A Network Medicine Approach to Investigation and Population-based Validation of Disease Manifestations and Drug Repurposing for COVID-19

Yadi Zhou[1†], Yuan Hou[1†], Jiayu Shen[1], Asha Kallianpur[1,2], Joe Zein[2,3], Daniel A. Culver[3,4], Samar Farha[3,4], Suzy Comhair[2,3], Claudio Fiocchi[2,3], Michaela U. Gack[2,3], Reena Mehra[2,3], Thaddeus Stappenbeck[2,3], Timothy Chan[2,3], Charis Eng[1,2,3,5,6], Jae U. Jung[2,3], Lara Jehi[2,3], Serpil Erzurum[2,3], Feixiong Cheng[1,2,6,*]

[1]Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA

[2]Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, OH 44195, USA

[3]Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA

[4]Department of Pulmonary Medicine, Respiratory Institute, Cleveland Clinic, Cleveland, OH 44195, USA

[5]Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

[6]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

[†]These authors contributed equally to this work.

*To whom correspondence should be addressed:
Feixiong Cheng, PhD
Lerner Research Institute, Cleveland Clinic
Tel: +1-216-444-7654; Fax: +1-216-636-0009
Email: chengf@ccf.org

Abstract

The global Coronavirus Disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has led to unprecedented social and economic consequences. The risk of morbidity and mortality due to COVID-19 increases dramatically in the presence of co-existing medical conditions while the underlying mechanisms remain unclear. Furthermore, there are no proven effective therapies for COVID-19. This study aims to identify SARS-CoV-2 pathogenesis, diseases manifestations, and COVID-19 therapies using network medicine methodologies along with clinical and multi-omics observations. We incorporate SARS-CoV-2 virus-host protein-protein interactions, transcriptomics, and proteomics into the human interactome. Network proximity measure revealed underlying pathogenesis for broad COVID-19-associated manifestations. Multi-modal analyses of single-cell RNA-sequencing data showed that co-expression of *ACE2* and *TMPRSS2* was elevated in absorptive enterocytes from the inflamed ileal tissues of Crohn's disease patients compared to uninflamed tissues, revealing shared pathobiology by COVID-19 and inflammatory bowel disease. Integrative analyses of metabolomics and transcriptomics (bulk and single-cell) data from asthma patients indicated that COVID-19 shared intermediate inflammatory endophenotypes with asthma (including *IRAK3* and *ADRB2*). To prioritize potential treatment, we combined network-based prediction and propensity score (PS) matching observational study of 18,118 patients from a COVID-19 registry. We identified that melatonin (odds ratio (OR) = 0.36, 95% confidence interval (CI) 0.22-0.59) was associated with 64% reduced likelihood of a positive laboratory test result for SARS-CoV-2. Using PS-matching user active comparator design, melatonin was associated with 54% reduced likelihood of SARS-CoV-2 positive test result compared to angiotensin II receptor blockers or angiotensin-converting enzyme inhibitors (OR = 0.46, 95% CI 0.24-0.86).

Introduction

The ongoing global Coronavirus Disease 2019 (COVID-19) pandemic has led to 5 million confirmed cases and 327,000 deaths worldwide as of May 21, 2020. The United States alone has 1.58 million confirmed cases with a death toll of more than 93,000. Several retrospective studies have reported the clinical characteristics of individuals with symptomatic COVID-19 infection, and an emerging theme has been the significantly higher risk of morbidity and mortality among individuals with one or more comorbid health conditions, such as hypertension, asthma, diabetes mellitus, cardiovascular or cerebrovascular disease, chronic kidney disease and malignancy (*1-6*). However, these retrospective clinical studies are limited by small sample sizes and unmeasured confounders, leaving the underlying patho-mechanisms largely unknown. More specifically, it is unclear whether associations of disease manifestations and COVID-19 severity are merely a reflection of poorer health in general, or a clue to shared pathobiological mechanisms.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes COVID-19, is an enveloped virus that carries a single-stranded positive-sense RNA genome (*6, 7*). SARS-CoV-2 is a newly discovered member of the coronavirus (CoV) family (*8*). SARS-CoV-2 enters host cells via binding with its spike protein to the angiotensin converting enzyme 2 (ACE2) receptor on the surfaces of many cell types (*9*). This binding is primed by transmembrane protease serine 2 (TMPRSS2) (*9*) and the host cell protease furin (*10*) (**Fig. 1A**). Studies have shown that *ACE2* and *TMPRSS2* are highly co-expressed in alveolar type II (AT2) epithelial cells in the lung (*11*), nasal mucosa (*12*), bronchial secretory cells (*13*), and absorptive enterocytes in the ileum (*14*). Yet, much remains to be learned on how these critical human proteins involved in the infection and replication of SARS-CoV-2 are associated with various

disease comorbidities and complications. Systematic identification of the host factors involved in the protein-protein interactions (PPIs) of SARS-CoV-2 and the human host will facilitate identification of therapeutic targets and advance understanding of the complications and comorbidities resulting from COVID-19 (*15-17*). Studies using transcriptomics (*18*), proteomics (*19*), and interactomics (*20*) (PPIs) methods have contributed to a better understanding of the SARS-CoV-2-host interactome, which enabled the investigation of the complications and comorbidities of SARS-CoV-2 and a facilitated search for effective treatment (**Fig. 1B**).

Major efforts are underway to develop safe and effective drugs to treat COVID-19: preventive and therapeutic strategies currently being explored include vaccination, SARS-CoV-2-specific antibody, novel nucleoside analogues such as remdesivir, and repurposed drugs (*21-25*). The development of vaccination and antibody therapies takes significantly longer compared to drug repurposing. Many existing drugs are currently or have been tested in clinical trials, such as the antimalarial drug hydroxychloroquine and protease inhibitor combination lopinavir/ritonavir (*26, 27*). However, early results from these trials have not yet shown significant clinical benefits for COVID-19 patients (*26, 27*). We recently evaluated more than 2,000 FDA approved/investigational drugs using network-based method and prioritized 16 drug candidates and 3 drug combinations for COVID-19 (*28*). Yet, the answer to the key question why an approved drug originally documented for other diseases can be repurposed for COVID-19 remains unclear. One possible explanation is that COVID-19 shares common disease pathobiology or functional pathways elucidated by the human PPIs (*24, 28, 29*). Systematic identification of common disease pathobiological pathways shared by COVID-19 and other diseases would offer novel targets and therapies for COVID-19.

In this study, we present an integrative network medicine platform that quantifies the association of COVID-19 with other diseases across six categories, including autoimmune, malignant, cardiovascular, metabolic, neurological, and pulmonary (**Fig. 1C**). The rationale for these analyses rests on the notions that (i) the proteins that functionally associate with disease (such as COVID-19) are localized in the corresponding subnetwork within the comprehensive human PPI network (*24, 29-31*); and (ii) proteins that are associated with a specific disease may be directly targeted by the virus or are in the close vicinity of virus host proteins. We first performed network analysis followed by single-cell RNA-Seq data analysis to identify the underlying pathobiological relationships between COVID-19 and its associated comorbidities. Secondly, we used our network-medicine findings and patient data from a large COVID-19 clinical registry database to identify and prioritize existing FDA-approved drugs as potential COVID-19 drug candidates (**Fig. 1D**).

Results

A global map of the SARS-CoV-2 virus-host interactome

We assembled four host gene/protein sets for SARS-CoV-2 (see Methods) (**Table S1**): (1) SARS2-DEG, representing the differentially expressed genes (DEGs) from the transcriptomic data of SARS-CoV-2 infected primary human bronchial epithelial cells; (2) SARS2-DEP, representing the differentially expressed proteins (DEPs) from the proteomic data of SARS-CoV-2 infected human Caco-2 cells; (3) HCoV-PPI, representing the literature-based virus-host proteins across multiple human coronaviruses (HCoVs), including the SARS-CoV-1 (from the 2002-2003 pandemic) and MERS-CoV; and (4) SARS2-PPI (SARS-CoV-2 specific virus-host

5

PPIs). Since HCoV-PPI and SARS2-PPI are both physical virus-host PPIs, we further combined them as the fifth data set, PanCoV-PPI.

We first performed functional enrichment analyses for the five different data sets. We found that while these data sets share several common pathways and ontology terms (**Fig. 2A**), such as phagosome, measles, apoptosis, NF-kappa B signaling pathway, neutrophil-related immunity, apoptotic processes, virus transport, viral genome replication, and response to interferon, they differ considerably in terms of their most significantly enriched pathways (**Fig. S1-S5**). This is especially noticeable for SARS2-DEP and SARS2-PPI. While SARS2-DEG (**Fig. S1**) and HCoV-PPI (**Fig. S3**) show more enrichment in immune responses and viral pathways; SARS2-DEP (**Fig. S2**) is more related to various cellular metabolic pathways; and SARS2-PPI (**Fig. S4**) is more enriched in DNA replication, RNA transcription, and protein translation. These observations suggest that these different SARS-CoV-2 data sets capture complementary aspects of the biological and cellular states of the viral life cycle and host immunity. Therefore, building a global virus-host map (including interactome, transcriptome, and proteome) that incorporates data from transcriptomics, proteomics, and physical virus-host PPIs are essential for a better understanding of the pathogenesis of COVID-19. This global virus-host map for SARS-CoV-2 can offer a more complete picture of the interconnected functional pathways involved in viral pathogenesis, thereby facilitating discovery of therapeutic targets.

Network and biological characteristics of virus-host interactome for SARS-CoV-2

In addition to identifying the functions that these data sets represent, we next characterized the network patterns (node degree in the human PPI network) and bioinformatics features of these SARS-CoV-2 data sets, including ratio of non-synonymous to synonymous substitutions

6

(*dN/dS*), evolutionary rate ratio, and lung expression specificity (**Fig. 2B-G** and **Fig. S6**). To find common as well as unique network and bioinformatic characteristics of SARS-CoV-2, we further compiled four additional virus-host gene/protein networks identified by different methods for comparisons (**Table S2**): (1) 900 virus-host interactions connecting 10 other viruses and 712 host genes identified by gene-trap insertional mutagenesis, (2) 2,855 known virus-host interactions connecting 2,443 host genes and 55 pathogens identified from RNA interference (RNAi), (3) 579 host proteins mediating translation of 70 innate immune-modulating viral open reading frames (viORFs), and (4) 1,292 host genes mediating influenza-host interactions identified by co-immunoprecipitation and liquid chromatography-mass spectrometry (Co-IP+LC/MS). We found that host proteins in PanCoV-PPI (**Fig. 2B**) and four other data sets (SARS2-DEG, SARS2-DEP, HCoV-PPI, and SARS2-PPI) (**Fig. S6**) were more likely to be highly connected (high degree or connectivity) in the human PPI network, including several hubs, such as JUN, XPO1, MOV10, NPM1, VCP, and HNRNPA1. PanCoV-PPI has a comparable degree distribution with host proteins/genes identified by viORFs and Co-IP+LC/MS, although being marginally lower than that identified by RNAi and gene-trap insertional mutagenesis assay (**Fig. 2F**).

Expression patterns of genes in a specific disease-related tissue play crucial role for elucidation of disease pathogenesis and drug discovery (*24, 29*). Given the major impact of SARS-CoV-2 on pulmonary function and lung injury (*1*), we inspected the lung-specific expression of genes in PanCoV-PPI using a Z score measure compared to other 30 tissues from the GTEx database (see Methods). We found that most host genes for SARS-CoV-2 have high expression in lung (**Fig. 2E**, **Table S1**) compared to other tissues; yet, *ACE2* has a low expression in lung compared to other tissues. A recent study showed that *ACE2* was primarily

7

expressed in the epithelial cells in lungs, and only 3.8% of alveolar type 2 pneumocytes expressed both *ACE2* and *TMPRSS2*, but it is significantly upregulated in smokers and 24-48 hours following SARS-CoV-2 infection (*11, 32*). Another study also showed that despite relatively low expression of *ACE2* in the lung, *ACE2* was expressed in multiple epithelial cell types along the airway (*12*). Therefore, it is important to understand the cell type-specific SARS-CoV-2 pathogenesis at single cell levels.

To inspect the evolutionary factors underlying the SARS-CoV-2-human PPIs, we investigated the selective pressure and evolutionary rates quantified by the non-synonymous versus synonymous substitution rate ratios (*dN/dS* ratios, see Methods) using human-mouse orthologous gene pairs. We found that PanCoV-PPI has a stronger purifying selection (lower *dN/dS* ratios (**Fig. 2C**) and evolutionary rate ratio (**Fig. 2D**)) compared to the same number of random genes. PanCoV-PPI is also comparable to other four virus-host genes/proteins identified by different assays (**Fig. 2F, G**) in terms of node degrees and *dN/dS* ratios. Altogether, these observations suggest that the virus-host PPIs assembled in this study offer a high-quality interactome map for SARS-CoV-2 for identifying pathogenesis and potential treatments for COVID-19.

To inspect shared viral pathways across different viruses and SARS-CoV-2, we further performed network overlap analysis of PanCoV-PPI with 712 host genes across 10 types of other viruses identified by gene-trap insertional mutagenesis assays (*25*). We found a significant overlap of the SARS-CoV-2 host proteins with non-coronaviruses ($P < 0.002$, Fisher's exact test, **Fig. 2H**). For example, BRD2, a transcriptional regulator which belongs to Bromodomain and Extra-Terminal motif family, is connected to SARS-CoV-2 and two other viruses, herpes simplex virus 2 (HSV-2) and reovirus. RHOA, encoding a small GTPase protein in the Rho

family of GTPases, is connected to SARS-CoV-2, cowpox, and HSV-2 as well. RHOA has been

reported to be involved in multiple human diseases, including cardiovascular disease (*33*) and

cancer (*34*). These observations indicate possible disease manifestations associated with SARS-

CoV-2.

SARS-CoV-2 cellular network perturbations of disease manifestations

Investigation of the relationships between human host proteins targeted by SARS-CoV-2 and

disease susceptibility genes may offer crucial information for identifying COVID-19-associated

disease manifestations. We thus inspected the overlap between SARS-CoV-2 host genes/proteins

and the susceptibility gene sets implicated in different diseases and biological events (**Fig. 2I**).

We found that host genes/proteins targeted by SARS-CoV-2 are significantly enriched in

Mendelian disease ($P = 0.002$, Fisher's exact test), orphan disease ($P = 0.044$), and cancer ($P <$

0.001). Mechanistically, SARS-CoV-2 target host genes are significantly enriched in cell cycle

genes ($P < 0.001$) and innate immune genes ($P < 0.001$).

Using the five SARS-CoV-2 host gene/protein sets, we next tried to identify potential

COVID-19 comorbidities. To achieve this, we assembled the disease-protein network of

COVID-19 and six disease categories (**Table S3**). For cancer, the driver genes for pan-cancer

and individual cancer types were retrieved from the Cancer Gene Census (*35*) and a previous

study (*36*). For autoimmune, pulmonary, neurological, cardiovascular, and metabolic categories,

we extracted their associated genes/proteins from the Human Gene Mutation Database (*37*).

Using the disease-protein network together with the SARS2-PPI (SARS-CoV-2 virus-host

interactome) and HCoV-PPI (HCoV-host interactome) sets, we examined their overall

connectivity in the human PPI network (**Fig. 3A**, **Data File S1**). To build the global network for

9

the disease comorbidities, we extracted the PPIs from the human interactome for the virus target

proteins and disease-associated proteins. Each small node indicates a virus target host protein

(blue) or a disease-associated protein (green). Some disease-associated proteins can be directly

targeted by the viruses, as shown in orange. Edges among these protein nodes indicate PPIs. For

SARS-CoV-2, the targets of its individual viral proteins are shown based on the SARS2-PPI data

set. Due to the tendency of having common disease-associated proteins, some disease categories

tend to cluster closely, e.g., cancer and neurological. Diseases from other categories, such as

autoimmune and pulmonary, are scattered. Most of the virus target proteins are connected with

the disease-associated proteins, which suggests shared pathobiological pathways of COVID-19

and these diseases. Various cancer types formed a relatively distant module from the virus

targets, compared to other disease categories.

Shown in **Fig. 3A**, these diseases can be targeted directly or interact with the targets of

SARS-CoV-2 or other HCoVs. For example, among the four chronic obstructive pulmonary

disease (COPD) associated proteins shown in the network, TGFB1 is the direct target of HCoV-

229E, and all four proteins (TGFB1, DEFB1, SNAI1, and ADAM33) interact with at least one

SARS-CoV-2 viral protein target. The risk of various cardiovascular diseases was found to be

increased in COVID-19 patients, including heart block, coronary artery disease, and congestive

heart failure, and arrhythmia, which is consistent with clinically reported myocardial injury (*38*)

and cardiac arrest (*39*). These observations reveal common network relationship between

COVID-19 and human diseases (**Fig. 3A**). We then performed meta-analysis of 34 COVID-19

clinical studies (**Table S4**) to evaluate the pooled risk ratios of 10 comorbidities among 4,973

COVID-19 positive patients (including 2,268 mild and 731 severe patients). The random effects

model was used to estimate the pooled risk ratio of disease severity. The tau^2 and I^2 statistics

were used to evaluate the heterogeneity among studies. We found that 8 comorbidities have significantly higher risks in severe COVID-19 patients (**Fig. 3B**). The overall pooled risk ratio of COPD was 4.33 (95% confidence interval (CI) 2.42-7.74, $P = 0.001$) in 12 low heterogeneous clinical studies ($I^2 = 0.0\%$, $P = 0.6$). The COVID-19 patients with cardiovascular diseases have a risk ratio of 3.87 (95% CI 1.97-7.59, $P = 0.001$), and there was a slightly higher heterogeneity across the 13 studies ($I^2 = 57.6\%$, $P = 0.005$). Patients with stroke, diabetes, chronic kidney disease, hypertension, cancer, and history of smoking were also found with higher risk ratios of severe COVID-19. We next turned to quantify the network-based relationships between COVID-19 and human diseases under the human interactome model using the network proximity measure.

Network-based measure of COVID-19-associated disease manifestations

We systematically evaluated the network-based relationships of the 64 diseases across the 6 categories to COVID-19 (**Fig. 4A**). We used the state-of-the-art network proximity measure to evaluate the connectivity and the closeness of the disease proteins and SARS-CoV-2 host proteins, taking the topology of the human interactome network into consideration. To test the significance of the proximity, Z scores and P values were calculated based on the permutation tests and are shown in **Fig. 4A**. We found that each disease-disease pair has a well-defined network-based footprint. If the footprint between COVID-19 module and another disease module is significantly close (low Z score and $P < 0.05$), the magnitude of the proximity is indictive of their biological relationship: closer network proximity (**Fig. 4A**) of SARS-CoV-2 host genes/proteins with a disease module indicates higher potential of manifestation between COVID-19 and a specific disease. We first noticed that immunological, pulmonary, and

neurological diseases show significant proximity to SARS-CoV-2 data sets more frequently than

do cancer, cardiovascular, and metabolic diseases. Some diseases have significant proximities to

more than one SARS-CoV-2 data sets, most notably inflammatory bowel disease (IBD),

attention-deficit/hyperactivity disorder, and stroke, which achieved significant *P* value for all

five SARS-CoV-2 protein sets. Pulmonary diseases, including COPD, lung injury, pulmonary

fibrosis, and respiratory failure, achieved four significant proximities. Some diseases have

significant proximities to certain SARS-CoV-2 data sets, indicating associations at certain levels,

e.g., asthma (transcriptomic), respiratory distress syndrome (proteomic), and hypertension.

Network visualization can further show the connections between SARS-CoV-2 and other

diseases, for example, respiratory distress syndrome (**Fig. 4B**), sepsis (**Fig. 4C**), and COPD (**Fig.

S7**). We found that multiple SARS-CoV-2 host proteins are directly connected with the disease-

associated proteins (**Fig. 4B**). ABCA3 is a lipid transporter located in the outer membrane of

lamellar bodies in AT2 cells; the mutations of the *ABCA3* gene can disrupt pulmonary surfactant

homeostasis and lead to inherited pulmonary diseases (*40*). Another membrane surface protein,

the pulmonary-associated surfactant protein C encoded by *SFTPC*, can cause lung injury when

misfolded (*41*). For sepsis, we noticed several inflammatory and immune-related proteins, such

as IRAK1, IRAK3, IKBKB, and STAT3, suggesting overlap of the inflammatory response

activated in COVID-19 and sepsis (**Fig. 4C**). It has been reported that an overzealous production

of certain cytokines, such as IL-6, caused by dysregulation of innate immune responses to

SARS-CoV-2 infection can result in a 'cytokine storm', better known as cytokine release

syndrome (CRS) (*42*). The potential prognoses of acute respiratory distress syndrome and sepsis

using IL-6 expression levels were also established (*43-45*). IL-6 also has a significantly increased

expression level in the human bronchial epithelial cells infected with SARS-CoV-2 from the

SARS2-DEG data set (*18*). In addition, it is also potentially affected by SARS-CoV-2 through multiple PPIs (**Fig. 4B, 4C**, IL-6 neighbors), such as IL-6R, endophilin A1 (encoded by *SH3GL2*), and parathyroid hormone like hormone (encoded by *PTHLH*). Our random effects meta-analysis of 5 clinical studies of COVID-19 revealed that there was an increase of IL-6 levels in severe COVID-19 patients compared to non-severe COVID-19 patients (**Fig. 4D**). The mean difference was 33.0 pg/ml (95% CI 0.58-65.3, **Fig. 4D**) with high literature heterogeneity ($I^2 = 94\%$, $P < 0.001$). These results indicate that IL-6 plays a critical role in COVID-19-associated respiratory distress syndrome and sepsis. Due to the importance of IL-6 in SARS-CoV-2 infection, IL-6 antagonists including tocilizumab (*46*) (NCT04315480) and sarilumab (NCT04327388) are being investigated in clinical trials for treatment of patients with severe COVID-19.

As an illustration of the shared pathobiology and inflammatory pathways of COVID-19 with various disease manifestations (**Fig. 4**), we next turned to focus on two inflammation-driven diseases, asthma and IBD.

Inflammatory endophenotypes shared by COVID-19 and asthma

Patients with severe COVID-19 symptoms showed a higher prevalence of dyspnea (**Fig. S8A**, $P < 0.001$). To understand associations between COVID-19 and respiratory disease (including asthma), we adopted a multi-modal analysis utilizing metabolomics and transcriptomics data from two previous asthma cohorts (see Methods) under the human interactome network model. **Fig. 5A** shows the subnetwork of the connections among SARS-CoV-2 target host proteins and asthma-associated proteins. Most of these proteins have enriched connections within the subnetwork (**Fig. 5B**, blue bars) (more connections in the subnetwork than in a random network

of the same size in the human interactome, see Methods). Six overlapped proteins (orange) from both groups were identified: NFKBIA, IRAK3, TNC, IL6, ADRB2, and CD86. A recent study showed that glucose metabolism plays a key role in influenza A-regulated cytokine storm (*47*). The unique plasma metabolome of asthmatics versus healthy controls also suggests activated inflammatory and immune pathways (*48*). Therefore, in addition to PPIs, we also integrated metabolomics data generated in a previously assembled asthma cohort (*49*). By matching the enzymes of the differential metabolites and the proteins in the PPI network, we found three key metabolites, including arachidonate, L-arginine, and L-citrulline. L-arginine and L-citrulline were decreased in the sera of COVID-19 patients (*50*). A previous study showed that these metabolites were also decreased in asthma patients (*49*). Arachidonate, the precursor of a variety of products that regulate inflammatory pathways (*51*), was found to have an increased level in the inflamed airways of asthmatics (*52*). Arachidonate can be converted by 5-lipoxygenase encoded by *ALOX5* to leukotriene, which is release during an asthma attack and is responsible for the bronchoconstriction (*53*).

We further examined the DEGs from two asthma cohorts from the Severe Asthma Research Program (*54, 55*). Utilizing two bulk RNA-Seq data sets from asthma patients and healthy controls, we identified elevated expression of *IRAK3* and *ADRB2* in SARS-CoV-2 infected human bronchial epithelial cells (**Fig. 5B**). IRAK-M, encoded by *IRAK3*, regulates the toll-like receptor/interleukin (IL)-1 receptor pathway and NF-κB pathway, and *IRAK3* was identified as an asthma susceptibility gene (*56*). *ADRB2* encodes the beta 2-adrenergic receptor. The polymorphisms of *ADRB2* (p.Arg16Gly and p.Gln27Glu) increased the risk of asthma occurrence, and p.Gln27Glu was associated with asthma severity (*57*). Altogether, altered *IRAK3*

14

and *ADRB2* may explain relationships between COVID-19 and asthma, though these findings require experimental and clinically validation in patients with these disorders.

To understand expressions of the proteins in the asthma-COVID-19 network across different cell types, especially those cells which express *ACE2*, we analyzed the single-cell RNA-Seq data from bronchial epithelium (**Fig. 5C**) and lung (**Fig. 5F**) (*13*). Consistent with previous studies, *ACE2* and *TMPRSS2* have a relatively higher expression in a subtype of the secretory cells (secretory 3 cells), compared to other bronchial epithelial cell types (**Fig. 5D**, **Fig. S8C-F**). In the lung, *ACE2* and *TMPRSS2* have a relatively higher expression in AT2 cells (**Fig. 5G**, **Fig. S8H-K**). We further examined the expression of the genes in the asthma-COVID-19 network (**Fig. 5A**) in these cell types (**Fig. S9**). Several genes were also more highly expressed in secretory 3 (**Fig. 5E**, *ALOX5, IRAK3, ADRB2, TNIP1, BID, CXCL5,* and *NFKBIZ*) and in AT2 (**Fig. 5H**, *CFTR, NFKBIA, NFKBIZ,* and *TNC*), than in other cell types. *IRAK3* and *ADRB2* are among the six overlapped genes, potentially implicating *IRAK3* and *ADRB2* in COVID-19-associated asthma at the single cell level as well.

Immune pathobiology shared by COVID-19 and inflammatory bowel disease

It has been shown that human small intestine is an additional SARS-CoV-2 target organ using confocal- and electron-microscopy (*58*). Diarrhea is now well-described as an occasional presenting symptom of COVID-19 (*59*). Our network proximity analysis showed a significant association of COVID-19 and IBD across all five SARS-CoV-2 data sets (**Fig. 4A**). In addition, severe COVID-19 patients had higher risks of abdominal pain and diarrhea (**Fig. 6A**, **Fig. S10**). To understand these associations at the cellular level, we integrated network analysis and single-cell RNA-Seq analysis using publicly available data (*60*). As shown in **Fig. 6B**, although only

one IBD-associated protein, HEATR3, was found to be the target of the SARS-CoV-2 protein

Orf7a, other IBD-associated proteins showed enriched number of connections to the SARS-CoV-

2 target proteins (**Fig. 6I**, blue bars).

Using single-cell data from the ileum (distal small bowel) in Crohn's disease patients

(*60*), we found that *ACE2* and *TMPRSS2* had low to undetectable expression in the non-epithelial

cells (**Fig. 6C, 6E**, **Fig. S11A-E**). However, they showed higher expression levels in the

epithelial cells, especially absorptive enterocytes (**Fig. 6D, 6F**, **Fig. S11F-J**). We further found

that both *ACE2* and *TMPRSS2* had elevated expression levels in inflamed cells compared to

uninflamed cells in the absorptive enterocytes (**Fig. 6G**, **Fig. S11K, L**). In absorptive enterocytes

expressing *ACE2*, the expression of *ACE2* was significantly increased (**Fig. 6H**, $P = 0.016$) in the

inflamed ileal tissues of Crohn's disease patients compared to uninflamed tissues. These

observations prompted us to investigate the co-expression of the network genes in the absorptive

enterocytes (**Fig. 6I**). Several genes showed elevated co-expression with *ACE2* in inflamed cells,

including *XIAP, SMAD3, DLG5, SLC15A1, RAC1, STOM, RAB18*, and *AKAP8*.

We next turned to highlight two potential associations between COVID-19 and IBD.

First, SARS-CoV-2 protein Orf7a can directly interact with HEATR3, whose variant was shown

to be associated with increased risks of IBD by genome-wide association study (*61*). Second,

SARS-CoV-2 infection may impact RAC1 signal transduction pathways. RAC proteins play

important roles in many inflammatory pathways, and their dysregulation can be pathogenic.

Increased RAC1 expression by single nucleotide polymorphisms promotes an inflammatory

response in the colon (*62*). Mercaptopurine, an effective treatment for IBD, was found to lower

RAC1 expression in IBD patients (*63*). Since our results show that *RAC1* and *ACE2* had higher

co-expression in inflamed enterocytes (**Fig. 6I**), it is highly possible that these inflamed cells are

more susceptible to SARS-CoV-2 infection, and that the infection could lead to an altered RAC1 expression level through PPIs with virus target proteins STOM, HDAC2, POLA2, CIT, and RAP1GDS1 (**Fig. 6B**). Notably STOM was also highly co-expressed with ACE2 in inflamed cells compared to uninflamed cells (**Fig. 6I**).

Network-based drug repurposing for COVID-19

Knowledge of the complex interplays between SARS-CoV-2 targets and human diseases indicate possibilities of drug repurposing, as the drugs that target other diseases could potentially target SARS-CoV-2 through the shared functional PPI networks. In addition, drug repurposing efforts may also reveal unrecognized biological connections between diseases that existing drugs treat and COVID-19. For example, the aforementioned anti-inflammatory drugs, tocilizumab and sarilumab that are now being tested for COVID-19 are originally used for rheumatoid arthritis. Although not significant, our network proximity results show that rheumatoid arthritis have small network proximities (negative Z scores) across all five SARS-CoV-2 data sets (**Fig. 4A**). Another drug, the thiopurine mercaptopurine which has been used to treat IBD (*64*), was one of the top repurposable drugs for COVID-19 proposed in our previous work (*28*).

Therefore, we next performed network-based drug repurposing using the existing knowledge of the drug-target network and the global map of the SARS-CoV-2 interactome built in this study. The basis for the network-based drug repurposing methodologies rests on the observation that for a drug with multiple targets to be effective against a disease, its target proteins should be within or in the immediate vicinity of the corresponding subnetwork of the disease in the human interactome, as we have demonstrated in multiple diseases previously (*24, 29*). Using our state-of-the-art network proximity framework, we measured the "closest"

17

proximities of nearly 3,000 drugs and the four SARS-CoV-2 host gene/protein profiles (SARS2-DEG, SARS2-DEP, HCoV-PPI, and SARS2-PPI). Additionally, we performed gene set enrichment analysis (GSEA) using five gene/protein expression data sets, including one SARS-CoV-2 transcriptomics, one SARS-CoV-2 proteomics, one MERS-CoV and two SARS-CoV-1 transcriptomics data sets (see Methods). GSEA was used to evaluate the individual drugs for their potential to reverse the expression at the transcriptome or proteome level altered by the viruses.

In total, we computationally identified 34 drugs that were associated ($Z < -1.5$ and $P < 0.05$, permutation test) with the SARS-CoV-2 data sets (SARS2-DEG, SARS2-DEP, HCoV-PPI, and SARS2-PPI). These drugs were significantly proximal to two or more SARS-CoV-2 host protein sets (**Fig. 7A**, **Table S5**). We manually curated their reported antiviral profiles. The disease categories that these drugs have been used to treat are also shown in the figure. Ten drugs have been used to treat respiratory-related diseases, and the most common categories for these drugs are antibiotic and β2 agonist. The next most popular disease category is cardiovascular diseases, for which seven drugs were predicted. Among the 34 drugs, three drugs achieved significant network proximity with all four SARS-CoV-2 data sets investigated here. These drugs are antibiotic drug cefdinir, which is a cephalosporin for the treatment of bacterial infections (*65*); antineoplastic drug toremifene, a selective estrogen receptor modulator shows striking activities in blocking various viral infections at low micromolar levels, including Ebola virus (*66*) (50% inhibitive concentration [IC_{50}] = ~1 µM), MERS-CoV (*67*) (50% effective concentration [EC_{50}] = 12.9 µM), SARS-CoV-1 (*68*) (EC_{50} = 11.97 µM), and SARS-CoV-2 (*69*) (IC_{50} = 3.58 µM); and antihypertensive drug irbesartan, an angiotensin receptor block that can inhibit viral entry by inhibiting sodium/bile acid cotransporters (*70*).

18

Validating drug-outcome relationships on COVID-19 using patient data

We next evaluated drug-outcome relationship using a large-scale patient data from the Cleveland Clinic COVID-19 patient registry (see Methods). We used subject matter expertise based on: (i) literature-reported antiviral effects for human coronaviruses or other types of human viruses; (ii) strength of network proximity prediction and GSEA analysis (**Fig. 7A**); and (iii) availability of sufficient patient data for meaningful evaluation (exclusion of infrequently used drugs). Applying these criteria resulted in identifying melatonin, a physiologic hormone common to many living organisms.

A retrospective COVID-19 cohort analysis was conducted to validate the potential prevention effect of melatonin (**Fig. 7B**). Among a total of 18,118 patients tested for COVID-19 in the Cleveland Clinic Health System in Ohio and Florida, 1,675 patients were diagnosed with COVID-19 between March 8 and April 16, 2020 (**Table 1**). We found that melatonin usage was associated with a 50-60% reduced likelihood of a positive laboratory test result for SARS-CoV-2 (OR = 0.36, 95% CI 0.22-0.59, **Fig. 7B**) after adjusting for age, sex, race, and various disease comorbidities (diabetes, hypertension, heart disease, and pulmonary conditions, see Methods) using a propensity score (PS) matching method. Angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin II receptor blockers (ARBs) are two common types of drugs for treatment of hypertension. A recent study showed that inpatient use of ACEI/ARB was associated with lower risk of all-cause mortality compared with ACEI/ARB non-users hospitalized COVID-19 patients with hypertension (*71*). We further performed an observational study for three cohorts using user active comparator design and PS adjustment for confounding factors as described in our previous study (*24*). We found that melatonin usage was associated

19

with reduced risk of likelihood of a positive laboratory test result for SARS-CoV-2 compared to

ARBs (OR = 0.35, 95% CI 0.20-0.63) and ACEIs (OR = 0.56, 95% CI 0.31-0.99), and the

pooled data of ARBs and ACEIs (OR = 0.46, 95% CI 0.24-0.86) (**Fig. 7B**). Altogether, network-

based prediction (**Fig. 7A**) and multiple observational analyses (**Fig. 7B**) suggest potential of

treatment for COVID-19 by melatonin; yet, randomized controlled clinically trials are urgently

needed to test meaningfully the effect of melatonin for COVID-19 further.

Discussion

Recent studies indicated that SARS-CoV-2 infection was detected in multiple organs in addition

to lungs, including heart, pharynx, liver, kidneys, brain, and intestine (*58, 72*). SARS-CoV-2

RNA was also found in patient stool (*73*). Therefore, investigation of how SARS-CoV-2

associates with other diseases could help reveal and understand its impact on systems and organs

in addition to lungs. In this study, we systematically evaluated 64 diseases across 6 categories for

their potential manifestations with COVID-19. We started with assembling and characterizing

five SARS-CoV-2 data sets representing different cellular event levels including transcriptome,

proteome, and interactome. Using state-of-the-art network proximity measure, we identified

broad manifestations (such as autoimmune, neurological, and pulmonary, **Fig. 4A**) associated

with COVID-19. Retrospective meta-analyses using the clinical data of 4,973 patients across 34

studies confirmed our network-based findings. We further investigated the molecular

mechanisms of the association of COVID-19 and the comorbidities by multi-modal analyses of

large-scale single cell data. We identified the cell types that have the highest expression levels of

ACE2 and *TMPRSS2*: the lung AT2 cells, secretory bronchial epithelial cells, and absorptive

enterocytes in ileum. We examined the expression of asthma- and IBD-associated genes in their

relevant cell types. Combining these results with the results of differential expression analysis,

network connectivity, and differential metabolites, we identified several key pathogenic

pathways (including *IRAK3* and *ADRB2*) that can be altered by the viral infection for asthma.

For asthma, our network-based findings suggest several possible shared pathobiological

pathways associated with COVID-19. First, SARS-CoV-2 infection might alter the expression of

several key inflammatory genes: *IRAK3*, which is associated with asthma (*56, 74, 75*); *ADRB2*,

which is an essential genetic factor for asthma (*57, 76*); and *NFKBIA*, which showed critical

transcriptional responses in childhood asthma (*77*). These genes show high expression in

secretory 3 and AT2 cell types, suggesting a higher susceptibility to be impacted by SARS-CoV-

2 infection through ACE2. SARS-CoV-2 increased the expression of *IRAK3* and *ADRB2*, which

lead to a higher risk of asthma (**Fig. 5B**). Second, decreased levels of L-arginine and L-citrulline

were found in SARS-CoV-2 infected patients (*50*), while it has been shown that higher levels of

these metabolites are protective against asthma (*49*). L-arginine can be converted to nitric oxide

by the nitric oxide synthases, and it was shown that nitric oxide protects against viral infection

through multiple potential mechanisms (*78*).

Our network proximity results show a strong connection between IBD and COVID-19

(**Fig. 4A**). We have also showed IBD-related pathways that can potentially be affected by SARS-

CoV-2 infection (**Fig. 6B**). However, we should also note that the meta-analysis does not show a

significant risk ratio of digestive system disease in COVID-19 patients (**Fig. 3B**), while two

specific symptoms, abdominal pain and diarrhea, showed increased risks in severe COVID-19

patients (**Fig. 6A**). Digestive system disease covers a broad range of gastrointestinal diseases.

Future studies are needed to reveal and validate the associations of COVID-19 and individual gastrointestinal diseases.

Finally, we computationally prioritized nearly 3,000 FDA approved/investigational drugs for their potential anti-SARS-CoV-2 effects using network proximity measure and GSEA. A list of 34 repurposable drugs with their reported antiviral profiles are highlighted, among which five drugs are currently in clinical trials (**Fig. 7B**). We further explored drug-disease outcome for melatonin using a large-scale COVID-19 registry. We found that among individuals who received testing for SARS-CoV-2, melatonin usage was associated with a 50-60% reduced likelihood of a positive laboratory test result for SARS-CoV-2 (**Fig. 7B**). Using user active comparator design, we further discovered that melatonin usage was associated with reduced risk of likelihood of a positive laboratory test result for SARS-CoV-2 compared to ARBs and ACEIs as well. Exogenous melatonin may be of benefit in older patients with COVID-19 given aging-related reduction of endogenous melatonin and vulnerability of the elderly to lethality of SARS-CoV-2 (*79*), the latter potentially due to declining immunity, i.e. immunosenescence (*80*). Moreover, melatonin suppresses NLRP3 inflammasome activation induced by cigarette smoking and attenuates pulmonary inflammation (*81*), not only via reduction of nuclear factor kappa-β p65 (NF-κB p65) and tumor necrosis factor-α (TNF-α) expression, but also via increase in anti-inflammatory cytokines such as IL-10 or IL-6 which can also have anti-inflammatory effects (*82, 83*). Thus, large-scale observational studies and randomized controlled trials are needed to validate the clinical benefit of melatonin for patients with COVID-19. It would be important, however, that the trials are designed with the understanding of the mechanism of the drug to be repurposed. For example, it would be obvious that drugs that decrease viral entry, e.g., part of

melatonin's action, would be beneficial in preventing infection or very early in the COVID-19 course, but may be inconsequential when utilized in severe or end-stage infection.

We acknowledge several potential limitations. First, although we integrated data from multiple sources to build the human interactome and the drug target network, they are still incomplete. Second, this study relied on the SARS-CoV-2 target host genes/proteins data sets, and their quality may influence the performance of our network analysis. Third, the patient-data analysis is retrospective and may have selection bias. Although we performed multiple types of PS-matching observations, residual confounding is possible despite high-dimensional covariate adjustment. For example, carvedilol, approved for both hypertension and heart failure, was identified for potential treatment of COVID-19 by network proximity measure (**Fig. 7A**). We found that carvedilol usage reduces the likelihood of a positive laboratory test result for SARS-CoV-2 (OR = 0.43, 95% CI 0.27-0.70, **Fig. S12**) after adjusting for age, gender, race, smoking history using PS method. Yet, this association does not meet statistical significance after further adjusting for multiple disease comorbidities (including coronary artery disease, diabetes, hypertension, and COPD) (OR = 0.58, 95% CI 0.32-1.05, **Fig. S12**). There are two possible explanations: (1) the small number of carvedilol usage may lack the statistical power for current observational study; (2) carvedilol may only reduce the likelihood of a positive laboratory test result for SARS-CoV-2 for patients having existing health conditions, such as hypertension. Replication of the associations and causal inference using large databases may rule out treatment effect heterogeneity and possible confounding further. Finally, although we made the intriguing observation that the use of melatonin was much less prevalent in individuals testing positive for SARS-CoV-2, we recognize that many asymptomatic or minimally symptomatic persons with the virus were not tested and therefore, their use of melatonin was not evaluable; melatonin use

23

in this latter group might also have been high. It should therefore be noted that all drugs we identified as therapeutic candidates in this study must be validated using experimental assays and randomized clinical trials before they can be recommended for use in patients with COVID-19.

In conclusion, our study provides a powerful, integrative, network medicine strategy for advancing understanding of COVID-19-associated comorbidities and facilitating the identification of drug candidates for COVID-19. This approach also promises to address the translational gap between genomic studies and clinical outcomes, which poses a significant problem when rapid development of effective therapeutic interventions is critical during a pandemic. From a translational perspective, if broadly applied, the network-medicine tools applied here could prove helpful in developing effective treatment strategies for other complex human diseases as well, including other emerging infectious diseases.

Methods and Materials

Building the data sets of SARS-CoV-2 target host genes/proteins

We assembled four SARS-CoV-2 data sets of target host genes/proteins: (1) 246 differentially expressed genes in human bronchial epithelial cells infected with SARS-CoV-2 (*18*) (GSE147507), denoted as SARS2-DEG; (2) 293 differentially expressed proteins in human Caco-2 cells infected with SARS-CoV-2 (*19*), denoted as SARS2-DEP; (3) 134 strong literature evidence-based pan-human coronavirus target host proteins from our recent study (*28*) with 15 newly curated proteins, denoted as HCoV-PPI; (4) 332 proteins involved in the protein-protein interactions with 26 SARS-CoV-2 viral proteins identified by affinity purification-mass spectrometry (*20*), denoted as SARS2-PPI. Finally, due to the interactome nature of HCoV-PPI

24

and SARS2-PPI, we combined these data sets as the fifth SARS-CoV-2 data set, which has 460

proteins and is denoted as PanCoV-PPI.

Building the disease gene profiles

We compiled the disease-associated gene sets from various sources. All databases were accessed

on March 26, 2020.

Cancer. The pan-cancer driver genes were retrieved from the Cancer Gene Census (*35*). Driver

genes for individual cancer types were from a previous study (*36*).

Mendelian disease genes (MDGs). A set of 2,272 MDGs were retrieved from the Online

Mendelian Inheritance in Man (OMIM) database (*84*).

Orphan disease-causing mutant genes (ODMGs). A set of 2,124 ODMGs were retrieved from a

previous study (*85*).

Cell cycle genes. 910 human cell cycle genes were downloaded from a previous study identified

by a genome-wide RNAi screening (*86*).

Innate immune genes. 1,031 human innate immunity genes were collected form InnateDB (*87*).

Genes associated with autoimmune, pulmonary, neurological, cardiovascular, and metabolic

diseases. The disease-associated genes/proteins were extracted from the Human Gene Mutation

Database (HGMD) (*37*). A systematic search was manually performed for each disease using

keywords such as disease name, alias, and symptoms to identify the disease terms in the HGMD.

All disease terms were then verified for their relevance.

Functional enrichment analysis

We performed Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO)

biological process enrichment analyses to reveal the biological relevance and functional

pathways of the five SARS-CoV-2 data sets. All functional enrichment analyses were performed

using Enrichr (*88*).

Selective pressure and evolutionary rates characterization

We calculated the *dN/dS* ratio (*89*) and the evolutionary rate ratio (*90*) as described in our

previous study (*91*).

Tissue specificity analysis

The RNA-Seq data (transcripts per million, TPM) of 31 tissues from the GTEx V8 release

(accessed on March 31, 2020, https://www.gtexportal.org/home/) were downloaded. Genes with

count per million (CPM) ≥ 0.5 in over 90% samples in a tissue were considered tissue-expressed

genes and otherwise as tissue-unexpressed. To quantify the expression specificity of gene i in

tissue t, we calculated the mean expression E_i and the standard deviation σ_i of a gene's

expression across all considered tissues. The significance of gene expression specificity in a

tissue is defined as:

$$z_{it} = \frac{E_{it} - E_i}{\sigma_i} \qquad (1)$$

Risk ratio analysis for COVID-19 patients

PubMed, Embase and Medrxiv databases were searched for publications as of April 25th, 2020

(**Fig. S13**). The search was limited to English articles describing the demographic and clinical

features of SARS-CoV-2 cases. We used the search term ("SARS-COV-2" OR "COVID-19" OR

"nCoV 19" OR "2019 novel coronavirus" OR "coronavirus disease 2019") AND ("clinical characteristics" OR "clinical outcome" OR "comorbidities"). Only research articles were included; reviews, case reports, comments, editorials, and expert opinions were excluded. Three criteria were used to select studies from a total of 1,054 initial hits: (1) studies that have ≥ 20 COVID-19 patients; (2) studies that grouped the outcomes with varying degree of severity of COVID-19 (e.g., severe vs. non-severe) according to the American Thoracic Society guidelines for community-acquired pneumonia; and (3) studies that were from different institutions. Two criteria were used for exclusion: (1) studies that focus on specific populations (e.g., only death cases, pregnant women, children or from family clusters); and (2) basic molecular biology research. Finally, 34 studies meeting these criteria were used for further analyses.

We performed random effects meta-analysis to estimate the pooled risk ratio with 95% CI of 10 comorbidities between severe versus non-severe COVID-19 patients. Mantel-Haenszel method was used to estimate the pooled effects of results (*92*). DerSimonian-Laird method was used to estimate the variance among studies (*93*). The continuous data such as IL-6 levels were transformed to mean and standard deviations first using Wan's approach based on sample size, median and interquartile range (*94*). Next, we used inverse variance method to estimate the pooled mean difference and estimated the variance among studies by DerSimonian-Laird method. We estimated the pooled prevalence of 3 COVID-19 symptoms (abdominal pain, diarrhea, dyspnea) and 1 comorbidity (COPD) in three COVID-19 patient groups (severe, non-severe, and all) respectively. Random intercept logistic regression model was used to estimate pooled prevalence and maximum-likelihood estimator was used to quantify the heterogeneity of studies (*95*). The *tau²* and I^2 statistics were calculated for the heterogeneity among studies. We considered $I^2 < 50\%$ as low heterogeneity among studies, $50\% < I^2 < 75\%$ as moderate

27

heterogeneity, and $I^2 > 75\%$ as high heterogeneity. All meta-analyses were conducted by meta and dmetar packages in the R v3.6.3 platform.

Building the human protein-protein interactome

A total of 18 bioinformatics and systems biology databases were assembled to build a comprehensive list of human PPIs with five types of experimental evidences: (1) binary, physical PPIs from protein three-dimensional (3D) structures; (2) binary PPIs tested by high-throughput yeast-two-hybrid (Y2H) systems; (3) kinase-substrate interactions by literature-derived low-throughput or high-throughput experiments; (4) signaling network by literature-derived low-throughput experiments; and (5) literature-curated PPIs identified by affinity purification followed by mass spectrometry (AP-MS), Y2H, or by literature-derived low-throughput experiments. Inferred PPIs based on gene expression data, evolutionary analysis, and metabolic associations were excluded. Genes were mapped to their Entrez ID based on the NCBI database (*96*). The official gene symbols were based on GeneCards (https://www.genecards.org/). The final human protein-protein interactome used in this study included 351,444 unique PPIs (edges or links) connecting 17,706 proteins (nodes). Detailed descriptions for building the human protein-protein interactome are provided in our previous studies (*24, 29, 30, 97*).

Network proximity measure

We used the "closest" network proximity measure throughout this study. For two gene/protein sets A and B, their "closest" distance d_{AB} was calculated as:

$$\langle d_{AB} \rangle = \frac{1}{\|A\| + \|B\|} \left(\sum_{a \in A} min_{b \in B} \, d(a, b) + \sum_{b \in B} min_{a \in A} \, d(a, b) \right) \qquad (2)$$

28

where $d(a, b)$ is the shortest distance of a and b in the human interactome. To evaluate the significance, we performed a permutation test using randomly selected proteins from the whole interactome that were representative of the two protein sets being evaluated in terms of their degree distributions. We then calculated the Z score as:

$$Z_{d_{AB}} = \frac{d_{AB} - \overline{d_r}}{\sigma_r} \qquad (3)$$

where $\overline{d_r}$ and σ_r were the mean and standard deviation of the permutation test. All network proximity permutation tests in this study were repeated 1,000 times.

Network-based comorbidity analysis

To reveal potential COVID-19 comorbidities, we computed the network proximity of the disease-associated proteins for each disease and the five SARS-CoV-2 data sets. SARS-CoV-2 target proteins with a non-negative tissue specificity in lung were used in the computation. The degree enrichment for protein i in a subnetwork was calculated as:

$$e_i = \frac{d_i/n}{D_i/N} \qquad (4)$$

where d_i is the degree of i in the subnetwork, n is number of nodes in the subnetwork, D_i is the degree in the complete human protein interactome, and N is the total number of nodes in the interactome.

Bulk and single-cell RNA-Seq data analysis

Bulk RNA-Seq data sets for asthma patients were retrieved from the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/) using the accession number GSE63142 (*54*) and GSE130499 (*55*). Differential expression of three comparisons, severe vs. control, mild vs.

control, and severe vs. mild were performed using the GEO2R function. All single-cell data

analyses and visualizations were performed with the R package Seurat v3.1.4 (*98*). GSE134809

(*60*) was downloaded from the NCBI GEO database. This data set contains 67,050 inflamed and

uninflamed cells from the ileal samples of 8 patients with Crohn's disease. Single-cell data of

normal lung and primary human bronchial epithelial cells were downloaded from

https://data.mendeley.com/datasets/7r2cwbw44m/1. These data sets contain 39,778 lung cells

and 17,451 bronchial epithelial cells. Qualifying cells based on the criteria from the original

paper were used for the analysis. "NormalizeData" was used to normalize the data.

"FindIntegrationAnchors" and "IntegrateData" functions were used to integrate cells from

different samples. UMAP was used as the dimension reduction method for visualization. Cell

type markers from the original paper were used to label the cell types.

Building the drug-target network

To evaluate whether a drug is closely associated with SARS-CoV-2 target proteins in the human

interactome, we gathered the drug-target interaction information from several databases:

DrugBank database (v4.3) (*99*), Therapeutic Target Database (TTD) (*100*), PharmGKB database,

ChEMBL (v20) (*101*), BindingDB (*102*), and IUPHAR/BPS Guide to PHARMACOLOGY

(*103*). We included the interactions that have binding affinities Ki, Kd, IC_{50} or $EC_{50} \leq 10$ μM and

a unique UniProt accession number with "reviewed" status. The details for building the

experimentally validated drug-target network can be found in our recent studies (*24, 29, 30*).

Network-based drug repurposing

We computed the "closest" network proximity as described before for 2,938 FDA approved or investigational drugs and the five SARS-CoV-2 data sets. For prioritization, we ranked the drugs by their distance to the data sets (D < 2) and Z score (Z < -1.5) from the network proximity analysis. The antiviral profiles of the highlighted drugs were manually curated. COVID-19 related clinical trials were retrieved on May 10, 2020.

Gene set enrichment analysis (GSEA)

The gene set enrichment analysis was conducted as described in our recent work (*28*) as an additional evidence for drug repurposing. Briefly, for each drug and coronavirus target gene set, we computed an enrichment score to indicate whether a drug can reverse the effect of SARS-CoV-2 at the transcriptome or proteome level. Gene expression profiles for the drugs were retrieved from the Connectivity Map (CMAP) database (*104*). Five gene sets were evaluated: (1) the differentially expressed genes in human bronchial epithelial cells infected with SARS-CoV-2 (*18*) (GSE147507); (2) the differentially expressed proteins in human Caco-2 cells infected with SARS-CoV-2 (*19*); (3, 4) two transcriptome data sets of SARS-CoV-1 infected samples from patient's peripheral blood (*105*) (GSE1739) and Calu-3 cells (*106*) (GSE33267) respectively; (5) the differentially expressed genes in MERS-CoV infected Calu-3 cells (*107*) (GSE122876).

The enrichment score *ES* was calculated for up- and down- regulated genes separately first. The overall *ES* was calculated as:

$$ES = \begin{cases} ES_{up} - ES_{down}, & sgn(ES_{up}) \neq sgn(ES_{down}) \\ 0, & else \end{cases} \qquad (5)$$

where ES_{up} and ES_{down} were calculated using $a_{up/down}$ and $b_{up/down}$ as:

$$a = \max_{1 \leq j \leq s} \left(\frac{j}{s} - \frac{V(j)}{r} \right) \qquad (6)$$

31

$$b = \max_{1 \le j \le s}\left(\frac{V(j)}{r} - \frac{j-1}{s}\right) \qquad (7)$$

$j = 1, 2, \cdots, s$ were the genes in the gene sets sorted in ascending order by their rank in the drug

profiles. $V(j)$ indicates the rank of gene j, where $1 \le V(j) \le r$, with r being the number of

genes from the drug profile. Then, $ES_{up/down}$ was set to $a_{up/down}$ if $a_{up/down} > b_{up/down}$, and

was set to $-b_{up/down}$ if $b_{up/down} > a_{up/down}$. Permutation test was performed to evaluate the

significance. Drugs were considered to have potential treatment effect if $ES > 0$ and $P < 0.05$.

Patient data validation of the network-identified drugs using a COVID-19 registry

We used the institutional review board–approved registry COVID-19 registry data, including

18,118 individuals (1,675 positive) tested during March 8 to April 16, 2020 from the Cleveland

Clinic Health System in Ohio and Florida. Data include COVID-19 test results, baseline

demographic information, medications, and all recorded disease conditions and others. We

conducted a series of retrospective case-control studies with an active comparator new user

design to test the drug-outcome relationships for COVID-19. Data were extracted from

electronic health records (EPIC Systems) and were manually checked by a study team trained on

uniform sources for the study variables. We collected and managed all patient data using

REDCap electronic data capture tools. The exposures of drugs (including carvedilol and

melatonin) were used as recorded in the medication list in the electronic medical records at the

time of testing for SARS-CoV-2. A positive laboratory test result for COVID-19 was used as the

primary outcome. A propensity score (PS) was calculated for melatonin and carvedilol intake by

multivariable logistic regression models with two sets of covariates: (1) PS-matched model1,

age, gender, race and smoking; (2) PS-matched model2, age, gender, race, smoking, coronary

artery disease, diabetes, hypertension and COPD. The nearest-neighbor algorithm was used to

32

match the melatonin and carvedilol subjects and references subjects on the basis of PS ratio of 1:4. Two odds ratios (OR) models were used to estimate the COVID-19 positive risk between melatonin or carvedilol intaking groups and non-intaking groups: (1) OR model1, COVID-19 ~ melatonin/carvedilol; (2) OR model2 (adjusted OR model), COVID-19 ~ melatonin/carvedilol + age, gender, race, smoking, coronary artery disease, diabetes, hypertension and COPD. All analyses were conducted by matchit package in the R v3.6.3 platform.

Statistical analysis and network visualization

Statistical tests were performed with the Python package SciPy v1.3.0 (https://www.scipy.org/). $P < 0.05$ was considered statistically significant throughout this study. Networks were visualized using Gephi 0.9.2 (https://gephi.org/).

Acknowledgements

A.K., D.A.C., S.F., S.C., C.F., R.M., M.U.G., T.C., T.S., C.E., J.U.J., L.J., and S.E. discussed

and interpreted all results. Y.Z., Y.H., S.C., and F.C. wrote and critically revised the manuscript

with contributions from other co-authors.

Competing interests: The authors declare that they have no conflict of interest. The content of

this publication does not necessarily reflect the views of the Cleveland Clinic.

Data and code availability: The bulk and single-cell RNA-Seq data used in this study were

downloaded from the NCBI GEO database with accession numbers GSE63142, GSE130499, and

GSE134809. The lung and human bronchial epithelial single-cell data were downloaded from

https://data.mendeley.com/datasets/7r2cwbw44m/1. Source code and other supporting data can

be downloaded from https://github.com/ChengF-Lab/COVID-19_Map.

References

1.   P. K. Bhatraju, B. J. Ghassemieh, M. Nichols, R. Kim, K. R. Jerome, A. K. Nalla, A. L. Greninger, S. Pipavath, M. M. Wurfel, L. Evans, P. A. Kritek, T. E. West, A. Luks, A. Gerbino, C. R. Dale, J. D. Goldman, S. O'Mahony, C. Mikacenic, Covid-19 in Critically Ill Patients in the Seattle Region - Case Series. *The New England journal of medicine*, (2020).

2.   W. J. Guan, Z. Y. Ni, Y. Hu, W. H. Liang, C. Q. Ou, J. X. He, L. Liu, H. Shan, C. L. Lei, D. S. C. Hui, B. Du, L. J. Li, G. Zeng, K. Y. Yuen, R. C. Chen, C. L. Tang, T. Wang, P. Y. Chen, J. Xiang, S. Y. Li, J. L. Wang, Z. J. Liang, Y. X. Peng, L. Wei, Y. Liu, Y. H. Hu, P. Peng, J. M. Wang, J. Y. Liu, Z. Chen, G. Li, Z. J. Zheng, S. Q. Qiu, J. Luo, C. J. Ye, S. Y. Zhu, N. S. Zhong, Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England journal of medicine*, (2020).

3.   D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *Jama*, (2020).

4.   X. Yang, Y. Yu, J. Xu, H. Shu, J. Xia, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu, Y. Wang, S. Pan, X. Zou, S. Yuan, Y. Shang, Clinical course and outcomes of critically

ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet. Respiratory medicine*, (2020).

5.   F. Zhou, T. Yu, R. Du, G. Fan, Y. Liu, Z. Liu, J. Xiang, Y. Wang, B. Song, X. Gu, L. Guan, Y. Wei, H. Li, X. Wu, J. Xu, S. Tu, Y. Zhang, H. Chen, B. Cao, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet (London, England)* **395**, 1054-1062 (2020).

6.   N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J. Xia, T. Yu, X. Zhang, L. Zhang, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet (London, England)* **395**, 507-513 (2020).

7.   P. Zhou, X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, H. R. Si, Y. Zhu, B. Li, C. L. Huang, H. D. Chen, J. Chen, Y. Luo, H. Guo, R. D. Jiang, M. Q. Liu, Y. Chen, X. R. Shen, X. Wang, X. S. Zheng, K. Zhao, Q. J. Chen, F. Deng, L. L. Liu, B. Yan, F. X. Zhan, Y. Y. Wang, G. F. Xiao, Z. L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).

8.   N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan, I. China Novel Coronavirus, T. Research, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *The New England journal of medicine* **382**, 727-733 (2020).

9.   M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M. A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e278 (2020).

10.   M. Hoffmann, H. Kleine-Weber, S. Pöhlmann, A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Molecular Cell*, (2020).

11.   C. G. K. Ziegler, S. J. Allon, S. K. Nyquist, I. M. Mbano, V. N. Miao, C. N. Tzouanas, Y. Cao, A. S. Yousif, J. Bals, B. M. Hauser, J. Feldman, C. Muus, M. H. Wadsworth, S. W. Kazer, T. K. Hughes, B. Doran, G. J. Gatter, M. Vukovic, F. Taliaferro, B. E. Mead, Z. Guo, J. P. Wang, D. Gras, M. Plaisant, M. Ansari, I. Angelidis, H. Adler, J. M. S. Sucre, C. J. Taylor, B. Lin, A. Waghray, V. Mitsialis, D. F. Dwyer, K. M. Buchheit, J. A. Boyce, N. A. Barrett, T. M. Laidlaw, S. L. Carroll, L. Colonna, V. Tkachev, C. W. Peterson, A. Yu, H. B. Zheng, H. P. Gideon, C. G. Winchell, P. L. Lin, C. D. Bingle, S. B. Snapper, J. A. Kropski, F. J. Theis, H. B. Schiller, L.-E. Zaragosi, P. Barbry, A. Leslie, H.-P. Kiem, J. L. Flynn, S. M. Fortune, B. Berger, R. W. Finberg, L. S. Kean, M. Garber, A. G. Schmidt, D. Lingwood, A. K. Shalek, J. Ordovas-Montanes, SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell*, (2020).

12.   W. Sungnak, N. Huang, C. Bécavin, M. Berg, R. Queen, M. Litvinukova, C. Talavera-López, H. Maatz, D. Reichart, F. Sampaziotis, K. B. Worlock, M. Yoshida, J. L. Barnes,

35

N. E. Banovich, P. Barbry, A. Brazma, J. Collin, T. J. Desai, T. E. Duong, O. Eickelberg, C. Falk, M. Farzan, I. Glass, R. K. Gupta, M. Haniffa, P. Horvath, N. Hubner, D. Hung, N. Kaminski, M. Krasnow, J. A. Kropski, M. Kuhnemund, M. Lako, H. Lee, S. Leroy, S. Linnarson, J. Lundeberg, K. B. Meyer, Z. Miao, A. V. Misharin, M. C. Nawijn, M. Z. Nikolic, M. Noseda, J. Ordovas-Montanes, G. Y. Oudit, D. Pe'er, J. Powell, S. Quake, J. Rajagopal, P. R. Tata, E. L. Rawlins, A. Regev, P. A. Reyfman, O. Rozenblatt-Rosen, K. Saeb-Parsy, C. Samakovlis, H. B. Schiller, J. L. Schultze, M. A. Seibold, C. E. Seidman, J. G. Seidman, A. K. Shalek, D. Shepherd, J. Spence, A. Spira, X. Sun, S. A. Teichmann, F. J. Theis, A. M. Tsankov, L. Vallier, M. van den Berge, J. Whitsett, R. Xavier, Y. Xu, L.-E. Zaragosi, D. Zerti, H. Zhang, K. Zhang, M. Rojas, F. Figueiredo, H. C. A. L. B. Network, SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nature Medicine*, (2020).

13. S. Lukassen, R. L. Chua, T. Trefzer, N. C. Kahn, M. A. Schneider, T. Muley, H. Winter, M. Meister, C. Veith, A. W. Boots, B. P. Hennig, M. Kreuter, C. Conrad, R. Eils, SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J*, e105114 (2020).

14. H. Zhang, Z. Kang, H. Gong, D. Xu, J. Wang, Z. Li, Z. Li, X. Cui, J. Xiao, J. Zhan, T. Meng, W. Zhou, J. Liu, H. Xu, Digestive system is a potential route of COVID-19: an analysis of single-cell coexpression pattern of key proteins in viral entry process. *Gut*, gutjnl-2020-320953 (2020).

15. S. Yang, C. Fu, X. Lian, X. Dong, Z. Zhang, Understanding Human-Virus Protein-Protein Interactions Using a Human Protein Complex-Based Analysis Framework. *mSystems* **4**, (2019).

16. C. Liu, Y. Ma, J. Zhao, R. Nussinov, Y. Zhang, F. Cheng, Z. Zhang, Computational network biology: Data, models, and applications. DOI: 10.1016/j.physrep.2019.12.004. *Physics Reports*, (2020).

17. A. Zumla, J. F. Chan, E. I. Azhar, D. S. Hui, K. Y. Yuen, Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* **15**, 327-347 (2016).

18. D. Blanco-Melo, B. E. Nilsson-Payant, W.-C. Liu, R. Møller, M. Panis, D. Sachs, R. A. Albrecht, B. R. tenOever, SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. *bioRxiv*, 2020.2003.2024.004655 (2020).

19. D. Bojkova, K. Klann, B. Koch, M. Widera, D. Krause, S. Ciesek, J. Cinatl, C. Munch, Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature*, (2020).

20. D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. O'Meara, V. V. Rezelj, J. Z. Guo, D. L. Swaney, T. A. Tummino, R. Huettenhain, R. M. Kaake, A. L. Richards, B. Tutuncuoglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B. J. Polacco, H. Braberg, J. M. Fabius, M. Eckhardt, M. Soucheray, M. J. Bennett, M. Cakir, M. J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z. Z. C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I. T. Kirby, J. E. Melnyk, J. S. Chorba, K. Lou, S. A. Dai, I. Barrio-Hernandez, D. Memon, C.

Hernandez-Armenta, J. Lyu, C. J. P. Mathy, T. Perica, K. B. Pilla, S. J. Ganesan, D. J. Saltzberg, R. Rakesh, X. Liu, S. B. Rosenthal, L. Calviello, S. Venkataramanan, J. Liboy-Lugo, Y. Lin, X. P. Huang, Y. Liu, S. A. Wankowicz, M. Bohn, M. Safari, F. S. Ugur, C. Koh, N. S. Savar, Q. D. Tran, D. Shengjuler, S. J. Fletcher, M. C. O'Neal, Y. Cai, J. C. J. Chang, D. J. Broadhurst, S. Klippsten, P. P. Sharp, N. A. Wenzell, D. Kuzuoglu, H. Y. Wang, R. Trenker, J. M. Young, D. A. Cavero, J. Hiatt, T. L. Roth, U. Rathore, A. Subramanian, J. Noack, M. Hubert, R. M. Stroud, A. D. Frankel, O. S. Rosenberg, K. A. Verba, D. A. Agard, M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdin, A. Ashworth, O. Schwartz, C. d'Enfert, S. Mukherjee, M. Jacobson, H. S. Malik, D. G. Fujimori, T. Ideker, C. S. Craik, S. N. Floor, J. S. Fraser, J. D. Gross, A. Sali, B. L. Roth, D. Ruggero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. Garcia-Sastre, K. M. Shokat, B. K. Shoichet, N. J. Krogan, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, doi: 10.1038/s41586-020-2286-9 (2020).

21. F. Cheng, In Silico Oncology Drug Repositioning and Polypharmacology. *Methods Mol Biol* **1878**, 243-261 (2019).

22. F. Cheng, H. Hong, S. Yang, Y. Wei, Individualized network-based drug repositioning infrastructure for precision oncology in the panomics era. *Brief Bioinform* **18**, 682-697 (2017).

23. F. Cheng, J. L. Murray, D. H. Rubin, Drug Repurposing: New Treatments for Zika Virus Infection? *Trends Mol Med* **22**, 919-921 (2016).

24. F. Cheng, R. J. Desai, D. E. Handy, R. Wang, S. Schneeweiss, A. L. Barabasi, J. Loscalzo, Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun* **9**, 2691 (2018).

25. F. Cheng, J. L. Murray, J. Zhao, J. Sheng, Z. Zhao, D. H. Rubin, Systems Biology-Based Investigation of Cellular Antiviral Drug Targets Identified by Gene-Trap Insertional Mutagenesis. *PLoS Comput Biol* **12**, e1005074 (2016).

26. B. Cao, Y. Wang, D. Wen, W. Liu, J. Wang, G. Fan, L. Ruan, B. Song, Y. Cai, M. Wei, X. Li, J. Xia, N. Chen, J. Xiang, T. Yu, T. Bai, X. Xie, L. Zhang, C. Li, Y. Yuan, H. Chen, H. Li, H. Huang, S. Tu, F. Gong, Y. Liu, Y. Wei, C. Dong, F. Zhou, X. Gu, J. Xu, Z. Liu, Y. Zhang, H. Li, L. Shang, K. Wang, K. Li, X. Zhou, X. Dong, Z. Qu, S. Lu, X. Hu, S. Ruan, S. Luo, J. Wu, L. Peng, F. Cheng, L. Pan, J. Zou, C. Jia, J. Wang, X. Liu, S. Wang, X. Wu, Q. Ge, J. He, H. Zhan, F. Qiu, L. Guo, C. Huang, T. Jaki, F. G. Hayden, P. W. Horby, D. Zhang, C. Wang, A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. *New England Journal of Medicine*, (2020).

27. L. D. CHEN Jun, LIU Li,LIU Ping,XU Qingnian,XIA Lu,LING Yun,HUANG Dan,SONG Shuli,ZHANG Dandan,QIAN Zhiping,LI Tao,SHEN Yinzhong,LU Hongzhou, A pilot study of hydroxychloroquine in treatment of patients with common coronavirus disease-19 (COVID-19). *J Zhejiang Univ (Med Sci)* **49**, 0-0 (2020).

28. Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, F. Cheng, Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov* **6**, 14 (2020).

29. F. Cheng, W. Lu, C. Liu, J. Fang, Y. Hou, D. E. Handy, R. Wang, Y. Zhao, Y. Yang, J. Huang, D. E. Hill, M. Vidal, C. Eng, J. Loscalzo, A genome-wide positioning systems network algorithm for in silico drug repurposing. *Nat Commun* **10**, 3476 (2019).

30. F. Cheng, I. A. Kovacs, A. L. Barabasi, Network-based prediction of drug combinations. *Nat Commun* **10**, 1197 (2019).

31. D. M. Gysi, Í. D. Valle, M. Zitnik, A. Ameli, X. Gan, O. Varol, H. Sanchez, R. M. Baron, D. Ghiassian, J. Loscalzo, A.-L. Barabási, Network Medicine Framework for Identifying Drug Repurposing Opportunities for COVID-19. *arXiv*, 2004.07229 (2020).

32. G. Li, X. He, L. Zhang, Q. Ran, J. Wang, A. Xiong, D. Wu, F. Chen, J. Sun, C. Chang, Assessing ACE2 expression patterns in lung tissues in the pathogenesis of COVID-19. *J Autoimmun*, 102463 (2020).

33. H. Shimokawa, S. Sunamura, K. Satoh, RhoA/Rho-Kinase in the Cardiovascular System. *Circ Res* **118**, 352-366 (2016).

34. K. Zaoui, Z. Boudhraa, P. Khalife, E. Carmona, D. Provencher, A. M. Mes-Masson, Ran promotes membrane targeting and stabilization of RhoA to orchestrate ovarian cancer cell invasion. *Nat Commun* **10**, 2666 (2019).

35. S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, P. A. Futreal, COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* **39**, D945-950 (2011).

36. C. Liu, J. Zhao, W. Lu, Y. Dai, J. Hockings, Y. Zhou, R. Nussinov, C. Eng, F. Cheng, Individualized genetic network analysis reveals new therapeutic vulnerabilities in 6,700 cancer genomes. *PLoS Comput Biol* **16**, e1007701 (2020).

37. P. D. Stenson, E. V. Ball, M. Mort, A. D. Phillips, J. A. Shiel, N. S. Thomas, S. Abeysinghe, M. Krawczak, D. N. Cooper, Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat* **21**, 577-581 (2003).

38. K. J. Clerkin, J. A. Fried, J. Raikhelkar, G. Sayer, J. M. Griffin, A. Masoumi, S. S. Jain, D. Burkhoff, D. Kumaraiah, L. Rabbani, A. Schwartz, N. Uriel, Coronavirus Disease 2019 (COVID-19) and Cardiovascular Disease. *Circulation*, (2020).

39. E. Baldi, G. M. Sechi, C. Mare, F. Canevari, A. Brancaglione, R. Primi, C. Klersy, A. Palo, E. Contri, V. Ronchi, G. Beretta, F. Reali, P. Parogni, F. Facchin, D. Bua, U. Rizzi, D. Bussi, S. Ruggeri, L. Oltrona Visconti, S. Savastano, Out-of-Hospital Cardiac Arrest during the Covid-19 Outbreak in Italy. *New England Journal of Medicine*, (2020).

40. M. F. Beers, S. Mulugeta, The biology of the ABCA3 lipid transporter in lung health and disease. *Cell Tissue Res* **367**, 481-493 (2017).

41. L. M. Nogee, Abnormal expression of surfactant protein C and lung disease. *Am J Respir Cell Mol Biol* **26**, 641-644 (2002).

42. M. Z. Tay, C. M. Poh, L. Rénia, P. A. MacAry, L. F. P. Ng, The trinity of COVID-19: immunity, inflammation and intervention. *Nature Reviews Immunology*, (2020).

43.   D. Swaroopa, K. Bhaskar, T. Mahathi, S. Katkam, Y. S. Raju, N. Chandra, V. K. Kutala, Association of serum interleukin-6, interleukin-8, and Acute Physiology and Chronic Health Evaluation II score with clinical outcome in patients with acute respiratory distress syndrome. *Indian J Crit Care Med* **20**, 518-525 (2016).

44.   T. Hou, D. Huang, R. Zeng, Z. Ye, Y. Zhang, Accuracy of serum interleukin (IL)-6 in sepsis diagnosis: a systematic review and meta-analysis. *Int J Clin Exp Med* **8**, 15238-15245 (2015).

45.   J. Song, D. W. Park, S. Moon, H.-J. Cho, J. H. Park, H. Seok, W. S. Choi, Diagnostic and prognostic value of interleukin-6, pentraxin 3, and procalcitonin levels among sepsis and septic shock patients: a prospective controlled study according to the Sepsis-3 definitions. *BMC Infectious Diseases* **19**, 968 (2019).

46.   X. Xu, M. Han, T. Li, W. Sun, D. Wang, B. Fu, Y. Zhou, X. Zheng, Y. Yang, X. Li, X. Zhang, A. Pan, H. Wei, Effective treatment of severe COVID-19 patients with tocilizumab. *Proceedings of the National Academy of Sciences*, 202005615 (2020).

47.   Q. Wang, P. Fang, R. He, M. Li, H. Yu, L. Zhou, Y. Yi, F. Wang, Y. Rong, Y. Zhang, A. Chen, N. Peng, Y. Lin, M. Lu, Y. Zhu, G. Peng, L. Rao, S. Liu, O-GlcNAc transferase promotes influenza A virus–induced cytokine storm by targeting interferon regulatory factor–5. *Science Advances* **6**, eaaz7086 (2020).

48.   S. A. Comhair, J. McDunn, C. Bennett, J. Fettig, S. C. Erzurum, S. C. Kalhan, Metabolomic Endotype of Asthma. *J Immunol* **195**, 643-650 (2015).

49.   F. Holguin, H. Grasemann, S. Sharma, D. Winnica, K. Wasil, V. Smith, M. H. Cruse, N. Perez, E. Coleman, T. J. Scialla, L. G. Que, L-Citrulline increases nitric oxide and improves control in obese asthmatics. *JCI Insight* **4**, (2019).

50.   B. Shen, X. Yi, Y. Sun, X. Bi, J. Du, C. Zhang, S. Quan, F. Zhang, R. Sun, L. Qian, W. Ge, W. Liu, S. Liang, H. Chen, Y. Zhang, J. Li, J. Xu, Z. He, B. Chen, J. Wang, H. Yan, Y. Zheng, D. Wang, J. Zhu, Z. Kong, Z. Kang, X. Liang, X. Ding, G. Ruan, N. Xiang, X. Cai, H. Gao, L. Li, S. Li, Q. Xiao, T. Lu, Y. J. Zhu, H. Liu, H. Chen, T. Guo, Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *medRxiv*, 2020.2004.2007.20054585 (2020).

51.   B. Samuelsson, Arachidonic acid metabolism: role in inflammation. *Z Rheumatol* **50 Suppl 1**, 3-6 (1991).

52.   C. Calabrese, M. Triggiani, G. Marone, G. Mazzarella, Arachidonic acid metabolism in inflammatory cells of patients with bronchial asthma. *Allergy* **55 Suppl 61**, 27-30 (2000).

53.   A. Berger, What are leukotrienes and how do they work in asthma? *BMJ* **319**, 90 (1999).

54.   B. D. Modena, J. R. Tedrow, J. Milosevic, E. R. Bleecker, D. A. Meyers, W. Wu, Z. Bar-Joseph, S. C. Erzurum, B. M. Gaston, W. W. Busse, N. N. Jarjour, N. Kaminski, S. E. Wenzel, Gene expression in relation to exhaled nitric oxide identifies novel asthma phenotypes with unique biomolecular pathways. *Am J Respir Crit Care Med* **190**, 1363-1372 (2014).

55. N. Weathington, M. E. O'Brien, J. Radder, T. C. Whisenant, E. R. Bleecker, W. W. Busse, S. C. Erzurum, B. Gaston, A. T. Hastie, N. N. Jarjour, D. A. Meyers, J. Milosevic, W. C. Moore, J. R. Tedrow, J. B. Trudeau, H. P. Wong, W. Wu, N. Kaminski, S. E. Wenzel, B. D. Modena, BAL Cell Gene Expression in Severe Asthma Reveals Mechanisms of Severe Disease and Influences of Medications. *Am J Respir Crit Care Med* **200**, 837-856 (2019).

56. L. Balaci, M. C. Spada, N. Olla, G. Sole, L. Loddo, F. Anedda, S. Naitza, M. A. Zuncheddu, A. Maschio, D. Altea, M. Uda, S. Pilia, S. Sanna, M. Masala, L. Crisponi, M. Fattori, M. Devoto, S. Doratiotto, S. Rassu, S. Mereu, E. Giua, N. G. Cadeddu, R. Atzeni, U. Pelosi, A. Corrias, R. Perra, P. L. Torrazza, P. Pirina, F. Ginesu, S. Marcias, M. G. Schintu, G. S. D. Giacco, P. E. Manconi, G. Malerba, A. Bisognin, E. Trabetti, A. Boner, L. Pescollderungg, P. F. Pignatti, D. Schlessinger, A. Cao, G. Pilia, IRAK-M Is Involved in the Pathogenesis of Early-Onset Persistent Asthma. *The American Journal of Human Genetics* **80**, 1103-1114 (2007).

57. A. C. Z. de Paiva, F. A. d. L. Marson, J. D. Ribeiro, C. S. Bertuzzo, Asthma: Gln27Glu and Arg16Gly polymorphisms of the beta2-adrenergic receptor gene as risk factors. *Allergy, Asthma & Clinical Immunology* **10**, 8 (2014).

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

59. F. D'Amico, D. C. Baumgart, S. Danese, L. Peyrin-Biroulet, Diarrhea During COVID-19 Infection: Pathogenesis, Epidemiology, Prevention, and Management. *Clin Gastroenterol Hepatol*, (2020).

60. J. C. Martin, C. Chang, G. Boschetti, R. Ungaro, M. Giri, J. A. Grout, K. Gettler, L. S. Chuang, S. Nayar, A. J. Greenstein, M. Dubinsky, L. Walker, A. Leader, J. S. Fine, C. E. Whitehurst, M. L. Mbow, S. Kugathasan, L. A. Denson, J. S. Hyams, J. R. Friedman, P. T. Desai, H. M. Ko, I. Laface, G. Akturk, E. E. Schadt, H. Salmon, S. Gnjatic, A. H. Rahman, M. Merad, J. H. Cho, E. Kenigsberg, Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell* **178**, 1493-1508 e1420 (2019).

61. W. Zhang, K. Y. Hui, A. Gusev, N. Warner, S. M. Ng, J. Ferguson, M. Choi, A. Burberry, C. Abraham, L. Mayer, R. J. Desnick, C. J. Cardinale, H. Hakonarson, M. Waterman, Y. Chowers, A. Karban, S. R. Brant, M. S. Silverberg, P. K. Gregersen, S. Katz, R. P. Lifton, H. Zhao, G. Nunez, I. Pe'er, I. Peter, J. H. Cho, Extended haplotype association study in Crohn's disease identifies a novel, Ashkenazi Jewish-specific missense mutation in the NF-kappaB pathway gene, HEATR3. *Genes Immun* **14**, 310-316 (2013).

62. A. M. Muise, T. Walters, W. Xu, G. Shen–Tu, C. H. Guo, R. Fattouh, G. Y. Lam, V. M. Wolters, J. Bennitz, J. van Limbergen, P. Renbaum, Y. Kasirer, B. Y. Ngan, D. Turner, L. A. Denson, P. M. Sherman, R. H. Duerr, J. Cho, C. W. Lees, J. Satsangi, D. C. Wilson, A. D. Paterson, A. M. Griffiths, M. Glogauer, M. S. Silverberg, J. H. Brumell, Single Nucleotide Polymorphisms That Increase Expression of the Guanosine Triphosphatase RAC1 Are Associated With Ulcerative Colitis. *Gastroenterology* **141**, 633-641 (2011).

63. M. L. Seinen, G. P. van Nieuw Amerongen, N. K. H. de Boer, C. J. J. Mulder, J. van Bezu, A. A. van Bodegraven, Rac1 as a Potential Pharmacodynamic Biomarker for Thiopurine Therapy in Inflammatory Bowel Disease. *Therapeutic Drug Monitoring* **38**, 621-627 (2016).

64. B. I. Korelitz, Expert opinion: experience with 6-mercaptopurine in the treatment of inflammatory bowel disease. *World J Gastroenterol* **19**, 2979-2984 (2013).

65. C. M. Perry, L. J. Scott, Cefdinir: A Review of its Use in the Management of Mild-to-Moderate Bacterial Infections. *Drugs* **64**, 1433-1464 (2004).

66. L. M. Johansen, J. M. Brannan, S. E. Delos, C. J. Shoemaker, A. Stossel, C. Lear, B. G. Hoffstrom, L. E. Dewald, K. L. Schornberg, C. Scully, J. Lehar, L. E. Hensley, J. M. White, G. G. Olinger, FDA-approved selective estrogen receptor modulators inhibit Ebola virus infection. *Sci Transl Med* **5**, 190ra179 (2013).

67. Y. Cong, B. J. Hart, R. Gross, H. Zhou, M. Frieman, L. Bollinger, J. Wada, L. E. Hensley, P. B. Jahrling, J. Dyall, M. R. Holbrook, MERS-CoV pathogenesis and antiviral efficacy of licensed drugs in human monocyte-derived antigen-presenting cells. *PLoS One* **13**, e0194868 (2018).

68. J. Dyall, C. M. Coleman, B. J. Hart, T. Venkataraman, M. R. Holbrook, J. Kindrachuk, R. F. Johnson, G. G. Olinger, Jr., P. B. Jahrling, M. Laidlaw, L. M. Johansen, C. M. Lear-Rooney, P. J. Glass, L. E. Hensley, M. B. Frieman, Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. *Antimicrob Agents Chemother* **58**, 4885-4893 (2014).

69. S. Jeon, M. Ko, J. Lee, I. Choi, Y. Byun, S. Park, D. Shum, K. S., Identification of antiviral drug candidates against SARS-CoV-2 from FDA-approved drugs. *bioRxiv*, (2020).

70. X. J. Wang, W. Hu, T. Y. Zhang, Y. Y. Mao, N. N. Liu, S. Q. Wang, Irbesartan, an FDA approved drug for hypertension and diabetic nephropathy, is a potent inhibitor for hepatitis B virus entry by disturbing Na(+)-dependent taurocholate cotransporting polypeptide activity. *Antiviral Res* **120**, 140-146 (2015).

71. P. Zhang, L. Zhu, J. Cai, F. Lei, J.-J. Qin, J. Xie, Y.-M. Liu, Y.-C. Zhao, X. Huang, L. Lin, M. Xia, M.-M. Chen, X. Cheng, X. Zhang, D. Guo, Y. Peng, Y.-X. Ji, J. Chen, Z.-G. She, Y. Wang, Q. Xu, R. Tan, H. Wang, J. Lin, P. Luo, S. Fu, H. Cai, P. Ye, B. Xiao, W. Mao, L. Liu, Y. Yan, M. Liu, M. Chen, X.-J. Zhang, X. Wang, R. M. Touyz, J. Xia, B.-H. Zhang, X. Huang, Y. Yuan, L. Rohit, P. P. Liu, H. Li, Association of Inpatient Use of Angiotensin Converting Enzyme Inhibitors and Angiotensin II Receptor Blockers with

Mortality Among Patients With Hypertension Hospitalized With COVID-19. *Circulation Research* **0**.

72. V. G. Puelles, M. Lutgehetmann, M. T. Lindenmeyer, J. P. Sperhake, M. N. Wong, L. Allweiss, S. Chilla, A. Heinemann, N. Wanner, S. Liu, F. Braun, S. Lu, S. Pfefferle, A. S. Schroder, C. Edler, O. Gross, M. Glatzel, D. Wichmann, T. Wiech, S. Kluge, K. Pueschel, M. Aepfelbacher, T. B. Huber, Multiorgan and Renal Tropism of SARS-CoV-2. *The New England journal of medicine*, (2020).

73. F. Xiao, M. Tang, X. Zheng, Y. Liu, X. Li, H. Shan, Evidence for Gastrointestinal Infection of SARS-CoV-2. *Gastroenterology* **158**, 1831-1833 e1833 (2020).

74. K. Nakashima, T. Hirota, K. Obara, M. Shimizu, A. Jodo, M. Kameda, S. Doi, K. Fujita, T. Shirakawa, T. Enomoto, F. Kishi, S. Yoshihara, K. Matsumoto, H. Saito, Y. Suzuki, Y. Nakamura, M. Tamari, An association study of asthma and related phenotypes with polymorphisms in negative regulator molecules of the TLR signaling pathway. *Journal of Human Genetics* **51**, 284-291 (2006).

75. M. Pino-Yanes, I. Sanchez-Machin, J. Cumplido, J. Figueroa, M. J. Torres-Galvan, R. Gonzalez, A. Corrales, O. Acosta-Fernandez, J. C. Garcia-Robaina, T. Carrillo, A. Sanchez-Palacios, J. Villar, M. Hernandez, C. Flores, IL-1 receptor-associated kinase 3 gene (IRAK3) variants associate with asthma in a replication study in the Spanish population. *J Allergy Clin Immunol* **129**, 573-575, 575 e571-510 (2012).

76. E. Reihsaus, M. Innis, N. MacIntyre, S. B. Liggett, Mutations in the gene encoding for the beta 2-adrenergic receptor in normal and asthmatic subjects. *Am J Respir Cell Mol Biol* **8**, 334-339 (1993).

77. S. Ali, A. F. Hirschfeld, M. L. Mayer, E. S. Fortuno, N. Corbett, M. Kaplan, S. Wang, J. Schneiderman, C. D. Fjell, J. Yan, L. Akhabir, F. Aminuddin, N. Marr, T. Lacaze-Masmonteil, R. G. Hegele, A. Becker, M. Chan-Yeung, R. E. W. Hancock, T. R. Kollmann, D. Daley, A. J. Sandford, P. M. Lavoie, S. E. Turvey, Functional Genetic Variation in NFKBIA and Susceptibility to Childhood Asthma, Bronchiolitis, and Bronchopulmonary Dysplasia. *The Journal of Immunology* **190**, 3949-3958 (2013).

78. T. Akaike, H. Maeda, Nitric oxide and virus infection. *Immunology* **101**, 300-308 (2000).

79. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, J. T. Griffin, M. Baguelin, S. Bhatia, A. Boonyasiri, A. Cori, Z. Cucunuba, R. FitzJohn, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y. Wang, X. Xi, C. A. Donnelly, A. C. Ghani, N. M. Ferguson, Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*, (2020).

80. D. Aw, A. B. Silva, D. B. Palmer, Immunosenescence: emerging challenges for an ageing population. *Immunology* **120**, 435-446 (2007).

81. X. Wang, Y. Bian, R. Zhang, X. Liu, L. Ni, B. Ma, R. Zeng, Z. Zhao, X. Song, C. Liu, Melatonin alleviates cigarette smoke-induced endothelial cell pyroptosis through inhibiting ROS/NLRP3 axis. *Biochem Biophys Res Commun* **519**, 402-408 (2019).

82. Y. Shang, S. P. Xu, Y. Wu, Y. X. Jiang, Z. Y. Wu, S. Y. Yuan, S. L. Yao, Melatonin reduces acute lung injury in endotoxemic rats. *Chin Med J (Engl)* **122**, 1388-1393 (2009).

83. J. Zhang, L. Wang, W. Xie, S. Hu, H. Zhou, P. Zhu, H. Zhu, Melatonin attenuates ER stress and mitochondrial damage in septic cardiomyopathy: A new mechanism involving BAP31 upregulation and MAPK-ERK pathway. *J Cell Physiol* **235**, 2847-2856 (2020).

84. A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, V. A. McKusick, Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* **33**, D514-517 (2005).

85. M. Zhang, C. Zhu, A. Jacomy, L. J. Lu, A. G. Jegga, The orphan disease networks. *Am J Hum Genet* **88**, 755-766 (2011).

86. R. Kittler, L. Pelletier, A.-K. Heninger, M. Slabicki, M. Theis, L. Miroslaw, I. Poser, S. Lawo, H. Grabner, K. Kozak, J. Wagner, V. Surendranath, C. Richter, W. Bowen, A. L. Jackson, B. Habermann, A. A. Hyman, F. Buchholz, Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nature Cell Biology* **9**, 1401-1412 (2007).

87. K. Breuer, A. K. Foroushani, M. R. Laird, C. Chen, A. Sribnaia, R. Lo, G. L. Winsor, R. E. W. Hancock, F. S. L. Brinkman, D. J. Lynn, InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res.* **41**, D1228-D1233 (2012).

88. M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90-W97 (2016).

89. A. E. Hirsh, H. B. Fraser, D. P. Wall, Adjusting for selection on synonymous sites in estimates of evolutionary distance. *Mol Biol Evol* **22**, 174-177 (2005).

90. A. Bezginov, G. W. Clark, R. L. Charlebois, V. U. Dar, E. R. Tillier, Coevolution reveals a network of human proteins originating with multicellularity. *Mol Biol Evol* **30**, 332-346 (2013).

91. F. Cheng, P. Jia, Q. Wang, C.-C. Lin, W.-H. Li, Z. Zhao, Studying Tumorigenesis through Network Evolution and Somatic Mutational Perturbations in the Cancer Interactome. *Molecular Biology and Evolution* **31**, 2156-2169 (2014).

92. N. Mantel, W. Haenszel, Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *JNCI: Journal of the National Cancer Institute* **22**, 719-748 (1959).

93. R. DerSimonian, N. Laird, Meta-analysis in clinical trials. *Controlled Clinical Trials* **7**, 177-188 (1986).

94. X. Wan, W. Wang, J. Liu, T. Tong, Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology* **14**, 135 (2014).

95. T. Stijnen, T. H. Hamza, P. Özdemir, Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine* **29**, 3046-3067 (2010).

96.   N. R. Coordinators, Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7-19 (2016).

97.   I. N. Smith, S. Thacker, M. Seyfi, F. Cheng, C. Eng, Conformational Dynamics and Allosteric Regulation Landscapes of Germline PTEN Mutations Associated with Autism Compared to Those Associated with Cancer. *Am J Hum Genet* **104**, 861-878 (2019).

98.   A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420 (2018).

99.   V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. Maciejewski, D. Arndt, M. Wilson, V. Neveu, A. Tang, G. Gabriel, C. Ly, S. Adamjee, Z. T. Dame, B. Han, Y. Zhou, D. S. Wishart, DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* **42**, D1091-1097 (2014).

100.   H. Yang, C. Qin, Y. H. Li, L. Tao, J. Zhou, C. Y. Yu, F. Xu, Z. Chen, F. Zhu, Y. Z. Chen, Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Res* **44**, D1069-1074 (2016).

101.   A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, J. P. Overington, ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100-D1107 (2012).

102.   T. Q. Liu, Y. M. Lin, X. Wen, R. N. Jorissen, M. K. Gilson, BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **35**, D198-D201 (2007).

103.   A. J. Pawson, J. L. Sharman, H. E. Benson, E. Faccenda, S. P. H. Alexander, O. P. Buneman, A. P. Davenport, J. C. McGrath, J. A. Peters, C. Southan, M. Spedding, W. Y. Yu, A. J. Harmar, I. Nc, The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res.* **42**, D1098-D1106 (2014).

104.   J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J. P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, T. R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929-1935 (2006).

105.   R. Reghunathan, M. Jayapal, L.-Y. Hsu, H.-H. Chng, D. Tai, B. P. Leung, A. J. Melendez, Expression profile of immune response genes in patients with Severe Acute Respiratory Syndrome. *BMC Immunology* **6**, 2 (2005).

106.   L. Josset, V. D. Menachery, L. E. Gralinski, S. Agnihothram, P. Sova, V. S. Carter, B. L. Yount, R. L. Graham, R. S. Baric, M. G. Katze, Cell host response to infection with novel human coronavirus EMC predicts potential antivirals and important differences with SARS coronavirus. *mBio* **4**, e00165-00113 (2013).

107.   S. Yuan, H. Chu, J. F.-W. Chan, Z.-W. Ye, L. Wen, B. Yan, P.-M. Lai, K.-M. Tee, J. Huang, D. Chen, C. Li, X. Zhao, D. Yang, M. C. Chiu, C. Yip, V. K.-M. Poon, C. C.-S.

Chan, K.-H. Sze, J. Zhou, I. H.-Y. Chan, K.-H. Kok, K. K.-W. To, R. Y.-T. Kao, J. Y.-N. Lau, D.-Y. Jin, S. Perlman, K.-Y. Yuen, SREBP-dependent lipidomic reprogramming as a broad-spectrum antiviral target. *Nature Communications* **10**, 120 (2019).

Table

Table 1. Baseline characteristics of the melatonin and carvedilol intaking patient groups from the COVID-19 registry.

Melatonin						
	All patients			**SARS-CoV-2 positive patients**		
	Melatonin -	**Melatonin +**	**P**	**Melatonin -**	**Melatonin +**	**P**
Total patients	17312	806		1635	40	
Age	48.94 ± 20.49	62.84 ± 19.88	< 0.001	53.63 ± 18.89	65.24 ± 18.44	< 0.001
Race other %	1092 (6.3)	43 (5.3)	0.298	109 (6.7)	1 (2.5)	0.467
Race black %	3317 (19.2)	202 (25.1)	< 0.001	380 (23.2)	13 (32.5)	0.239
Race white %	11922 (68.9)	538 (66.7)	0.219	1044 (63.9)	23 (57.5)	0.51
Gender (male) %	6833 (39.5)	396 (49.1)	< 0.001	785 (48.0)	21 (52.5)	0.688
Smoking %	2222 (12.8)	129 (16.0)	0.258	91 (5.6)	3 (7.5)	0.743
Diabetes %	3222 (18.6)	333 (41.3)	< 0.001	336 (20.6)	13 (32.5)	0.456
Hypertension %	6524 (37.7)	594 (73.7)	< 0.001	662 (40.5)	27 (67.5)	0.029
Coronary artery disease %	1897 (11.0)	268 (33.3)	< 0.001	152 (9.3)	12 (30.0)	0.001
COPD & emphysema %	740 (4.3)	119 (14.8)	< 0.001	76 (4.6)	5 (12.5)	0.07
ACEIs (%)	1182 (6.8)	156 (19.4)	< 0.001	93 (5.7)	5 (12.5)	0.141
ARBs (%)	900 (5.2)	123 (15.3)	< 0.001	96 (5.9)	8 (20.0)	0.001
Carvedilol						
	All patients			**SARS-CoV-2 positive patients**		
	Carvedilol -	**Carvedilol +**	**P**	**Carvedilol -**	**Carvedilol +**	**P**
Total patients	17567	551		1647	28	
Age	49.06 ± 20.59	65.46 ± 16.22	< 0.001	53.67 ± 18.94	68.16 ± 13.52	< 0.001
Race other %	1102 (6.3)	33 (6.0)	0.856	109 (6.6)	1 (3.6)	0.794
Race black %	3333 (19.0)	186 (33.8)	< 0.001	379 (23.0)	14 (50.0)	0.002
Race white %	12134 (69.1)	326 (59.2)	< 0.001	1054 (64.0)	13 (46.4)	0.086
Gender (male) %	6916 (39.4)	313 (56.8)	< 0.001	789 (47.9)	17 (60.7)	0.248
Smoking %	2243 (12.8)	108 (19.6)	0.001	90 (5.5)	4 (14.3)	0.108
Diabetes %	3235 (18.4)	320 (58.1)	< 0.001	333 (20.2)	16 (57.1)	< 0.001
Hypertension %	6601 (37.6)	517 (93.8)	< 0.001	662 (40.2)	27 (96.4)	< 0.001
Coronary artery disease %	1871 (10.7)	294 (53.4)	< 0.001	148 (9.0)	16 (57.1)	< 0.001
COPD & emphysema %	791 (4.5)	68 (12.3)	< 0.001	74 (4.5)	7 (25.0)	0.001
ACEIs (%)	1190 (6.8)	148 (26.9)	< 0.001	93 (5.6)	5 (17.9)	0.02
ARBs (%)	884 (5.0)	139 (25.2)	< 0.001	93 (5.6)	11 (39.3)	< 0.001

COPD, chronic obstructive pulmonary disease. ACEIs, angiotensin-converting enzyme inhibitors. ARBs, angiotensin II receptor blockers. Ages are shown as mean ± standard deviation. All other characteristics are shown as number of cases (percentage). *P* values were calculated by Fisher's exact test. +: drug usage; -: non-drug usage.

Figure Legends

Fig. 1. Overall workflow of this study. (**A**) A diagram illustrating the basic pathogenesis of SARS-CoV-2. (**B**) A diagram illustrating how to build a global interactome map for SARS-CoV-2. We compiled the SARS-CoV-2 human target gene/protein sets from multi-omics data from transcriptome, proteome, and human interactome, and validated network-based findings using patient data from the COVID-19 registry. (**C**) A diagram illustrating network-based measure of disease manifestations associated with COVID-19. We systematically evaluated the network proximities of the SARS-CoV-2 human target genes/proteins with 64 diseases across six main categories: autoimmune, cancer, cardiovascular, metabolic, neurological, and pulmonary. (**D**) A workflow illustrating validation of network-based findings. We performed single-cell analyses to further investigate the underlying mechanisms of COVID-19 with asthma and inflammatory bowel disease. We prioritized nearly 3,000 FDA approved/investigational drugs for their potential anti-SARS-CoV-2 effects from network-based findings and validated drug-COVID-19 outcomes using the institutional review board-approved COVID-19 patient registry.

Fig. 2. Network and biological characteristics of the SARS-CoV-2 interactome map. (**A**) Pathway and gene ontology (biological process) enrichment analysis results of the SARS-CoV-2 host genes/proteins across five different data sets. We assembled five gene/protein data sets from SARS-CoV-2 host protein-protein interactions, transcriptomics, and proteomics (**Table S1**). (**B, C, D**) Network and biological characteristics of the SARS-CoV-2 host genes/proteins. The proteins in PanCoV-PPI have higher node degrees (**B**), lower non-synonymous to synonymous substitutions (*dN/dS*) ratios (**C**), and lower evolutionary ratios (**D**) compared to randomly

selected proteins (grey, mean ± standard deviation of 100 repeats). (**E**) Among the 460 proteins in PanCoV-PPI, 450 (98%) are expressed in lungs, and 317 (69%) are lung specific expression ($Z > 0$). (**F, G**) The distribution of the node degrees in the human interactome and *dN/dS* ratios of PanCoV-PPI and four published virus-related host protein sets. (**H**) The shared target human proteins (blue) of SARS-CoV-2 (red) and other viruses (green). The detailed data are provided in **Table S2**. (**I**) SARS-CoV-2 target proteins overlap significantly with disease-associated genes (Mendelian disease and Orphan disease), cancer genes, cell cycle genes, and innate immune genes.

Fig. 3. A global network illustrating disease manifestations associated with human coronavirus. (**A**) The target human proteins of SARS-CoV-2 are connected to the disease-associated proteins. Blue links (edges) indicate physical protein-protein interactions. For SARS-CoV-2, its viral proteins are shown by light red node. The target human proteins (blue) of the viruses are intricately connected to the disease-associated proteins (green). Human disease nodes are colored by different disease categories: autoimmune, cancer, cardiovascular, metabolic, neurological, and pulmonary. (**B**) Estimation of the pooled risk ratio using random effects meta-analysis for 10 comorbidities between severe versus non-severe COVID-19 patients. The *tau²* and I^2 statistics were calculated to quantify the heterogeneity among studies. $I^2 < 50\%$ was considered as low heterogeneity among studies, $50\% < I^2 < 75\%$ was considered as moderate heterogeneity, and $I^2 > 75\%$ was considered as high heterogeneity.

Fig. 4. A landscape of disease manifestations associated with COVID-19 quantified by network proximity measure. (**A**) Heatmaps showing the network proximities of COVID-19

with 64 diseases across 6 categories. The network proximities of the disease modules and the

five SARS-CoV-2 data sets were evaluated using the "closest" network proximity measure (see

Methods). The magnitude of the proximity is indictive of their biological relationship: closer

network proximity of SARS-CoV-2 host genes/proteins of a disease indicates higher potential of

manifestation between COVID-19 and a specific disease. $P < 0.05$ computed by permutation test

was considered significant (indicated by horizontal bars). Three categories, autoimmune,

pulmonary, and neurological frequently show significant proximities to COVID-19.

Inflammatory bowel disease, attention-deficit/hyperactivity disorder, and stroke achieved

significance with all five SARS-CoV-2 data sets. (**B, C**) Highlighted subnetworks between

SARS-CoV-2 host genes/proteins with the disease-associated proteins of respiratory distress

syndrome (**B**) and sepsis (**C**). (**D**) Clinical data analyses showed an association of COVID-19

severity with IL-6 expression levels in patients. Meta-analysis of random effects model was

performed using the mean difference in IL-6 (pg/ml). There was a high heterogeneity among

these studies ($I^2 = 94\%$, $P < 0.001$).

Fig. 5. Endophenotypes between asthma and COVID-19. (**A**) A highlighted subnetwork

between the asthma-associated genes, differential metabolites in asthma, and SARS-CoV-2 host

genes, under the human interactome network model. (**B**) A heatmap highlighting differential

gene expression analyses for the genes identified in asthma and COVID-19 subnetwork analysis

(**A**). We performed differential gene expression analysis using two existing asthma cohorts

(GSE63142 and GSE130499). SvsC, severe versus control; MvsC mild versus control; SvsM,

severe versus mild. Blue bars show the node degree enrichment in the subnetwork (**A**) compared

to a random network of the same size (see Methods). Black dotted line indicates enrichment = 1.

49

(**C**) UMAP visualization for human bronchial epithelial cells. (**D**) Cell type-specific expression levels of *ACE2* and *TMPRSS2* across 14 cell types in human bronchial epithelial cells. (**E**) Cell type-specific expression levels of seven highlighted inflammatory genes (**A** and **B**) show elevated expression levels in secretory 3 cells compared to other cell types. (**F**) UMAP visualization for lung cells. (**G**) Cell type-specific expression levels of *ACE2* and *TMPRSS2* across 9 cell types in lung cells. (**H**) Cell type-specific expression levels of four highlighted inflammatory genes (**A** and **B**) show elevated expression levels in alveolar type II cells compared to other cell types.

Fig. 6. Inflammatory endophenotype between inflammatory bowel disease (IBD) and COVID-19. (**A**) Severe COVID-19 patients have higher risks of abdominal pain and diarrhea by meta-analysis. (**B**) A highlighted subnetwork between the IBD-associated genes and the SARS-CoV-2 virus proteins and virus-host (human) proteins under the human interactome network model. (**C**) UMAP visualization of non-epithelial cells from the ileal tissues of patients with Crohn's disease. (**D**) UMAP visualization of epithelial cells from the ileal tissues of patients with Crohn's disease. (**E**) Cell type-specific expression of *ACE2* and *TMPRSS2* in non-epithelial cells (**C**). (**F**) Cell type-specific expression of *ACE2* and *TMPRSS2* in epithelial cells (**D**). (**G**) The co-expression of *ACE2* and *TMPRSS2* are elevated in absorptive enterocytes of the inflamed ileal tissues compared to uninflamed tissues in patients with Crohn's disease. (**H**) Box plot showing the expression of *ACE2* and *TMPRSS2* in absorptive enterocytes expressing *ACE2* and *TMPRSS2*, respectively. (**I**) Co-expression analysis for the genes in the subnetwork with *ACE2* and *TMPRSS2*. Heatmap shows the Pearson correlation coefficients (PCC) of *ACE2* and *TMPRSS2* with other genes (labeled in **B**) in the absorptive enterocytes. Blue bars show the

50

degree enrichment of the genes in the subnetwork compared to a random network of the same size. Black dotted line indicates enrichment = 1.

Fig. 7. Network-based prediction and patient-based validation of drug repurposing for COVID-19. (**A**) Thirty-four drugs from the top predicted list are highlighted with the disease category they are approved by U.S. FDA. We highlighted three types of evidences: (i) network proximities of drug's targets across the four SARS-CoV-2 data sets (SARS2-DEG, SARS2-DEP, HCoV-PPI, and SARS2-PPI) in the human interactome; (ii) gene set enrichment analysis (GSEA) scores across five coronavirus transcriptomics and proteomics data sets (see Methods), and (iii) literature-reported antivirus profiles. GSEA scores shown in grey indicate that these drugs cannot be evaluated due to the lack of data. Five drugs that are currently being or have been tested in clinical trials are highlighted. Horizontal bars indicate $P < 0.05$. ES, enrichment score by GESA analysis (see Methods). HIV, human immunodeficiency virus. EBOV, Zaire ebolavirus. RV, rhinovirus. HAV, hepatitis A virus. HBV, hepatitis B virus. HCV, hepatitis C virus. HDV, hepatitis D virus. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2. SARS-CoV, severe acute respiratory syndrome coronavirus. MERS-CoV, Middle East respiratory syndrome coronavirus. HCoV-229E, human coronavirus 229E. MHV, mouse hepatitis virus. IAV, influenza A virus. IBV, influenza B virus. WNV, West Nile virus. ZIKV, Zika virus. ANDV, Andes virus. CHIKV, chikungunya virus. CMV, cytomegalovirus. DENV, dengue virus. EMCV, encephalomyocarditis virus. MV, measles virus. RSV, respiratory syncytial virus. SVCV, spring viremia of carp virus. (**B**) Observational studies highlighting that patients taking melatonin were associated with reduced risk of likelihood of a positive laboratory test result for SARS-CoV-2 from the COVID-19 registry. Melatonin usage also reduced the risk

compared to angiotensin-converting enzyme inhibitors (ACEIs), angiotensin II receptor blockers (ARBs), and the pooled patient data of ARBs/ACEIs. PS, propensity score. OR, odds ratio. PS-matched model1 was matched using age, gender, race, and smoking; PS-matched model2 was matched using age, gender, race, smoking, coronary artery disease, diabetes, hypertension, and chronic obstructive pulmonary disease (COPD). OR model1, the odds ratios of COVID-19 in melatonin exposure; OR model2, the odds ratios of COVID-19 in melatonin exposure adjusted by age, gender, race, smoking, coronary artery disease, diabetes, hypertension, and COPD.
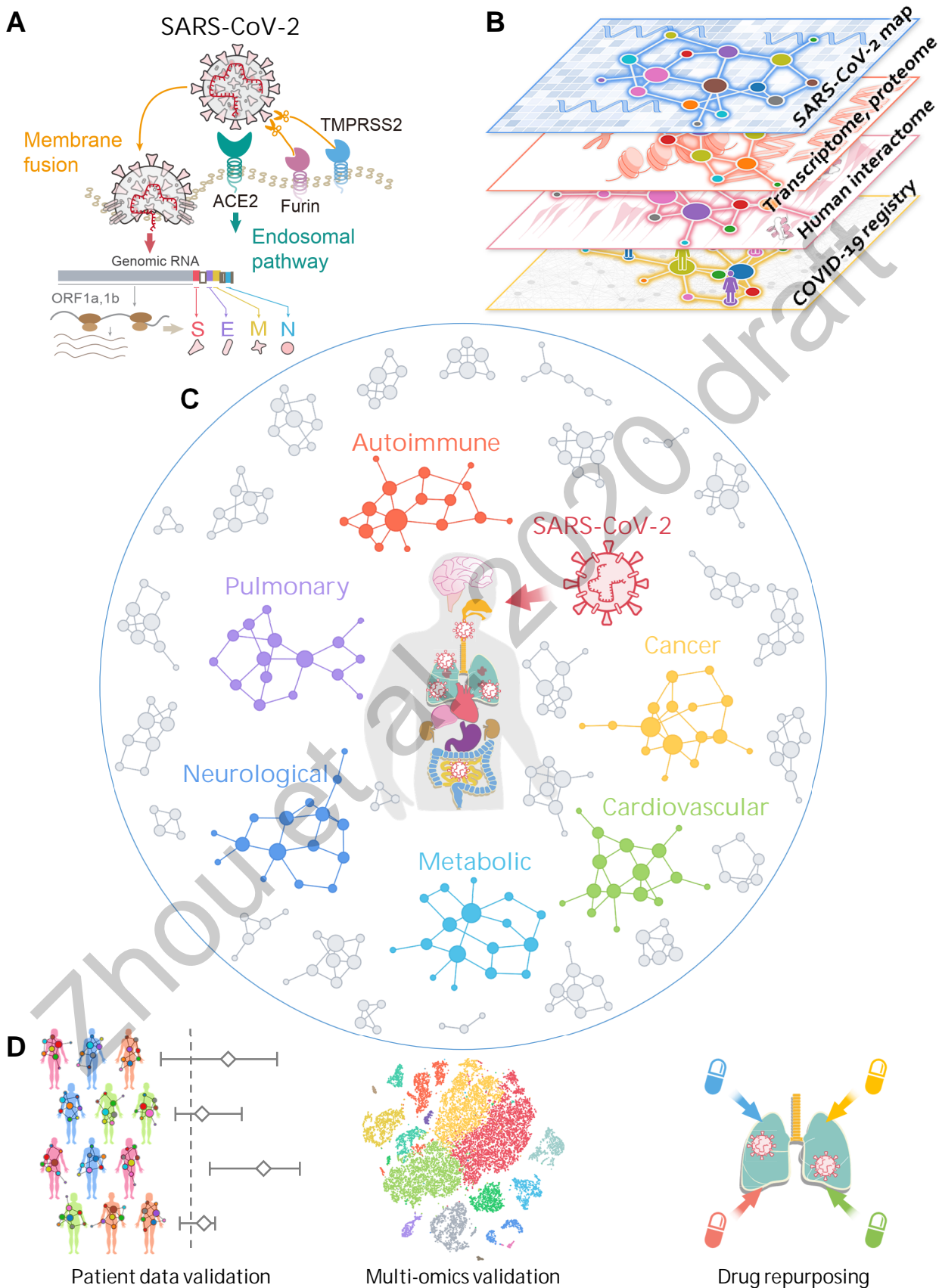
Figure 1

Figure 2

Figure 3

A

Legend:
- Autoimmune (pink)
- Cancer (gray)
- Cardiovascular (green)
- Metabolic (purple)
- Neurological (yellow)
- Pulmonary (brown)
- Human coronaviruses (HCoVs) (dark red)
- SARS-CoV-2 proteins (salmon)
- Target host proteins of HCoVs (blue)
- Disease associated proteins (green)
- Overlap of (blue) and (green) (orange)
- Human PPIs
- Virus-human PPIs
- Disease-protein associations

B

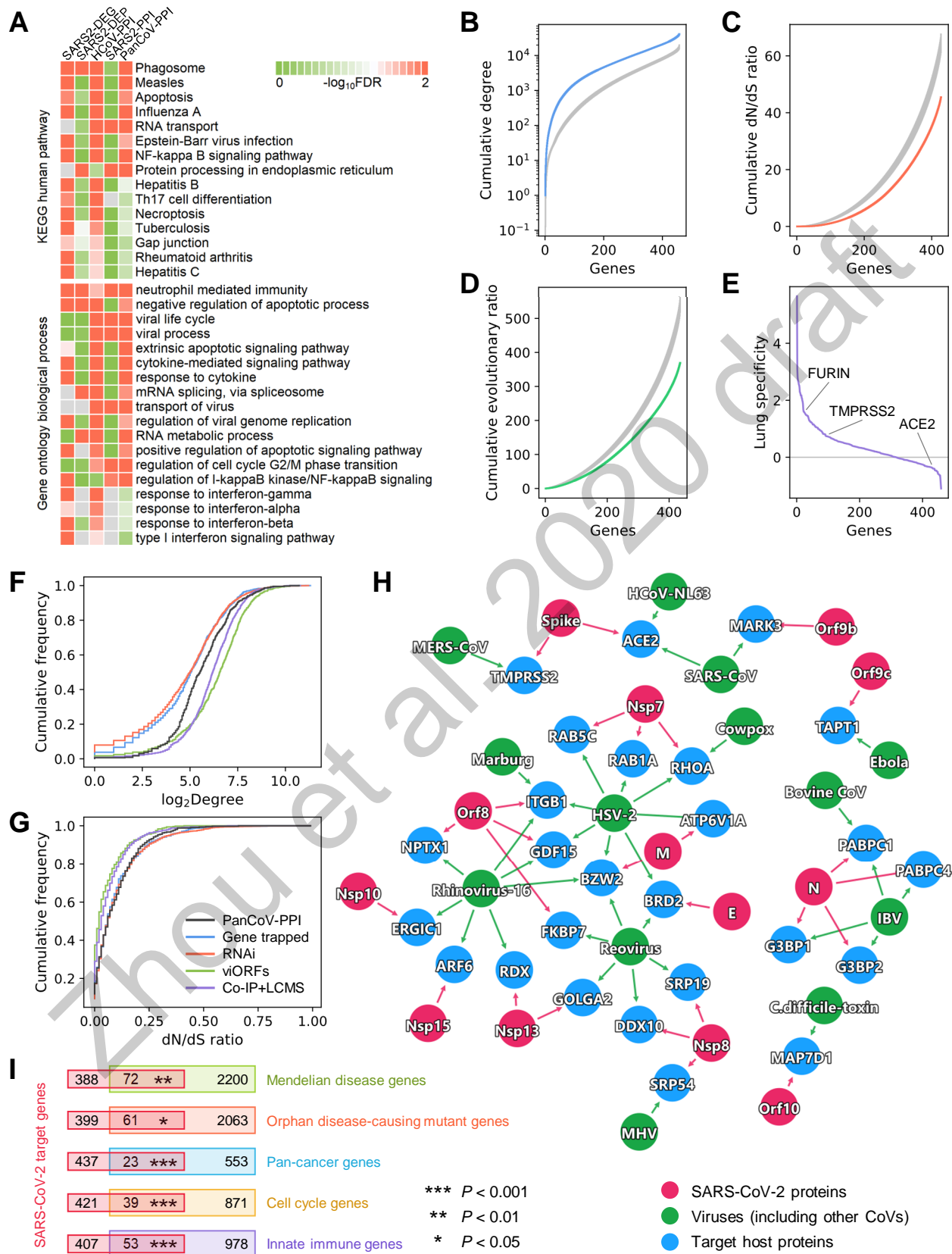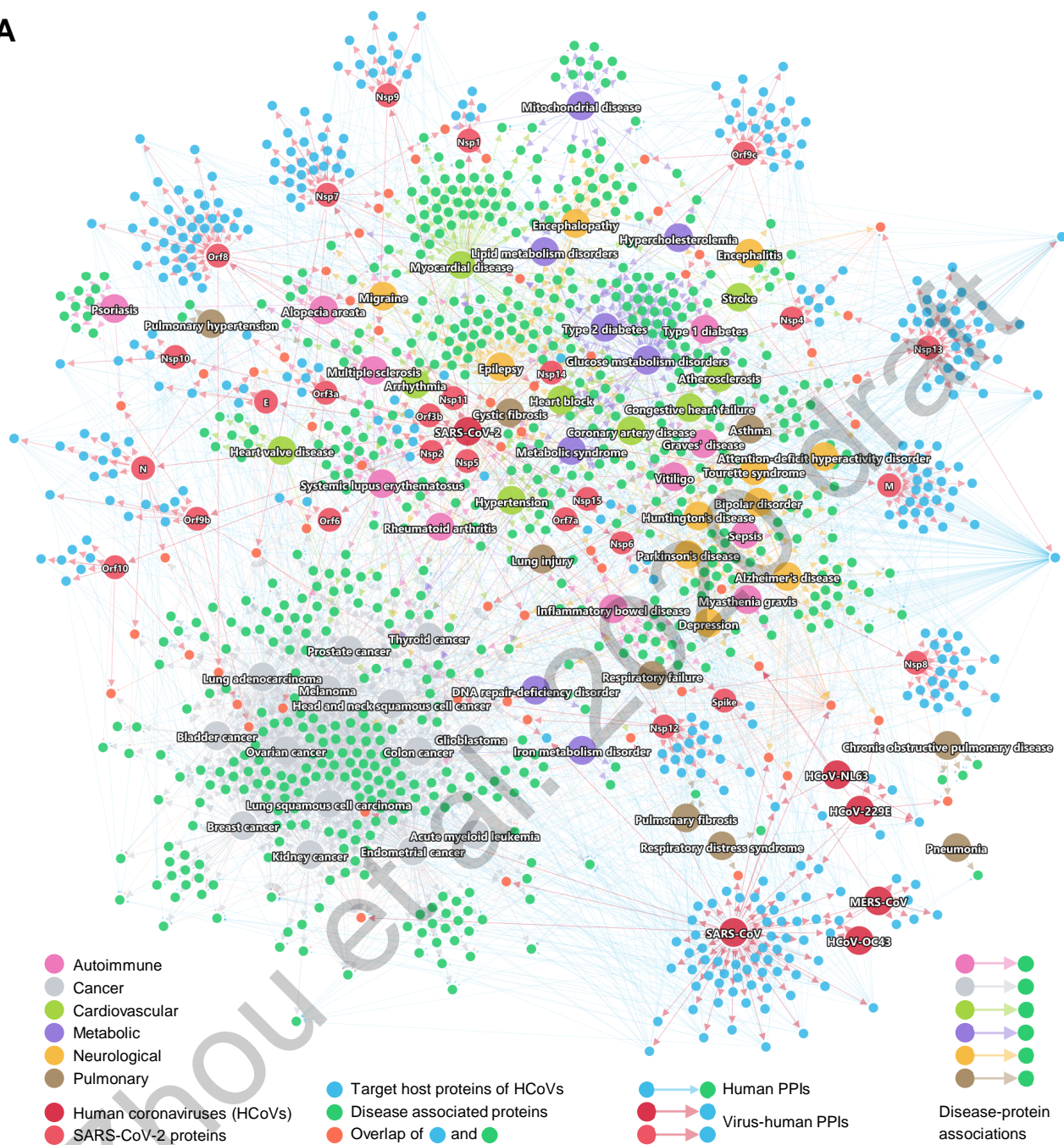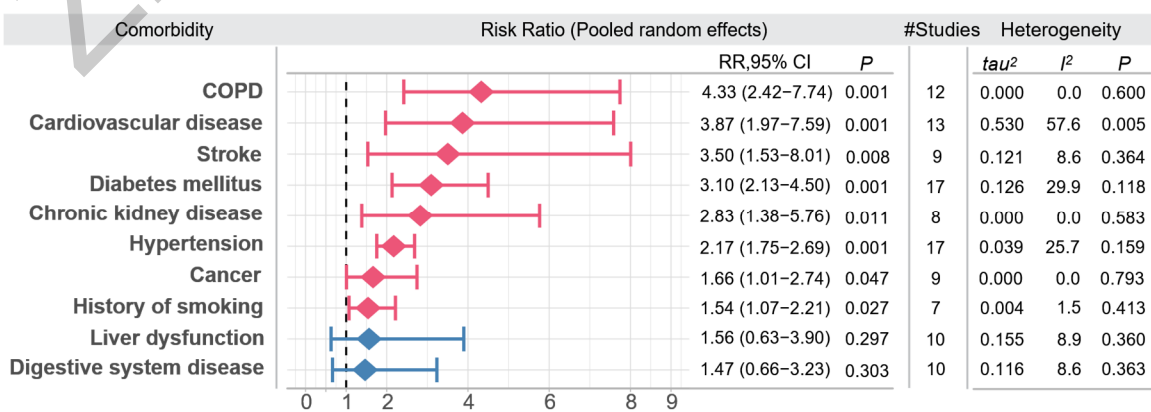| Comorbidity	Risk Ratio (Pooled random effects)		#Studies	Heterogeneity		
	RR,95% CI	P		tau²	I²	P
COPD	4.33 (2.42–7.74)	0.001	12	0.000	0.0	0.600
Cardiovascular disease	3.87 (1.97–7.59)	0.001	13	0.530	57.6	0.005
Stroke	3.50 (1.53–8.01)	0.008	9	0.121	8.6	0.364
Diabetes mellitus	3.10 (2.13–4.50)	0.001	17	0.126	29.9	0.118
Chronic kidney disease	2.83 (1.38–5.76)	0.011	8	0.000	0.0	0.583
Hypertension	2.17 (1.75–2.69)	0.001	17	0.039	25.7	0.159
Cancer	1.66 (1.01–2.74)	0.047	9	0.000	0.0	0.793
History of smoking	1.54 (1.07–2.21)	0.027	7	0.004	1.5	0.413
Liver dysfunction	1.56 (0.63–3.90)	0.297	10	0.155	8.9	0.360
Digestive system disease	1.47 (0.66–3.23)	0.303	10	0.116	8.6	0.363

Figure 4

A

Autoimmune
- Type 1 diabetes
- Inflammatory bowel disease
- Multiple sclerosis
- Rheumatoid arthritis
- Alopecia areata
- Graves' disease
- Myasthenia gravis
- Psoriasis
- Sepsis
- Systemic lupus erythematosus
- Vitiligo

Pulmonary
- Asthma
- Chronic obstructive pulmonary disease
- Cystic fibrosis
- Lung injury
- Pneumonia
- Pulmonary fibrosis
- Pulmonary hypertension
- Respiratory distress syndrome
- Respiratory failure

Neurological
- Alzheimer's disease
- Autism
- Bipolar disorder
- Depression
- Encephalitis
- Epilepsy
- Parkinson's disease
- Attention-deficit hyperactivity disorder
- Encephalopathy
- Huntington's disease
- Migraine
- Tourette syndrome

Cancer
- Pan-cancer
- BLCA (Bladder)
- BRCA (Breast)
- COAD (Colon)
- GBM (Glioblastoma)
- HNSC (Head and neck squamous cell)
- KIRC (Kidney)
- LAML (Acute myeloid leukemia)
- LUAD (Lung adenocarcinoma)
- LUSC (Lung squamous cell carcinoma)
- OV (Ovarian)
- PRAD (Prostate)
- SKCM (Skin cutaneous melanoma)
- THCA (Thyroid)
- UCEC (Endometrial)

Cardiovascular
- Arrhythmia
- Atherosclerosis
- Coronary artery disease
- Heart block
- Congestive heart failure
- Heart valve disease
- Hypertension
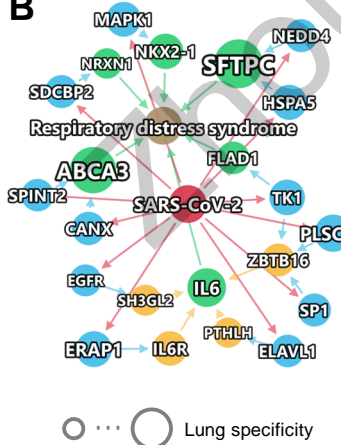- Myocardial disease
- Stroke

Metabolic
- Type 2 diabetes
- DNA repair-deficiency disorder
- Glucose metabolism disorders
- Hypercholesterolemia
- Iron metabolism disorder
- Lipid metabolism disorders
- Metabolic syndrome
- Mitochondrial disease

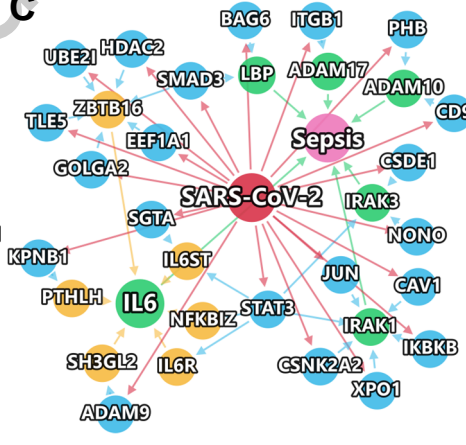Columns: SARS2-DEG, SARS2-DEP, HCoV-PPI, SARS2-PPI, PanCoV-PPI

Z score: -2 to 2, P < 0.05

B Lung specificity

C

D Favours severe | Mean difference (95% CI)

Study	Mean difference (95% CI)
Zhou, F et al.	4.57 (3.12~6.02)
Cai, Q et al.	26.9 (20.0~33.7)
Chen, G et al.	43.2 (−3.33~89.7)
Mo, P et al.	57.0 (33.9~80.1)
Wang, Z et al.	74.7 (19.7~129.7)
Pooled effects	33.0 (0.58~65.3)

- Target host proteins of SARS-CoV-2
- Disease associated proteins
- IL6 neighbors
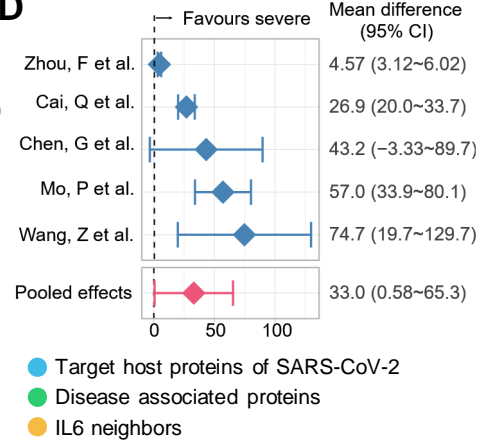
Figure 5

Figure 6

Figure 7