 16, e0245258. <https://doi.org/10.1371/journal.pone.0245258>.
15. Saha, S., Halder, A.K., Bandyopadhyay, S.S., Chatterjee, P., Nasipuri, M., Bose, D., and Basu, S. (2022). Drug repurposing for COVID-19 using computational screening: Is Fostamatinib/R406 a potential candidate? *Methods* 203, 564–574. <https://doi.org/10.1016/j.ymeth.2021.08.007>.
16. Overmyer, K.A., Shishkova, E., Miller, I.J., Balnis, J., Bernstein, M.N., Peters-Clarke, T.M., Meyer, J.G., Quan, Q., Muehlbauer, L.K., Trujillo, E.A., et al. (2021). Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Syst.* 12, 23–40.e7. <https://doi.org/10.1016/j.cels.2020.10.003>.
17. Arunachalam, P.S., Wimmers, F., Mok, C.K.P., Perera, R.A.P.M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., et al. (2020). Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* 369, 1210–1220. <https://doi.org/10.1126/science.abc6261>.
18. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. <https://doi.org/10.1093/nar/28.1.27>.
19. Cahill, K.M., Huo, Z., Tseng, G.C., Logan, R.W., and Seney, M.L. (2018). Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. *Sci. Rep.* 8, 9588. <https://doi.org/10.1038/s41598-018-27903-2>.
20. Xu, G., Qi, F., Li, H., Yang, Q., Wang, H., Wang, X., Liu, X., Zhao, J., Liao, X., Liu, Y., et al. (2020). The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing. *Cell Discov.* 6, 73. <https://doi.org/10.1038/s41421-020-00225-2>.
21. Plaisier, S.B., Taschereau, R., Wong, J.A., and Graeber, T.G. (2010). Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res.* 38, e169. <https://doi.org/10.1093/nar/gkq636>.
22. Liao, M., Liu, Y., Yuan, J., Wen, Y., Xu, G., Zhao, J., Cheng, L., Li, J., Wang, X., Wang, F., et al. (2020). Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* 26, 842–844. <https://doi.org/10.1038/s41591-020-0901-9>.
23. Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., Quan, S., Zhang, F., Sun, R., Qian, L., et al. (2020). Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* 182, 59–72.e15. <https://doi.org/10.1016/j.cell.2020.05.032>.
24. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 7, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
25. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>.
26. Schaefer, C.F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K.H. (2009). PID: the Pathway Interaction Database. *Nucleic Acids Res.* 37, D674–D679. <https://doi.org/10.1093/nar/gkn653>.
27. Croft, D., O’Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., Caudy, M., Garapati, P., Gopinath, G., Jassal, B., et al. (2011). Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39, D691–D697. <https://doi.org/10.1093/nar/gkq1018>.
28. Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D.N., Hanspers, K., A Miller, R., Digles, D., Lopes, E.N., Ehrhart, F., et al. (2021). WikiPathways: connecting communities. *Nucleic Acids Res.* 49, D613–D621. <https://doi.org/10.1093/nar/gkaa1024>.
29. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. <https://doi.org/10.1038/75556>.
30. Gene Ontology Consortium (2021). The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res.* 49, D325–D334. <https://doi.org/10.1093/nar/gkaa1113>.
31. Godec, J., Tan, Y., Liberzon, A., Tamayo, P., Bhattacharya, S., Butte, A.J., Mesirov, J.P., and Haining, W.N. (2016). Compendium of Immune Signatures Identifies Conserved and Species-Specific Biology in Response to Inflammation. *Immunity* 44, 194–206. <https://doi.org/10.1016/j.immuni.2015.12.006>.
32. Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. <https://doi.org/10.1038/ng1180>.
33. Canman, C.E., Lim, D.S., Cimprich, K.A., Taya, Y., Tamai, K., Sakaguchi, K., Appella, E., Kastan, M.B., and Siliciano, J.D. (1998). Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science* 281, 1677–1679. <https://doi.org/10.1126/science.281.5383.1677>.
34. West, E.E., Youngblood, B., Tan, W.G., Jin, H.-T., Araki, G., Alexe, G., Koniczny, B.T., Calpe, S., Freeman, G.J., Terhorst, C., et al. (2011). Tight Regulation of Memory CD8+ T Cells Limits Their Effectiveness during Sustained High Viral Load. *Immunity* 35, 285–298. <https://doi.org/10.1016/j.immuni.2011.05.017>.
35. Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* 171, 1437–1452.e17. <https://doi.org/10.1016/j.cell.2017.10.049>.
36. Group, R.C., Horby, P., Lim, W.S., Emberson, J.R., Mafham, M., Bell, J.L., Linsell, L., Staplin, N., Brightling, C., Ustianowski, A., et al. (2021). Dexamethasone in Hospitalized Patients with Covid-19. *N. Engl. J. Med.* 384, 693–704. <https://doi.org/10.1056/NEJMoa2021436>.
37. Tomazini, B.M., Maia, I.S., Cavalcanti, A.B., Berwanger, O., Rosa, R.G., Veiga, V.C., Avezum, A., Lopes, R.D., Bueno, F.R., Silva, M.V.A.O., et al. (2020). Effect of Dexamethasone on Days Alive and Ventilator-Free in Patients With Moderate or Severe Acute Respiratory Distress Syndrome and COVID-19: The CoDEX Randomized Clinical Trial. *JAMA* 324, 1307–1316. <https://doi.org/10.1001/jama.2020.17021>.

38. Pickard, A., Calverley, B.C., Chang, J., Garva, R., Gago, S., Lu, Y., and Kadler, K.E. (2021). Discovery of re-purposed drugs that slow SARS-CoV-2 replication in human cells. *PLoS Pathog.* 17, e1009840. <https://doi.org/10.1371/journal.ppat.1009840>.
39. Takahashi, Y., Hayakawa, A., Sano, R., Fukuda, H., Harada, M., Kubo, R., Okawa, T., and Kominato, Y. (2021). Histone deacetylase inhibitors suppress ACE2 and ABO simultaneously, suggesting a preventive potential against COVID-19. *Sci. Rep.* 11, 3379. <https://doi.org/10.1038/s41598-021-82970-2>.
40. Zhang, Q., Chen, C.Z., Swaroop, M., Xu, M., Wang, L., Lee, J., Wang, A.Q., Pradhan, M., Hagen, N., Chen, L., et al. (2020). Heparan sulfate assists SARS-CoV-2 in cell entry and can be targeted by approved drugs in vitro. *Cell Discov.* 6, 80. <https://doi.org/10.1038/s41421-020-00222-5>.
41. Somasekharan, S.P., and Gleave, M. (2021). SARS-CoV-2 nucleocapsid protein interacts with immunoregulators and stress granules and phase separates to form liquid droplets. *FEBS Lett.* 595, 2872–2896. <https://doi.org/10.1002/1873-3468.14229>.
42. Mysinger, M.M., Carchia, M., Irwin, J.J., and Shoichet, B.K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* 55, 6582–6594. <https://doi.org/10.1021/jm300687e>.
43. Lähti, S., Niinivehmas, S., and Pentikäinen, O.T. (2016). Rocker: Open source, easy-to-use tool for AUC and enrichment calculations and ROC visualization. *J. Cheminform.* 8, 45. <https://doi.org/10.1186/s13321-016-0158-y>.
44. Oyagbemi, A.A., Ajibade, T.O., Aboua, Y.G., Gbadamosi, I.T., Adedapo, A.D.A., Aro, A.O., Adejumo, O.A., Thamahane-Katengua, E., Omobowale, T.O., Falayi, O.O., et al. (2021). Potential health benefits of zinc supplementation for the management of COVID-19 pandemic. *J. Food Biochem.* 45, e13604. <https://doi.org/10.1111/jfbc.13604>.
45. Erten, S., Bebek, G., Ewing, R.M., and Koyutürk, M. (2011). DADA: Degree-Aware Algorithms for Network-Based Disease Gene Prioritization. *BioData Min.* 4, 19. <https://doi.org/10.1186/1756-0381-4-19>.
46. Shaban, M.S., Müller, C., Mayr-Buro, C., Weiser, H., Meier-Soelch, J., Albert, B.V., Weber, A., Linne, U., Hain, T., Babayev, I., et al. (2021). Multi-level inhibition of coronavirus replication by chemical ER stress. *Nat. Commun.* 12, 5536. <https://doi.org/10.1038/s41467-021-25551-1>.
47. Fung, T.S., Huang, M., and Liu, D.X. (2014). Coronavirus-induced ER stress response and its involvement in regulation of coronavirus-host interactions. *Virus Res.* 194, 110–123. <https://doi.org/10.1016/j.virusres.2014.09.016>.
48. Echavarría-Consuegra, L., Cook, G.M., Busnadiego, I., Lefèvre, C., Keep, S., Brown, K., Doyle, N., Dowgier, G., Franaszek, K., Moore, N.A., et al. (2021). Manipulation of the unfolded protein response: A pharmacological strategy against coronavirus infection. *PLoS Pathog.* 17, e1009644. <https://doi.org/10.1371/journal.ppat.1009644>.
49. Cheng, F., Kovács, I.A., and Barabási, A.L. (2019). Network-based prediction of drug combinations. *Nat. Commun.* 10, 1197. <https://doi.org/10.1038/s41467-019-09186-x>.
50. Santoro, F., Nuñez-Gil, I.J., Vitale, E., Viana-Llamas, M.C., Reche-Martinez, B., Romero-Pareja, R., Feltez Guzman, G., Fernandez Rozas, I., Urbarrí, A., Becerra-Muñoz, V.M., et al. (2022). Antiplatelet therapy and outcome in COVID-19: the Health Outcome Predictive Evaluation Registry. *Heart* 108, 130–136. <https://doi.org/10.1136/heartjnl-2021-319552>.
51. Cho, J., Lee, Y.J., Kim, J.H., Kim, S.I., Kim, S.S., Choi, B.S., and Choi, J.H. (2020). Antiviral activity of digoxin and ouabain against SARS-CoV-2 infection and its implication for COVID-19. *Sci. Rep.* 10, 16200. <https://doi.org/10.1038/s41598-020-72879-7>.
52. Choy, K.T., Wong, A.Y.L., Kaewpreedee, P., Sia, S.F., Chen, D., Hui, K.P.Y., Chu, D.K.W., Chan, M.C.W., Cheung, P.P.H., Huang, X., et al. (2020). Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Res.* 178, 104786. <https://doi.org/10.1016/j.antiviral.2020.104786>.
53. Chen, Y., Lear, T.B., Evankovich, J.W., Larsen, M.B., Lin, B., Alfaras, I., Kennerdell, J.R., Salminen, L., Camarco, D.P., Lockwood, K.C., et al. (2021). A high-throughput screen for TMPRSS2 expression identifies FDA-approved compounds that can limit SARS-CoV-2 entry. *Nat. Commun.* 12, 3907. <https://doi.org/10.1038/s41467-021-24156-y>.
54. Bansbach, J., Wenz, J., Kaufmann, K., Heinrich, S., and Kalbhenn, J. (2022). Sevoflurane in combination with esketamine is an effective sedation regimen in COVID-19 patients enabling assisted spontaneous breathing even during prone positioning. *Anaesthesiol. Intensive Ther.* 54, 23–29. <https://doi.org/10.5114/ait.2022.113950>.
55. Han, H.J., Nwagwu, C., Anyim, O., Ekwereamad, C., and Kim, S. (2021). COVID-19 and cancer: From basic mechanisms to vaccine development using nanotechnology. *Int. Immunopharmacol.* 90, 107247. <https://doi.org/10.1016/j.intimp.2020.107247>.
56. Schmidt, A.L., Labaki, C., Hsu, C.Y., Bakouny, Z., Balanchivadze, N., Berg, S.A., Blau, S., Daher, A., El Zarif, T., Friese, C.R., et al. (2022). COVID-19 vaccination and breakthrough infections in patients with cancer. *Ann. Oncol.* 33, 340–346. <https://doi.org/10.1016/j.annonc.2021.12.006>.
57. Lemos, A.E.G., Silva, G.R., Gimba, E.R.P., and Matos, A.D.R. (2021). Susceptibility of lung cancer patients to COVID-19: A review of the pandemic data from multiple nationalities. *Thorac. Cancer* 12, 2637–2647. <https://doi.org/10.1111/1759-7714.14067>.
58. Saini, K.S., Tagliamento, M., Lambertini, M., McNally, R., Romano, M., Leone, M., Curigliano, G., and de Azambuja, E. (2020). Mortality in patients with cancer and coronavirus disease 2019: A systematic review and pooled analysis of 52 studies. *Eur. J. Cancer* 139, 43–50. <https://doi.org/10.1016/j.ejca.2020.08.011>.
59. Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218. <https://doi.org/10.1038/nature12213>.
60. Huttlin, E.L., Bruckner, R.J., Navarrete-Perea, J., Cannon, J.R., Baltier, K., Gebreab, F., Gygi, M.P., Thorncroft, A., Zarraga, G., Tam, S., et al. (2021). Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* 184, 3022–3040.e28. <https://doi.org/10.1016/j.cell.2021.04.011>.
61. Mullen, P.J., Garcia, G., Purkayastha, A., Matulionis, N., Schmid, E.W., Momcilovic, M., Sen, C., Langerman, J., Ramaiah, A., Shackelford, D.B., et al. (2021). SARS-CoV-2 infection rewires host cell metabolism and is potentially susceptible to mTORC1 inhibition. *Nat. Commun.* 12, 1876. <https://doi.org/10.1038/s41467-021-22166-4>.
62. Cheng, F., Zhao, J., Wang, Y., Lu, W., Liu, Z., Zhou, Y., Martin, W., Wang, R., Huang, J., Hao, T., et al. (2021). Comprehensive characterization of protein-protein interactions perturbed by disease mutations. *Nat. Genet.* 53, 342–353. <https://doi.org/10.1038/s41588-020-00774-y>.
63. Gardinassi, L.G., Souza, C.O.S., Sales-Campos, H., and Fonseca, S.G. (2020). Immune and Metabolic Signatures of COVID-19 Revealed by Transcriptomics Data Reuse. *Front. Immunol.* 11, 1636. <https://doi.org/10.3389/fimmu.2020.01636>.
64. Wilk, A.J., Rustagi, A., Zhao, N.Q., Roque, J., Martínez-Colón, G.J., McKechnie, J.L., Ivison, G.T., Ranganath, T., Vergara, R., Hollis, T., et al. (2020). A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* 26, 1070–1076. <https://doi.org/10.1038/s41591-020-0944-y>.
65. Tan, L., Wang, Q., Zhang, D., Ding, J., Huang, Q., Tang, Y.-Q., Wang, Q., and Miao, H. (2020). Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct. Target. Ther.* 5, 33. <https://doi.org/10.1038/s41392-020-0148-4>.
66. Tavakolpour, S., Rakhshandehroo, T., Wei, E.X., and Rashidian, M. (2020). Lymphopenia during the COVID-19 infection: What it shows and what can be learned. *Immunol. Lett.* 225, 31–32. <https://doi.org/10.1016/j.imlet.2020.06.013>.

67. Xiong, Y., Liu, Y., Cao, L., Wang, D., Guo, M., Jiang, A., Guo, D., Hu, W., Yang, J., Tang, Z., et al. (2020). Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerg. Microbes Infect.* 9, 761–770. <https://doi.org/10.1080/22221751.2020.1747363>.
68. Manjili, R.H., Zarei, M., Habibi, M., and Manjili, M.H. (2020). COVID-19 as an Acute Inflammatory Disease. *J. Immunol.* 205, 12–19. <https://doi.org/10.4049/jimmunol.2000413>.
69. Shin, H., and Wherry, E.J. (2007). CD8 T cell dysfunction during chronic viral infection. *Curr. Opin. Immunol.* 19, 408–415. <https://doi.org/10.1016/j.coi.2007.06.004>.
70. Bae, E.Y., Sanders, J.M., Johns, M.L., Lin, K., Ortwine, J.K., Wei, W., Mang, N.S., and Cutrell, J.B. (2021). Therapeutic Options for Coronavirus Disease 2019 (COVID-19): Where Are We Now? *Curr. Infect. Dis. Rep.* 23, 28. <https://doi.org/10.1007/s11908-021-00769-8>.
71. Wrobel, A.G., Benton, D.J., Roustian, C., Borg, A., Hussain, S., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2022). Evolution of the SARS-CoV-2 spike protein in the human host. *Nat. Commun.* 13, 1178. <https://doi.org/10.1038/s41467-022-28768-w>.
72. Leal-Esteban, L.C., and Fajas, L. (2020). Cell cycle regulators in cancer cell metabolism. *Biochim. Biophys. Acta. Mol. Basis Dis.* 1866, 165715. <https://doi.org/10.1016/j.bbadis.2020.165715>.
73. Zong, Z., Wei, Y., Ren, J., Zhang, L., and Zhou, F. (2021). The intersection of COVID-19 and cancer: signaling pathways and treatment implications. *Mol. Cancer* 20, 76. <https://doi.org/10.1186/s12943-021-01363-1>.
74. Robilotti, E.V., Babady, N.E., Mead, P.A., Rolling, T., Perez-Johnston, R., Bernardes, M., Bogler, Y., Caldararo, M., Figueroa, C.J., Glickman, M.S., et al. (2020). Determinants of COVID-19 disease severity in patients with cancer. *Nat. Med.* 26, 1218–1223. <https://doi.org/10.1038/s41591-020-0979-0>.
75. Kim, E.S., and Salgia, R. (2009). MET Pathway as a Therapeutic Target. *J. Thorac. Oncol.* 4, 444–447. <https://doi.org/10.1097/JTO.0b013e31819d6f91>.
76. Peruzzi, B., and Bottaro, D.P. (2006). Targeting the c-Met Signaling Pathway in Cancer: Fig. 1. *Clin. Cancer Res.* 12, 3657–3660. <https://doi.org/10.1158/1078-0432.CCR-06-0818>.
77. Birchmeier, C., Birchmeier, W., Gherardi, E., and Vande Woude, G.F. (2003). Met, metastasis, motility and more. *Nat. Rev. Mol. Cell Biol.* 4, 915–925. <https://doi.org/10.1038/nrm1261>.
78. Eckschlag, T., Plch, J., Stiborova, M., and Hrabeta, J. (2017). Histone Deacetylase Inhibitors as Anticancer Drugs. *Int. J. Mol. Sci.* 18, 1414. <https://doi.org/10.3390/ijms18071414>.
79. Zhang, M., Zhang, L., Hei, R., Li, X., Cai, H., Wu, X., Zheng, Q., and Cai, C. (2021). CDK inhibitors in cancer therapy, an overview of recent development. *Am. J. Cancer Res.* 11, 1913–1935.
80. Liu, J., Qu, L., Meng, L., and Shou, C. (2019). Topoisomerase inhibitors promote cancer cell motility via ROS-mediated activation of JAK2-STAT1-CXCL1 pathway. *J. Exp. Clin. Cancer Res.* 38, 370. <https://doi.org/10.1186/s13046-019-1353-2>.
81. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. <https://doi.org/10.1093/bioinformatics/btp616>.
82. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. <https://doi.org/10.1093/bioinformatics/bts034>.
83. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. <https://doi.org/10.1093/nar/gkv007>.
84. Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*, pp. 1–260.
85. Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., et al. (2025). gplots: Various R programming tools for plotting data. <https://github.com/talgalili/gplots>.
86. Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGA-biolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44, e71. <https://doi.org/10.1093/nar/gkv1507>.
87. Neese, F., Wennmohs, F., Becker, U., and Riplinger, C. (2020). The ORCA quantum chemistry program package. *J. Chem. Phys.* 152, 224108. <https://doi.org/10.1063/5.0004608>.
88. McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. <https://doi.org/10.1093/nar/gks042>.
89. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. <https://doi.org/10.1186/gb-2010-11-3-r25>.
90. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15, R29. <https://doi.org/10.1186/gb-2014-15-2-r29>.
91. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 57, 289–300.
92. Blighe, K., Rana, S., and Lewis, M. (2018). EnhancedVolcano: Publication-Ready Volcanoplots with Enhanced Colouring and Labeling. <https://github.com/kevinblighe/EnhancedVolcano>.
93. Sturm, G., Finotello, F., Petitprez, F., Zhang, J.D., Baumbach, J., Fridman, W.H., List, M., and Aneichyk, T. (2019). Comprehensive evaluation of transcriptome-based cell-type quantification methods for immunology. *Bioinformatics* 35, i436–i445. <https://doi.org/10.1093/bioinformatics/btz363>.
94. Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., Krogsdam, A., Loncova, Z., Posch, W., Wilflingseder, D., et al. (2019). Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.* 11, 34. <https://doi.org/10.1186/s13073-019-0638-6>.
95. Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. <https://doi.org/10.1093/nar/28.1.27>.
96. Murtagh, F., and Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *J. Classif.* 31, 274–295. <https://doi.org/10.1007/s00357-014-9161-z>.
97. Ward, J.H. (1963). Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* 58, 236–244. <https://doi.org/10.1080/01621459.1963.10500845>.
98. Silva, T.C., Colaprico, A., Olsen, C., D'Angelo, F., Bontempi, G., Ceccarelli, M., and Noushmehr, H. (2016). TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor packages. *F1000Res.* 5, 1542. <https://doi.org/10.12688/f1000research.8923.2>.
99. Mounir, M., Lucchetta, M., Silva, T.C., Olsen, C., Bontempi, G., Chen, X., Noushmehr, H., Colaprico, A., and Papaleo, E. (2019). New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput. Biol.* 15, e1006701. <https://doi.org/10.1371/journal.pcbi.1006701>.
100. Shang, J., Mosure, S.A., Zheng, J., Brust, R., Bass, J., Nichols, A., Solt, L.A., Griffin, P.R., and Kojetin, D.J. (2020). A molecular switch regulating transcriptional repression and activation of PPARgamma. *Nat. Commun.* 11, 956. <https://doi.org/10.1038/s41467-020-14750-x>.
101. Case, D.A., Aktulga, H.M., Belfon, K., Cerutti, D.S., Cisneros, G.A., Cruzeiro, V.W.D., Forouzesh, N., Giese, T.J., Götz, A.W., Gohlke, H., et al. (2023). AmberTools. *J. Chem. Inf. Model.* 63, 6183–6191. <https://doi.org/10.1021/acs.jcim.3c01153>.

102. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* *46*, D1074–D1082. <https://doi.org/10.1093/nar/gkx1037>.
103. Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C.C., Pettersen, E.F., Goddard, T.D., Meng, E.C., Sali, A., and Ferrin, T.E. (2012). UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J. Struct. Biol.* *179*, 269–278. <https://doi.org/10.1016/j.jsb.2011.09.006>.
104. Roe, D.R., and Cheatham, T.E., 3rd. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theor. Comput.* *9*, 3084–3095. <https://doi.org/10.1021/ct400341p>.
105. Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* *14*, 33. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
106. Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F., and Schroeder, M. (2015). PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res.* *43*, W443–W447. <https://doi.org/10.1093/nar/gkv315>.
107. Morselli Gysi, D., do Valle, Í., Zitnik, M., Ameli, A., Gan, X., Varol, O., Ghiassian, S.D., Patten, J.J., Davey, R.A., Loscalzo, J., and Barabási, A.L. (2021). Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc. Natl. Acad. Sci. USA* *118*, e2025581118. <https://doi.org/10.1073/pnas.2025581118>.
108. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* *583*, 459–468. <https://doi.org/10.1038/s41586-020-2286-9>.
109. Köhler, S., Bauer, S., Horn, D., and Robinson, P.N. (2008). Walking the interactome for prioritization of candidate disease genes. *Am. J. Hum. Genet.* *82*, 949–958. <https://doi.org/10.1016/j.ajhg.2008.02.013>.
110. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* *13*, 2498–2504. <https://doi.org/10.1101/gr.1239303>.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE                                                                                                             | SOURCE                                                                 | IDENTIFIER                                                                                                        |
|---------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| <b>Deposited data</b>                                                                                                           |                                                                        |                                                                                                                   |
| Estimated Counts Data for RNA-seq analysis of peripheral blood in COVID-19                                                      | Overmyer et al. <sup>16</sup> 2021                                     | GEO: GSE157103                                                                                                    |
| Single cell RNA-seq counts data of BALF in COVID-19                                                                             | Liao et al. <sup>22</sup> 2020                                         | GEO: GSE145926                                                                                                    |
| Hallmark, BioCarta, Kyoto Encyclopedia of Genes and Genomes, Pathway Interaction Database, Reactome, and WikiPathways gene sets | MSigDB (v7.4)                                                          | N/A                                                                                                               |
| The Cancer Genome Atlas                                                                                                         | National Cancer Institute                                              | N/A                                                                                                               |
| Expanded CMap LINCS Resource 2020 (CMap2020)                                                                                    | Connectivity Map                                                       | N/A                                                                                                               |
| COVID-19 repurposing screen (LINCS_COVID_PUBLIC)                                                                                | Connectivity Map                                                       | N/A                                                                                                               |
| BioPlex 3.0 Interactions Network                                                                                                | Huttlin et al. <sup>60</sup> 2021                                      | N/A                                                                                                               |
| Onco-PPI Network                                                                                                                | Cheng et al. <sup>62</sup> 2020                                        | <a href="https://github.com/ChengF-Lab/oncoPPIs">https://github.com/ChengF-Lab/oncoPPIs</a>                       |
| Crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2                                                   | Protein Data Bank                                                      | PDB: 6M0J                                                                                                         |
| <b>Software and algorithms</b>                                                                                                  |                                                                        |                                                                                                                   |
| R (v4.4.2)                                                                                                                      | R Development Core Team                                                | <a href="https://www.r-project.org/">https://www.r-project.org/</a>                                               |
| Python (v3.14)                                                                                                                  | Python Software Foundation                                             | <a href="https://www.python.org/">https://www.python.org/</a>                                                     |
| edgeR                                                                                                                           | Robinson et al. <sup>81</sup> 2010                                     | N/A                                                                                                               |
| sva                                                                                                                             | Leek et al. <sup>82</sup> 2012                                         | N/A                                                                                                               |
| limma                                                                                                                           | Ritchie et al. <sup>83</sup> 2015                                      | N/A                                                                                                               |
| EnhancedVolcano                                                                                                                 | Bioconductor                                                           | N/A                                                                                                               |
| immunedeconv                                                                                                                    | Omnideconv                                                             | <a href="https://github.com/omnideconv/immunedeconv">https://github.com/omnideconv/immunedeconv</a>               |
| ggplot2                                                                                                                         | Wickham <sup>84</sup> 2016                                             | N/A                                                                                                               |
| Gene Set Enrichment Analysis Software                                                                                           | Broad Institute <sup>25,32</sup>                                       | N/A                                                                                                               |
| gplots                                                                                                                          | Warnes et al. <sup>85</sup> 2015                                       | N/A                                                                                                               |
| RRHO2                                                                                                                           | Cahill et al. <sup>19</sup> 2018<br>Plaisier et al. <sup>21</sup> 2010 | <a href="https://github.com/RRHO2/RRHO2">https://github.com/RRHO2/RRHO2</a>                                       |
| TCGAbiolinks                                                                                                                    | Colaprico et al. <sup>86</sup> 2016                                    | N/A                                                                                                               |
| seaborn                                                                                                                         | Michael Waskom                                                         | <a href="https://seaborn.pydata.org/">https://seaborn.pydata.org/</a>                                             |
| umap-learn                                                                                                                      | Leland McInnes                                                         | <a href="https://umap-learn.readthedocs.io/en/latest/">https://umap-learn.readthedocs.io/en/latest/</a>           |
| FAISS                                                                                                                           | Meta Platforms, Inc.                                                   | <a href="https://github.com/facebookresearch/faiss">https://github.com/facebookresearch/faiss</a>                 |
| RepurNet (GitHub)                                                                                                               | This paper                                                             | <a href="https://github.com/Stephen-Yi-Laboratory/RepurNet">https://github.com/Stephen-Yi-Laboratory/RepurNet</a> |
| RepurNet (Zenodo)                                                                                                               | This paper                                                             | <a href="https://zenodo.org/records/18881178">https://zenodo.org/records/18881178</a>                             |
| Glycam                                                                                                                          | Dr. Robert J. Woods                                                    | <a href="https://glycam.org/">https://glycam.org/</a>                                                             |
| ORCA                                                                                                                            | Neese et al. <sup>87</sup> 2020                                        | <a href="https://www.faccts.de/orca/">https://www.faccts.de/orca/</a>                                             |
| MOE (2020.09)                                                                                                                   | Chemical Computing Group                                               | <a href="https://www.chemcomp.com/en/Products.htm">https://www.chemcomp.com/en/Products.htm</a>                   |
| Cytoscape (v3.8.2)                                                                                                              | Cytoscape Consortium                                                   | <a href="https://cytoscape.org/">https://cytoscape.org/</a>                                                       |
| BioRender                                                                                                                       | BioRender                                                              | <a href="https://www.biorender.com/">https://www.biorender.com/</a>                                               |

## METHOD DETAILS

### Regulatory network gene expression signature analysis

An estimated counts mRNA expression matrix was obtained from the GEO dataset GSE157103, which included 126 samples of primary leukocytes from whole blood, each from a different hospitalized patient.<sup>16</sup> Of these 126 samples, 50 were from ICU patients who had COVID, 50 from non-ICU patients who had COVID, 17 samples from ICU patients who did not have COVID, and 9 samples from non-ICU patients who did not have COVID. Three comparisons were constructed based on the phenotypes of patient samples: COVID ICU vs. nonCOVID ICU, COVID ICU vs. COVID nonICU, and COVID vs. nonCOVID. Expression data was prepared using the “edgeR” package in R.<sup>81,88</sup> Genes that were not expressed at more than 5 counts per million across at least the sample size of the smallest group within each comparison were excluded from differential gene expression analysis, and expression counts were normalized with trimmed means of M-values (TMM) normalization.<sup>89</sup> Batch effect correction was conducted using the “sva” package in R.<sup>82</sup> Differential gene expression analysis using a limma-voom pipeline in R on these three comparisons produced their respective DEG profiles.<sup>83,90</sup> Genes were only considered differentially expressed if they had a false discovery rate (FDR) adjusted p-value < 0.05 and a  $|\log_2(\text{fold-change})| \geq 1$ . Volcano plots were constructed using the “EnhancedVolcano” package in R.<sup>91,92</sup>

### Immune cell type deconvolution

TPM normalized mRNA expression data from the GEO dataset GSE157103 was deconvoluted for immune cell signals using the “quantiseq” method in the “immunedeconv” package to test for differences in immune cell type proportions across each of our differential gene expression analysis comparisons.<sup>93,94</sup> Wilcoxon rank-sum tests and Bonferroni p-value correction for multiple comparisons were used to determine significant differences in immune cell type proportions using a Bonferroni corrected p-value < 0.05 as a cutoff for statistical significance. Jitter plots were constructed using the “ggplot2” package in R.<sup>32,84</sup>

### Gene set enrichment analysis (GSEA)

GSEA software developed by Subramanian et al. was used to test for significantly enriched gene sets from each of our DEG profiles.<sup>25,32</sup> TMM normalized expression data from every differential gene expression analysis were used to test for enrichment in the Hallmark, BioCarta, KEGG, PID, Reactome, WikiPathways, gene ontology, oncogenic signature, and ImmuneSigDB collections available on MSigDB (v7.4).<sup>24–31,95</sup> The “Signal2Noise” metric was used to rank genes, and the “phenotype” option was used as the permutation type in calculating statistical significance if there were more than 7 samples in all groups tested. Otherwise, the “gene\_set” option was used as the permutation type in calculating statistical significance. Gene sets with an FDR adjusted p-value < 0.05 were considered statistically significant. Barplots were constructed using the “ggplot2” package in R.<sup>84</sup> The “gplots” package in R was used to construct heatmaps.<sup>85</sup> For heatmaps, mRNA expression matrices from GSE157103 were  $\log_2(\text{TMM}+1)$  transformed. For clustering, distance was calculated using Pearson’s correlation distance, and hierarchical clustering was performed using the “ward.D2” method.<sup>96,97</sup> Hypergeometric overlap analysis was conducted using the 376 human genes included in the COVID disease module to test for gene set enrichment in the KEGG database.

### Comparison of blood-derived DEG profiles to BALF-Derived DEG profiles

Single-cell mRNA expression data was obtained from the GEO dataset GSE145926, which included samples from BALF of 6 patients with severe COVID, 3 patients with mild COVID, and 3 healthy controls.<sup>22</sup> Before running differential gene expression analysis, single-cell expression data was collapsed to produce an expression matrix equivalent to that from a bulk RNA-seq analysis. DEG profiles were obtained from comparisons based on the severity of COVID in patient samples: Severe vs. Healthy, Severe vs. Mild, and COVID vs. Healthy, which were used to validate the COVID ICU vs. nonCOVID ICU, COVID ICU vs. COVID nonICU, and COVID vs. nonCOVID DEG profiles, respectively. Differential gene expression analysis was conducted in the same fashion previously described. Using the Rank-Rank Hypergeometric Overlap (RRHO) analysis from the “RRHO2” package in R, concordance and/or discordance of DEG profiles were tested between the COVID ICU vs. nonCOVID ICU and the Severe vs. Healthy profiles; the COVID ICU vs. COVID non-ICU and the Severe vs. Mild profiles; and the COVID vs. nonCOVID and the COVID vs. Healthy profiles.<sup>19,21</sup> GSEA was also run on BALF-derived DEG profiles (as previously described) to identify any shared enriched gene sets between BALF-derived DEG profiles and blood-derived DEG profiles.

### Gene regulatory network comparison of COVID to TCGA cancer profiles

HTSeq gene expression counts matrices were obtained for each of the 33 TCGA cancers using the “TCGAbiolinks” package in R.<sup>86,98,99</sup> Analysis was focused on TCGA-LUAD and TCGA-LUSC, but differential gene expression analysis was applied to all 21 TCGA cancers with 3 or more normal tissue samples. For each of these 21 TCGA cancers, solid primary tumor samples were compared to solid tissue normal samples to generate DEG profiles. Differential gene expression analysis was conducted on each of the 21 TCGA cancers in the same fashion previously described. Each of the 21 TCGA cancers’ DEG profiles were then compared to the previously generated COVID vs. nonCOVID DEG profile using RRHO analysis from the “RRHO2” package in R.<sup>19,21</sup> The 19 TCGA cancers that had concordant DEG profiles with a max  $-\log_{10}(\text{p-value}) > 3$  were analyzed further with GSEA using the previously described GSEA parameters. The dotplot was constructed using the “ggplot2” package in R.<sup>84</sup>

### Drug reversal calculations with the LINCS1000 platform

The LINCS1000 platform measures expression for a core 978 gene panel using the Luminex FlexMap 3D Scanner.<sup>35</sup> The core panel was then used to extrapolate gene expression estimates for the core panel as well as 11,350 genes with closely correlated expression. These data were then processed to account for variability among replicates before normalizing the values across all treatments in the database (Level 5 data, ModZ values). The final processed data was interpretable as a z-score indicating the strength of up-regulation or downregulation in a gene after a certain treatment, where the treatment was defined by the cell line, compound, dosage, time duration, and batch.

We used the 2020 expanded LINCS release on the Broad Institute's Connectivity Map to search for profiles of interest and conduct the first part of the RepurNet analysis. After downloading profiles with a 24 hr screen duration and subset to the top 60 upregulated and downregulated genes, we calculated the Pearson's correlation coefficient (R) between log<sub>2</sub>(fold-change) expression values observed in COVID infection against the ModZ values for the corresponding genes in each profile. For further analysis to assess if the number of genes significantly affected the composition of the compound reversal rankings, the full 60 gene subset was compared to possible subsets of the top 10, 20, 30, 40, and 50 genes (Table S11). The pipeline was repeated for these different numbers and assessed for similarity in final rank order of all available compounds using Kendall's Tau. Our analysis result showed the pipeline to be generally robust to the number of genes selected in terms of which compounds are identified as recapitulating and reversing the disease signature.

The Pearson's correlation coefficient was selected due to a greater sensitivity to outliers, which would prove beneficial in selecting for compounds that can cause a significant reversal in a subset of genes even without significant impact on the rest of the genes. An  $R > 0$  indicates a positive correlation while an  $R < 0$  indicates a negative correlation. Consequently, we filtered for the 3000 treatments with the lowest R values as a starting point before examining the 200 treatments with the lowest R values in greater detail. Clustering analyses were conducted using Ward's linkage. Heatmaps were built with the Python seaborn package.

### Clustering on UMAP embeddings and nearest-neighbor based scoring

We used the COVID specific LINCS release on the Broad institute's connectivity map ([https://clue.io/data/COVID#LINCS\\_COVID\\_PUBLIC](https://clue.io/data/COVID#LINCS_COVID_PUBLIC)) to select the lung-cancer specific cell line A549. The matrix consisting of the gene expression and drug-treatments with 24h exposure was used to calculate the top 40 UMAP embeddings of the gene expression values for every drug treatment with number of neighbors set to 15 and the minimum distance of 0.1 in the UMAP function of the umap-learn package (<https://umap-learn.readthedocs.io/en/latest/>) in Python. To perform k-means nearest neighbor clustering, we used the FAISS package in Python (<https://github.com/facebookresearch/faiss>), which was used for efficient similarity search and clustering of dense vectors. The number of centroids for the clustering was set to 10 and number of iterations were set to 20. The resulting cluster memberships were visualized in the principal component space and FAISS's index search function was used to generate neighborhood matrices comprising of the nearest fifty neighbors for every drug treatment in the aforementioned dataset. To find the clusters enriched with candidate drug treatment, we used the hypergeometric test using the phyper function in R with the False Discovery Rate (FDR) controlled at 1%. Neighborhood information was used to compute a treatment-associated nearest neighbor score, where one was added to the score if a CDT was in the top twenty neighbors of the treatment or nothing was added. Spearman correlation coefficient was calculated between the coefficient values indicating effective reversal of observed gene signature by perturbagens of differentially expressed genes between COVID ICU vs. COVID nonICU patients and the above-mentioned neighborhood-based scoring method. The Spearman correlation coefficient was selected over the Pearson correlation coefficient to better assess for a monotonic negative association without an emphasis on linearity. The scoring process was repeated for 5, 10, and 50 nearest neighbors to test for the robustness of the scoring method.

### Cheminformatic structural docking analysis

Molecular modeling was done to interrogate the binding of the compounds to the SARS-CoV-2 spike protein interface with ACE2 receptor. First, the complex structure of the SARS-CoV-2 spike protein was obtained from the Protein Data Bank (PDB: 6M0J).<sup>100</sup> Glycans (DGlcNAc1-OH) were added to the structure using the Glycoprotein Builder tool from Glycam (<http://glycam.org>). Then, N-glycosidic bonds between glycans and asparagine residues and disulfide bonds were built using tLeap, which is implemented in AmberTools23.<sup>101</sup> Doxorubicin (DRC), Copansilib (CPS), Rapamycin (RPC), and Tyrphostin AG (TAG) structures were obtained from Drugbank,<sup>102</sup> and geometry optimization of these structures were performed with ORCA<sup>87</sup> using Density Functional Theory (DFT) at the B3LYP level of theory with def2-SV(P) basis set, additional relativistic correction (def2/J), and a D3 dispersion correction. Ligands' charges were calculated and incorporated through UCSF Chimera using M1-BCC charges.<sup>103</sup> The receptor-ligand complex was then prepared in tLeap. The systems were solvated in a cubic TIP3PBOX water box with a 10 Å box size, and counterions (25 Na<sup>+</sup>) were added to neutralize net charge. Topology and coordinate files were then generated for molecular dynamics simulations.

The molecular modeling suite MOE 2020.09 was used for the docking studies. The complex was prepared for docking by protonating the system at pH 7.4. A SiteFinder algorithm was used to identify a possible binding pocket between the two proteins, and subsequently used for the docking protocol. An induced-fit docking was done, and the top pose of each compound in complex with the protein pocket used for further analysis.

The parameters for modeling the molecular docking are as follows. Initially, with the protein and ligands fixed with a restraint force constant of 50.0 kcal mol<sup>-2</sup>, the systems were minimized for 500 cycles of steepest descent followed by 500 cycles of conjugate gradient. Then, with the protein's and ligands' heavy atoms fixed with a restraint force of 10.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>, the systems were minimized again for 50,000 cycles of steepest descent followed by 50,000 cycles of conjugate gradient. Next, with the protein's and ligands' heavy atoms fixed with a restraint force of 50.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>, the systems were heated from 0 K to 300 K for 1 nanosecond in two steps. For the first step, the temperature of the systems was raised from 0 K to 300 K in 200 picoseconds, and on the second step, the temperature of the systems was held constant at 300 K for the remaining time (800 picoseconds). With the protein's and ligands' heavy atoms fixed with a restraint force of 1.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>, the systems were further equilibrated using NVT ensemble (constant volume and temperature) for 1 nanosecond at 300 K then using NPT (constant pressure and temperature) ensemble for 1 nanosecond at 300 K and 1.0 bar using Monte Carlo barostat. Finally, two independent NPT productions were performed without any constraints for 100 nanoseconds each. Frames were saved every 1 picosecond for simulation analysis using CPPTRAJ<sup>104</sup> (Process TRAJectory in C++) and VMD<sup>105</sup> (Visual Molecular Dynamics). The relative free energy of binding of the ligands to the protein were estimated using the Generalized Born (GB) model on the last 20 nanoseconds of both productions. The last 20 nanoseconds of the productions were clustered using a hierarchical agglomerative algorithm with a minimum distance between each cluster of 3 Å. Protein-ligand interaction profiles for both ligands were obtained using the centroid of the most populated cluster as a representative binding pose through PLIP (Protein-Ligand Interaction Profiler).<sup>106</sup>

### COVID disease module construction

The COVID disease module was built by combining and constraining two large protein-protein interaction networks. The first network, a human interactome file, was retrieved from Morselli-Gysi et al. due to their comprehensive combination of protein-protein interactions with additional regulatory relationships.<sup>107</sup> The second network featured the host-viral protein interactions identified for the SARS-CoV-2 viral proteins.<sup>108</sup> After connecting the viral proteins to the human interactome, the viral proteins were treated as seed genes in a random walk with restart model to predict the risk of perturbation in all the genes across the interactome using an  $\alpha = 0.7$ .<sup>109</sup> The gene scores were then adjusted using the eigenvector centrality method to account for variations in node degree across the interactome.<sup>45</sup> The genes were prioritized according to the adjusted scores and added one at a time until all SARS-CoV-2 viral proteins were successfully connected in a subgraph of the human interactome. This smallest subgraph, featuring 376 human genes, was treated as the COVID disease module.

### Network-guided rational design of drug combinations

To execute the second part of the RepurNet analysis, we aimed to use the COVID-19 as an emblematic disease module and assessed for feasible drug combinations from the individual compounds identified in the *in silico* screen. From prior studies, effective combinations were found to have close drug-disease proximity and stronger drug-drug separation, indicating the two drugs should practically target complementary portions of the module.<sup>49</sup> To estimate this, we first simulated diffusion of each compound through the human interactome using the DADA eigenvector centrality approach.<sup>45</sup> For each possible pair of compounds, we then subtracted the diffusion profiles to produce a difference vector. An L1 normalization score was calculated for each vector by adding up the absolute value of all differences and dividing by 100. A larger L1 score could therefore be interpreted as a higher difference in the diffusion profiles, indicating the pair of compounds target different parts of the disease module.

For network visualizations, the nodes were colored based on the calculated difference in diffusion scores at each node. Given a difference profile for diffusion (Compound 1) – diffusion (Compound 2), large positive differences indicate a node is targeted strongly by Compound 1 whereas large negative differences support targeting by Compound 2.

### Gene regulatory network analysis of similarities between cancers and COVID

The BioPlex 3.0 Interactions Network (293T cells) of the human interactome was obtained to examine interactions of interest that are relevant to both COVID and the TCGA cancers LUAD and LUSC.<sup>60</sup> The genes comprising the 54 enriched GOBP terms shared between COVID and either TCGA-LUAD or TCGA-LUSC were subsetted from the BioPlex 3.0 Interactions Network. The subsetted network was visualized using Cytoscape (v3.8.2).<sup>110</sup> Genes were annotated by whether they were significantly upregulated or downregulated in the COVID vs. nonCOVID DEG profile and by the number of GOBP terms in which they are categorized.

A pan-cancer network of putative oncogenic protein-protein interactions (onco-PPIs) was obtained from Cheng et al. (2020) to investigate mutated genes relevant to cancer that are dysregulated in COVID.<sup>62</sup> The onco-PPI network was subsetted for all differentially expressed genes in the COVID vs. nonCOVID DEG profile and visualized using Cytoscape (v3.8.2).<sup>110</sup> Genes were annotated by whether they were significantly upregulated or downregulated in the COVID vs. nonCOVID DEG profile and by the number of TCGA cancers' individual onco-PPI networks in which they were identified by Cheng et al. (2020).<sup>62</sup>

### QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical details for each analysis can be found in the [method details](#) section and in all figure legends. Generally, statistical analysis was conducted in R or Python. P-value adjustment for multiple hypothesis testing was implemented as needed and as described above.