Large scale profiling of SARS-CoV-2-infected patients identified potential therapeutic host targets and drug candidates for COVID-19

Sidharth Jain[1,#], Samantha Rego[1,#] & Sivanesan Dakshanamurthy[1,*]

[1]Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington DC USA.

[#]equal contribution authors

[*]Corresponding author

Dr. Sivanesan Dakshanamurthy, PhD., MBA.

Professor, & Director of Computational Chemistry Shared Resources,

3970 Reservoir Road,

Lombardi Comprehensive Cancer Center,

Georgetown University Medical center,

Washington DC 20057 USA

Email: sd233@georgetown.edu

Abstract

Given the rapid spread of SARS-CoV-2 and rising death toll of COVID-19 in the current absence of effective treatments, it is imperative that therapeutics are developed and made available to patients as quickly as possible. Publicly available COVID-19 patient data can be used to identify host therapeutic targets, tailoring treatments to the disease signatures observed in patients. In this study, we identify potential host therapeutic targets based on gene expression alterations observed in COVID-19 patients. We analyzed RNAseq data from airway samples of COVID-19 patients and healthy controls to detect significantly differentially expressed genes and pathways that present potential therapeutic targets. Our analysis revealed expression changes in key genes involved in activation of immune pathways, as well as genes targeted by SARS-CoV-2 to interfere with normal host cell functioning. Critical changes were observed in a number of genes, including EIF2AK2, which was shown to play important roles in activating the interferon response and interfering with host cell translational machinery in SARS-CoV-2 infection, presenting a prospective therapeutic target. We also identified drugs with potential to modulate multiple therapeutic targets within the most significant pathways. Our results both validate key genes, pathways, and drug candidates that have been reported by other studies and suggest others that have not been well-characterized and warrant further investigation by future studies. Further investigation of these therapeutic targets and their drug interactions may lead to effective therapeutic strategies to combat the current COVID-19 pandemic and protect against future outbreaks.

Keywords: COVID-19 patient analysis; SARS-CoV-2; host target identification; drug repurposing; gene expression analysis; proteomics; bioinformatics

Introduction

The exponential spread of SARS-CoV-2 is a significant socioeconomic and public health threat[1]. A novel coronavirus was initially discovered in Wuhan, China in late 2019[2] and was later classified as Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2). Coronavirus disease 2019 (COVID-19), the disease caused by SARS-CoV-2, rapidly spread globally and was declared a pandemic by the WHO in March of 2020. The COVID-19 pandemic continues to grow, causing over 100 million confirmed cases and over two million deaths worldwide as of January 26, 2020[3]. The broad reach of this highly contagious and, in some cases, lethal virus makes it imperative for effective treatments to be discovered as quickly as possible.

While drug discovery efforts often take years, drug repurposing offers a more efficient approach to identifying potential treatments through alternative uses of already approved or clinical trial drugs. Drug repurposing can accelerate drug development timelines by several years, as well as substantially reducing risks and costs, since the drugs have already undergone clinical and safety trials in the initial approval process[4]. High-throughput approaches to drug repurposing offer the opportunity for quick and efficient identification of drug-target associations from large amounts of data. Systems biology can then be used to contextualize these associations within the virus-host-drug interactome, and thus can assist in predicting novel drug targets.

Few studies so far have combined these approaches to identify potential therapeutic targets and drug candidates for COVID-19 treatment. Fagone et al. identified potential therapeutic targets from the transcriptomic profiles of primary human lung epithelium upon SARS-CoV-2 infection and used computational analysis to predict potential drug candidates[5]. Xing et al. analyzed gene expression profiles from SARS-CoV and MERS-CoV infected samples for disease signatures and predicted drug candidates for SARS-CoV-2[6]. They then experimentally validated

these drugs in SARS-CoV-2 infected cell lines. These studies analyzed data from in vitro models of SARS-CoV-2, SARS-CoV, and MERS-CoV infections rather than COVID-19 patient samples, which could potentially limit the translational capacity of their findings. The use of patient data to identify disease signatures allows for an unbiased search for drug candidates based on their ability to reverse the gene expression changes observed in COVID-19 patients. As disease severity in COVID-19 has been found to be largely dependent on host factors[7], it is essential that studies of disease pathogenesis and therapeutic strategies do not only focus on the virus itself, but also take host factors into account. This host-focused approach to identifying therapeutic targets can also be used to tailor treatment strategies for COVID-19 patients so that mild or severe COVID-19 may be treated with different drugs that can more effectively target their differing disease signatures.

In this study, we sought to identify host therapeutic targets for COVID-19 treatment from gene expression changes in respiratory tract tissue of COVID-19 patients. RNA-seq data of airway samples from COVID-19 patients were obtained from the NCBI Gene Expression Omnibus (GEO) database[8]. We then employed *Partek*® *Flow*® software (v10.0)[9] to perform differential gene expression analysis in SARS-CoV-2 infected and healthy samples from these datasets. From this analysis, we identified disease-specific gene expression signatures which were investigated further to identify potential therapeutic targets, either through individual gene expression changes or overall pathway alterations. Using these targets, we were able to identify potential drug candidates for COVID-19 treatment.

Materials and Methods

Acquisition and Processing of GEO Data

To understand the host transcriptomic changes that occur upon SARS-CoV2 infection, we first gathered relevant publicly available gene expression datasets. We queried GEO for RNAseq datasets, focusing on patient samples obtained from the pulmonary tract (**Supplementary Fig. 1**). We found four datasets (**Table 1**) that satisfied the query criteria and obtained FASTQ-level data for each sample. Using the standard RNAseq pipeline in *Partek*® *Flow*® software, v10.0., samples were processed, and gene-level counts were quantified. For GSE152075, FASTQ data could not be obtained due to patient privacy concerns, but gene-level counts were provided, which was used for further analysis.

Differential Expression Analysis

Data acquired from GEO was processed using the Gene Set Analysis algorithm in *Partek*® *Flow*® software, v10.0., to identify differentially expressed genes in each experiment. Samples were labeled based on the presence or absence of detected SARS-CoV2, and only respiratory samples were compared. Differentially expressed genes from each dataset were identified at a threshold of FDR < 0.05, and up- and down-regulated genes were defined using a fold-change cutoff of ∓2 (**Supplementary Tables 1 and 2**). Genes found as up-regulated in two or more experiments were selected for further analysis.

Pathway Analysis and Interaction Mapping

Differentially expressed genes identified by Gene Set Analysis in *Partek® Flow®* software, v10.0., for each dataset were separated into up-regulated and down-regulated based on fold change and uploaded into Ingenuity Pathway Analysis (IPA)[10]. Core analysis with cutoffs FDR < 0.05 and fold change ∓ 2 were run in IPA separately for up-regulated and down-regulated genes for each dataset, to identify the significant pathways implicated by the patterns of differential gene expression in each dataset (**Supplementary Figs. 2 and 3**). IPA determines the significance of the association between the data set and the canonical pathway based on the ratio of the number of molecules from the data set that map to the pathway divided by the total number of molecules that map to the canonical pathway, as well as by the p-value of the association, calculated by a right-tailed Fisher's Exact Test.

The Comparison Analysis function in IPA was used to compare the up-regulated pathways and down-regulated pathways across datasets. Pathway diagrams for pathways identified as most significantly up- or down-regulated in both datasets were analyzed for drug and gene interactions, as well as downstream effects of the observed differential gene expression.

Identifying Drug-Gene Interactions

Differentially expressed genes were queried in Drug Gene Interaction Database (DGIdb) to identify existing drugs with experimentally determined associations with our DEGs[11]. Any genes with known drug interactions were annotated using DGIdb. Drugs were also annotated based on known action (e.g., agonist, antagonist, etc), and interaction with SARS-CoV-2[12] (**Supplementary Table 4**).

Results

Our search of the GEO database revealed four RNAseq datasets that met the inclusion criteria for the study (**Supplementary Fig. 1**), however one was excluded from analysis due to poor read quality (**Table 1**). RNAseq analysis was performed on the three remaining datasets using the Partek Flow genomics suite with a standard RNAseq pipeline. Differential gene expression analysis was performed using the Gene Set Analysis tool, and differentially expressed genes (DEGs) were identified comparing samples from SARS-CoV-2 infected patients and uninfected controls in each of these datasets. Differentially expressed genes in COVID-19 patients were identified with a significance threshold of FDR < 0.05 and significantly up- or down-regulated genes were defined as having a fold change of +/-2. This analysis identified 457 significantly up-regulated genes and 2380 significantly down-regulated genes in GSE152075 (**Figure 2a**). 203 significantly up-regulated genes and 55 significantly down-regulated genes were identified from GSE147507 (**Figure 2b**). Only one significant DEG was identified from GSE150316 using our parameters, so this dataset was excluded from further analysis (**Figure 2c**).

Two datasets, GSE152075 and GSE147057, were further analyzed for commonly up- and down-regulated genes through Comparison Analysis in Ingenuity Pathway Analysis, which revealed 67 commonly upregulated and 12 commonly down-regulated genes (**Supplementary Table 3**). A significant number of ribosomal proteins was found to be down-regulated in both datasets. Many interferons and cytokines such as chemokines were upregulated in both datasets, reflecting an activated inflammatory immune response in COVID-19 patients.

To investigate the pathways significantly altered in COVID-19 patient samples, we used Ingenuity Pathway Analysis (IPA) to analyze the patterns of differential gene expression revealed by the lists of DEGs generated by Partek Flow. Core analysis of the differentially expressed genes

in IPA identified the most significant molecules and pathways involved in the patterns of differential expression in each dataset. These molecules and pathways represent potential therapeutic targets for COVID-19 treatment, as they are most highly involved in COVID-19 pathogenesis. Up-regulated genes may be targeted by inhibition, while down-regulated genes may be targets of agonism. Pathway analysis identified several key pathways that were significantly up- and down-regulated across datasets (**Supplementary Figure 4**). These included the coronavirus pathogenesis pathway and replication cycle, immunomodulatory pathways such as communication between innate and adaptive immune cells, NF-kB signaling pathway, and interferon signaling. Up-regulated pathways represented in the dataset demonstrated immune response activation, including interferon signaling, chemokine signaling, and coronavirus pathogenesis pathway. Using the Drug-Gene Interaction Database (DGIdb), identified pathways were also mapped to existing pharmaceutical agents and drugs from DGIdb were identified that bind to the most significantly differentially expressed genes.

Within the pathways found to be most significantly affected in our datasets, key genes were significantly up- or down-regulated, presenting potential drug targets. Analysis of the SARS-CoV-2 replication cycle identified key host genes involved in viral replication (**Figure 3**). ACE2 and the replication complex were downregulated in one dataset. TMPRRS2 was upregulated in one dataset and tubulins were upregulated in both datasets. Drug-gene interaction analysis through DGIdb revealed multiple existing compounds with potential to target host mechanisms of viral proliferation, specifically viral entry, replication, and assembly within host cells (**Supplementary Table 4**). Some of the identified drugs, such as ribavirin and colchicine, have been suggested by other studies and are already being tested or used in COVID-19 patients[13,14], validating the relevance of our results. Other compounds we identified with relevant targets have not been extensively studied in the context of COVID-19 but warrant further investigation (**Supplementary Table 4**).

The eIF2 signaling pathway was the most significant down-regulated pathway in GSE152075 and one of the most significantly down-regulated pathways in GSE147507 (**Figure 4, Supplementary Figures 2 and 3**). Multiple ribosomal proteins were among the most significantly downregulated in each dataset (**Supplementary Tables 1 and 2**). Genes for components of the 40S and 60S ribosomal subunits were significantly downregulated in both datasets, suggesting that that viral proteins interact with and modulate the expression of host translation machinery (**Figure 4)**. The consistent down-regulation of multiple ribosomal proteins in SARS-CoV-2 infected samples suggests that translational functions are diminished by SARS-CoV-2 infection.

The interferon signaling pathway was entirely up-regulated, revealing many key genes responsible for activating the interferon response in SARS-CoV-2 infection (**Figure 5**). An important hub of this pathway is EIF2AK2, a kinase that is responsible for activating the inflammasome. JAK/STAT expression is also significantly upregulated, as is expression of the JAK/STAT signaling pathway. DGIdb analysis identified many compounds with potential to modulate these upregulated key activators of host interferon response. The potential role of anti-TNF therapy in treating COVID-19 has been suggested by other studies, and some TNF inhibitors such as infliximab are being studied in clinical trials[15]. Many of the drugs interacting with the interferon signaling pathway are already being studied in COVID-19 clinical trials, although TLR4 inhibitors have not yet been studied in the context of COVID-19.

Immunomodulatory pathways, specifically the communication between innate and adaptive immune cells, were also significantly up-regulated and revealed key up-regulated host genes (**Figure 6**). Multiple TLRs were shown to be upregulated, and many compounds to target these genes were identified by DGIdb. Drugs that bind nonspecifically to TLR proteins have been tested extensively, including such drugs as hydroxychloroquine and ritonavir[16,17,18]. Because this pathway is important for coordinating both adaptive and innate immune responses, there may

have been some logic in this strategy, but in practice these drugs did not perform well, as reported by the Recovery trial[19] and the WHO "Solidarity" clinical trial[20]. Further study is necessary to understand how these pathways may be exploited to address the symptoms of COVID-19. Our data suggests that interactions with B cells occur primarily through TNFSF13B, whereas IL-10 signaling seems to mediate the interaction with T-cells (**Figure 6**). TNFSF13B was significantly upregulated across datasets, presenting a potential target for inhibitors to modulate the host B cell response. CXCL10, a chemokine ligand, may also be targeted by inhibitors to regulate its effects on the adaptive immune response.

Discussion

Our results further validate key genes, pathways, and potential therapies that have been reported by other studies. ACE2 and TMPRSS2, known to be used by SARS-CoV-2 for cell entry, were significantly differentially expressed in our analysis of COVID-19 patients[21]. This emphasizes the importance of further trials studying drug candidates that target these genes, such as those identified by our analysis (**Supplementary Table 4**). Our results also further support the important role of inflammatory signaling and interferon response in COVID-19 pathogenesis suggested by other studies. Many of the significantly up-regulated genes identified in our analysis correspond to interferon response and chemokine/cytokine signaling, suggesting their importance as potential therapeutic targets for COVID-19 treatment. Specifically, IFIT1, IFIT2, and IFIT3, which are known to be effectors of Type I IFN response, were significantly upregulated[22]. Type I IFN activity has been suggested to play an important role in SARS-CoV-2 infection and to be a marker of disease severity[23]. IFITM1 and IFITM3 were also significantly upregulated in both datasets. These interferon-induced transmembrane proteins have been found to enhance SARS-CoV-2 infection, possibly by promoting viral invasion[24]. Our results suggest that therapies modulating IFN signaling may be an effective treatment strategy, as infected cells

have a strongly active interferon response pathway. Viral-host interactions may also occur in this pathway, especially as SARS-CoV-2 has been shown to modulate interferon signaling using orf1ab[25]. Further studies of the virus-human interactome may help elucidate how the virus impacts the interferon signaling pathway and reveal more potential drug targets.

Both datasets also had significant upregulation of many cytokines, as well as the chemokine signaling pathway, suggesting an activation of the inflammatory response in COVID-19 patients. Excessive release of cytokines by the dysregulated immune system has been found to be characteristic of severe COVID-19 infection, leading to a cytokine release storm[26]. This leads to the acute respiratory distress syndrome (ARDS), a major cause of morbidity and mortality in COVID-19 patients, making management of excessive cytokine release a critical treatment strategy for improving patient outcomes[27]. Our results affirm the important role of cytokines in COVID-19 pathogenesis, and suggest specific cytokine drug targets, such as CXCL10 and TNFSF13B, and potential therapies whose effect on COVID-19 should be studied further (**Supplementary Table 4**).

Analysis of significantly downregulated genes revealed suppression of many ribosomal proteins in both datasets analyzed, suggesting the importance of the interaction of viral proteins with host ribosomes in SARS-CoV-2 infection (**Figure 4)**. Instead of entering the nucleus, SARS-CoV-2 accesses and takes over the ribosomes of the host cell directly. Yuan et al. reported that SARS-CoV-2 uses its nonstructural protein 1 (Nsp1) to bind to ribosomes and suppress host cell protein production while promoting viral protein synthesis, enabling the virus to redirect host protein synthesis machinery towards viral proteins[28]. Specifically, Nsp1 has been found to target 40S, and Schubert et al. used cryo-EM to show how Nsp1 binds to 40S, blocking the mRNA entry channel to inhibit mRNA binding and translation initiation[29]. This mechanism may explain why 40S and other ribosomal proteins were found to be significantly downregulated in COVID-19

patients in the datasets we analyzed (**Supplementary Tables 1 and 2**), indicating translation inhibition consistent with the effects of Nsp1 reported by Schubert et al[29]. The prevalence of significant decreased expression of ribosomal proteins observed in our analysis suggests the importance of this mechanism of viral infection and production in COVID-19. While our DGIdb analysis suggests EIF2AK2 as an upstream drug target of eIF2 signaling, this was the only drug target identified in the pathway (**Figure 4**). Our analysis did not investigate viral proteins, such as Nsp1, as drug targets. More studies investigating drugs with potential to block this mechanism are needed.

Several target genes, pathways, and potential therapies identified by our analyses have not been studied extensively in the context of COVID-19. While the importance of the interferon response in COVID-19 pathogenesis has been characterized by other studies, our results suggest specific key genes in this pathway affected in COVID-19 patients and specific compounds that can target these genes that have not been well-studied. Our analysis reveals the importance of the kinase EIF2AK2 as a hub of the interferon signaling pathway up-regulated in COVID-19 patients, presenting a viable drug target for therapies to modulate the host interferon response (**Figure 5**). As EIF2AK2 plays an important role in two of the pathways identified as most significant in our analyses, further study of the role of EIF2AK2 in COVID-19 pathogenesis and the immune response, as well as further trials of PKR inhibitors in COVID-19 treatment are recommended by our results. The important role of JAK/STAT signaling in interferon response is also suggested by our results, but has also not been extensively studied in the context of COVID-19. Many of the JAK and STAT inhibitors we identified as potential drug repurposing candidates are already under study in COVID-19 treatment[30,31]. Our results help illuminate the mechanisms underlying their impact on the host immune response in SARS-CoV-2 infection and further validate the importance of studies of these potential therapies. We also suggest TLR4 inhibitors as novel drug repurposing candidates for study in COVID-19 treatment. Although the effect of

TLR4 inhibition in COVID-19 has not been studied, our results suggest potential therapeutic efficacy of such a strategy. The significance of TLRs in COVID-19 pathogenesis demonstrated by our results is supported by evidence that SARS-CoV-2 spike proteins interact with extracellular domains of multiple TLRs, including TLR4, demonstrating their important role in SARS-CoV-2 infection[32].

Our data is restricted by the limitations of the GEO datasets available at the time of our analysis, including low quality reads and insufficient annotations providing patient sample information. We did not have access to information about potential underlying conditions or disease severity in COVID-19 patients that could have impacted their gene expression profiles. While the limitation of our inclusion criteria to only include COVID-19 patient samples rather than in vitro SARS-CoV-2 infection limited our sample size, it allowed us to more accurately assess how COVID-19 alters gene expression in patients and identify host therapeutic targets. As our analysis was limited to host gene expression alterations, we could only identify drug candidates that target host genes rather than viral genes. Future studies of virus-host protein interactions may help illuminate the mechanisms underlying how SARS-CoV-2 creates the patterns of differential gene expression we observed in infected samples, and identify drug candidates to target these mechanisms.

Conclusion

Publicly available data has allowed us to perform an unbiased search for host therapeutic targets indicated by gene expression alterations observed in COVID-19 patients, as well as potential drug candidates for these targets. While many studies have focused on potential therapeutics to target viral proteins, our analysis of important host factors contributing to disease pathogenesis provides necessary insight for more patient-focused therapeutic strategies based

on host factors and disease severity. Our results both validate key targets, pathways, and drug candidates that have been reported by other studies, and suggest others that have not been well-characterized and warrant further investigation. Our lists of key targets and pathways provide researchers with many potential avenues for studying major mechanisms of COVID-19 pathogenesis. Additionally, these host therapeutic targets can help guide drug discovery efforts towards those most significantly indicated by COVID-19 disease signatures. Information about drug toxicity, efficacy, and clinical applicability is needed to further validate the drugs identified in our analysis, but our drug list provides researchers with many potential existing FDA-approved drug candidates for studies of COVID-19 treatments in a time when this search for improved therapeutics is urgent. Even in the context of a functioning and widely distributed vaccine, identifying improved treatments for COVID-19 will be relevant, as distribution and compliance with vaccine administration may significantly hinder efforts to control the global pandemic. Additionally, expanding our arsenal of effective treatments is necessary to escape novel SARS-CoV-2 variants now arising and spreading, as well as to prepare for potential future outbreaks.

Acknowledgments.

The author Sivanesan Dakshanamurthy wishes to acknowledge the support by Georgetown University Medical Center COVID19 Pilot Grant, CCSG grant P30 CA051008/CA/NCI NIH HHS/United States, GUMC Lombardi Comprehensive Cancer Center, and GUMC Computational Chemistry Shared Resources (CCSR).

GEO Accession	Platform	Study design	Tissue	Sample size	Infected	Controls
GSE152075	**GPL18573**	Examined host gene expression across infection status, viral load, age, and sex among RNA-sequencing profiles of nasopharyngeal swabs.	Nasopharyngeal swab from upper respiratory tract	484	430	54
GSE150316	**GPL18573**	Autopsy samples from patients deceased due to SARS-Cov2 infection were collected for RNA-seq analysis to assess viral load and immune response.	Lung, heart, jejunum, liver, kidney, bowel, fat, skin, and marrow biopsies	37 (21 lung tissue)	32 (16)	5
GSE147507	**GPL18573**	In humans, primary human lung epithelium (NHBE), lung alveolar cells (A549), and Calu-3 cells were mock treated or infected with SARS-CoV-2, IAV, or RSV. Ferrets were mock treated or infected with pH1N1 virus or SARS-CoV-2. Lung biopsies from two healthy participants and one COVID-19 patient were also analyzed.	Primary human lung epithelium (NHBE), lung alveolar cells (A549), Calu-3 cells, lung biopsies Ferret nasal washes and trachea	110 (4 human lung biopsies)	67 (2)	43 (2)
Total				509	448	61

Table 1. Datasets from the NCBI Gene Expression Omnibus (GEO) database included in analysis. Datasets were identified from the query in the GEO database as of July 28, 2020.

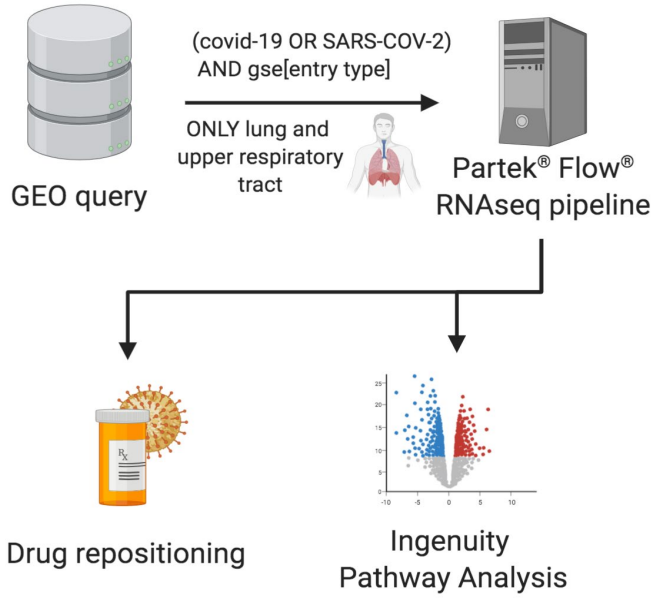
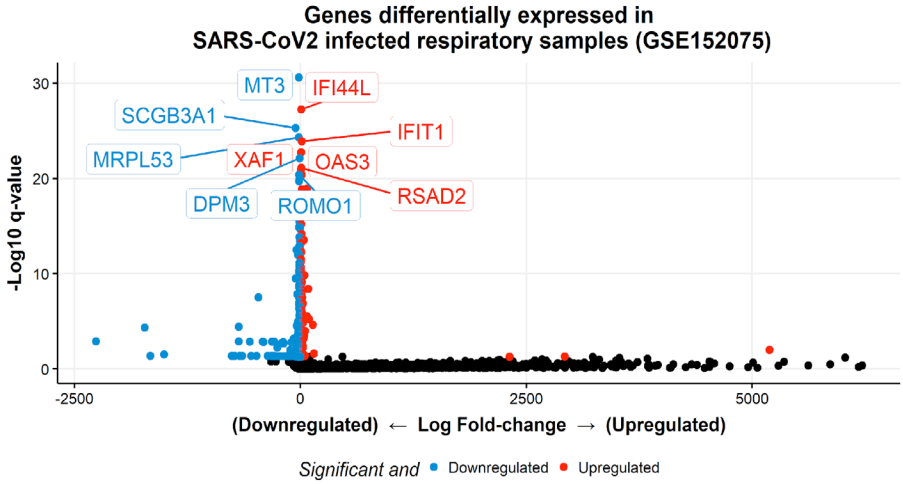
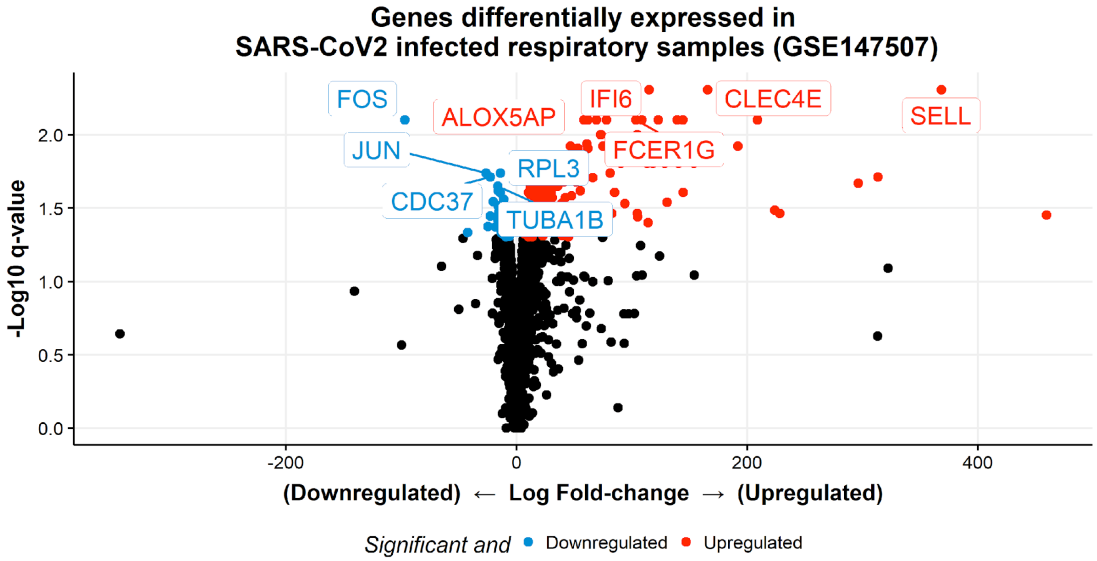


Figure 1. Analysis workflow for therapeutic target identification and drug repositioning based on RNAseq gene expression data. Our analysis began with identifying RNAseq datasets from GEO to differential gene expression analysis of datasets, pathway analysis based on DEGs, and identification of drug repositioning candidates to target significant genes and pathways. Figure created with BioRender.com[33].

(**A**)


Genes differentially expressed in
SARS-CoV2 infected respiratory samples (GSE152075)
MT3
IFI44L
SCGB3A1
IFIT1
MRPL53
XAF1
OAS3
RSAD2
DPM3
ROMO1
-Log10 q-value
30
20
10
0
-2500
0
2500
5000
(Downregulated) ← Log Fold-change → (Upregulated)
Significant and
Downregulated
Upregulated

(**B**)


Genes differentially expressed in
SARS-CoV2 infected respiratory samples (GSE147507)
FOS
ALOX5AP
IFI6
CLEC4E
SELL
JUN
RPL3
FCER1G
CDC37
TUBA1B
-Log10 q-value
2.0
1.5
1.0
0.5
0.0
-200
0
200
400
(Downregulated) ← Log Fold-change → (Upregulated)
Significant and
Downregulated
Upregulated

(C)

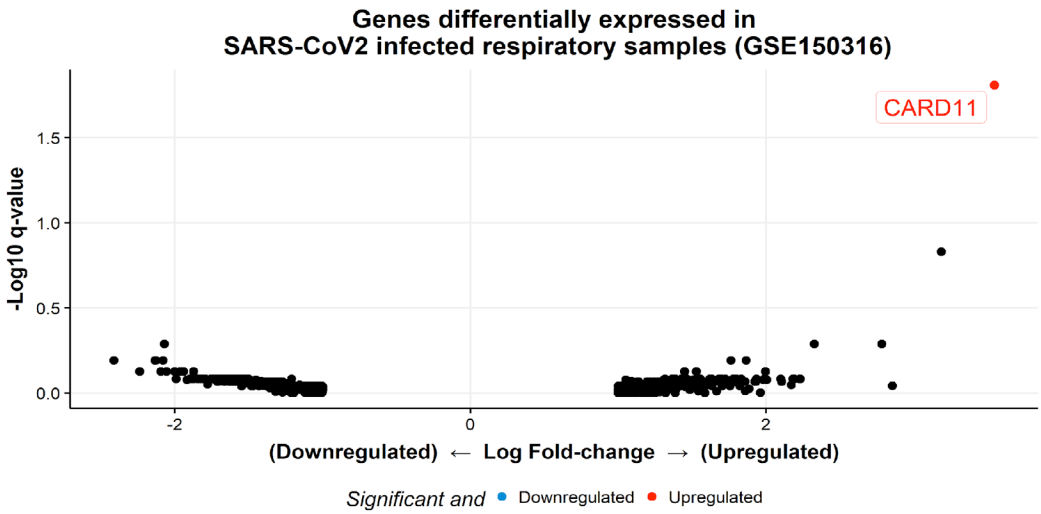


Figure 2. Volcano plots show significantly up- and down-regulated genes in SARS-CoV-2 infected respiratory samples in each of the three analyzed datasets.

Significantly upregulated genes are indicated by red, significantly downregulated genes are indicated by blue. Differentially expressed genes were identified using *Partek*® *Flow*® at a threshold of FDR<0.05 and log2 fold-change ∓1. Volcano plots were generated using R. (**A**) Analysis of GSE152075 (n=484) found 440 significantly up-regulated and 2443 significantly down-regulated genes. (**B**) Analysis of GSE147507 (n=4) found 3 significantly up-regulated and 55 significantly down-regulated genes. (**C**) Analysis of GSE150316 (n=21) found one significantly up-regulated gene and no significantly down-regulated genes.

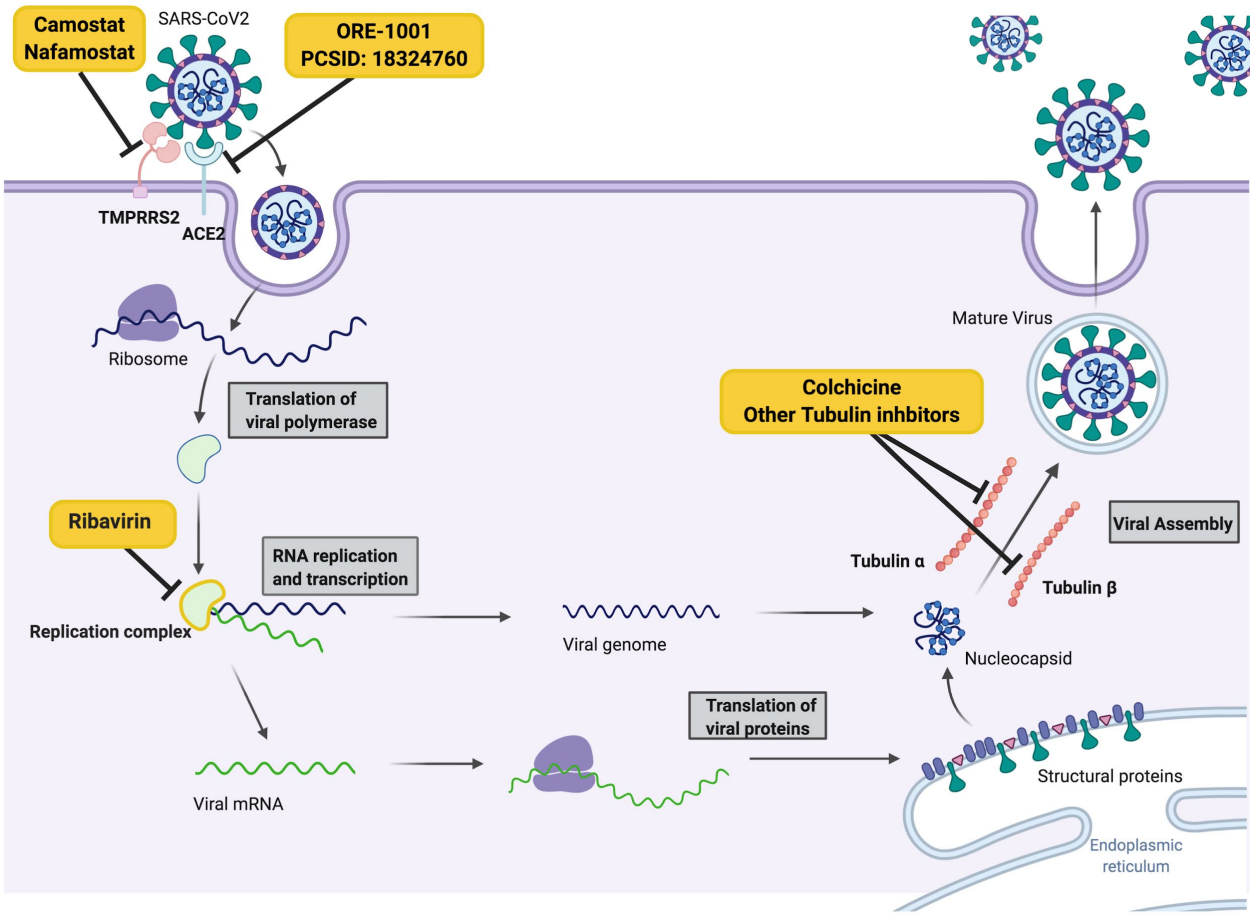


Figure 3. Differentially expressed genes indicate potential drug targets for blocking viral entry, replication, and assembly in the SARS-CoV-2 replication cycle. A simplified representation of the SARS-CoV-2 replication cycle was generated in BioRender.com to highlight the key genes and pathways indicated as significant in our analyses, especially potential drug targets. Compounds identified from the DGIdb analysis whose effects on COVID-19 have been suggested by other studies are shown modulating their respective targets. See **Supplementary Figure 5** for figure legend. Adapted from "Coronavirus Replication Cycle", by BioRender.com (2020). Retrieved from https://app.biorender.com/biorender-templates

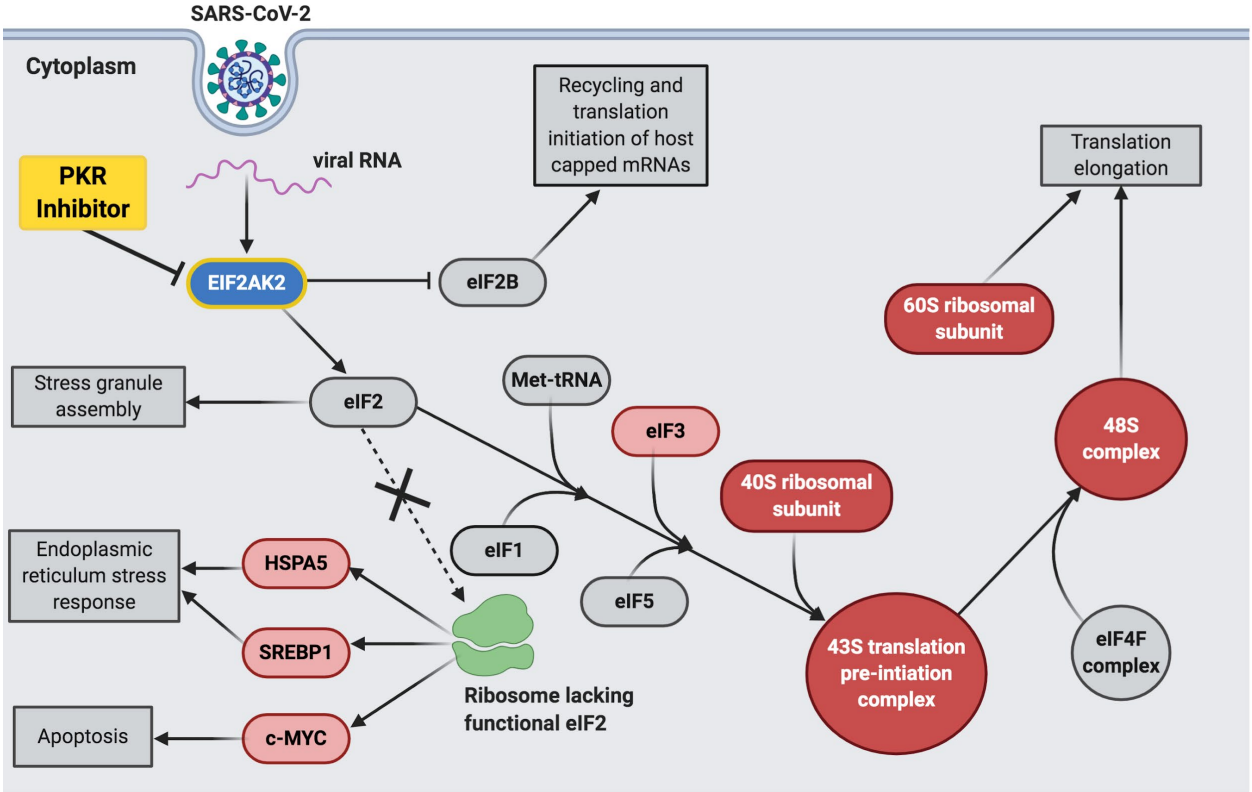


Figure 4. Downregulation of key genes in the eIF2 signaling pathway reveals how SARS-CoV-2 interferes with host cell translation machinery to inhibit host protein synthesis. The eIF2 signaling pathway indicated as significant in our datasets by IPA was recreated in BioRender.com to highlight the elements of the pathway most significantly represented in our data. Genes that were not significantly differentially expressed in our datasets and did not interact with key genes were excluded from the figure for clarity. EIF2AK2 was the only potential drug target identified by DGIdb analysis within this pathway, shown modulated by PKR inhibitors. See **Supplementary Figure 5** for figure legend.

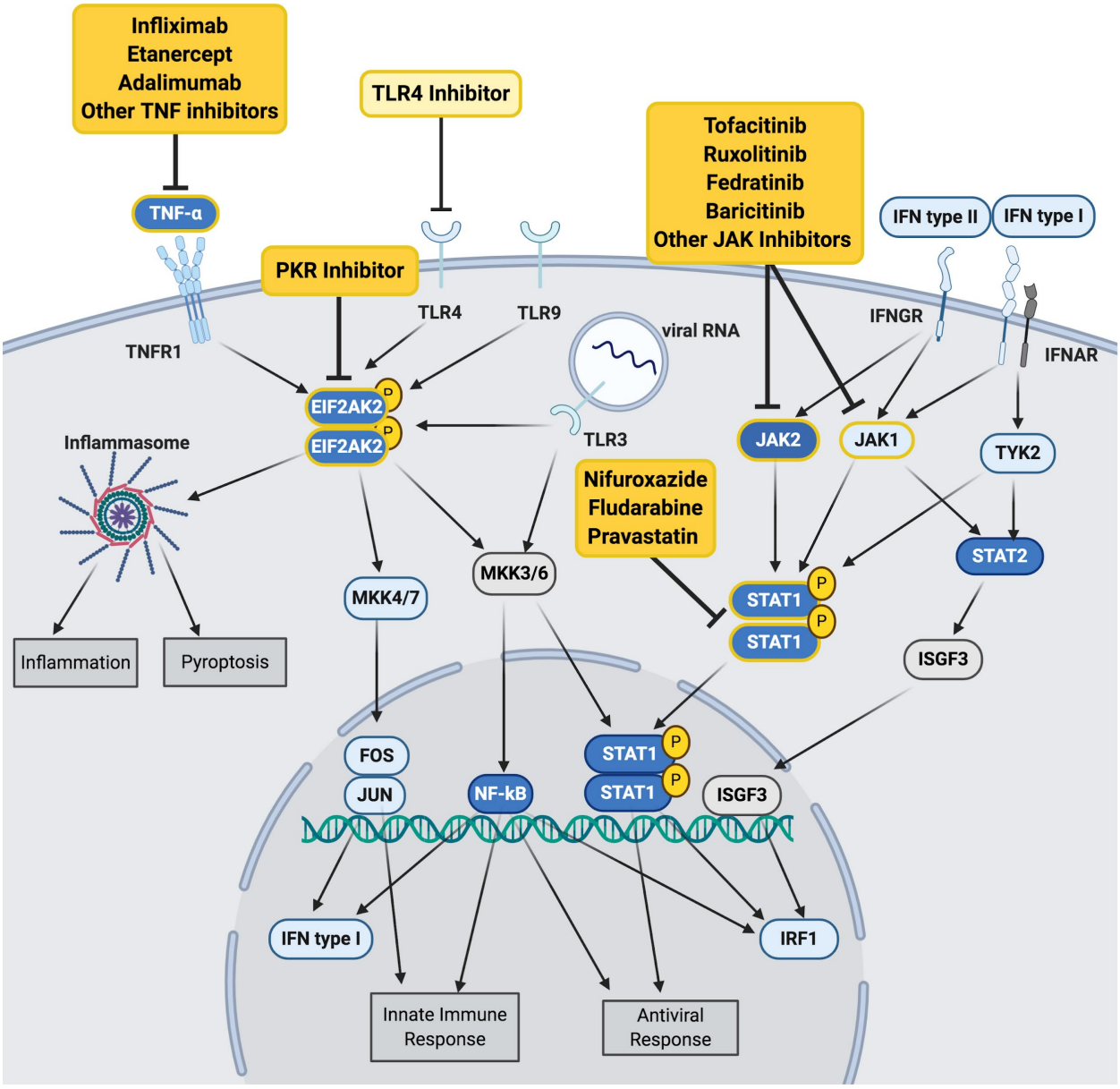


Figure 5. Upregulation of the interferon signaling pathway in SARS-CoV-2 infected samples reveals many key genes and potential drug targets. This pathway figure was designed in BioRender.com based on the EIF2AK2 pathway diagram indicated as significantly represented in our datasets by IPA. Elements of the pathway not significant in the datasets analyzed were excluded for clarity. Compounds identified from the DGIdb analysis whose effects on COVID-19 have been suggested by other studies are shown modulating their respective

targets. TLR4 inhibitors have not been well-studied in the context of COVID-19, indicated by the lighter yellow color. See **Supplementary Figure 5** for figure legend.

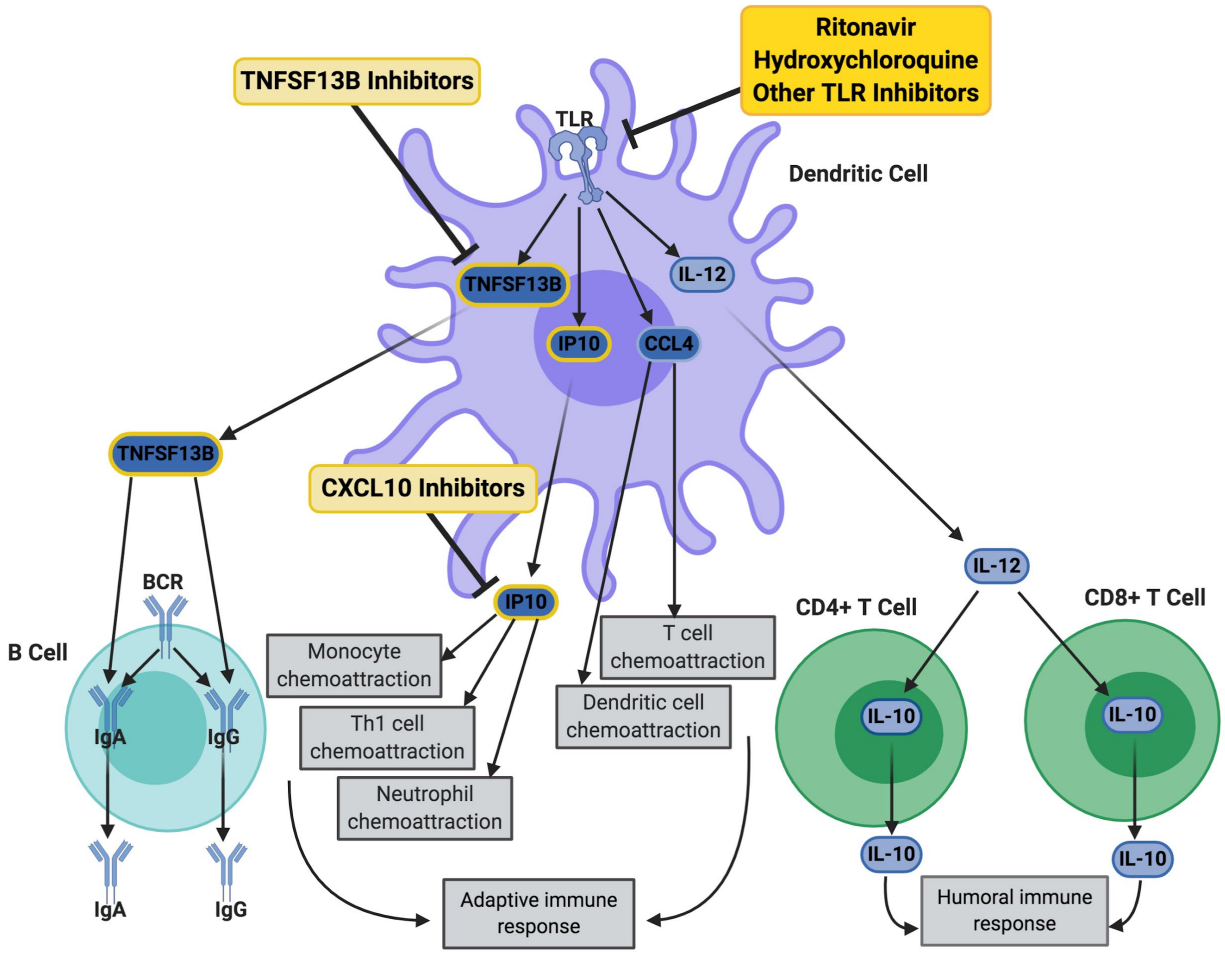


Figure 6. Upregulation of key genes involved in the communication between innate and adaptive immune cells in SARS-CoV-2 infected samples identifies potential repurposing targets for existing drugs. This pathway figure was designed in BioRender.com based on the communication between innate and adaptive immune cells pathway diagram indicated as significantly represented in our datasets by IPA. Elements of the pathway not significant in the datasets analyzed were excluded for clarity. Compounds identified from the DGIdb analysis whose effects on COVID-19 have been suggested by other studies are shown modulating their respective targets. TNFSF13B and CXCL10 inhibitors have not been well-studied in the context

of COVID-19, indicated by their lighter yellow color. See **Supplementary Figure 5** for figure legend.

References

(1) Wu, F.; Zhao, S.; Yu, B.,; Chen, Y.; Wang, W.; Song, Z.; Hu, Y.; Tao, Z.; Tian, J.; Pei, Y.; Yuan, M.; Zhang, Y.; Dai, F.; Liu, Y.; Wang, Q.; Zheng, J.; Xu, L.; Holmes, E. C.; & Zhang, Y. A new coronavirus associated with human respiratory disease in China. Nature. 2020, 579, 265–9. https ://doi.org/10.1038/s4158 6-020-2008-3

(2) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; & Tan, W. A novel coronavirus from patients with pneumonia in China, 2019. The New England Journal of Medicine. 2020, 382(8), 727–733. https://doi.org/10.1056/NEJMoa2001017

(3) COVID-19 Coronavirus Pandemic. 2021. Retrieved January 27, 2021, from https://www.worldometers.info/

(4) Ashburn, T. & Thor, K. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov. 2004, 3, 673–683. https://doi.org/10.1038/nrd1468

(5) Fagone, P.; Ciurleo, R.; Lombardo, S. D.; Iacobello, C.; Palermo, C. I.; Shoenfeld, Y.; Bendtzen, K.; Bramanti, P.; & Nicoletti, F. Transcriptional landscape of SARS-CoV-2 infection dismantles pathogenic pathways activated by the virus, proposes unique sex-specific differences and predicts tailored therapeutic strategies. Autoimmunity Reviews. 2020, 19(7), 102571. https://doi.org/10.1016/j.autrev.2020.102571

(6) Xing, J.; Shankar, R.; Drelich, A.; Paithankar, S.; Chekalin, E.; Dexheimer, T.; Chua, M.; Rajasekaran, S.; Tseng, C. K.; & Chen, B. Analysis of infected host gene expression reveals repurposed drug candidates and time-dependent host response dynamics for COVID-19. BioRxiv. 2020, 2020.04.07.030734. https://doi.org/10.1101/2020.04.07.030734.

(7) Brodin, P. Immune determinants of COVID-19 disease presentation and severity. Nature Medicine. 2021, 27(1), 28–33. https://doi.org/10.1038/s41591-020-01202-8.

(8) Barrett, T.; Wilhite, S. E.; Ledoux, P.; Evangelista, C.; Kim, I. F.; Tomashevsky, M.; Marshall, K. A.; Phillippy, K. H.; Sherman, P. M.; Holko, M.; Yefanov, A.; Lee, H.; Zhang, N.; Robertson, C. L.; Serova, N.; Davis, S.; & Soboleva, A. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Research. 2013, 41(Database issue):D991-5.

(9) Partek Inc. Partek® Flow® (Version 10.0) [Computer software]. 2020. https://www.partek.com/partek-flow/

(10) Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014,

30(4), 523-30.

(11) Freshour, S.; Kiwala, S.; Cotto, K. C.; Coffman, A. C.; McMichael, J. F.; Song, J.; Griffith, M.; Griffith, O. L.; & Wagner, A. H. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Research. 2020, https://doi.org/10.1093/nar/gkaa1084. PMID: 33237278

(12) Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K. M. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020, 583, 459–468. https://doi.org/10.1038/s41586-020-2286-9

(13) Schlesinger, N.; Firestein, B.L.; & Brunetti, L. Colchicine in COVID-19: An old drug, new use. Curr Pharmacol Rep. 2020, 6, 137–145 https://doi.org/10.1007/s40495-020-00225-6

(14) Zhaori, G.; Lu, L.; Liu, C.; & Guo, Y. Progresses in clinical studies on antiviral therapies for COVID-19—Experience and lessons in design of clinical trials. Pediatr Invest. 2020, 4, 263- 274. https://doi.org/10.1002/ped4.12227

(15) Robinson, P. C.; Richards, D.; Tanner, H. L.; & Feldmann, M. Accumulating evidence suggests anti-TNF therapy needs to be given trial priority in COVID-19 treatment. The Lancet Rheumatology. 2020, 2(11), e653–55. https://doi.org/10.1016/S2665-9913(20)30309-X.

(16) Bansal, P.; Goyal, A.; Cusick IV, A.; Lahan, S.; Dhaliwal, H. S.; Bhyan, P.; Brijmohan Bhattad, P.; Aslam, F.; Ranka, S.; Dalia, T.; Chabbra, L.; Sanghavi, D.; Sonani, B.; & Davis III, J. M. Hydroxychloroquine: A comprehensive review and its controversial role in Coronavirus Disease 2019. Annals of Medicine. 2021, 53(1), 117–34. https://doi.org/10.1080/07853890.2020.1839959.

(17) RECOVERY Collaborative Group. Lopinavir–Ritonavir in patients admitted to hospital with COVID-19 (RECOVERY): A randomised, controlled, open-Label, platform trial. *The* Lancet. 2020, 396(10259), 1345–52. https://doi.org/10.1016/S0140-6736(20)32013-4.

(18) Ungogo, M. A.; Mohammed, M.; Umar, B. N.; Bala, A. A.; & Khalid, G. M. Review of pharmacologic and immunologic agents in the management of COVID-19. Biosafety and Health. 2021. https://doi.org/10.1016/j.bsheal.2021.01.001.

(19) RECOVERY Collaborative. Effect of Hydroxychloroquine in hospitalized patients with COVID-19. New England Journal of Medicine. 2020. https://doi.org/10.1056/NEJMoa2022926.

(20) WHO Solidarity Trial Consortium. Repurposed antiviral drugs for COVID-19 —

Interim WHO Solidarity Trial results. New England Journal of Medicine. 2020. https://doi.org/10.1056/NEJMoa2023184.

(21) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T.S.; Harrier, G.; Wu, N.; Nitsche, A.; Muller, M.A.; Drosten, C.; & Pohlmann, S. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020, 181(2), 271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052.

(22) Siegfried, A.; Berchtold, S.; Manncke, B.; Deuschle, E.; Reber, J.; Ott, T.; Weber, M.; Kalinke, U.; Hofer, M. J.; Hatesuer, B.; Schughart, K.; Gailus-Durner, V.; Fuchs, H.; Hrabe de Angelis, M.; Weber, F.; Hornef, M.W.; Autenrieth, I.B.; & Bohn, E. IFIT2 is an effector protein of type I IFN-mediated amplification of lipopolysaccharide (LPS)-induced TNF-α secretion and LPS-induced endotoxin shock. Journal of Immunology. 2013, 191(7), 3913–3921. https://doi.org/10.4049/jimmunol.1203305

(23) Hadjadj, J.; Yatim, N.; Barnabei, L.; Corneau, A.; Boussier, J.; Smith, N. et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science. 2020, 369(6504), 718. https://doi.org/10.1126/science.abc6027.

(24) Bozzo, C.P.; Nchioua, R.; Volcic, M.; Krüger, J.; Heller, S.; Stürzel, C.M. et al. IFITM proteins promote SARS-CoV-2 infection and are targets for virus inhibition. BioRxiv. 2020, 2020.08.18.255935. https://doi.org/10.1101/2020.08.18.255935.

(25) Kumar, R.; Verma, H.; Singhvi, N.; Sood, U.; Gupta, V.; Singh, M.; Kumari, R.; Hira, P.; Nagar, S.; Talwar, C.; Nayyar, N.; Anand, S.; Rawat, C.D.; Verma, M.; Negi, R.M.; Singh, Y.; & Lal, R. Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogeographical distribution. Edited by Ileana M. Cristea. MSystems. 2020, 5(4), e00505-20. https://doi.org/10.1128/mSystems.00505-20.

(26) Moore, J. B. & June, C. H. Cytokine release syndrome in severe COVID-19. Science. 2020, 368(6490), 473. https://doi.org/10.1126/science.abb8925.

(27) Zhang, B.; Zhou, X.; Zhu, C.; Song, Y.; Feng, F.; Qiu, Y.; Feng, J.; Jia, Q.; Song, Q.; Zhu, B.; & Wang, J. Immune phenotyping based on the neutrophil-to-lymphocyte ratio and IgG level predicts disease severity and outcome for patients with COVID-19. Frontiers in Molecular Biosciences. 2020. 7, 157. https://doi.org/10.3389/fmolb.2020.00157

(28) Yuan, S.; Peng, L.; Park, J.J.; Hu, Y.; Devarkar, S.C.; Dong, M.B.; Shen, Q.; Wu, S.; Chen, S.; Lomakin, I.B.; & Xiong, Y. Nonstructural protein 1 of SARS-CoV-2 is a potent pathogenicity factor redirecting host protein synthesis machinery toward viral RNA. Molecular Cell. 2020, 80(6), 1055-1066.e6. https://doi.org/10.1016/j.molcel.2020.10.034

(29) Schubert, K.; Karousis, E.D.; Jomaa, A.; Scaiola, A.; Echeverria, B.; Gurzeler, L.; Leibundgut, M.; Thiel, V.; Muhlemann, O.; & Ban, N. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. Nat Struct Mol Biol. 2020, 27, 959–966 https://doi.org/10.1038/s41594-020-0511-8

(30) Goker Bagca, B. & Biray Avci, C. The potential of JAK/STAT pathway inhibition by ruxolitinib in the treatment of COVID-19. Cytokine & Growth Factor Reviews. 2020, *54*, 51–62. https://doi.org/10.1016/j.cytogfr.2020.06.013

(31) Satarker, S.; Tom, A.A.; Shaji R.A.; Alosious, A.; Luvis, M.; & Nampoothiri, M. JAK-STAT pathway inhibition and their implications in COVID-19 therapy. Postgraduate Medicine. 2020, 0(0), 1–19. https://doi.org/10.1080/00325481.2020.1855921.

(32) Choudhury, A. & Mukherjee, S. In silico studies on the comparative characterization of the interactions of SARS-CoV-2 spike glycoprotein with ACE-2 receptor homologs and human TLRs. J Med Virol. 2020, 1-9. https://doi.org/10.1002/jmv.25987

(33) BioRender. (n.d.). Retrieved from https://biorender.com/

SUPPORTING INFORMATION:

Supplementary Figure 1. PRISMA diagram

Supplementary Figure 2. Bar graph showing significantly up- and down-regulated canonical pathways in SARS-CoV- 2 infected samples from dataset 147507.

Supplementary Figure 3. Bar graph showing significantly up- and down-regulated canonical pathways in SARS-CoV- 2 infected samples from dataset 152075.

Supplementary Figure 4. Heat map of canonical pathways most significantly and consistently up- and down-regulated in SARS-CoV-2 infected samples across both datasets analyzed.

Supplementary Figure 5. Figure legend for pathway figures.

Supplementary Table 1. Significantly differentially expressed genes in SARS-CoV-2 infected samples from dataset GSE147507.

Supplementary Table 2. Significantly differentially expressed genes in SARS-CoV-2 infected samples from dataset GSE152075.

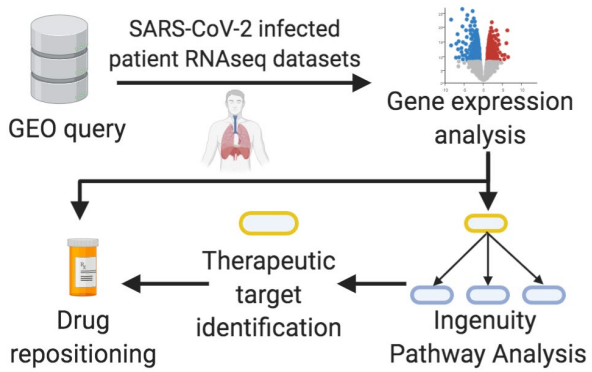
Supplementary Table 3. Genes significantly upregulated across both datasets.

Supplementary Table 4. Genes significantly downregulated across both datasets.

Supplementary Table 5. Drugs with experimentally determined associations with genes significantly upregulated in our datasets.

Supplementary Table 6. Drugs with experimentally determined associations with genes significantly downregulated in our datasets.

For TOC only:


SARS-CoV-2 infected
patient RNAseq datasets
GEO query
Gene expression
analysis
Therapeutic
target
identification
Drug
repositioning
Ingenuity
Pathway Analysis