3-9-2021

COVID-19: disease pathways and gene expression changes predict methylprednisolone can improve outcome in severe cases.

Sorin Draghici

Tuan-Minh Nguyen

Larry A Sonna

Cordelia Ziraldo

Radu Vanciu

See next page for additional authors

Authors

Sorin Draghici, Tuan-Minh Nguyen, Larry A Sonna, Cordelia Ziraldo, Radu Vanciu, Raef Fadel, Austin Morrison, Rachel M. Kenney, George Alangaden, Mayur Ramesh, and Gil Mor

Subject Section

COVID-19: disease pathways and gene expression changes predict methylprednisolone can improve outcome in severe cases

Sorin Draghici [1,2,5]*, Tuan-Minh Nguyen [1], Larry A. Sonna [3], Cordelia Ziraldo [4], Radu Vanciu [4], Raef Fadel [5], Austin Morrison [6], Rachel M. Kenney [6], George Alangaden [7], Mayur Ramesh [7], Gil Mor [2,*]

[1]Department of Computer Science, Wayne State University, Michigan 48202, USA,
[2]Department of Obstetrics and Gynecology, Wayne State University, Michigan 48202, USA,
[3]Bon Air Consulting, USA,
[4]Advaita Bioinformatics, USA,
[5]Department of Internal Medicine, Henry Ford Health System, USA,
[6]Department of Pharmacy, Henry Ford Health System, USA, and
[7]Division of Infectious Diseases, Henry Ford Health System, USA

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: COVID-19 has several distinct clinical phases: a viral replication phase, an inflammatory phase, and in some patients, a hyper-inflammatory phase. High mortality is associated with patients developing cytokine storm syndrome. Treatment of hyper-inflammation in these patients using existing, approved therapies with proven safety profiles could address the immediate need to reduce mortality.
Results: We analyzed the changes in the gene expression, pathways and putative mechanisms induced by SARS-CoV2 in NHBE, and A549 cells, as well as COVID-19 lung vs. their respective controls. We used these changes to identify FDA approved drugs that could be repurposed to help COVID-19 patients with severe symptoms related to hyper-inflammation. We identified methylprednisolone (MP) as a potential leading therapy. The results were then confirmed in five independent validation data sets including Vero E6 cells, lung and intestinal organoids, as well as additional patient lung sample vs. their respective controls. Finally, the efficacy of MP was validated in an independent clinical study. Thirty-day all-cause mortality occurred at a significantly lower rate in the MP-treated group compared to control group (29.6% vs. 16.6%, $p = 0.027$). Clinical results confirmed the *in silico* prediction that MP could improve outcomes in severe cases of COVID-19. A low number needed to treat ($NNT = 5$) suggests MP may be more efficacious than dexamethasone or hydrocortisone.
Availability: iPathwayGuide is available at https://ipathwayguide.advaitabio.com/
Contact: sorin@wayne.edu and gmor@med.wayne.edu
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Most current efforts related to COVID-19 span a number of areas as follows: i) antivirals, ii) vaccine development, iii) diagnostic tests, and iv) patient-supporting interventions. Without reducing the significance and impact of any of the areas above, there is an important aspect

that has not been elucidated: the identification and treatment of patients developing critical conditions and risk of mortality. Recently, Mehta *et al.* stated that "Accumulating evidence suggests that a subgroup of patients with severe COVID-19 might have a cytokine storm syndrome" that correlates with high mortality (Mehta *et al.*, 2020). Therefore, identification and appropriate management of the patients developing cytokine storm syndrome is critical for successful outcomes. Treatment of hyper-inflammation in these patients using existing, approved therapies with proven safety profiles could address the immediate need to reduce the rising mortality.

COVID-19 has several distinct clinical phases: an infection phase, a viral replication phase, an inflammatory phase, and in some patients, a hyper-inflammatory phase or cytokine storm (Siddiqi and Mehra, 2020; Ayres, 2020). After the initial viral phase of the illness, some patients will develop a cytokine storm which has being associated with the acute respiratory distress syndrome (ARDS) and mortality. Therefore, in order to decrease the risk of mortality is necessary to distinguish the phase where the viral pathogenicity is dominant versus when the host inflammatory response overtakes the pathology (Siddiqi and Mehra, 2020; Ayres, 2020). A potential approach is to develop interventions that could inhibit/prevent the hyper-inflammatory process leading to the cytokine storm. A strong argument in favor of also targeting the host response is offered by the data on influenza. Even though influenza patients receive optimal anti-viral therapy, approximately 25% of the critically ill influenza patients still die (Ayres, 2020; Louie *et al.*, 2012). This suggests that anti-viral therapy alone will not be sufficient for COVID-19 either, and the host response to the virus still needs to be taken into consideration.

However, approaches aiming at modulating the immune response face some concerns. In particular, it may seem counter-intuitive to try to diminish the immune response in a patient whose immune system is fighting against a virus. Modulating the immune system is likely unnecessary and counter-productive for patients whose immune system is doing a good job at resolving the infection, while it could potentially be life-saving for those whose inflammatory response has become dysregulated. If a patient has developed severe respiratory symptoms and is hypoxic, the host response that lead to ARDS, sepsis, and organ failure has already been initiated (Mehta *et al.*, 2020). At this point, the focus should shift to supporting the patient's systems and preventing collapse triggered by hyper-inflammation (Ayres, 2020).

In order to identify the best potential therapeutic approach, we performed a transcriptome analysis of tissues and cell samples infected with SARS-CoV-2 in order to understand the main mediators of the inflammatory process. Once characterized the inflammatory pathways we identified drugs that would mitigate or alleviate some of the devastating over-reactions of the host's immune system (e.g. cytokine storm). Finally, we evaluated the efficacy of the identified drug in a small cohort of COVID-19 patients.

2 Approach

We used data from cell lines, cell cultures as well as human patients to understand the changes induced by the infection with SARS-CoV-2.

We started by analyzing transcriptomic data to compare the A549 lung cell line infected with SARS-CoV-2 vs. mock infection (henceforth A549CoV2vsControl), A549 infected with seasonal influenza A virus vs. mock infection (A549IAVvsControl), and A549 infected with human respiratory syncytial virus vs. mock infection (A549RSVvsControl). We also compared the transcriptional response in primary human bronchial epithelial (NHBE) between cells infected with SARS-CoV2 and mock infection (NHBECoV2vsControl). Finally, we compared the transcriptional response in COVID-19 lung tissues vs. healthy lung tissue (COVID19vsControl). These data were collected at Mount Sinai and are available in GEO as the GSE147507 data set (Blanco-Melo *et al.*, 2020).

The motivation behind studying these contrasts was to be able to differentiate a specific cellular response, as observed when using cell lines, versus the response of the organism as it is reflected in a particular tissue, as observed when using patient samples. Also, by comparing the IAV or RSV infections with the SARS-CoV-2 infection, we can differentiate between a general response to a viral infection versus a specific response to the corona virus. The approach used can be summarized as follows:

1. We first used a GO analysis to see what biological processes appear to be involved in the SARS-CoV-2 infection. This was based first on an enrichment analysis (Tavazoie *et al.*, 1999; Draghici *et al.*, 2003) followed by a more sophisticated analysis that takes into consideration the relationships between the GO terms and eliminates the redundancy (Alexa *et al.*, 2006). Both were followed by an FDR correction for multiple comparisons (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001).
2. We then used a pathway analysis to identify the impacted pathways. The approach used here, the impact analysis (Draghici *et al.*, 2007; Tarca *et al.*, 2009b), uses not only the measured fold changes but also the position of every gene on every pathway, as well as the direction and type of every signal from one gene to another.
3. This was followed by an upstream analysis aiming to identify any specific upstream regulators that may play a role in this infection and/or the immune response to it.
4. The next step was a mechanism inference aiming to identify the mechanisms likely to be involved on these pathways or linking the genes involved in the key biological processes identified. This was based on the pathways and biological processes identified above, the measured fold changes in the genes participating in these, and all known protein-protein interactions (PPIs), both from existing pathways, as well as from known PPIs databases such as STRING.
5. The last step was the drug-target analysis which took the processes, pathways, and genes identified above and aimed to estimate the ability of each known drug to reverse the most relevant gene expression changes induced by the SARS-CoV-2 infection. This step was based on known interactions between drugs and genes or proteins.

The existing FDA-approved drug that resulted from the process above was validated in five additional independent datasets coming from different laboratories, as well as in a clinical study.

3 Results

3.1 Disrupted genes and biological processes.

We evaluated the biological processes that are affected by SARS-Cov-2 in lung epithelial cells. We performed a comparison of the affected biological processes in COVID19vsControl, NHBECoV2vsControl, A549CoV2vsControl, A549IAVvsControl, and A549RSVvsControl (Fig. S1). The biological processes (BPs) are shown ordered by their significance in COVID19vsControl. In spite of a larger number of differentially expressed (DE) genes in the SARS-Cov-2-infected lung (815), there are only 7 significant biological processes involved, which may indicate a more coordinated, systemic response. In contrast, the changes in the NHBE cells are characterized by fewer DE genes (only 223) but span more uncoordinated biological processes. This is illustrated in Fig. S2 which shows the BPs ordered in the order of significance from NHBECoV2vsControl.

3.2 Putative mechanisms of disease.

We performed an analysis aiming to identify putative mechanisms of disease. As part of this analysis we identified four genes that were predicted to be activated upstream regulators based on the observed changes in their downstream genes. These were IRF9, STAT2, IFNG, and IFNB1.

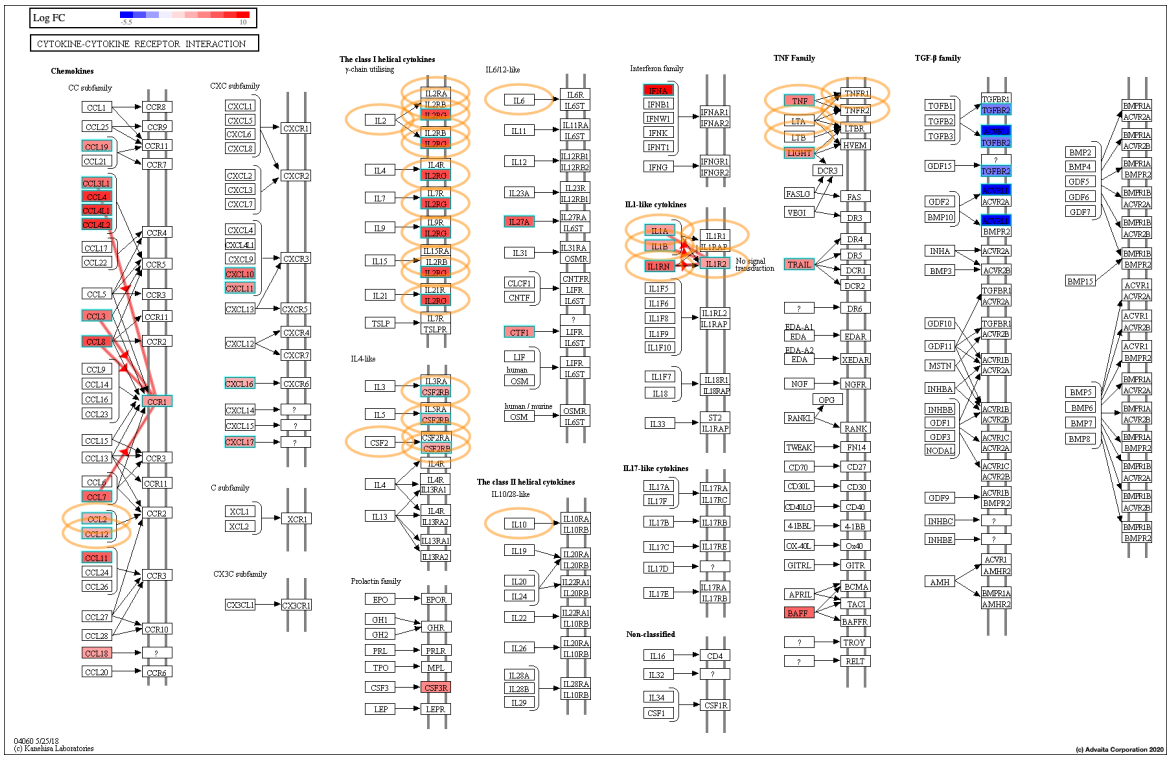

Fig. 1: The cytokine-cytokine receptor interactions pathway is the most significantly impacted pathway in COVID19vsControl and the second most impacted in NHBECoV2vsControl. This is due mainly to the large number of dis-regulated cytokines. The image shows up-regulated genes (red), down-regulated genes (blue), as well as genes that are targeted by existing FDA-approved drugs (in ovals). Note that all up-regulated genes are pro-inflammatory while the down-regulated genes are anti-inflammatory, supporting the idea that the severe symptoms may be caused by a cytokine storm.

These suggest two different potential mechanisms. The first appears to be triggered by STAT2 and IRF9, which have 16 common target genes that are also all significantly up-regulated (IFI6, IFIT1, IFIT2, IFIT3, IFITM1, IFITM3, OAS1, OAS3, OAS S32, MX1, MX2, RSAD2, OASL, XAF1, IRF2, and IRF7). This mechanism is also known to be involved in the response to influenza A (see influenza A pathway in Fig. S4).

The second putative mechanism involves interferon beta and gamma, which are targeting 5 common downstream genes: CXCL10, IDO1, DOX58, STAT1, which are up-regulated, and HMOX2 which is down-regulated. Interferon regulatory factors (IRFs) are subdivided into the interferonic IRFs (IRF2-3-7 and 9), the stress responsive IRFs (IRF1 and 5), the hematopoietic IRFs (IRF4 and 8) and morphogenic IRF6. IRF9 is a regulator of type I IFN signaling and is known to interact with STAT2 (Horvath *et al.*, 1996) and STAT1 to form the heterotrimeric transcription factor complex (ISGF3) that binds to interferon-stimulated response elements (ISREs) to induce the expression of interferon stimulated genes (ISG). During viral infections, ISGs perform two key functions: 1) directly limit viral replication by shutting down protein synthesis and triggering apoptosis; 2) activate key components of the innate and adaptive immune system, including antigen presentation and production of cytokines. The genes triggered by the STAT2 and IRF9 pathway include genes responsible for limiting viral replication (IFI6, IFIT1, IFIT2, IFIT3, IFITM1, IFITM3, OAS1, OAS3, OAS2, MX1, MX2, RSAD2, OASL) and inducers of apoptosis (XAF1, IRF2, IRF7). CXCL10, IDO1, DOX58, and STAT1 are genes associated with lymphocyte recruitment and immune regulation.

Interestingly, STAT2 and IRF9 together are also identified as activated upstream regulators due to 15 downstream targets even in the NHBECoV2vsControl (see Fig. S5). However, in the NHBE cells, the interferon activators were replaced by an interleukin based mechanism centered around IL1B, IL6, IL17A, adiponectin (ADIPOQ) and tumor necrosis factor (TNF).

We also looked at genes that are known to modulate or inhibit the inflammatory response such as IL1RN IL10, and IL13. In the COVID19vsControl, IL1RN was up with a log2 fold change of 6.2 fold (a 78-fold increase, FDR-corrected $p = 10^{-6}$), IL10 was up 2.8 fold (FDR-corrected $p = 0.55$), while the measurement for IL13 was not available. In the NHBECoV2vsControl, IL1RN was up only 1.26 fold (FRD-corrected $p = 0.035$), while measurements for IL10 and IL13 were not available. However, in this contrast, 14 out of 15 DE genes immediately downstream of IL10 and usually inhibited by IL10 were up-regulated which strongly supports the hypothesis that IL10 is inhibited (FDR-corrected $p = 5.17 \times 10^{-9}$).

3.3 Impacted pathways.

The significantly impacted pathways are shown in Fig. S6 and Fig. S7 ordered by their significance in COVID19vsControl and NHBECoV2vsControl, respectively. The p-values represent a combination of enrichment and perturbation p-values (Draghici *et al.*, 2007) subsequently corrected with FDR (see Supp. Materials for details).

Fig. 1 shows the most impacted pathway, the *Cytokine-cytokine interactions*. The gene expression changes on this pathway show a very large number of pro-inflammatory cytokines and cytokine receptors being up-regulated, while several of the anti-inflammatory genes in the TGF-beta family are down-regulated. This corroborates the GO analysis finding above that there is a very strong positive regulation of inflammatory response.

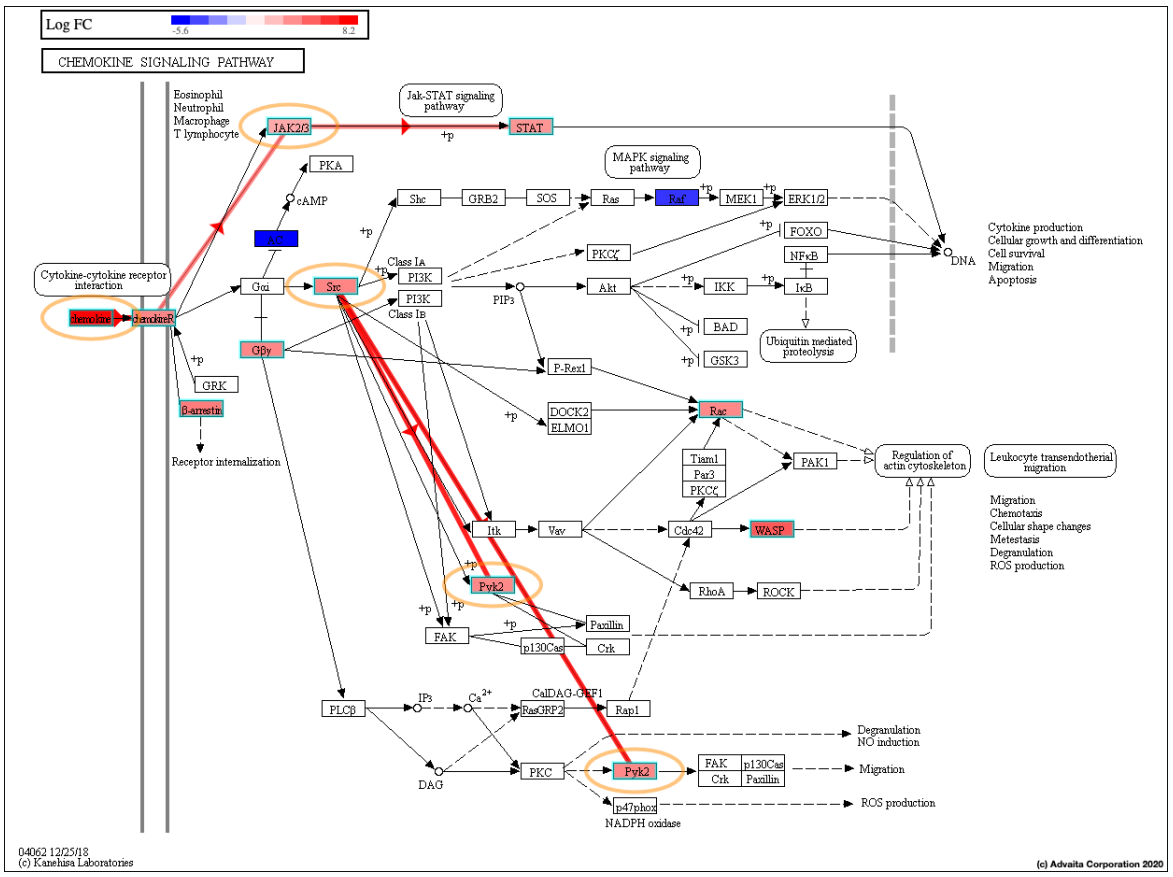

Fig. 2: The chemokine signaling pathway is the second most significantly impacted pathway in COVID19vsControl. The red arrows represent chains of coherent perturbation propagation, i.e. sequence of steps for which the observed expression changes are coherent with the expected changes according to the phenomena described by the pathway. The node labelled "chemokines" represents 11 chemokines measured to be up-regulated (CCL2, CCL3, CCL4, CCL7, CCL8, CCL11, CCL18, CCL19, CXCL10, CXCL11, CXCL16). The CCR1 receptor is also up-regulated, as well as JAK3 and STAT1. The ovals represent drug targets for which FDA-approved drugs already exist. On this pathway, the impact is due both to the large number of DE genes (26 out of 130), as well as to the signal propagation as shown by the red arrows.

Fig. 2 shows the *Chemokine signaling pathway*. On this pathway, the impact is due both to the large number of DE genes (26 out of 130), as well as to the clear signal propagation from the chemokines outside the cell (11 chemokines up-regulated), through the chemokine receptor and via the JAK and STAT mechanism. Note that the same mechanism is also identified on the Influenza A pathway shown in Fig. S4 in the Supplementary Materials. Fig. S9 shows another view of the mechanism involving the genes on this pathway and all their known interactions.

Together, the GO analysis, pathway analysis and the putative mechanisms identified by the analysis above strongly suggest a hyper-inflammation/cytokine storm.

3.4 Screening of potential therapeutic approaches: Proposed drugs.

Once we identified the main regulatory pathways potentially associated with hyper-inflammation, we evaluated *in silico* FDA-approved drugs that could show activity on multiple components of inflammation and consequently could be used for the management of severe COVID-19 cases. We considered the number of DE genes that would be reverted by each drug, as well as calculated a Bonferroni-corrected p-value indicating the suitability of each drug for repurposing in COVID-19 based on two different approaches (see "Methods" section). We looked for drugs that have both small Benforroni-corrected p-values as well as revert a larger number of DE genes. The top five drugs drug identified by our analysis are shown in Fig. 3. Methylprednisolone (MP) and prednisolone are corticosteroids currently used to modulate the immune response in rheumatoid arthritis. Diclofenac is a non-steroidal anti-inflammatory drug (NSAID). Tofacitinib is a JAK inhibitor (see the JAK-STAT mechanism identified in Fig. 2). Gold sodium thiomalate is an older anti-inflammatory drug, also used in the treatment of rheumatoid arthritis.

Methylprednisolone (MP) is the drug that was identified as the most likely to work by inhibiting the inflammatory pathway. This drug targets 27 genes that are found to be DE in COVID19vsControl. Out of these 27 genes, the drug would revert the changes in 25 of them. The drug also had an extremely significant p-value even after a Bonferroni correction which is the most stringent correction available ($p = 5.72 \times 10^{-10}$). MP also reverted 22 out of 22 genes found to be DE in NHBECoV2vsControl, and 25 out of 26 genes found to be DE in A549CoV2vsControl. Fig. 4 shows the putative mechanism through which MP acts on the DE genes in COVID19vsControl, and how these genes influence the BPs found to be significantly impacted.

3.5 Validation on independent data sets

The initial results obtained on from the data above were subsequently confirmed using additional data, spanning again all three types of samples: cell lines, tissue cultures, and patient samples. The additional cell line data

Chemical name	COVID19 vs. Healthy- mRNA (RNA-seq)		lung epithelium (NHBE) with CoV vs. Control- mRNA (RNA-seq)		lung alveolar (A549) with IAV vs. Control- mRNA (RNA-seq)		lung alveolar (A549) with CoV2 vs. Control- mRNA (RNA-seq)		lung alveolar (A549) with RSV vs. Control- mRNA (RNA-seq)	
	consistent (-)/ DE targets	p-value	consistent (-)/ DE targets	p-value	consistent (-)/ DE targets	p-value	consistent (-)/ DE targets	p-value	consistent (-)/ DE targets	p-value
Methylprednisolone	25/27	5.725e-10	22/22	8.996e-14			25/26	6.029e-14	35/37	3.183e-15
Gold Sodium Thiomalate	22/24	5.973e-8	22/22	8.996e-14			24/25	9.998e-14	35/36	8.245e-16
Prednisolone	27/34	1.737e-7	22/24	1.373e-12			22/23	2.754e-13	35/41	2.930e-13
tofacitinib	18/19	8.804e-12	15/15	3.238e-12			14/14	3.935e-12	21/22	9.635e-13
Diclofenac	28/35	5.307e-10	15/20	1.103e-8	1/3	1.000	20/20	2.702e-12	31/36	2.225e-11

Fig. 3: The top five drugs proposed for repurposing. The table shows both p-values corrected with Bonferroni, as well as the number of DE genes that would be reverted out of the total number of DE genes immediately downstream of each drug (annotated as “consistent (-)/DE targets" in the table). Methylprednisolone (MP) and prednisolone are corticosteroids currently used to modulate the immune response in rheumatoid arthritis. Gold Sodium Thiomalate is an older drug, not currently in use in the US. Diclofenac is a NSAID and tofacitinib is a JAK inhibitor. The column for A549IAVvsControl is empty because there are no DE genes targeted by these drugs in this contrast.

include data from Vero E6 cells infected with SARS-CoV-2 vs. controls. Additional tissue cultures include lung organoids infected with SARS-CoV-2 vs. controls. In addition, even though the SARS-CoV-2 virus is primarily thought to infect the lungs with transmission through the respiratory route, it has been suggested that the intestine may present another viral target organ (Lamers *et al.*, 2020). For this reason, we also compared the expression profiles of intestine organoids infected with SARS-CoV-2 vs. mock infection in differentiation and expansion media. These data are available in GEO as the GSE153940 (Vero E6 cells) (Riva *et al.*, 2020), GSE160435 (lung organoids), and GSE149312 (intestinal organoids) (Lamers *et al.*, 2020) data sets.

Finally, twenty nine additional samples from the lung of deceased COVID-19 patients with a high viral load and five controls were included from the Massachusetts General Hospital and Columbia University Irving Medical Center. These are available in GEO as the GSE150316 data set (Desai *et al.*, 2020).

Fig. S10 in Supplementary Materials show the biological processes common across the five additional independent data sets according to the the classical enrichment analysis. All of these are consistent with a viral response. In particular, the Type I interferon pathway is significant in every single data set analyzed, both initially, as well as in the additional validation data sets. The upstream regulators identified in the initial analysis, STAT2 and IRF9, were also found to be significant up-stream regulators in every single validation data set (see Fig. S11).

Fig. S12 shows existing FDA-approved drugs identified as suitable candidates for repurposing based on these five additional data sets. These additional and independent data sets span across cell lines, cell cultures and patient data, as detailed above. Note that the top five drugs obtained on these additional data sets are matching perfectly with those shown in Fig. 3 even though the data sets analyzed were completely independent.

3.6 In vivo effect of methylprednisolone: Clinical validation.

In an independent study, 213 patients diagnosed with COVID-19 were enrolled. 81 (38%) received conventional therapy (control group) while 132 patients (62%) received MP (MP group). The clinical characteristics and treatments received by the patients are shown in Table S2 and Table S3, respectively. As shown in Table S4 thirty day all-cause mortality occurred at a significantly lower rate in the MP group compared to control group (29.6% vs. 16.6%, $p = 0.027$). No statistical difference was detected in the proportion of patients prescribed empiric antibiotics or the time to empiric therapy. Kaplan Meier survival curve for 30-day mortality

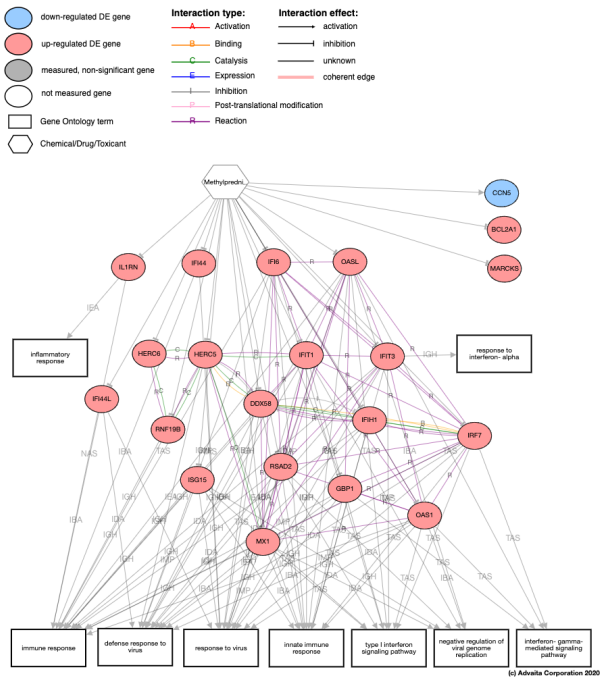

Fig. 4: The putative mechanism through which MP acts on the genes measured to be DE, and how these genes influence the biological processes found to be significantly impacted in the COVID19vsControl. This figure shows that: i) MP is known to revert the measured changes in all these 21 DE genes; ii) 20 out of these 21 DE genes are up-regulated suggesting a very strong immune response; ii) many of the DE gene targeted directly by MP are directly involved in the top biological processes identified as significantly perturbed by the disease.

demonstrated increased probability of survival at 30-days in the MP group as compared to the control group ($p = 0.0204$) (Fig. 5).

When comparing the two groups, those patients treated with a 3-day methylprednisolone protocol spent less time in the hospital (5 vs 8 days) and were less likely to be admitted to the ICU (27% vs 44%), being placed on a ventilator (22% vs 37%) or dying (14% vs 26%).

For a composite end point of preventing ICU admission, need for mechanical ventilator or mortality, the number needed to treat (NNT) to benefit a single patient was only 5 when methylprednisolone was used early in hospitalization. To prevent mortality, the NNT to benefit a single patient was only 8 for all hospitalized patients. This is in contrast to the RECOVERY trial (NCT04323592) for dexamethasone, where NNT was 8 for patients on mechanical ventilation and 25 for patients needed oxygen to prevent mortality.

3.7 Other drugs investigated.

We also looked at other drugs that have already been proposed as repurposing candidates for COVID-19 including: **chloroquine**, **hydroxychloroquine**, **erythromycin**, **prednisone**, **dexamethasone**, **ibuprofen**, **ritonavir**, **aspirin**, and **clopidogrel**. Most or all of these drugs are currently under clinical trials (Sanders *et al.*, 2020).

Chloroquine was found to revert only 2 out of 4 genes found to be differentially expressed in the COVID19vsControl ($p = 1$) and only 3 out of 6 genes found to be differentially expressed in NHBECoV2vsControl only ($p = 1$). Furthermore, this drug was not found to be potentially effective in reversing the changes in A549RSVvsControl ($p = 1$) or A549CoV2vsControl ($p = 1$). These results suggest that chloroquine

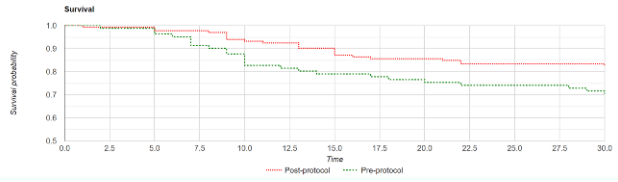

Fig. 5: Kaplan Meier survival curve for 30-day mortality demonstrating increased probability of survival at 30-days in the post methylprednisolone cohort as compared to the pre-methylprednisolone cohort (p= 0.0204).

would not be a good potential candidate for repurposing for the goal of modulating the immune response.

Hydroxychloroquine did not appear as a good candidate for repurposing in any of the phenotypes and contrasts studied here. Chloroquine and hydroxychloroquine do not target any of the 21 genes that are both severely dysregulated, and also targeted by the proposed drugs. Note that while these drugs do not reverse observed gene expression changes, they may act as anti-virals by potentially inhibiting the viral replication (Sanders *et al.*, 2020).

Erythromycin targets only three DE genes in COVID19vsControl and would revert only 2 of those. This yields an insignificant p value (Bonferroni-corrected $p = 1$ and FDR-corrected $p = 0.75$).

Ibuprofen was also not found to be a good candidate for use in COVID-19, having the potential to revert only 4 out of 10 DE genes ($p = 1$). This suggests a phenomenon in the class of NSAIDs similar with that observed within the corticosteroids: while one or two specific drugs may be effective, these effects cannot be generalized to the entire class. In other words, not all NSAIDs may be equally helpful in modulating the over-inflammation induced by COVID-19.

Ritonavir was found to be significant ($p = 0.002$) in reverting changes in 8 out of its 10 targets that were measured to be differentially expressed in NHBECoV2vsControl. However, ritonavir was not found to be effective in reversing the gene changes induced in COVID19vsControl.

Clopidogrel targets only one DE gene in COVID19vsControl, one DE gene in A549CoV2vsControl, and no DE gene in NHBECoV2vsControl. These results suggest that this drug is unlikely to be effective in COVID-19 with a p-value of 1 across all experiments.

Finally, **aspirin** is targeting 44 DE genes in COVID19vsControl but reversing only 22 of them. In NHBECoV2vsControl samples, aspirin was found to target 18 genes and revert 12 on them. In NHBE, aspirin has a raw p-value of 0.028 but after an FDR correction this becomes 0.187. In the COVID19vsControl contrast, even the raw p-value is 0.939 and becomes 1 after any correction.

4 Discussion

In the present study we described an initial characterization of the main pro-inflammatory pathways induced by SARS-Cov-2 infection on human lung epithelial cells and the identification of the most effective therapeutic approach to inhibit this cytokine storm.

In this study we have identified MP as the most effective, FDA approved, drug that targets critical components of the inflammatory pathway responsible for ARDS. Furthermore, we demonstrated its efficacy in a clinical trial in which MP decreased the incidence of mortality in COVID-19 patients. An important finding of this study is that drugs in the same class might not necessarily have similar effects. For instance, MP and prednisolone were predicted to be effective in reverting many of the changes triggered by COVID-19, while other closely-related corticosteroids such as prednisone were not. MP and prednisolone are corticosteroids currently used to modulate the immune response in rheumatoid arthritis. Interestingly, we observed that the putative mechanisms through which these drugs would revert the genes dysregulated in COVID-19 are different (Fig. 6). MP for example inhibits STAT1, IFT3 and HERC5 while prednisolone has an impact on IFIT genes such as IFT1, IFIT3, IFI6, and IFI4L.

We also looked at other corticosteroids such as prednisone, and hydrocortisone. However, prednisone was found to target only 3 DE genes in the COVID19vsControl and only 2 DE genes in the NHBECoV2vsControl. From those, prednisone would revert only 1 of the 3 DE genes in the COVID19vsControl and 0 out of 2 DE genes in the NHBECoV2vsControl. Both yielded insignificant Bonferroni-corrected p-values ($p = 1$) suggesting that prednisone is not expected to be a highly effective treatment. Prednisolone, dexamethasone, and hydrocortisone belong to the same family of corticosteroid anti-inflammatory agents and there is also a structural similarity between them (Fig. S13). In spite of this structural similarity, hydrocortisone is known to revert only 8 out of 10 DE genes in the COVID19vsControl (FDR-corrected $p = 0.57$) and 5 out of 8 DE genes in the NHBECoV2vsControl (FDR-corrected $p = 0.038$, Bonferroni-corrected $p = 1$). Dexamethasone was found to revert 33 out of 69 DE genes in the COVID19vsControl (FDR-corrected $p = 1$) and 27 out of 45 DE genes in the NHBECoV2vsControl (FDR-corrected $p = 0.002$, Bonferroni-corrected $p = 0.066$). Dexamethasone is significant in the NHBECoV2vsControl but not in the COVID19vsControl. Hydrocortisone appears as significant in COVID19vsControl, but only marginally so in the NHBECoV2vsControl.

Such differences between drugs in the same class can be potentially explained in two ways. First, different drugs can be associated with a different number of annotations. For instance, the number of genes that a given drug is known to be targeting can influence its significance. Table S1 in Supplementary Materials shows the number of known targets associated with some drugs relevant to COVID-19. Second, there could be genuine differences between the effectiveness of different corticosteroids, potentially due to a different number of genes truly impacted by each drug. The results of these study, based on all annotations available to date, suggest that MP would revert the largest number of the gene perturbed by COVID-19, followed by dexamethasone and, as shown in the outcome of COVID-19 infected patients, have a major impact on their clinical outcome. Prednisone and hydrocortisone revert much fewer known genes and consequently, could have an effect but it is expected to be less effective than the one observed with MP. Future clinical trials comparing the efficacy of these different corticosteroids are necessary to confirm our findings

The host inflammatory response in the lungs lead to acute lung injury and ARDS. This constitutes the main rationale for the use of corticosteroids. However, administration of corticosteroids is associated with multiple side effects, such as an increased risk of secondary infection and delayed viral clearance. A recent article in Lancet reports that clinical evidence does not support corticosteroid treatment for COVID-19 (Russell *et al.*, 2020). However, this report looks at corticosteroids as an entire class of drugs. A recent retrospective study of 201 patients with COVID-19 in China found that treatment with MP for those who developed ARDS was associated effective in decreasing the risk of death. Among patients with ARDS, treatment with MP decreased the risk of death (HR, 0.38; 95% CI, 0.20-0.72). In this study, 23 of 50 [46%] patients with MP treatment died compared to 21 deaths out of 34 patients without MP treatment (Wu *et al.*, 2020a). Both reports are entirely consistent with our findings: corticosteroids in general are NOT expected to help as a class of drugs, but rather each steroid should be assessed individually.

Methylprednisolone has been also the focus on several other recent clinical studies. Wu *et al.* studied the effect of MP in a cohort of 201 COVID-19 patients, of which 84 developed ARDS (Wu *et al.*, 2020b). They report that "among patients with ARDS, treatment with MP decreased the risk of death (HR, 0.38; 95% CI, 0.20-0.72)." The percentage of people

Fig. 6: The putative mechanisms through which methylprednisolone would revert the changes triggered by COVID-19 in the lung tissue.

who developed ARDS and eventually died was reduced from 61.8% (21 of 34) to 46% (23 of 50).

Salton *et al.* conducted a multi-center, observational study to explore the association between a prolonged, low-dose of MP and a composite endpoint including the need for ICU referrals, intubation, or death within 28 days (www.clinicaltrials.gov, NCT04323592) (Salton *et al.*, 2020). They found that in a cohort of 83 patients treated with MP and 90 controls, the treatment with MP significantly lowered the hazard of death (71%). The composite end point was met by 19 vs. 40 (adjusted hazard ratio (HR) 0.41; 95% confidence interval (CI): 0.24-0.72). Transfer to ICU and need for invasive MV was necessary in 15 vs. 27 ($p = 0.07$) and 14 vs. 26 ($p = 0.10$), respectively. By day 28, the MP group had fewer deaths (6 vs. 21, adjusted HR=0.29; 95% CI: 0.12-0.73) and more days off invasive MV (24.0 +/- 9.0 vs. 17.5 +/- 12.8; $p = 0.001$).

A similar multi-center study focus on MP was undertaken in Spain (European Clinical Trials Register: 2020-001934-37). This study involved 85 patients of which 34 were randomized to MP, 22 were assigned to MP by the clinician and 29 constituted the control group. The composite endpoint requirement of non-invasive ventilation, admission to ICU, and death. This study reported that the use of MP was associated with a reduced risk of the composite endpoint (RR 0.55 95% CI 0.33-0.91) (https://www.medrxiv.org/content/10.1101/2020.06.17.20133579v1).

Finally, preliminary results from a large randomized study undertaken in the UK (NCT04381936) show that dexamethasone is also able to significantly reduce the number of deaths. In this study the steroid group included 2,100 participants and was compared with the standard of care group including 4,300 participants. Unfortunately, this study placed dexamethasone in the same arm with MP (MP to be administered to pregnant women instead of dexamethasone), so this study will not be able to elucidate whether MP and dexamethasone have any differences in efficacy. However, the results of the clinical study presented here suggest that MP has a lower NNT than dexamethasone. The potential higher efficacy of MP needs to be confirmed in larger cohorts. However, in a recent meta-analysis of randomized studies investigating prolonged corticosteroid treatment in non-viral ARDS, methylprednisolone treatment achieved a reduction in duration of mechanical ventilation and mortality superior to that of dexamethasone or hydrocortisone (Meduri *et al.*, 2020).

We also looked at other drugs that have already been proposed as repurposing candidates for COVID-19 including: chloroquine, hydroxychloroquine, erythromycin, prednisone, ibuprofen, ritonavir, aspirin, and clopidogrel. None of these was predicted to be effective in reverting SARS-CoV2 gene expression changes (see Supplementary Materials).

In summary, we describe the inflammatory pathways associated with the cytokine storm found in COVID-19 patients and describe the characterization of the potential effective therapeutic approaches that, by targeting these pathways could effectible reduce the hyper-inflammatory response responsible for the development of ARDS, a main cause of mortality in COVID-19 patients. Clinical results confirmed the efficacy of the *in silico* prediction that indicated MP could improve outcomes in severe COVID-19.

5 Methods

Gene Ontology (GO) analysis method. For each GO term (Ashburner and Lewis, 2002; Consortium *et al.*, 2004), the number of DE genes annotated to the term is compared to the number of DE genes expected just by chance. A p-value is computed using the hypergeometric distribution (Draghici *et al.*, 2003; Draghici, 2011) and corrected with FDR and Bonferroni. We also used an intelligent pruning approach inspired by the *elim* and *weight* pruning methods (Alexa *et al.*, 2006). The algorithm constructs a custom cut through the GO hierarchy by starting with the most specific nodes and calculating their p-value with all genes assigned directly to each such node. If a node is significant, it is reported as such. If the node is not significant, the genes associated to the given node are propagated to its direct ancestors and a p-value is calculated for each of those. See Supplementary Materials for full details.

Pathway analysis method. The pathways analysis was performed using the Impact Analysis method (Draghici *et al.*, 2007; Tarca *et al.*, 2009a; Khatri *et al.*, 2007). The impact analysis uses two types of evidence: i) the over-representation of differentially expressed (DE) genes in a given pathway and ii) the perturbation of that pathway computed by propagating the measured expression changes across the pathway topology. These aspects are captured by two independent probability values, *pORA* and *pAcc*, that are then combined in a unique pathway-specific p-value. More details are provided in the Supplementary Materials and elsewhere (Draghici *et al.*, 2007; Tarca *et al.*, 2009b).

Putative mechanism inference. The impact analysis allows the computation of the perturbation at the level of every gene on every pathway. Putative mechanisms are identified as sequences of pathway signals for which the measured gene expression changes are consistent with the sequence of events described by the pathway. An example would be the chemokine-chemokine receptor-JAK-STAT sequence identified on the Chemokine Signaling Pathway in Fig. 2. This approach is able to identify candidates for potentially active mechanisms on existing pathways and has also been shown to have an increased accuracy (Nguyen *et al.*, 2019).

The prediction of upstream Chemicals, Drugs, Toxicants (CDTs) is based on two types of information: i) the enrichment of DE genes from the experiment and ii) a network of interactions from the Advaita Knowledge Base (AKB v1910, www.advaitabio.com). The network is a directed graph in which the source node represents either a chemical substance or compound, a drug, or a toxicant (CDT). We focused our work on FDA-approved drugs that could be repurposed. The edges represent known increase or decrease expression effects that these CDTs have on various genes. The analysis considers the hypothesis that a drug could revert the measured gene expression changes. This analysis is related but

substantially different from existing drug repurposing approaches, such as (Sirota *et al.*, 2011; Peyvandipour *et al.*, 2020; Saberian *et al.*, 2019). More details are included in Supplementary Materials.

Clinical Validation. We evaluated the MP protocol with a single pretest, single post-test quasi-experiment from March 12-March 27, 2020 at a 5 hospital health system in Michigan. Patients were compared before and after implementation of the MP protocol on March 20th. The clinical characteristics of the patients are shown in Table S2. The primary endpoint was 30 day all-cause mortality.

The methylprednisolone protocol. Patients with PCR confirmed COVID-19 who required 4 liters or more of oxygen per minute on admission, or who had escalating oxygen requirements from baseline, were recommended to receive IV methylprednisolone 0.5 to 1 mg/kg/day in 2 divided doses for 3 days. Patients who required ICU admission were eligible to extend the IV methylprednisolone course to a maximum of 7 days at the discretion of the medical team. Institutional guidelines also recommended hydroxychloroquine 400 mg twice daily for 2 doses on day 1, followed by 200 mg twice daily on days 2–5.

Statistical Analysis of clinical data. Survival analysis was performed using the Kaplan-Meier method and log-rank test. More details about the statistical analysis and characteristics of the patient population are included in the Supplementary Materials.

6 Conclusion

This paper presents an approach for drug repurposing based on identifying drugs that could revert gene expression changes associated with most perturbed biological processes and pathways. Results from a clinical study undertaken in a cohort of 213 patients in a multi-center hospital system confirmed the efficacy of the *in silico* prediction that indicated MP could improve outcomes in severe COVID-19. This prediction is also supported by the results of independent clinical studies with the same drug undertaken in Italy (173 patients) and Spain (85 patients). The drug repurposing approach described here, as well as the drugs identified, might be important for any future pandemic involving hyper-inflammation.

7 Authors Contributions

SD conceived and designed the study, performed the data analysis, interpreted the RNA-Seq data, and wrote the manuscript. RV, TMN performed the RNA-Seq data analysis, interpreted data and assisted with the manuscript preparation. LAS helped with clinical interpretation and assisted with the manuscript preparation. CZ revised the manuscript. RF, AM, RMK GA and MR conducted the clinical study and analyzed the clinical data. GM interpreted data and wrote the manuscript.

8 Funding

This work was supported in part by NIH grant NIAID 1R01AI145829-01 (GM) and NSF grant 2029572 (CZ and SD).

References

Alexa, A., Rahnenfuhrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, **22**(13), 1600–7.

Ashburner, M. and Lewis, S. (2002). On ontologies for biologists: the Gene Ontology—untangling the web. In *'In Silico' Simulation of Biological Processes: Novartis Foundation Symposium 247*, volume 247, pages 66–83. Wiley Online Library.

Ayres, J. S. (2020). Surviving COVID-19: A disease tolerance perspective. *Science Advances*, **6**(18).

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of The Royal Statistical Society B*, **57**(1), 289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, **29**(4), 1165–1188.

Blanco-Melo, D., Nilsson-Payant, B. E., Liu, W.-C., Uhl, S., Hoagland, D., Møller, R., Jordan, T. X., Oishi, K., Panis, M., Sachs, D., Wang, T. T., Schwartz, R. E., Lim, J. K., Albrecht, R. A., and tenOever, B. R. (2020). Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*, **181**(5), 1036–1045.

Consortium, G. O. *et al.* (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Research*, **32**(suppl 1), D258–D261.

Desai, N., Neyaz, A., Szabolcs, A., Shih, A. R., Chen, J. H., Thapar, V., Nieman, L. T., Solovyov, A., Mehta, A., Lieb, D. J., *et al.* (2020). Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. *medRxiv*.

Draghici, S. (2011). *Statistics and Data Analysis for Microarrays using R and Bioconductor*. Chapman and Hall/CRC Press.

Draghici, S., Khatri, P., Martins, R. P., Ostermeier, G. C., and Krawetz, S. A. (2003). Global functional profiling of gene expression. *Genomics*, **81**(2), 98–104.

Draghici, S., Khatri, P., Tarca, A. L., Amin, K., Done, A., Voichiţa, C., Georgescu, C., and Romero, R. (2007). A systems biology approach for pathway level analysis. *Genome Research*, **17**(10), 1537–1545.

Horvath, C. M., Stark, G. R., Kerr, I. M., and Darnell, J. E. (1996). Interactions between STAT and non-STAT proteins in the interferon-stimulated gene factor 3 transcription complex. *Molecular and Cellular Biology*, **16**(12), 6957–6964.

Khatri, P., Draghici, S., Tarca, A. L., Hassan, S. S., and Romero, R. (2007). A system biology approach for the steady-state analysis of gene signaling networks. In *Progress in Pattern Recognition, Image Analysis and Applications*, pages 32–41. Springer.

Lamers, M. M., Beumer, J., Vaart, J. v. d., Knoops, K., Puschhof, J., Breugem, T. I., Ravelli, R. B. G., Schayck, J. P. v., Mykytyn, A. Z., Duimel, H. Q., Donselaar, E. v., Riesebosch, S., Kuijpers, H. J. H., Schippers, D., Wetering, W. J. v. d., Graaf, M. d., Koopmans, M., Cuppen, E., Peters, P. J., Haagmans, B. L., and Clevers, H. (2020). SARS-CoV-2 productively infects human gut enterocytes. *Science (New York, N.Y.)*, page eabc1669.

Louie, J. K., Yang, S., Acosta, M., Yen, C., Samuel, M. C., Schechter, R., Guevara, H., and Uyeki, T. M. (2012). Treatment with neuraminidase inhibitors for critically ill patients with influenza A (H1N1) pdm09. *Clinical Infectious Diseases*, **55**(9), 1198–1204.

Meduri, G. U., Annane, D., Confalonieri, M., Chrousos, G. P., Rochwerg, B., Busby, A., Ruaro, B., and Meibohm, B. (2020). Pharmacological principles guiding prolonged glucocorticoid treatment in ards. *Intensive care medicine*, pages 1–13.

Mehta, P., McAuley, D. F., Brown, M., Sanchez, E., Tattersall, R. S., Manson, J. J., HLH Across Speciality Collaboration, *et al.* (2020). COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet (London, England)*, **395**(10229), 1033.

Nguyen, T.-M., Shafi, A., Nguyen, T., and Draghici, S. (2019). Identifying significantly impacted pathways: a comprehensive review and assessment. *Genome Biology*, **20**(1), 1–15.

Peyvandipour, A., Shafi, A., Saberian, N., and Draghici, S. (2020). Identification of cell types from single cell data using stable clustering. *Scientific reports*, **10**(1), 1–12.

Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Pache, L., Burgstaller-Muehlbacher, S., Jesus, P. D. D., Teriete, P., Hull, M. V., Chang, M. W., Chan, J. F.-W., Cao, J., Poon, V. K.-M., Herbert, K. M., Cheng, K., Nguyen, T.-T. H., Rubanov, A., Pu, Y., Nguyen, C., Choi, A., Rathnasinghe, R., Schotsaert, M., Miorin, L., Dejosez, M., Zwaka, T. P., Sit, K.-Y., Martinez-Sobrido, L., Liu, W.-C., White, K. M., Chapman, M. E., Lendy, E. K., Glynne, R. J., Albrecht, R., Ruppin, E., Mesecar, A. D., Johnson, J. R., Benner, C., Sun, R., Schultz, P. G., Su, A. I., García-Sastre, A., Chatterjee, A. K., Yuen, K.-Y., and Chanda, S. K. (2020). Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature*, **586**(7827), 113–119.

Russell, C. D., Millar, J. E., and Baillie, J. K. (2020). Clinical evidence does not support corticosteroid treatment for 2019-nCoV lung injury. *The Lancet*, **395**(10223), 473–475.

Saberian, N., Peyvandipour, A., Donato, M., Ansari, S., and Draghici, S. (2019). A new computational drug repurposing method using established disease–drug pair knowledge. *Bioinformatics*, **35**(19), 3672–3678.

Salton, F., Confalonieri, P., Meduri, G. U., Santus, P., Harari, S., Scala, R., Lanini, S., Vertui, V., Oggionni, T., Caminati, A., *et al.* (2020). Prolonged low-dose methylprednisolone in patients with severe covid-19 pneumonia. In *Open forum infectious diseases*, volume 7, page ofaa421. Oxford University Press US.

Sanders, J. M., Monogue, M. L., Jodlowski, T. Z., and Cutrell, J. B. (2020). Pharmacologic treatments for coronavirus disease 2019 (covid-19): a review. *Jama*, **323**(18), 1824–1836.

Siddiqi, H. K. and Mehra, M. R. (2020). COVID-19 illness in native and immunosuppressed states: A clinical–therapeutic staging proposal. *The Journal of Heart and Lung Transplantation*, **39**(5), 405.

Sirota, M., Dudley, J. T., Kim, J., Chiang, A. P., Morgan, A. A., Sweet-Cordero, A., Sage, J., and Butte, A. J. (2011). Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Science Translational Medicine*, **3**(96).

Tarca, A. L., Draghici, S., Khatri, P., Hassan, S. S., Mittal, P., Kim, J.-s., Kim, C. J., Kusanovic, J. P., and Romero, R. (2009a). A novel signaling pathway impact analysis. *Bioinformatics*, **25**(1), 75–82.

Tarca, A. L., Draghici, S., Khatri, P., Hassan, S. S., Mittal, P., Kim, J.-S., Kim, C. J., Kusanovic, J. P., and Romero, R. (2009b). A novel signaling pathway impact analysis (SPIA). *Bioinformatics*, **25**(1), 75–82.

Tavazoie, S., Hughes, J. D., Campbell, M. J., Cho, R. J., and Church, G. M. (1999). Systematic determination of genetic network architecture. *Nature Genetics*, **22**, 281–285.

Wu, C., Chen, X., Cai, Y., Zhou, X., Xu, S., Huang, H., Zhang, L., Zhou, X., Du, C., Zhang, Y., *et al.* (2020a). Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Internal Medicine*.

Wu, C., Chen, X., Cai, Y., Xia, J., Zhou, X., Xu, S., Huang, H., Zhang, L., Zhou, X., Du, C., Zhang, Y., Song, J., Wang, S., Chao, Y., Yang, Z., Xu, J., Zhou, X., Chen, D., Xiong, W., Xu, L., Zhou, F., Jiang, J., Bai, C., Zheng, J., and Song, Y. (2020b). Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Internal Medicine*, **180**(6).