Article

SARS-CoV-2 Infections - Gene Expression Omnibus (GEO) Data Mining, Pathway Enrichment Analysis, and Prediction of Repurposable Drugs/Compounds

Srilakshmi Chaparala, Carrie L Iwema and Ansuman Chattopadhyay *

Health Sciences Library System, University of Pittsburgh, Pittsburgh, PA, USA
*   Correspondence: ansuman@pitt.edu

Abstract: The COVID-19 global pandemic has created dire consequences with an alarming rate of morbidity and mortality. There are not yet vaccine or efficacious treatment options to combat the causative SARS-CoV-2 infection. This paper describes the identification of potentially repurposable drugs for COVID-19 treatment by conducting pathway enrichment analysis on publicly available Gene Expression Omnibus datasets. We first determined SARS-CoV-2 infection-induced alterations of host gene expressions and pathways. We then identified drugs or compounds that target and counter virus-triggered cellular perturbations, suggesting their potential repurposing for COVID-19 treatment. The key findings are that SARS-CoV-2 infection in host cells induces mitochondrial dysfunction, inhibits oxidative phosphorylation, and activates several immune response and pro-inflammatory pathways. Triptolide, the major bioactive component of a traditional Chinese medicine herb, may rescue mitochondrial dysfunction by activating oxidative phosphorylation. Further *in vitro* and *in vivo* studies are necessary to verify these results prior to clinical application.

Keywords: COVID-19; coronavirus; drug repositioning; pathway enrichment

Introduction

The novel Coronavirus Disease, COVID-19, was declared a pandemic by the World Health Organization on March 11, 2020 [1]. The causative virus, Severe Acute Respiratory Syndrome Coronavirus-2 (SARS- CoV-2), first emerged at the end of 2019, and rapidly spread throughout the world. The morbidity and mortality statistics are overwhelming; as of September 15, 2020, 29.3 million people worldwide are reported to be infected with the virus, with a mortality of 930 thousand people [2].

SARS-CoV-2 is a positive-sense single-stranded RNA virus belonging to the genus beta group of the Orthocoronavirinae family that shares genome sequence similarity with two other highly pathogenic members of the group, Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and the Middle East Respiratory Syndrome Coronavirus (MERS-CoV). All three members are thought to have emerged from zoonotic events that triggered global outbreaks during the past two decades [3].

The SARS-CoV-2 genome is organized into two segments. The 5' end covering two-thirds of the genome encodes 16 non-structural proteins (Nsp 1-16), and the 3' end encodes four structural proteins: spike (S), membrane (M), envelope (E), and nucleocapsid (N) [4]. Through its S protein, SARS-CoV-2 binds to the Angiotensin-Converting Enzyme-2 (ACE2) receptor expressed on the host cell surface and enters the cell through receptor-mediated endocytosis.

Since SARS-CoV-2 emerged recently, much of the pathophysiology of COVID-19 is not fully understood. Scientists from all over the world rapidly engaged in studying this virus and began publishing results as well as posting findings in preprint servers for immediate knowledge dissemination. As of September 15, 2020, a search for "COVID-19 OR SARS-CoV-2" in the Europe PMC database results in over 103,000 research articles and 13,000 preprints, which illustrates both the high volume and rapid publication timeline of COVID-19 research [5].

The clinical manifestation of SARS-CoV-2 infections varies and includes a broad range of symptoms, from none to mild such as fever, cough, and general malaise in the majority of cases, to acute respiratory distress syndrome (ARDS) and acute lung injury in severe cases [6,7]. Currently there is no cure, vaccine, or standard treatment protocol for COVID-19. The available therapeutic approach is symptomatic treatment with previously approved medications used to treat other disease indications [8], including the anti-malarial drug Hydroxychloroquine (HCQ) alone or in conjunction with the antibiotic Azithromycin [9,10], the antiviral drug Remdesivir [11], and the synthetic corticosteroid Dexamethasone [12].

Unfortunately, the lack of efficacy and occurrence of severe side effects in some cases limits the utility of HCQ and Azithromycin, and they have since been discontinued from clinical trials [13]. A recent clinical study revealed that Remdesivir is superior to placebo in reducing the recovery time of hospitalized COVID-19 patients, but ineffective on reducing the mortality rate [11]. Despite some success in reducing the mortality rate [12,14,15], using Dexamethasone to treat COVID-19 is clinically challenging [16]. Drug administration timing and dosage are critical as too early administration of Dexamethasone, a known immune suppressor, could prevent initiation of the body's immune defense mechanism. This results in increased viral loads leading to severe adverse consequences, as previously reported with SARS-CoV infection [17].

The paucity of efficacious treatment options for COVID-19 prompted us to search for additional potential repurposable drugs. To achieve this, we adopted an *in silico* approach based on transcriptomics data analysis. Since cellular state is commonly determined by studying a cell's transcriptomic profile, comparison of the transcriptomic signature between two conditions (e.g., disease vs. normal healthy state, virus-infected vs. uninfected status, drug vs. mock treatment) captures a snapshot of the altered cellular state induced by the perturbing agent or condition [18]. To uncover the molecular mechanism of virus infections, disease progressions, and a drug's mode of action, researchers routinely conduct transcriptomic profile studies and share their data via deposition into public repositories such as the NCBI Gene Expression Omnibus (GEO) [19]. As a result, gene expression databases are a rich resource for scientific discovery, providing functional genomics

data from over 3.5 million samples. The transcriptomic profile of SARS-CoV-2 infections, as well as several drug-induced gene expression signatures, are readily available within these databases.

In this paper we discover potentially repurposable drugs for COVID-19 treatment by conducting pathway enrichment analysis on publicly available GEO datasets. First, we describe SARS-CoV-2 infection-induced alterations of host gene expression and pathways. Then we identify drugs that target and counter virus-induced cellular perturbations, suggesting their potential repurposing for COVID-19 treatment. Further *in vitro* and *in vivo* studies will be necessary to verify these results prior to clinical application.

Methods

Workflow

Figure 1 outlines the 5 steps taken in this study. They are described in detail below.
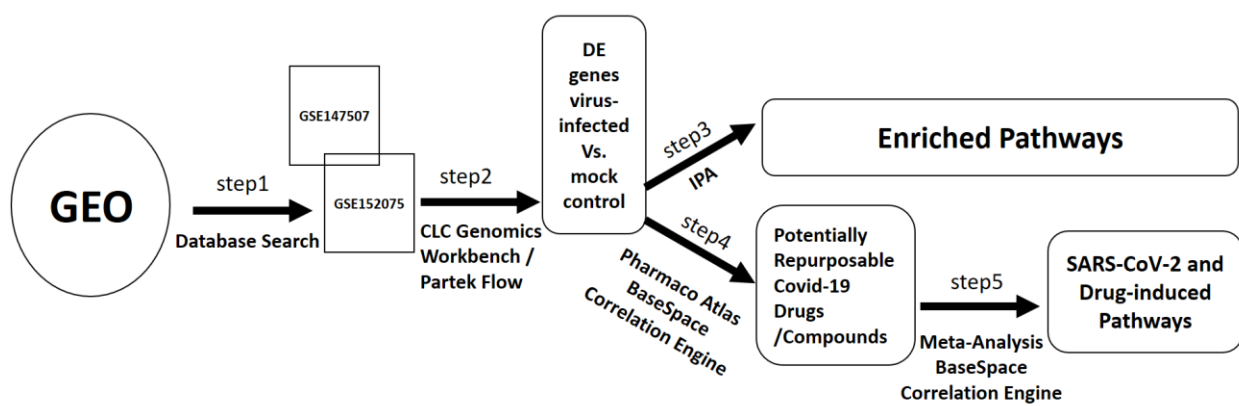

Figure 1. Methods workflow

1. Identify SARS-CoV-2-induced host-response datasets in NCBI Gene Expression Omnibus (GEO).
2. Identify differentially expressed (DE) genes (virus-infected vs. mock control) using either CLC Genomics Workbench or Partek Flow software.
3. Identify statistically enriched pathways associated with SARS-CoV-2-induced host genes using Ingenuity Pathway Analysis software.
4. Identify a list of drugs/compounds revealing the opposite pattern of SARS-CoV-2-induced host gene expression signatures using the Pharmaco Atlas application of BaseSpace Correlation Engine software.
5. Compare pathway activity states (up- or down-regulated) between SARS-CoV-2-induced and drug-induced human genes using the Meta-Analysis tool of BaseSpace Correlation Engine software.

GEO data identification (Step 1)

We searched the NCBI GEO datasets [20] with the terms "SARS-CoV-2 OR COVID-19" and the filters "human" for organism and "expression profiling by high throughput sequencing" as study type. We sought datasets where SARS-CoV-2 infection was studied on established laboratory cell lines and on cell cultures derived from cells isolated from SARS-CoV-2 infected humans. Out of 30 search results (July 25, 2020) we identified GSE147507 [21] (bulk RNA-Seq data generated from various cell lines) and GSE152075 [22] (RNA-Seq data from nasopharyngeal cells of SARS-CoV-2 positive and

negative individuals). More information on the GEO dataset search results is available in supplementary methods (SM1).

RNA-Seq analysis (Step 2)

For GSE147507, CLC Genomics Workbench v20 (QIAGEN Digital Insights) [23] was used to perform RNA-Seq data analysis on raw sequencing reads and for computing differentially expressed (DE) genes between virus-infected vs. mock controls. For GSE152075, DE genes were identified from gene expression count data using Partek Flow software (Partek Inc.) [24]. A detailed description of RNA-Seq analysis methods by CLC Genomics Workbench is available in supplementary methods (SM2).

Pathway enrichment analysis (Step 3)

The Ingenuity Pathway Analysis (IPA) suite of tools (QIAGEN Digital Insights) [25] was used on DE genes to identify statistically enriched biological pathways induced by virus infection. We uploaded DE genes between the virus-infected vs. mock-control samples and applied the IPA Core Analysis tool to identify statistically significant metabolic and signaling pathways. Statistical significance was calculated using the right-tailed Fisher Exact Probability Tests; biological pathways showing p-value < 0.05 were considered statistically significant. The activity status of pathways was determined by calculating the activity z-score [26], which is a statistical measure of how closely the gene expression pattern present in the query dataset compares to the expected pattern based on the literature findings. A positive score indicates an overall increase in the pathway activity, whereas a negative value indicates an overall decrease in activity. The IPA Comparison Analysis tool was used to compare pathway enrichment analysis results generated from the multiple datasets used in our study.

Identification of potentially repurposable drugs and drug target pathways (Steps 4 and 5)

DE genes between SARS-CoV-2-infected vs. mock-treated samples from series 16 of GSE147507 were uploaded to BaseSpace Correlation Engine (BSCE) software (Illumina Inc.) [27]. The Pharmaco Atlas tool was used to identify drugs exhibiting a negative correlation to SARS-CoV-2-induced gene expression profiles.

The top five drugs from the Pharmaco Atlas results (sorted by the negative correlation score) were used to search BSCE datasets comprising DE genes between drug treatment vs. control experiments using the filters "human" for organism, "RNA expression" for data type, and "lung" as a keyword. The identified datasets are available in supplementary methods (SM3).

All identified datasets along with the SARS-CoV-2-induced host DE genes from series 16 of GSE147507 were then added to the BSCE *Meta-Analysis* tool. The filter term "Canonical pathways | Broad MSigDB" available under the "BioGroup Results" option was used to compare the activity patterns of statistically enriched pathways across SARS-CoV-2 and drug-induced genes. The BSCE datasets and analyses are available in supplementary methods (SM4).

Results

The number of SARS-CoV-2-induced genes correlates with increased viral load and level of ACE2 expression.

To understand the host viral response mechanism, we (1) measured gene expression, (2) identified the DE genes between virus-infected (positive) vs. mock-control (negative) samples from each dataset, and (3) performed pathway enrichment analysis.

The number of DE genes was used as a measure of the virus-induced host response (Table 1).

Table 1. RNA-seq analysis results of GSE147507 samples showing host gene expressions.

Sample	Cell Line	Treatment	MOI	Mapped to Human genome %(SD)	ACE2 TPM Mean	TMPRSS2 TPM Mean	NRP1 TPM Mean	CTSB TPM Mean	CTSL TPM Mean	DEG Virus infected Vs. Mock
				GSE147507						
Series 1	NHBE	Mock	-	95.17(1.45)	1.78	5.53	16.19	669.93	153.38	147
		SARS-Cov2	2.0	95.39(0.32)	2.11	4.18	12.13	705.77	152.22	
Series 2	A549	Mock	-	92.38(2.51)	0.00	0.00	70.53	170.66	918.86	63
		SARS-Cov2	0.2	92.21(2.48)	0.00	0.00	69.93	161.44	958.51	
Series 5	A549	Mock	-	96.79(0.23)	0.00	0.00	59.89	142.81	613.25	2430
		SARS-Cov2	2.0	97.52(0.03)	0.23	0.05	87.54	220.09	733.14	
Series 6	A549 + ACE2	Mock	-	82.34(0.79)	17587.88	0.00	26.47	187.00	347.40	1586
		SARS-Cov2	0.2	36.07(2.55)	25240.34	0.00	35.90	232.19	353.42	
Series 16	A549 + ACE2	Mock	-	74.73(2.34)	22722.79	0.00	37.90	95.65	421.10	4684
		SARS-Cov2	2.0	27.69(0.63)	28486.18	0.00	36.53	138.44	240.33	
Series 7	Calu3	Mock	-	98.16(0.14)	26.66	8.85	22.46	232.19	76.45	2424
		SARS-Cov2	2.0	82.43(1.61)	20.65	18.95	28.09	262.51	60.59	
Series 8	A549	Mock	-	98.22(0.32)	0.03	0	49.18	119.61	639.92	2048
		RSV	2.0	80.6(2.48)	0.08	0.04	115.06	315.16	1903.66	
		HPIV3	3.0	87.96(1.12)	0.29	0	99.93	132.46	1322.23	1024

In agreement with Blanco-Melo et al. (contributors of the GSE140507 dataset), we observed that the number of DE genes increased in parallel to the multiplicity of infection (MOI) used for virus infection. Between series 2 and 5, a 10-fold MOI from 0.2 to 2.0 increased the number of DE genes from 63 to 2430. Similarly, in A549 cells treated with exogenous ACE2 (series 6 and 16), a 10-fold MOI from 0.2 to 2.0 increased the number of DE genes from 1586 to 4648.

We also checked the expression of several genes associated with SARS-CoV-2 pathogenesis. For example, the ACE2 and TMPRSS2 genes facilitate virus entry and protein priming, respectively [28]. We observed that both ACE2 and TMPRSS2 are expressed in NHBE and Calu3 cells but not in A549 cells. As expected ACE2 is highly expressed in the A549 cells when supplemented with the exogenous ACE2 (series 6 and 16).

Since the A549 wild type cells exhibited an absence of ACE2 and TMPRSS2 expression, we searched for an alternative gene that the virus might use to gain cellular entry. A recent finding suggests that Neuropilin Receptor 1 (NRP1) may serve as a host factor for SARS-CoV-2 infection [29]. We observed NRP1 expression in all cell lines we examined, including A549. Two endosomal cysteine proteases, Cathepsin L/B (CATL and CATB), exhibit the potential to compensate for TMPRSS2 [30] and are also expressed in A549 cells.

SARS-CoV-2 infection results in mitochondrial dysfunction, activation of immune response pathways, inhibition of oxidative phosphorylation, and protein synthesis initiation.

To ascertain the SARS-CoV-2-induced alterations of signaling and metabolic pathways in host cells, we selected the GSE147507 series 16 dataset, as with 4684 DE genes it had the most robust host response. We performed pathway enrichment analysis using Ingenuity Pathway Analysis (IPA) software and show the top ten statistically enriched pathways in Figure 2. The activity status of a pathway is calculated by activity z-score algorithm [26]. The full list of statistically enriched pathways is available in a supplementary document (SD1).
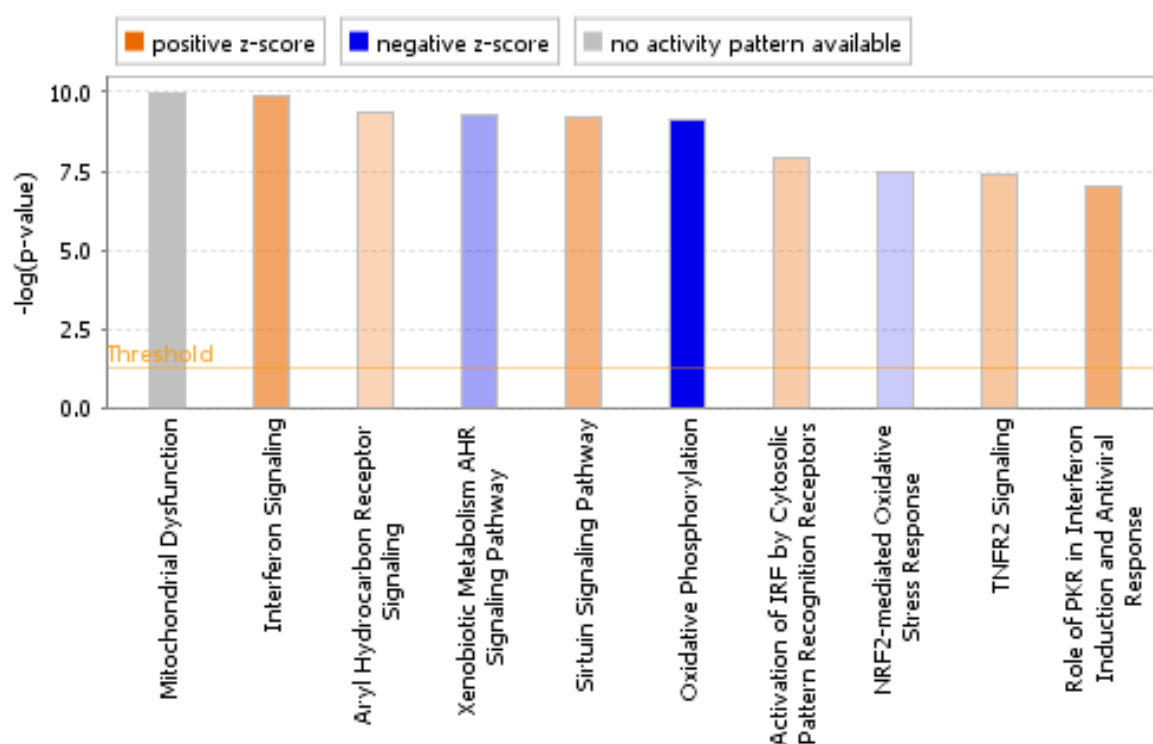

Figure 2. Top ten significantly enriched biological pathways for SARS-CoV-2-induced human gene expression. Ingenuity Pathway Analysis software was used to determine statistically significant pathways (x-axis = name of pathway; y-axis = –log of the p-value). The significance value (p- value of overlap) of the pathways was calculated by the right-tailed Fisher's Exact Test. Color intensity is proportional to the magnitude of the IPA z-score algorithm and reflects the degree of pathway activation (orange) or inhibition (blue). The gray bar indicates a pathway for which no prediction can be made due to insufficient evidence in the Ingenuity Knowledgebase.

The most statistically significant enriched pathway was mitochondrial dysfunction. The most inhibited pathway was oxidative phosphorylation, which is a mitochondrial metabolic pathway. The activated pathways include several immune-response and pro-inflammatory signaling pathways: Interferon, Sirtuin, Tumor Necrosis Factor 2 (TNFR2), and cytosolic pattern recognition receptor-mediated Interferon Regulatory Factor (IRF) activation. Supplementary figures provide examples of IPA pathway diagrams overlaid with DE genes observed between SARS-CoV-2-infected vs. mock controls for the most activated (Interferon signaling; SF1) and most inhibited (oxidative phosphorylation; SF2) pathways.

To ensure IPA analysis validity, we compared the IPA results with the BSCE-computed pathway enrichment results for the series 16 DE genes (SARS-CoV-2 infected vs. mock control). Whereas IPA searches a proprietary pathway knowledgebase and uses the right-tailed Fisher Exact test to identify pathway enrichments, BSCE uses an open-access Molecular Signatures Database (MSigDB) collection [31] and applies a modified gene set enrichment analysis algorithm to compute pathway enrichment

scores[27]. The full list of BSCE-computed statistically enriched pathways is available in a supplementary document (SD2). We determined that the two bioinformatics tools provide similar results, particularly when comparing the statistical enrichment of immune response and pro-inflammatory pathways activation and mitochondrial pathways inhibition.

We also performed pathway enrichment analysis of DE genes from two additional datasets (GSE140507, GSE 152075) comprising four cell types (NHBE, A549, Calu3, NSP) and three virus infection types (SARS-CoV-2, RSV, HIPV). To understand which biological processes are relevant to each dataset, we applied the IPA Comparison Analysis tool. We used hierarchical clustering for sorting and activity z-scores for displaying the comparison analysis results. Figure 3 shows the top 20 pathway analysis comparison results. The entire comparison analysis results are available in a supplementary document (SD3).
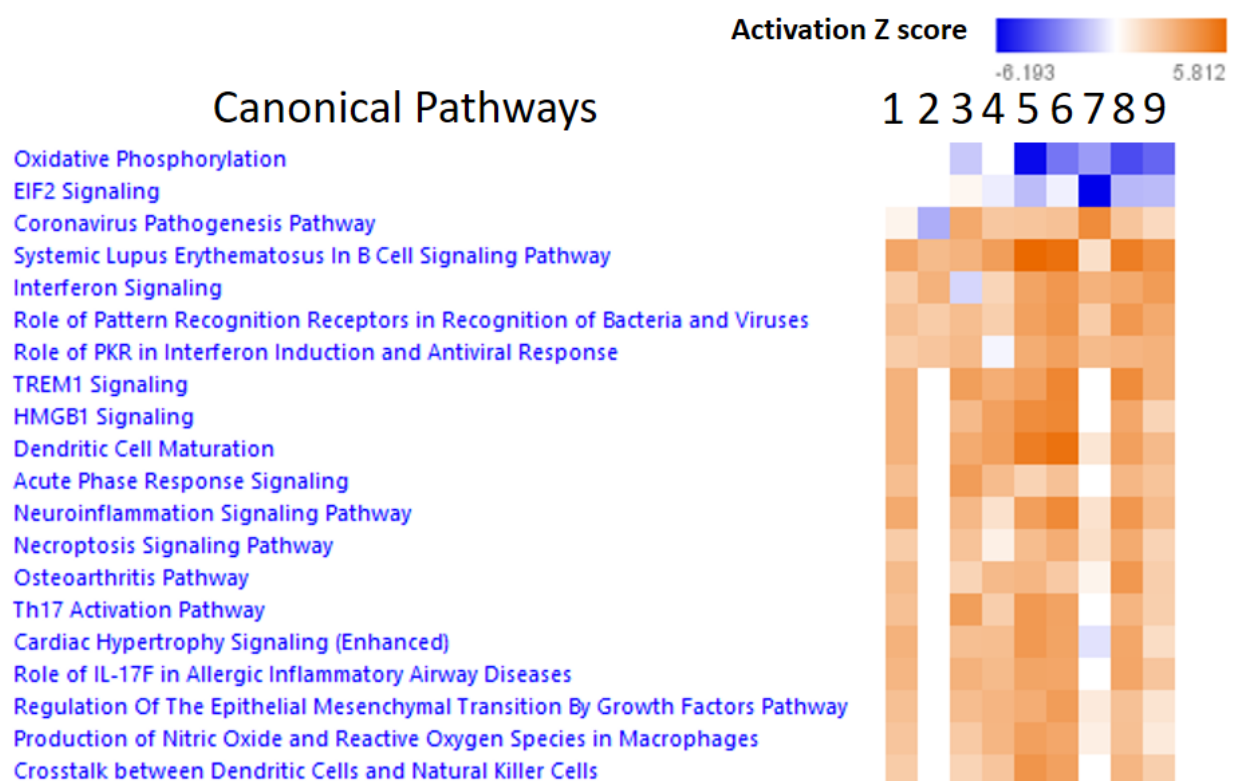

Figure 3. Comparison of biological pathways associated with differentially expressed genes between SARS-CoV-2, RSV, and HIPV-infected vs. mock control. (1) NHBE cells + SARS-CoV-2 (MOI 2.0) vs. mock; (2) A549 cells + SARS-CoV-2 (MOI 0.02) vs. mock; (3) A549 cells + SARS-CoV-2 (MOI 2.0) vs. mock; (4) A549 cells + ACE2 + SARS-CoV-2 (MOI 0.02) vs. mock; (5) A549 cells + ACE2 + SARS-CoV-2 (MOI 2.0) vs. mock; (6) Calu3 cells + SARS-CoV-2 (MOI 2.0) vs. mock; (7) NSP cells from SARS-CoV-2 positive vs. negative individuals; (8) A549 cells + RSV (MOI 2.0) vs. mock; (9) A549 cells + HIPV (MOI 3.0) vs. mock. Color intensity is proportional to the magnitude of the IPA z-score algorithm and reflects the degree of pathway activation (orange) or inhibition (blue).

For A549 cells in GSE140507 (Figure 3, lanes 2- 5), both oxidative phosphorylation and the Eukaryotic Initiation Factor 2 (EIF2) mediated protein synthesis pathway demonstrate SARS-CoV-2 dose-dependent inhibition. Oxidative phosphorylation exhibits increased inhibition corresponding to a 10-fold SARS-CoV-2 MOI increase in the A549 cells (lanes 2 vs 3), and both pathways exhibit increased inhibition in A549 cells + ACE2 (lanes 4 vs 5). A similar inhibitory trend is observed in Calu3 cells (lane 6), nasopharyngeal (NSP) cell swabs from SARS-CoV-2 positive vs negative individuals (lane 7), and A549 cells infected with RSV and HIPV (MOI 2.00) (lanes 8, 9).

Our NSP cell results are in agreement with the contributors of GSE152075, who reported that host responses to SARS-CoV-2 are dependent on viral load and observed a strong inhibition of

oxidative phosphorylation in SARS-CoV-2 infected cells through gene set enrichment analysis [32]. The activated pathways shown in Figure 3 are primarily involved in several immune-response and inflammatory processes, including interferon and cytokine signaling pathways, and reveal a dose-dependent up-regulation in A549 infected cells (lanes 2-5). Taken together, the results from *in vitro* experiments with the cell lines (A549 and Calu3) and the *in vivo* results from NSP cells directly isolated from virus infected humans suggest that the inhibition of EIF2 and oxidative phosphorylation along with activation of immune response and pro-inflammatory pathways are the characteristic signatures of SARS-CoV-2, RSV, and HIPV infections.

Identification of drugs and compounds negatively correlated to SARS-CoV-2 gene expression patterns.

After identifying both up- and down-regulated genes and pathways in SARS-CoV-2 transcriptomic data, we next sought to identify drugs and small molecules that may reverse these virus-induced host gene expression alterations and subsequently be further investigated for the possible treatment of COVID-19. This concept is central to the connectivity map (CMap) approach of repurposing established drugs for off-label diseases or indications [18].

We searched for drugs and compounds corresponding to DE genes from the GSE147507 Series 16 datasets using the Pharmaco Atlas application of BSCE software. BSCE searches GEO datasets and returns a list of drugs and compounds that produce a negatively correlated gene expression profile for the input query. The top 25 drugs, sorted by decreasing negative correlation score, are shown with their correlation scores in Table 2.

Table 2. Top 25 potentially repurposable drugs or compounds for Covid-19 sorted by the negative correlation score.

Drug	Score	Correlation
Dactinomycin	72	negative
Triptolide	54	negative
Dexamethosone	49	negative
Estradiol	48	negative
Hydroxychloroquine	47	negative
Isoniazid	46	negative
Gefitinib	46	negative
Ticlopidine	46	negative
Adalimumab	46	negative
Echinomycin	46	negative
Immunoglobulin Fc Fragments	45	negative
Phenobarbital	45	negative
Methylnitrosourea	45	negative
Choloroform	45	negative
Antibodies, Monoclonal	45	negative
4'-epidaunomycin	44	negative
Yorinostat	44	negative
phenothiazine	44	negative
Methylene Chloride	44	negative
bisphenol A	44	negative
Y 27632	44	negative
Infliximab	44	negative
Valproic acid	43	negative
Genistein	43	negative
Azithromycin	43	negative

We noted the presence of currently prescribed COVID-19 drugs including Dexamethasone, Hydroxychloroquine (HCQ), and Azithromycin in the list at positions three, five, and twenty-five, respectively (Table 2) [8]. The entire list of drugs is available as a supplementary document (SD4).

8

Triptolide is a potent activator of mitochondrial metabolic pathways, i.e., oxidative phosphorylation and TCA cycle and respiratory electron transport, whereas Hydroxychloroquine downregulates interferon and cytokine signaling.

We focused on the top five drugs from Table 2 for further analysis. Dactinomycin is a DNA binding compound that inhibits transcription [33]. Triptolide is a diterpenoid epoxide and the major bioactive component of *Tripterygium wilfordii Hook.f.* (TWHF), a Traditional Chinese Medicine (TCM) herb [34]. Dexamethasone is an anti-inflammatory 9-fluoro-glucocorticoid [35]. Estradiol is a naturally occurring hormone circulating endogenously in females [36]. HCQ is a chemotherapeutic agent that acts against erythrocytic forms of malarial parasites [37].

To ascertain the effect of these five drugs on SARS-CoV-2 triggered host pathways, we first searched for drug-induced transcriptomics datasets in BSCE. Then we compared pathway enrichment profiles between the SARS-CoV-2 and drug-induced gene expression datasets.

Since SARS-CoV-2 is a respiratory virus, we focused on BSCE datasets in which drug treatment induced gene expression was studied in cell lines derived from lung tissue. We did not find any lung cell-related gene expression data in two of the top five drugs, Estradiol and HCQ. We therefore excluded Estradiol from further analysis, but continued with HCQ since some studies indicate that it has potential to treat SARS-CoV-2 [38]. However, it is important to note that the HCQ data is from experiments performed in human Peripheral Blood Mononuclear Cells (PBMC), not lung tissue.

Of the SARS-CoV-2-induced pathways, we selected three up-regulated (generic transcription, cytokine signaling, and interferon signaling) and three down-regulated (translation, oxidative phosphorylation and TCA cycle and respiratory electron transport) for comparative analysis in BSCE. In Figure 4, panels A and B display the up and down-regulated pathways, respectively. In each graph, lane 1 is the input query dataset representing the result from SARS-CoV-2 vs. mock-infected samples from the GSE140507 series 16 dataset. Lanes 2-17 are the pathway enrichment analysis results of several drug treatments in different cell lines, including A549, airway smooth muscle, and PBMC.
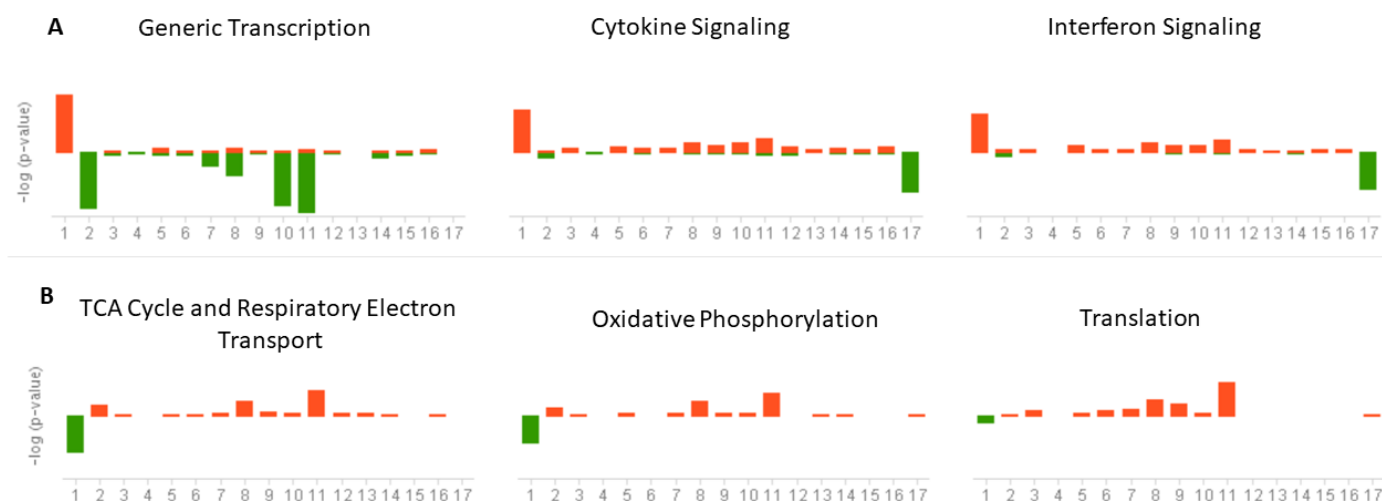

Figure 4. Pathway activity comparison – SARS-CoV-2 infection vs. drug treatments. Each plot represents the activity patterns of a pathway across multiple datasets. The x-axis lists the datasets and the y-axis shows the correlation score. The red bar above the midline represents the strength of overlap or enrichment between the up-regulated genes and the pathway genes. In contrast, the green bar below the midline represents the strength of overlap or enrichment between the down-regulated genes and the pathway genes. The absence of a colored bar means that the correlation is insignificant. Panel A: SARS-CoV-2-induced up-regulated pathways. Panel B: down-regulated pathways.

SARS-Cov-2 infection: (1) A549 cells + ACE2 + SARS-CoV-2 (MOI 2.0) vs. Mock (GSE147507); **Dactinomycin: (2)** A549 + Dactinomycin vs. Mannitol (control) (GSE6400);   **Triptolide: (3)** A549 cells + Triptolide (0.05 μM, 1 hour) vs. control (GSE16760); **(4)** A549 cells + Triptolide (0.05 μM, 2 hour) vs. control (GSE16760); **(5)** A549 cells + Triptolide (0.05 μM, 4 hour) vs. control (GSE16760); **(6)** A549 cells + Triptolide (0.15 μM, 1 hour) vs. control (GSE16760); **(7)** A549 cells + Triptolide (0.05 μM, 2 hour) vs. control (GSE16760); **(8)** A549 cells + Triptolide (0.15 μM, 4 hour) vs. control (GSE16760); **(9)** A549 cells + Triptolide (0.45 μM, 1 hour) vs. control (GSE16760); **(10)** A549 cells + Triptolide (0.45 μM, 2 hour) vs. control (GSE16760); **(11)** A549 cells + Triptolide (0.45 μM, 4 hour) vs. control (GSE16760);   **Dexamethasone: (12**) Airway Smooth Muscle Cells + Dexamethasone (1.0 μM, 4 hours) vs. control (GSE34313);); **(13)** Airway Smooth Muscle Cells + Dexamethasone (1.0 μM, 24 hours) vs. control (GSE34313);); **(14)** Airway Smooth Muscle Cells + Dexamethasone (1.0 μM, 18 hours) vs. control (GSE52778);); **(15)** A549 cells + Dexamethasone (1.0 μM, 5 hours) vs. control (GSE52778); (**16,** A549 cells + Dexamethasone (1.0 μM, 6 hours) vs. control (GSE52778); **Hydroxychloroquine: (17)** PBMC + Hydroxychloroquine (20 μM, 24 hour) + heat killed group A Streptococcus (MOI 10.0) vs. heat killed group A Streptococcus (MOI 10.0) only (GSE74235).

In panel A, generic transcription shows SARS-CoV-2-induced gene activation (lane 1). The known transcription inhibitor Dactinomycin (lane 2) displays the most potent inhibitory effect. Triptolide (lanes 3-11) reveals a dose- and time of treatment-dependent down-regulation. In contrast, Dexamethasone (lanes 12-16) suggests minimal impact, and HCQ (lane 17) was ineffective. Both cytokine and interferon signaling exhibit similar pathway enrichment profiles. SARS-CoV-2 infection increases pathway activity (lane 1). In contrast, the most potent inhibitory effect was observed when HCQ activity was measured in human PBMC, stimulated with rheumatogenic, heat-killed Group A Streptococcus (GAS) (lane 17). Dactinomycin also reveals down-regulation to a much lesser extent (lane 2). Triptolide and Dexamethasone show no inhibitory effects (lanes 3-16).

In Panel B, all three pathways reveal parity amongst pathway profiles. SARS-CoV-2 infection-induced inhibition is observed in the translation, oxidative phosphorylation, and TCA cycle and respiratory electron transport pathways. Triptolide (lanes 3-11) displays a dose- and time of treatment-dependent up-regulation. Lanes 5, 8, and 11 represent 4-hour treatment of Triptolide with incremental increases in concentration from 0.05 to 0.15 to 0.45 μM, resulting in concomitant step-wise pathway activation. Among other tested drugs, only Dactinomycin (lane 2) reveals a modest up-regulation of all three pathways.

Discussion

Our project aim was to identify genes and pathways that might be involved in SARS-CoV-2 pathogenesis as well as discover potential drugs or compounds that could be repurposed for treatment of COVID-19. To achieve this goal, we took an *in silico* approach and leveraged the rich collection of publicly available transcriptomic data. To gain mechanistic insight into the SARS-CoV-2 triggered host response, we identified virus-induced host gene expression patterns from two GEO datasets and performed pathway enrichment analysis. We discovered that the number of virus-induced DE genes positively correlates with the dose of virus infection and level of ACE2 expression. We performed in depth pathway enrichment analysis of virus-induced host response using two bioinformatics tools. The key findings from pathway analysis are mitochondrial dysfunction, activation of immune-inflammatory pathways, and strong inhibition of oxidative phosphorylation.

Connecting established drugs with off-label diseases through transcriptomic data analysis is becoming a common practice in the field of bioinformatics, and has been successfully used to identify repurposed drugs through connectivity map (CMap) projects [18]. Repositioning drugs for obesity [39], influenza [40], ovarian cancer [41], and gastric cancer [42] are a few prime examples of drug discovery using the transcriptomic data analysis approach. Motivated by this, we used BSCE software to identify drugs by searching drug/compound-related GEO datasets to reveal a reverse pattern of SARS-CoV-2-promoted gene expression signature. The current prescribed treatments for COVID-19 include Dexamethasone, Remdesivir, and HCQ alone or with Azithromycin. Our drug discovery method

successfully identified Dexamethasone, HCQ, and Azithromycin. Our method failed to identify the other common drug, Remdesivir, due to the lack of available and relevant GEO datasets.

The BSCE-generated comparative pathway analysis shed light on the mode of action of four drugs by revealing which SARS-CoV-2-prompted pathway activity patterns could be reversed. The activation of cytokine and interferon signaling may lead to a hyperactive immune-inflammatory response, aka "cytokine storm" [16]. In severe COVID-19 cases, ARDS and multiple-organ failure occurs rapidly, resulting in death within a short time [43]. A cytokine storm is considered to be one of the major causes of ARDS and multiple-organ failure. In our study, the most potent inhibition of cytokine and interferon signaling was observed by transcriptomic data from human PBMC, where the HCQ effect was measured against heat killed GAS infection. Amongst other drugs we studied, only Dactinomycin showed a modest inhibitory effect against the two pro-inflammatory pathways.

Our pathway analysis results suggest that mitochondrial dysfunction and strong inhibition of mitochondrial oxidative phosphorylation are the prominent signatures of several viral infections, including SARS-CoV-2, RSV, and HIPV. The role of mitochondria in a viral infection host response has been previously described [44]. The SARS-CoV encoded protein, ORF-9b, localizes into the mitochondria and manipulates its function to evade host immune response [45]. The role of mitochondrial dysfunction in COVID-19 pathogenesis has also been postulated [45–48]. Some studies show that coronaviruses manipulate host mitochondria and cause mitochondrial dysfunction leading to an inflammatory response [49]. A link between mitochondrial dysfunction and COVID-19-related sepsis is also described [50]. Considering that mitochondria are cellular powerhouses and play the central role in maintaining oxidative homeostasis, rescuing mitochondrial function through therapeutic interventions could be pivotal in reducing COVID-19 morbidity and mortality.

In our analysis, we found that Triptolide might have the potential to intervene therapeutically with mitochondrial dysfunction, as it exhibited the most potent dose and time of treatment dependent activation on both oxidative phosphorylation and TCA cycle and respiratory electron transport pathways. Triptolide, a diterpene triepoxide, is a pharmacologically active compound isolated from *Tripterygium wilfordii Hook F* (TwHF), a perennial vine (Thunder God Vine) commonly grown in southeast China, and used in Traditional Chinese Medicine to treat autoimmune and inflammatory conditions [51]. It covers a broad spectrum of biological profiles including anti-inflammatory, immunosuppressive, antifertility, antitumor activities, and neuroprotective effects. The therapeutic potential of Triptolide have been tested for nephritis, asthma, arthritis, Alzheimer's Disease, Parkinson's Disease, Pancreatic cancer, and HIV [34,51,52]. Despite its therapeutic potential, its application in clinical practice is limited due to poor water solubility and multi-organ toxicity [53]. However, several Triptolide derivatives that improve water solubility and reduce the toxicity have been developed [51], including Minnelide, which exhibits promising results for pancreatic cancer in preclinical models [54]. These Triptolide derivatives could be tested for clinical efficacy against COVID-19.

Responding to an urgent need to find COVID-19 treatments, many research groups have taken a variety of approaches to identify repurposable drugs, including viral protein structure-based molecular docking [55], network-based [56,57] and connectivity map-based [58] drug repurposing methods. In this paper we describe an additional approach, the application of anti-correlation based transcriptomic data analysis methods using publicly available GEO datasets to identify drugs that may be repositioned for COVID-19 treatments.

There are limitations with this *in silico* data mining approach. The absence of drug or compound-induced transcriptomic data in the GEO database, such as in the case of Remdesivir, prevents its identification as a potentially repurposable drug. We mined data generated by different research groups using a variety of experimental protocols. Microarray data were compared with the results of RNA-Seq data. The inherent heterogeneity of experimental platforms might lead to erroneous conclusions. Therefore, all drug predictions described here warrant experimental and clinical verifications. However, we hope that these findings will be helpful to understand the pathogenesis

of SARS-CoV-2 and can motivate researchers to test these compounds that eventually could help to find better treatment options for COVID-19.

Authors Contributions: A.C. conceived the project, performed pathway enrichment analysis, and wrote the initial manuscript draft. S.C. performed RNA-Seq data analysis, created figures and tables, and deposited the data. C.I. provided editorial suggestions and reworked the manuscript to its final form.

Supplementary Information:

Methods

SM1—GEO datasets

GSE147507 contains bulk RNA-Seq data from both virus-infected and mock-infected experiments performed on a variety of cell lines and animal models. Out of 110 samples, we selected data from three cell lines for our analysis: Normal Human Bronchial Epithelial (NHBE), A549—derived from explanted cultures of human lung cancer tissue, and Calu3—epithelial cells derived from lung adenocarcinoma. The three virus strains used for infection were: SARS-CoV-2 (USA-WA1/2020) with Multiplicity of Infection (MOI) 0.2 or 2.0, RSV (A2 strain; MOI 2.0), and human parainfluenza virus-3 (HPIV-3; MOI 3. 0). PolyA RNA was harvested 24 hours post-infection for SARS-CoV-2, RSV, and HPIV-3. GSE152075 contains RNA-Seq profiles of nasopharyngeal swabs from 430 individuals with SARS-CoV-2 and 54 negative controls.

SM2—CLC Genomics data

Raw reads for all human samples (GSE147507; series 1-9, 15, 16) were downloaded from the NCBI Sequence Read Archive [59] and data analysis was performed using CLC Genomics Workbench v20 (QIAGEN Digital Insights). Raw reads were aligned to the human reference genome (hg38) with annotations (Ensemble v91) using default settings, and unmapped reads were collected and aligned against corresponding virus reference genome sequence. The differential expression of genes between virus-infected vs mock-treated samples were computed for each series. Genes with FDR-adjusted p-value <0.05 and log fold change +/- 1.0 were considered DE genes.

SM3—BaseSpace Correlation Engine drug-related datasets

Dactinomycin: results from a microarray experiment in which A549 human lung cancer cells were seeded eight days before drug treatment (GSE6400) [60]. At four hours before RNA isolation, Dactinomycin (5 μg/mL final concentration) or control (5% mannitol) solution was added to the cell cultures.

Triptolide: results from a microarray experiment describing identification of genes affected by a short treatment of A549 cells with increased concentrations of Triptolide (GSE16760) [61]. The dataset includes DE genes from Triptolide-treated vs. untreated samples in which three concentrations (0.05, 0.15, 0.45 μM) of the drug were introduced at three treatment time points (1, 2, 4 hours).

Dexamethasone: results from four GEO datasets measuring gene expression in human lung tissue-derived cells treated with Dexamethasone. GSE34313 is a microarray dataset of airway smooth

muscle cell responses to Dexamethasone at 4 and 24 hours [62]. GSE52778 is an RNA-Seq dataset after 18 hours of Dexamethasone treatment [63]. GSE96649 is an RNA-seq dataset from A549 cells with Dexamethasone (1uM) for treatment at 5 hours vs. untreated controls [64] . GSE17307 is a microarray dataset of Dexamethasone (1uM) treatment at 6 hours vs. untreated controls [65].

Hydroxychloroquine: GSE74235 is an RNA-Seq dataset of human Peripheral Blood Mononuclear Cells from three healthy participants, stimulated with rheumatogenic, heat-killed Group A Streptococcus (GAS) for 24 hours, MOI 10 [66] . The effect of Hydroxychloroquine (20 µM) was measured in combination with GAS.

SM4—BaseSpace Correlation Engine pathway comparison

BSCE was used to mine the GEO datasets and compute DE genes between two conditions. For microarray experiments, statistical significance was determined using the Welch t-test. Genes with p-value cutoff 0.05 and fold change cutoff +/- 1.20 were considered DE genes. For RNA-Seq experiments, raw sequencing reads were aligned to the human genome sequence using STAR 2.3 and RefSeq annotations. Differential gene expression results between the control and test sample groups were calculated using DESeq2. Genes with Benjamini-Hochberg adjusted q-value cutoff 0.05 and a fold change cutoff of +/-1.2 were considered DE. BSCE software uses a ranked-based non-parametric Running Fisher test to compare gene expression signatures between two DE gene lists and calculate a correlation score [27].

Documents

SD1. IPA canonical pathway enrichment analysis results for DE genes from SARS-CoV-2-infected vs. mock control from series 16 (GSE147507). A spreadsheet containing pathway names, p-values, activity z-scores, and overlapping genes. DOI: 10.6084/m9.figshare.12899387

SD2. BSCE pathway enrichment analysis results for DE genes from SARS-CoV-2-infected vs. mock control from series 16 (GSE147507). A spreadsheet containing pathway names, pathway direction, and p-values. DOI: 10.6084/m9.figshare.12899399

SD3. IPA canonical pathway comparison across multiple samples. A spreadsheet containing pathway names and activity z-scores. DOI: 10.6084/m9.figshare.12925724

SD4. Predicted COVID-19 repurposable drugs/compound. The full list of BaseSpace Correlation Engine computed drugs or compound revealing a negative correlation with the gene expression profile of series16_GSE147507 (differentially expressed genes from A549 cells transfected with exogenous ACE2, infected with SARS-CoV-2 (USA-WA1/2020) vs. mock control). DOI: 10.6084/m9.figshare.12899408
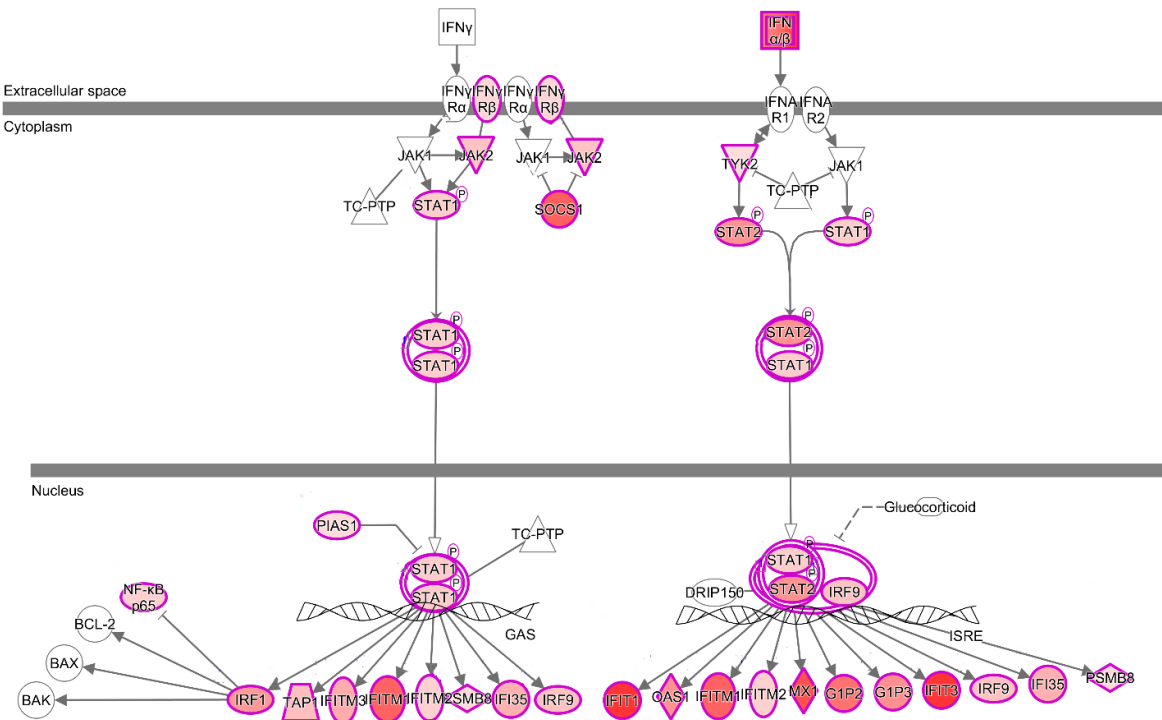
Table

ST1. RNA-Seq analysis results (GSE147507) showing SARS-CoV-2 viral gene expression.

Gene	S1_SARSCOV2_TPM Mean	S2_SARSCOV2_TPM Mean	S5_SARSCOV2_TPM Mean	S6_SARSCOV2_TPM Mean	S7_SARSCOV2_TPM Mean	S16_SARSCOV2_TPM Mean
ORF1ab_1	0	0	0	5847.45	1413.60	5146.64
ORF1ab_2	1734.47	1820.60	4256.43	5782.00	832.73	1981.77
S	14069.43	29658.22	34891.84	19707.84	7974.92	12356.82
ORF3a	13245.19	53848.57	58476.78	35125.22	19570.48	28092.09
E	1086.75	19383.58	36610.71	15730.09	9225.75	14808.52
M	201586.32	119051.8	112120.11	99627.36	86629.40	73195.24
ORF6	3407.41	10348.67	12259.08	32704.57	10697.02	4592.52
ORF7a	24082.43	113181.1	118683.19	103757.22	95575.52	122842.32
ORF7b	0	315.0067	3462.13	5360.12	1578.68	580.01
ORF8	23198.55	105271.3	146962.49	117131.70	131925.11	155547.37
N	714286.99	545871.2	471327.45	519120.58	632843.03	539498.95

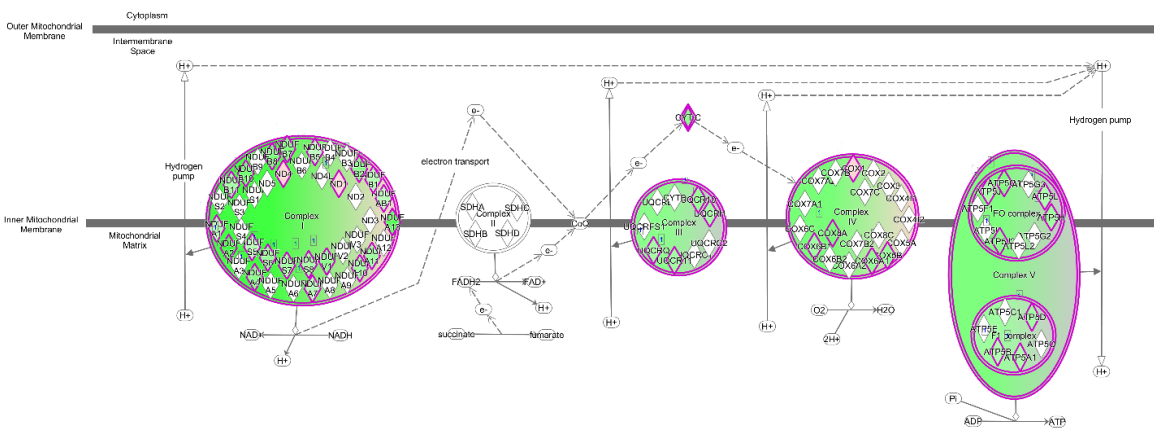| ORF10 | 3302.41 | 1250.03 | 949.80 | 40105.84 | 1733.76 | 41357.76 |

Figures

SF1. Interferon signaling pathway overlaid with differentially expressed genes observed between SARS-CoV-2-infected vs. mock control.

The fuchsia color indicated up-regulated genes with $\log_2$Fold change > 1. The intensity of the color is proportional to the magnitude of fold change.

SF2. Mitochondrial oxidative phosphorylation pathway overlaid with differentially expressed genes observed between SARS-CoV-2-infected vs. mock control.

The green color indicates down-regulated genes with $\log_2$Fold change < -1. The fuchsia color indicates up-regulated genes with $\log_2$Fold change > +1. The intensity of the color is proportional to the magnitude of fold change.

References

1. Coronavirus (COVID-19) events as they happen. Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen [Accessed September 9, 2020].

2. Coronavirus Resource Center *Johns Hopkins University*. Available at: https://coronavirus.jhu.edu/ [Accessed September 9, 2020].

3. Amanat F, Krammer F (2020) SARS-CoV-2 Vaccines: Status Report. *Immunity* 52(4):583–589.

4. Chen Y, Liu Q, Guo D (2020) Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol* 92(4):418–423.

5. Europe PMC. Available at: https://europepmc.org [Accessed September 9, 2020].

6. Wölfel R, et al. (2020) Virological assessment of hospitalized patients with COVID-2019. *Nature* 581(7809):465–469.

7. Xu Z, et al. (2020) Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir Med* 8(4):420–422.

8. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB (2020) Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. *JAMA*.

9. Cavalcanti AB, et al. (2020) Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate Covid-19. *N Engl J Med*.

10. Yu B, Wang DW, Li C (2020) Hydroxychloroquine application is associated with a decreased mortality in critically ill patients with COVID-19. *medRxiv*. doi:10.1101/2020.04.27.20073379

11. Beigel JH, et al. (2020) Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med*.

12. Horby P, et al. (2020) Effect of Dexamethasone in Hospitalized Patients with COVID-19: Preliminary Report. *medRxiv*. doi:10.1101/2020.06.22.20137273

13. Lane JC., et al. (2020) Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study. *medRxiv*. doi:10.1101/2020.04.08.20054551

14. The Writing Committee for the REMAP-CAP Investigators, et al. (2020) Effect of Hydrocortisone on Mortality and Organ Support in Patients with Severe COVID-19. *JAMA*.

15. Association Between Administration of Systemic Corticosteroids and Mortality Among Critically Ill Patients with COVID-19: A Meta-analysis | Critical Care Medicine | JAMA | JAMA Network Available at: https://jamanetwork.com/journals/jama/fullarticle/2770279 [September 9, 2020].

16. Ye Q, Wang B, Mao J (2020) The pathogenesis and treatment of the `Cytokine Storm' in COVID-19. *J Infect* 80(6):607–613.

17. Auyeung TW, et al. (2005) The use of corticosteroid as treatment in SARS was associated with adverse outcomes: a retrospective cohort study. *J Infect* 51(2):98–102.

18. Ra K, et al. Connectivity Mapping: Methods and Applications.

19. Barrett T, et al. (2013) NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 41(Database issue): D991–5.

20. Home - GEO - NCBI. Available at: https://www.ncbi.nlm.nih.gov/geo/ [Accessed September 9, 2020].

21.   Transcriptional response to SARS-CoV-2 infection *Series GSE147507*. Available at: https://geo.metadataplus.biothings.io/geo/query/acc.cgi?acc=GSE147507 [Accessed September 9, 2020].

22.   In vivo antiviral host response to SARS-CoV-2 by viral load, sex, and age [dataset I] *Series GSE152075*. Available at: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE152075 [Accessed September 9, 2020].

23.   QIAGEN CLC Genomics Workbench | QIAGEN Digital Insights Available at: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/ [Accessed September 9, 2020].

24.   Partek Flow Genomic Analysis Software - Lab and Enterprise Solutions – Partek Inc Available at: https://www.partek.com/partek-flow/ [Accessed September 9, 2020].

25.   Ingenuity Pathway Analysis | QIAGEN Digital Insights Available at: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/ [Accessed September 9, 2020].

26.   Krämer A, Green J, Pollard J, Tugendreich S (2014) Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* 30(4):523–530.

27.   Kupershmidt I, et al. (2010) Ontology-based meta-analysis of global collections of high-throughput public data. *PLoS One* 5(9).

28.   Hou Y, et al. (2020) New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC Med* 18(1):216.

29.   Daly JL, et al. (2020) Neuropilin-1 is a host factor for SARS-CoV-2 infection. *BioRxiv*. doi:10.1101/2020.06.05.134114

30.   Hoffmann M, et al. (2020) SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181(2):271–280.e8.

31.   MSigDB Collections (Broad Institute and UC San Diego) *MSigDB Collections - C2 Canonical Pathways*. Available at: https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp#C2 [Accessed September 9, 2020].

32.   Lieberman NAP, et al. (2020) In vivo antiviral host response to SARS-CoV-2 by viral load, sex, and age. *BioRxiv*. doi:10.1101/2020.06.22.165225

33.   Dactinomycin *DrugBank*. Available at: https://www.drugbank.ca/drugs/DB00970 [Accessed September 9, 2020].

34.   Zheng Y, Zhang W-J, Wang X-M (2013) Triptolide with potential medicinal value for diseases of the central nervous system. *CNS Neurosci Ther* 19(2):76–82.

35.   Dexamethasone *DrugBank*. Available at: https://www.drugbank.ca/drugs/DB01234 [Accessed September 9, 2020].

36.   Estradiol *DrugBank*. Available at: https://www.drugbank.ca/drugs/DB00783 [Accessed September 9, 2020].

37.   HydroxyChloroquine *DrugBank*. Available at: https://www.drugbank.ca/drugs/DB01611 [Accessed September 9, 2020].

38.   Ip A, et al. (2020) Hydroxychloroquine in the treatment of outpatients with mildly symptomatic COVID-19: A multi-center observational study. *medRxiv*. doi:10.1101/2020.08.20.20178772

39.   Liu J, Lee J, Salazar Hernandez MA, Mazitschek R, Ozcan U (2015) Treatment of obesity with celastrol. *Cell* 161(5):999–1011.

40.   Josset L, et al. (2010) Gene expression signature-based screening identifies new broadly effective influenza a antivirals. *PLoS One* 5(10).

41.   Raghavan R, et al. (2016) Drug discovery using clinical outcome-based Connectivity Mapping: application to ovarian cancer. *BMC Genomics* 17(1):811.

42.   Claerhout S, et al. (2011) Gene expression signature analysis identifies vorinostat as a candidate therapy for gastric cancer. *PLoS One* 6(9):e24662.

43.   Zhang B, et al. (2020) Clinical characteristics of 82 cases of death from COVID-19. *PLoS One* 15(7):e0235458.

44.   Koshiba T (2013) Mitochondrial-mediated antiviral immunity. *Biochim Biophys Acta* 1833(1):225–232.

45.   Shi C-S, et al. (2014) SARS-coronavirus open reading frame-9b suppresses innate immunity by targeting mitochondria and the MAVS/TRAF3/TRAF6 signalosome. *J Immunol* 193(6):3080–3089.

46.   Saleh J, Peyssonnaux C, Singh KK, Edeas M (2020) Mitochondria and microbiota dysfunction in COVID-19 pathogenesis. *Mitochondrion* 54:1–7.

47.   Lane RK, Hilsabeck T, Rea SL (2015) The role of mitochondrial dysfunction in age-related diseases. *Biochim Biophys Acta* 1847(11):1387–1400.

48.   Singh KK, Chaubey G, Chen JY, Suravajhala P (2020) Decoding SARS-CoV-2 hijacking of host mitochondria in COVID-19 pathogenesis. *Am J Physiol Cell Physiol* 319(2):C258–C267.

49.   Clementz MA, Kanjanahaluethai A, O'Brien TE, Baker SC (2008) Mutation in murine coronavirus replication protein nsp4 alters assembly of double membrane vesicles. *Virology* 375(1):118–129.

50.   Shenoy S (2020) Coronavirus (Covid-19) sepsis: revisiting mitochondrial dysfunction in pathogenesis, aging, inflammation, and mortality. *Inflamm Res*.

51.   Ziaei S, Halaby R (2016) Immunosuppressive, anti-inflammatory and anti-cancer properties of triptolide: A mini review. *Avicenna J Phytomed* 6(2):149–164.

52.   Wan Z, Chen X (2014) Triptolide inhibits human immunodeficiency virus type 1 replication by promoting proteasomal degradation of Tat protein. *Retrovirology* 11:88.

53.   Li X-J, Jiang Z-Z, Zhang L (2014) Triptolide: progress on research in pharmacodynamics and toxicology. *J Ethnopharmacol* 155(1):67–79.

54.   Banerjee S, Saluja A (2015) Minnelide, a novel drug for pancreatic and liver cancer. *Pancreatology* 15(4 Suppl):S39–43.

55.   Alamri MA, et al. (2020) Discovery of human coronaviruses pan-papain-like protease inhibitors using computational approaches. *J Pharm Anal*.

56.   Kumar S, et al. (2020) Identification of Drugs Targeting Multiple Viral and Human Proteins Using Computational Analysis for Repurposing Against COVID-19. doi:10.26434/chemrxiv.12366938.v1

57.   Karunakaran KB, Balakrishnan N, Ganapathiraju M (2020) Potentially repurposable drugs for COVID-19 identified from SARS-CoV-2 Host Protein Interactome.

58.   Mousavi SZ; Rahmanian M; Sami A (2020): A Connectivity Map-Based Drug Repurposing Study and Integrative Analysis of Transcriptomic Profiling of SARS-CoV-2 Infection. *ChemRxiv*, doi: 10.26434/chemrxiv.12469316.v1

59.   Home - SRA - NCBI Available at: https://www.ncbi.nlm.nih.gov/sra [Accessed September 3, 2020].

60.   Wang Z, et al. (2007) Synthesis and biologic properties of hydrophilic sapphyrins, a new class of tumor-selective inhibitors of gene expression. *Mol Cancer* 6:9.

61.   Vispé S, et al. (2009) Triptolide is an inhibitor of RNA polymerase I and II-dependent transcription leading predominantly to down-regulation of short-lived mRNA. *Mol Cancer Ther* 8(10):2780–2790.

62.   Masuno K, et al. (2011) Expression profiling identifies Klf15 as a glucocorticoid target that regulates airway hyperresponsiveness. *Am J Respir Cell Mol Biol* 45(3):642–649.

63.   Himes BE, et al. (2014) RNA-Seq transcriptome profiling identifies CRISPLD2 as a glucocorticoid responsive gene that modulates cytokine function in airway smooth muscle cells. *PLoS One* 9(6):e99625.

64.   Dendoncker K, et al. (2017) The nature of the GRE influences the screening for GR-activity enhancing modulators. *PLoS One* 12(7):e0181101.

65.   Muzikar KA, Nickols NG, Dervan PB (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci USA* 106(39):16598–16603.

66.   Hydroxychloroquine inhibits responses to group A streptococcus in peripheral blood mononuclear cells Available at: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE74235 [Accessed September 9, 2020].