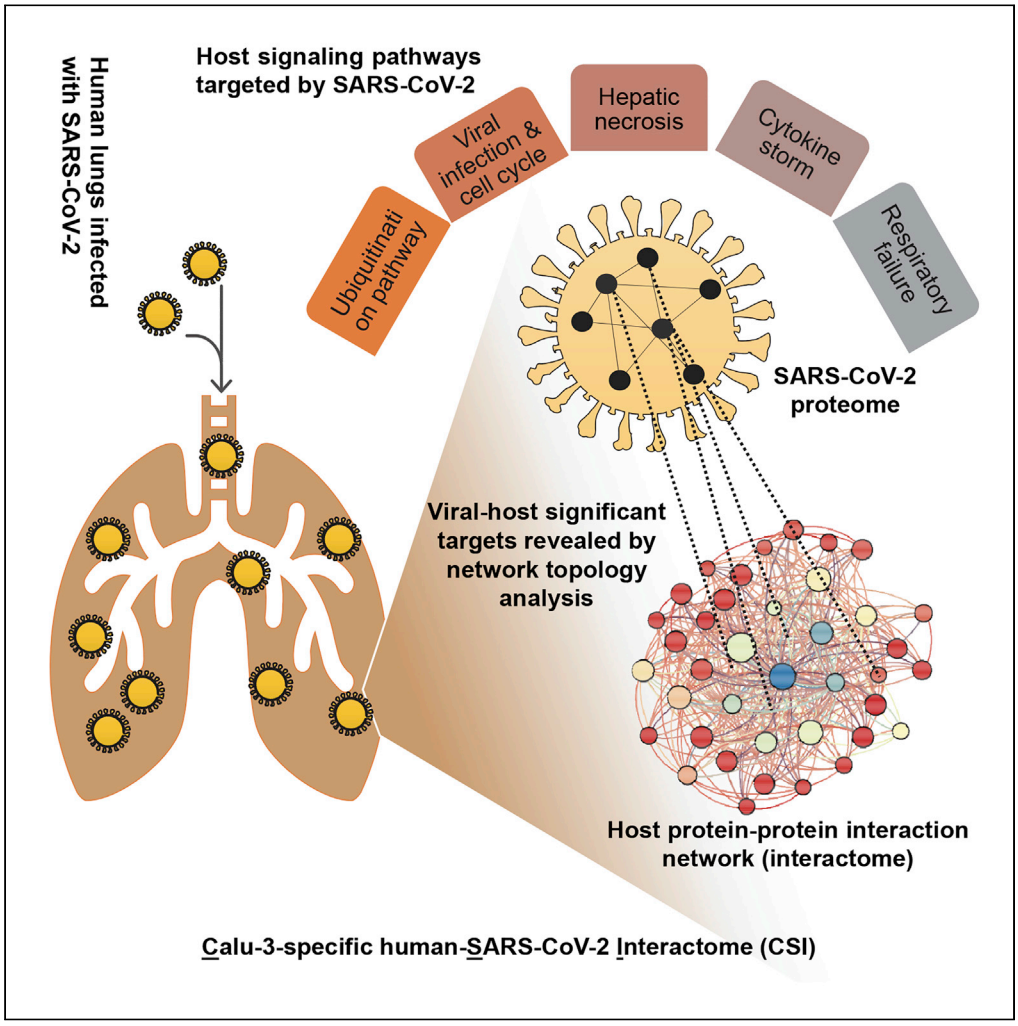
iScience

CellPress
OPEN ACCESS

Article

Integrative Network Biology Framework Elucidates Molecular Mechanisms of SARS-CoV-2 Pathogenesis



Nilesh Kumar,
Bharat Mishra,
Adeel Mehmood

mohammadathar@uabmc.edu (M.A.)
smukhtar@uab.edu (M.S.M.)

HIGHLIGHTS

A network-biology-fueled approach to integrate multiomics data of host-viral network

Network centrality and pathway enrichment analyses reveal 33 high-value targets

Dynamic regulatory circuitry and molecular events pertinent to COVID-19

Underlying molecular mechanisms and pathways of SARS-CoV-2 pathogenesis

Kumar et al., iScience *23*, 101526
September 25, 2020
https://doi.org/10.1016/j.isci.2020.101526

iScience

CellPress
OPEN ACCESS

Article

Integrative Network Biology Framework Elucidates Molecular Mechanisms of SARS-CoV-2 Pathogenesis

Nilesh Kumar,[1,6] Bharat Mishra,[1,6] Adeel Mehmood,[1,2] Mohammad Athar,[3,*] and M Shahid Mukhtar[1,4,5,7,*]

SUMMARY

COVID-19 (coronavirus disease 2019) is a respiratory illness caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Although the pathophysiology of this virus is complex and largely unknown, we employed a network-biology-fueled approach and integrated transcriptome data pertaining to lung epithelial cells with human interactome to generate Calu-3-specific human-SARS-CoV-2 interactome (CSI). Topological clustering and pathway enrichment analysis show that SARS-CoV-2 targets central nodes of the host-viral network, which participate in core functional pathways. Network centrality analyses discover 33 high-value SARS-CoV-2 targets, which are possibly involved in viral entry, proliferation, and survival to establish infection and facilitate disease progression. Our probabilistic modeling framework elucidates critical regulatory circuitry and molecular events pertinent to COVID-19, particularly the host-modifying responses and cytokine storm. Overall, our network-centric analyses reveal novel molecular components, uncover structural and functional modules, and provide molecular insights into the pathogenicity of SARS-CoV-2 that may help foster effective therapeutic design.

INTRODUCTION

From the epicenter of the COVID-19 (coronavirus disease 2019) outbreak in China, the disease has spread globally with over 16.8 million confirmed cases and almost 662,000 fatalities as of July 29, 2020, and the World Health Organization (WHO) warned that the pandemic is accelerating worldwide (Cascella et al., 2020; Hsu et al., 2020). Apart from the human tragedy, COVID-19 has a growing detrimental impact on the global economy and will likely cause trillions of dollars in financial losses worldwide in 2020 alone. COVID-19 is an infectious respiratory illness caused by a highly contagious and pathogenic SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2). This single-stranded RNA virus belongs to the family *Coronaviridae* and is closely related to another human coronavirus SARS-CoV with 89.1% nucleotide similarity (Cascella et al., 2020; Wu et al., 2020). SARS-CoV and another human coronavirus MERS-CoV (Middle East respiratory syndrome-CoV) caused two previous global epidemics in 2003 and 2012, respectively, both characterized by high fatality rates (Cascella et al., 2020; Wu et al., 2020). These coronaviruses mainly spread from a contagious individual to a healthy person through respiratory droplets derived from an infected person's cough or sneeze and from direct contact with contaminated surfaces or objects, where the virus can maintain its viability for a period ranging from hours to days (Cascella et al., 2020; Wu et al., 2020). Unlike other coronaviruses, SARS-CoV-2 transmits more efficiently and sustainably in the community according to the Centers for Disease Control and Prevention (CDC) (Center for Disease Control, 2020). Although the majority of patients infected with SARS-CoV-2 develop a mild-to-moderate self-resolving respiratory illness, infants and older adults ($\geq$65 years) as well as patients with underlying medical conditions such as cardiovascular disease, diabetes, chronic respiratory disease, renal dysfunction, obesity, and cancer are more vulnerable (Cascella et al., 2020; Wu et al., 2020). The pathophysiology of SARS-CoV-2 is complex and largely unknown but is associated with an extensive immune reaction referred to as "cytokine storm" triggered by the excessive production of interleukin 1 beta (IL-1β), interleukin 6 (IL-6), and others. The cytokine release syndrome leads to widespread tissue damage and multiple organ failure (Cascella et al., 2020). Although no vaccine or antiviral drugs are currently available to prevent or treat COVID-19, identifying molecular targets of the virus could help uncover effective treatment. Toward this, the generation of a human-SARS-CoV-2 interactome, integration of virus-related transcriptome to interactome,

[1]Department of Biology, University of Alabama at Birmingham, 464 Campbell Hall, 1300 University Boulevard, AL 35294, USA

[2]Department of Computer Science, University of Alabama at Birmingham, 1402 10th Avenue S., Birmingham, AL 35294, USA

[3]Department of Dermatology, School of Medicine, University of Alabama at Birmingham, 1720 University Boulevard, AL 35294, USA

[4]Nutrition Obesity Research Center, University of Alabama at Birmingham, 1675 University Boulevard, Birmingham, AL 35294, USA

[5]Department of Surgery, University of Alabama at Birmingham, 1808 7th Avenue S, Birmingham, AL 35294, USA

[6]These authors contributed equally

[7]Lead Contact

*Correspondence: mohammadathar@uabmc.edu (M.A.), smukhtar@uab.edu (M.S.M.)

https://doi.org/10.1016/j.isci.2020.101526

the discovery of disease-related structural and functional modules, and dynamic transcriptional modeling will provide insights into the virulence mechanisms of this deadly virus.

Networks encompass a set of nodes and edges, also referred to as vertices and links, respectively. Nodes are systems components, whereas edges represent the interactions or relationships among the nodes (Vidal et al., 2011). In biological systems, genes and their products perform their functions by interacting with other molecular components within the cell. For instance, proteins directly or indirectly interact with each other under both steady-state and different stress conditions to form static and dynamic complexes and participate in diverse signaling cascades, distinct cellular pathways, and a wide spectrum of biological processes. Proteome scale maps of such protein-protein interactions are referred to as interactomes (McCormack et al., 2016). Meanwhile, specialized pathogens including viruses, bacteria, and eukaryotes employ a suite of pathogenic or virulent proteins, which interact with high-value targets in host interactomes to extensively rewire the flow of information and cause disease (Mukhtar et al., 2011; Pan et al., 2016; Vidal et al., 2011; Wessling et al., 2014). Therefore, analyzing the network architecture and deciphering the structural properties of host-pathogen interactomes may reveal novel components in virus pathogenicity. Such analyses indicate that diverse cellular networks are governed by universal laws and exhibit scale-free network topology, whose degree distribution follows a power-law distribution with a few nodes harboring increased connectivity (Mishra et al., 2019). Given that diverse biological systems display similar network architecture and topology, several structural features and physical characteristics within a cellular network may act as indicators of important nodes. These include degree (the number of edges of a node), and betweenness (the fraction of the shortest paths that include a node) (Garbutt et al., 2014; McCormack et al., 2016; Mishra et al., 2019). Indeed, it has been shown that hubs (high-degree nodes) and bottlenecks (high-betweenness nodes) are targets of numerous human viral, human bacterial, and other diseases (Abreu et al., 2012; Arabidopsis Interactome Mapping, 2011; Calderwood et al., 2007; de Chassey et al., 2008; Gulbahce et al., 2012; Huttlin et al., 2017; Mukhtar et al., 2011; Pan et al., 2016; Pfefferle et al., 2011; Roohvand et al., 2009; Rozenblatt-Rosen et al., 2012; Shapira et al., 2009; Simonis et al., 2012; Vidal et al., 2011; Wessling et al., 2014). In addition, host protein targets of diverse pathogens were demonstrated to be in close proximity (shortest path) with differentially expressed genes (DEGs) (Mishra et al., 2017). Specific to the viral-host pathosystem, the network analysis of several interactomes including human T cell lymphotropic viruses, Epstein-Barr virus, hepatitis C virus, influenza virus, and human papillomavirus indicate that several of the above-described topological features are associated with viral targets. Recently, in SARS-CoV human interactome, nodes corresponding to hubs and bottlenecks including respiratory chain complex I proteins were identified as targets of SARS-CoV. This system-wide analysis also identified several immunophilins as direct physical interacting partners of the CoV nonstructural protein 1 (Nsp1) (Pfefferle et al., 2011). Importantly, using affinity-purification mass spectrometry (AP-MS), a proteome-scale mapping recently identified 332 SARS-CoV-2 interacting proteins in humans (Gordon et al., 2020). Although this groundbreaking study paved new avenues to investigate novel therapeutic targets using a systems pharmacology approach, integration of spatiotemporal coronavirus-specific expression data is necessary to reveal the cell- and organ-type-specific human-viral interaction architecture. In addition, network biology presents a next-generation, integrative approach for drug repurposing that can predict individual or, more likely, combinatorial sets of drugs with high efficacy against SARS-CoV-2 in the cell-type context (Nabirotchkin et al., 2020; Vitali et al., 2016).

Here, we generated a comprehensive human-SARS-CoV-2 interactome encompassing 12,852 nodes and 84,100 edges. By integrating expressed SARS-CoV-2 and SARS-CoV genes, we created Calu-3-specific human-SARS-CoV-2 and SARS-CoV interactome (CSI) containing 4,176 nodes and 18,630 edges. Our network analysis indicates that the average degree, betweenness, and information centrality of SARS-CoV-2 and SARS-CoV interacting proteins (SIPs) are enriched in CSI. Module-based functional pathway analyses discovered a number of disease-related clusters that are enriched in several signaling pathways and biological processes including eIF2 signaling/translation, striated muscle contraction, protein ubiquitination pathway, viral infection/mRNA transport, T cell receptor regulation of apoptosis, cell cycle, p38 MK2 pathway, and amino acid degradation pathways. Network topology analyses identified 33 high-value targets of SARS-CoV-2, which can form complexes with other highly influential nodes within CSI. These most important nodes are possibly involved in the viral entry, proliferation, and survival in the host tissue as well as required to induce a conducive environment for viral sustenance and pathogenesis. Moreover, we incorporated transcriptome data of COVID-19 patients derived from bronchoalveolar lavage fluid (BALF) cells and peripheral blood mononuclear cells (PBMC) with our CSI data. Subsequently, we performed dynamic

gene regulation modeling on the CSI nodes to decipher the intricate relationships between important transcription factors (TFs) and their target genes upon SARS-CoV-2 infection. Of particular interest is the TF-regulatory relationships involved in host modifying processes such as protein translation, ubiquitination, interferon signaling, and the cytokine storm. In summary, our integrative network topology analyses led us to elucidate the underlying molecular mechanisms and pathways of SARS-CoV-2 pathogenesis.

RESULTS

Integrated Interactome-Transcriptome Analysis to Generate Calu-3-Specific Human-SARS-CoV-2 Interactome

Likely, the outcome of SARS-CoV-2 infection can largely be determined by the interaction patterns of host proteins and viral factors. To build the human-SARS-CoV-2 interactome, we first assembled a comprehensive human interactome encompassing experimentally validated PPIs from the STRING database (Szklarczyk et al., 2015). Because the STRING database is not fully updated, we manually curated PPIs from four additional proteomes-scale interactome studies, i.e., Human Interactome I and II, BioPlex, QUBIC, and CoFrac (reviewed in (Luck et al., 2017). This yielded an experimentally validated interactome containing 18,906 nodes and 444,633 edges (Figure 1A). Subsequently, we compiled an exhaustive list of 394 host proteins interacting with the novel human coronavirus that was referred to as SIPs (Table S1). This comprises 332 human proteins associated with the peptides of SARS-CoV-2 (Gordon et al., 2020), whereas the remaining 62 host proteins interacted with the viral factors of other human coronaviruses including SARS-CoV and MERS-CoV (Zhou et al., 2020), which could also be of significance in understanding the molecular pathogenesis of SARS-CoV-2. By querying these 394 SIPs in the human interactome, we generated a subnetwork of 12,852 nodes and 84,100 edges that covers first and second neighbors of 373 SIPs (Figure 1A). Given that the SIPs-derived PPI subnetwork may not operate in all spatial or temporal conditions, coronavirus-specific expression data are used to filter the interactions in the context of COVID-19. Toward this, we took advantage of a high-resolution temporal transcriptome derived from human airway epithelial cells (Calu-3) treated with SARS-CoV and SARS-CoV-2 over time *in vitro* in culture (Table S1). By integrating this Calu-3 expression data with SIPs-derived PPI subnetwork, we generated CSI that contains 214 SIPs interacting with their first and second neighbors, forming a network of 4,176 nodes and 18,630 edges (Figure 1B and Table S1). Moreover, we performed edge- and node-based comparative analysis to identify the shared and distinct interaction patterns regulating SARS-CoV and SARS-CoV-2 (Figure 1C and Table S1). Finally, we showed that CSI follows a power-law degree distribution with a few nodes harboring increased connectivity compared with a random network and thus exhibits properties of a scale-free network ($r^2 = 0.92$) (Figure 1D and Table S1), similar to other previously generated human-viral interactomes (Ackerman et al., 2018; Becerra et al., 2017; Bosl et al., 2019). Taken together, we constructed a robust, high-quality CSI that was further utilized for network-aided architectural and functional pathway analyses.

Network Topology and Module-Based Functional Analyses Reveal that SARS-CoV-2 Targets Core Signaling Pathways of the Host Network

From the network biology standpoint, a viral infection as well as other pathogen attacks can be viewed as a set of strategic perturbations, at least in part, within the core components of the host interactome (Ackerman et al., 2018; Becerra et al., 2017; Bosl et al., 2019). Because such central nodes correspond to proteins that exhibit increased connectivity and/or central positions within a network, we addressed a question of whether SARS-CoV-2 also attacks such important nodes within CSI. Toward this, we calculated the average degree (number of connections), betweenness (the fraction of all shortest paths that include a node within a network), load centrality (the fraction of all shortest paths that pass through a node), information centrality (the harmonic mean of all the information measures for a node in a connected network), and PageRank index (counting incoming and outgoing connections considering the weight of the edge) for SIPs and compared them with their first and second neighbors. We demonstrated that these five topological features of SIPs were significantly higher than the other nodes within CSI (Figures 2A, B, and C, S1A and S1B, and Table S2; t test $P < 0.01$). We also showed that SIPs were significantly enriched in CSI compared with the human interactome (Figure 2D and Table S2; hypergeometric $P < 5.20 \times 10^{-11}$). These results indicate that SARS-CoV-2 targets core structural components of the human-viral interactome and prompted another question as to whether CSI also activates common biological processes in response to viral infection. Because nodes within CSI form protein complexes among themselves and are transcriptionally regulated, we reasoned that densely connected nodes within this network may participate in similar biological functions. Toward this, we investigated the underlying modular structures (protein clusters ≥ 5 nodes) in CSI followed by Ingenuity Pathway Analysis (IPA). This approach allowed us to identify 25 modules ranging

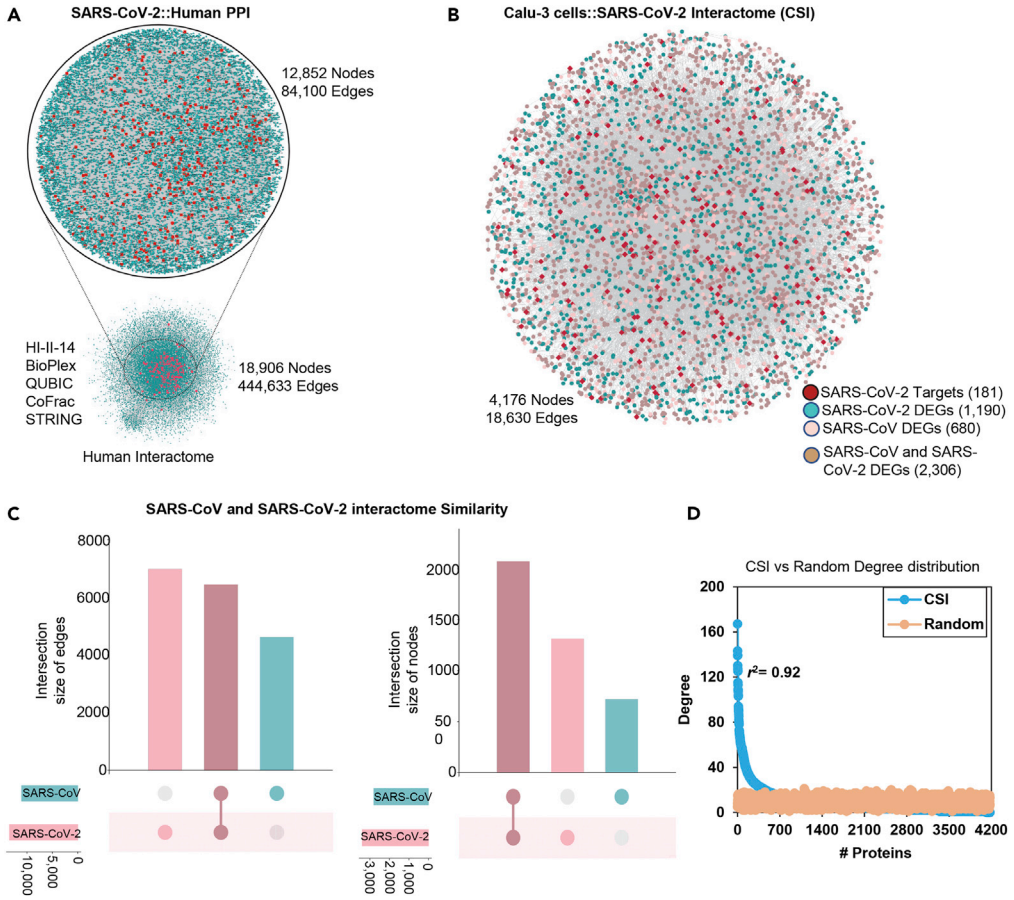


Figure 1. Integrative Multi-Omics Analysis Identified Calu-3-Specific Human-SARS-CoV-2 Interactome (CSI)

(A) Human interactomes (HI-II-14, BioPlex, QUBIC, CoFrac, and STRING) connections and 373 SARS-CoV-2 interacting proteins (SIPs) were used to extract the "SARS-CoV-2:Human PPI" (12,852 nodes and 84,100 edges) including all possible interactions.

(B) Calu-3-specific human-SARS-CoV-2 interactome (CSI) with 4,176 Nodes and 18,630 Edges (Red: 214 SIPs, Turquoise: SARS-CoV-2 DEGs, Pink: SARS-CoV DEGs, LightBrown: SARS-CoV-2 and SARS-CoV DEGs).

(C) The UpSet plot of SARS-CoV and SARS-CoV interactome overlapping nodes and edges.

(D) The degree of CSI nodes displays power-law (r^2 = 0.92) distribution and follow scale-free property.

"See also Table S1."

from 5 to 100 nodes for the smallest and largest modules, respectively. Subsequently, we examined the biological processes, cellular pathways, and signaling cascades that are modulated in the top 8 modules (Figures 2E–L and S1C and Table S2). Significantly enriched signaling pathways and biological processes included eIF2 signaling/translation representing protein translation control, striated muscle contraction, protein ubiquitination pathway, viral infection/mRNA transport, T cell receptor regulation of apoptosis, cell cycle, p38 MK2 pathway, and amino acid degradation (−log(p-value) ≥2) (Figures 2E–L and S1C and Table S2). Finally, we performed a human phenotype ontology analysis that identifies phenotypic abnormalities encountered in human diseases. Significantly enriched terms included loss of speech, hypoventilation, respiratory failure, and abnormality of the common coagulation pathway (Figure S1D, Table S2). Collectively, we showed that SARS-CoV-2 proteins interact with central nodes of CSI, and these proteins are implicated in core molecular and cellular pathways to establish infection and continue disease progress.

Network Topology Framework Identifies the Most Influential Nodes in CSI

Human-viral interactome landscapes of several viruses have previously shown that viral proteins interact with nodes corresponding to high degree (hubs) and high betweenness (bottlenecks), and such structural

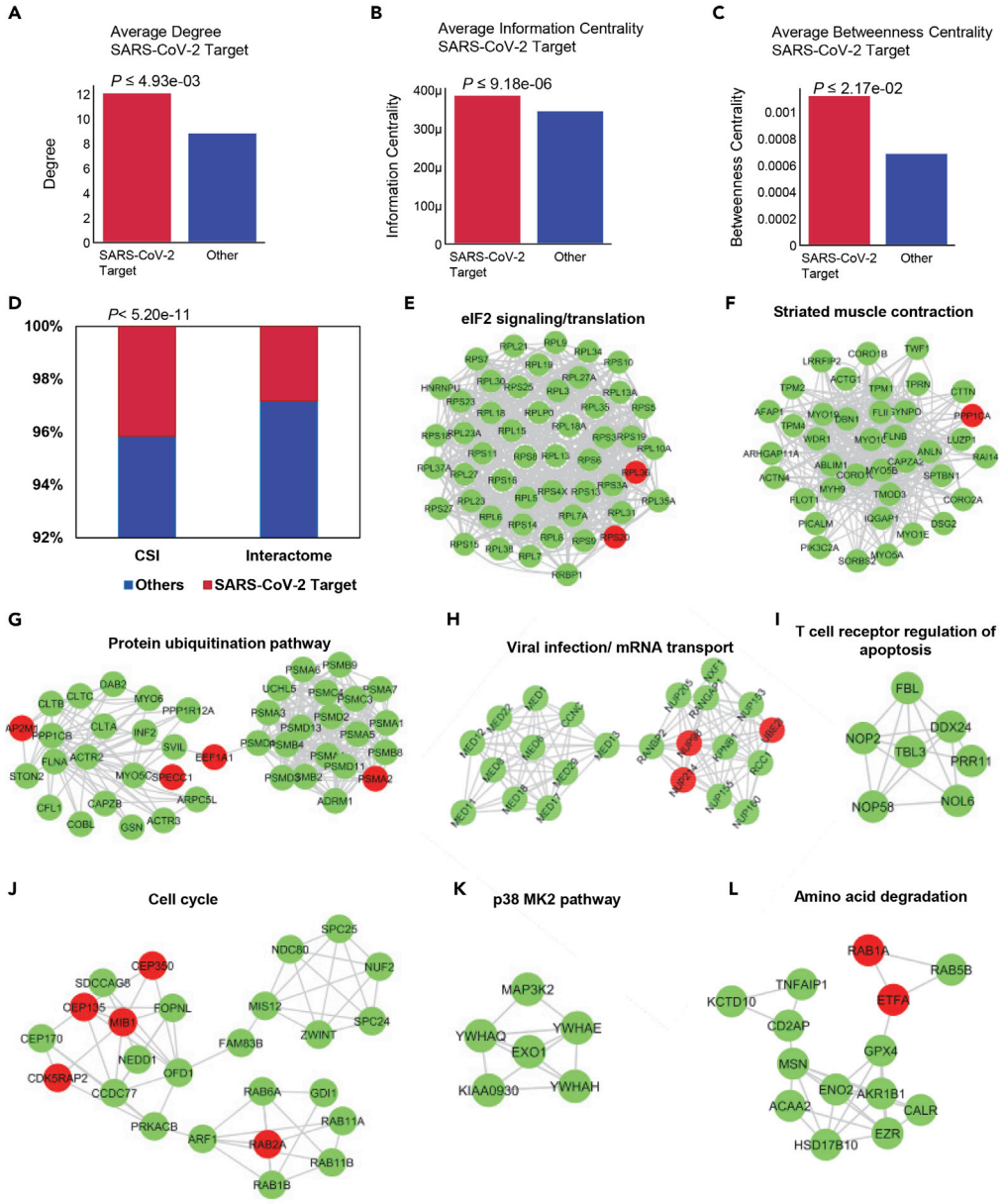


Figure 2. SARS-CoV-2 Interacting Proteins (SIPs) Structural and Functional Properties in Calu-3-Specific Human-SARS-CoV-2 Interactome (CSI)

(A) The average degree of SIPs (12.01) is significantly higher than other interacting proteins (8.78) in the CSI network (*t* test, $P < 4.93 \times 10^{-3}$).

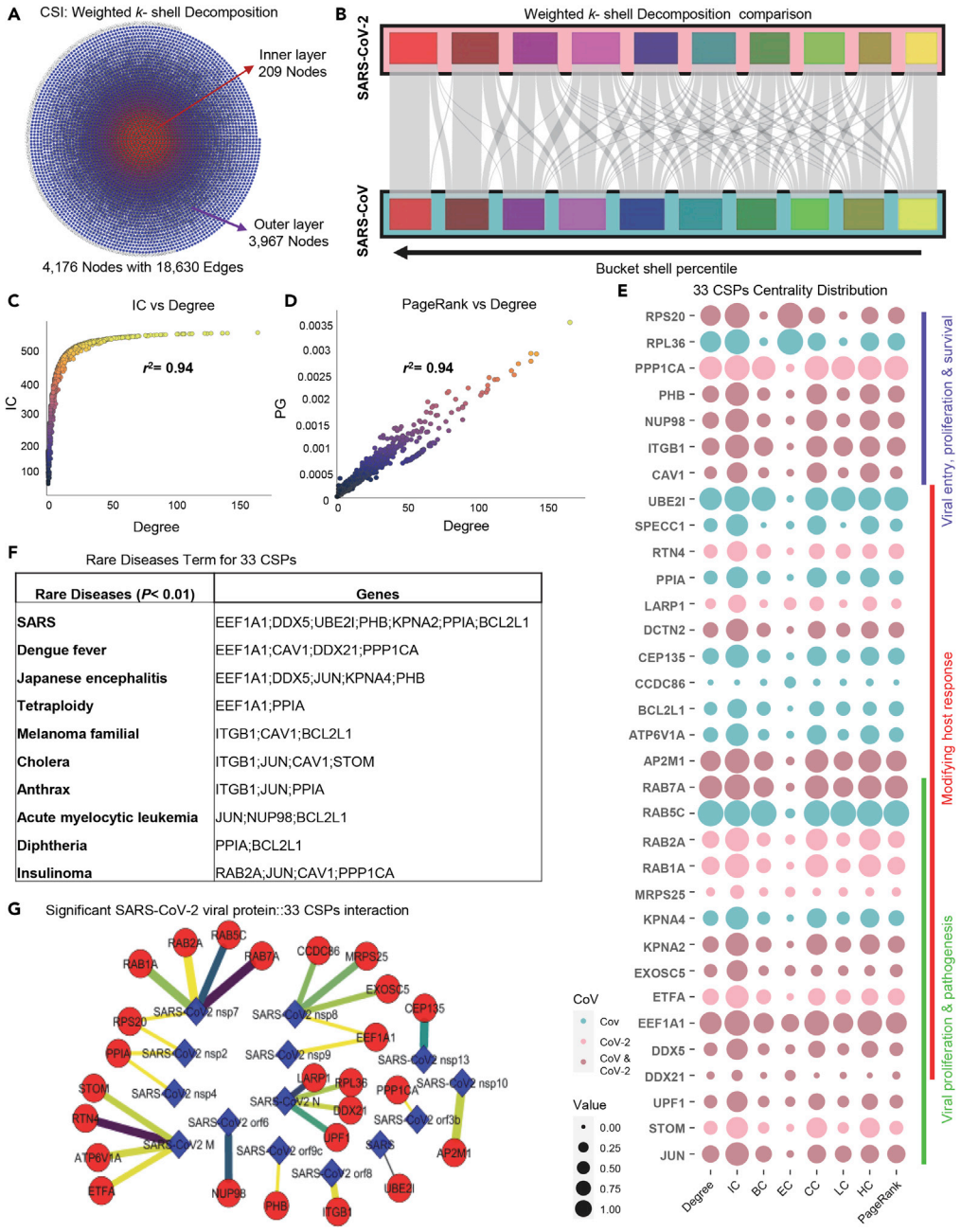
(B) Average information centrality (IC) of SIPs (0.00038) is significantly heightened compared with remaining proteins (0.00034) in CSI network (t test, $P < 9.18 \times 10^{-6}$).

(C) SIPs exhibit significantly increased average betweenness centrality (BW; 0.0011) compared with other interacting proteins (0.00068) in the CSI network (t test, $P < 2.17 \times 10^{-2}$).

(D) SIPs are significantly enriched in CSI network than the human interactome (hypergeometric test, $P < 5.20 \times 10^{-11}$).

(E–L) CSI subnetworks of highly clustered modules were obtained with the application of MCODE Cytoscape app and *k*-means clustering. The significant functional annotation was done by ingenuity pathway analysis (IPA) (Red = SARS-CoV-2 targets, Green = CSI nodes).

"See also Figure S1 and Table S2."



Rare Diseases (P< 0.01)	Genes
SARS	EEF1A1;DDX5;UBE2I;PHB;KPNA2;PPIA;BCL2L1
Dengue fever	EEF1A1;CAV1;DDX21;PPP1CA
Japanese encephalitis	EEF1A1;DDX5;JUN;KPNA4;PHB
Tetraploidy	EEF1A1;PPIA
Melanoma familial	ITGB1;CAV1;BCL2L1
Cholera	ITGB1;JUN;CAV1;STOM
Anthrax	ITGB1;JUN;PPIA
Acute myelocytic leukemia	JUN;NUP98;BCL2L1
Diphtheria	PPIA;BCL2L1
Insulinoma	RAB2A;JUN;CAV1;PPP1CA

Figure 3. Identification of Most Influential Nodes in Calu-3-Specific Human-SARS-CoV-2 Interactome (CSI) Using Network Biology Framework

(A) Weighted *k*-shell decomposition identifies inner and peripheral layers of CSI. Thirty-three percent of all shells are considered as inner layers. Top 5% of the inner layer proteins are considered as significant (Red = inner layer, Blue = peripheral layer).

(B) SARS-CoV-2 and SARS-CoV interactome network similarity based on weighted *k*-shell decomposition. Shells numbers were normalized to percentile and binned into 10 buckets represented as color blocks in the Sankey diagrams. The red block is for 100 to 90 percentile and percentile bucket in yellow represents 10 to zero percentile. The arrow represents the increase in the bucket shell percentile size from zero to 100 with bins of 10 percentile.

(C and D) Correlation between information centrality and degree (C, r^2 = 0.94) and PageRank and degree (D, r^2 = 0.94).

(E) Thirty-three CSI significant proteins (CSPs) that are inner layer SIPs and exhibit more than one high significant centrality measures (degree, information centrality (IC), betweenness centrality (BW), eigenvector centrality (EV), clustering

Figure 3. *Continued*
coefficient (CC), load centrality (LC), harmonic centrality (HC), and PageRank). The size of the blue spot determines the significant central node in centrality indices.
(F) Significantly enriched rare disease terms of 33 CSPs in SARS, Dengue fever, and Japanese encephalitis ($p < 0.0001$) using Enrichr tool.
(G) Network representation of significant SARS-CoV-2 viral protein interaction with 33 CSPs (Nodes: Red = viral proteins, Blue = CSPs significantly targeted by viral protein; edge width = MIST score, edge color = AvgSpec).
"See also Figure S1 and Table S3."

features have been previously used to predict viral targets (Ackerman et al., 2018; Becerra et al., 2017; Bosl et al., 2019). In addition to hubs and bottlenecks, the PageRank algorithm was also effectively used to identify viral targets (Devkota et al., 2018). Moreover, these physical characteristics can also be used to prioritize the most influential genes in CSI for biological relevance and drug target discovery. Here, we used nine different centrality indices to identify the most influential nodes referred to as CSI significant proteins (CSPs). This includes the above-described degree, betweenness, information centrality, PageRank index, and load centrality as well as additional features such as eigenvector centrality (a measure of the influence of a node in a network), clustering coefficient (reciprocal of the sum of the length of the shortest paths between the node and network), harmonic centrality (reverses the sum and reciprocal operations of closeness centrality), and weighted *k*-shell decomposition (an edge weighting method based on adding the degree of two nodes in network partition). Weighted *k*-shell decomposition analysis was recently performed to increase the predictability of host targets of bacterial pathogens (Ahmed et al., 2018). Here, we showed that the top 5% (209) of nodes reside in the inner layers of CSI (Figure 3A and Table S2). Although our qualitative comparative analysis revealed distinct components of SARS-CoV and SARS-CoV-2 in CSI (Figure 1C), we further delineated the varied interaction patterns of common nodes based on weight *k*-shell centrality measure (Figure 3B). This comparative analysis may provide insights into the differential importance of overlapping components of SARS-CoV and SARS-CoV-2 in CSI (Figures 1C and 3B). For other centrality measures, we also maintained a stringent threshold of top 5% to be considered as a highly influential node or CSP. Evidently, we can expect overlapping topological features for the same set of nodes. Noticeably, we observed a strong positive correlation between information centrality and degree (Figure 3C; $r^2 = 0.94$), betweenness and degree (Figure S1E; $r^2 = 0.68$), and PageRank and degree (Figure 3D; $r^2 = 0.94$, Table S2).

Collectively, we identified 33 CSPs that exhibit more than one high centrality measure (Figures 3E and Table S3). For instance, DDX21, DDX5, EEF1A1, EXOSC5, KPNA2, JUN, UPF1, RAB7A, AP2M1, DCTN2, CAV1, ITGB1, NUP98, PHB, and RPS20 that have been implicated in both SARS-CoV and SARS-CoV-2 are enriched in all the centrality measures tested in our study (Figures 3E and F, and Table S3). In addition, RAB5C, PPP1CA, UBE2I, EEF1A1, RAB7A, and AP2M1 are enriched in more than five centrality measures (Figures 3E and F, and Table S3). We categorized these 33 CSPs into three major groups based on their potential roles in COVID-19. Althoguh we expect some, if not all, of these proteins to have more than one function(s), the group-1 CSPs might be largely relevant to modifying host response following SARS-CoV-2 infection (Figure 3E). Moreover, the proteins in the other two groups might be involved in viral entry, proliferation, survival, and pathogenesis as well as cytokine storm (Figure 3E; see details in discussion). Furthermore, we found that these 33 CSPs are targets of some of the well-known SARS-CoV-2 viral proteins. SARS-CoV-2 nsp7 targets most of the CSPs (i.e. five in total), SARS-CoV2 nsp8 targets four CSPs, and SARS-CoV-2 M has four CSPs targets, whereas other SARS-CoV-2 nsps' (2, 4, and 10) and SARS-CoV-2 orfs' (3b, 6, 8, and 9c) possess relatively fewer targets. Intriguingly, two of our CSPs (PPIA and RPS20) are targets of more than one SARS-CoV-2 protein (Figures 3G and S1F), whereas PHB is the target of several viral proteins tested as bait at a low threshold. It is also important to note that PHB is targeted by viral proteins of SARS-CoV (Zhou et al., 2020). These data support previous findings that an individual viral factor can target multiple host nodes and several viral proteins can interact with the same host protein (Ackerman et al., 2018; Becerra et al., 2017; Bosl et al., 2019). Collectively, these data strengthen our notion that centrality measures can be an effective method to predict highly influential nodes, leading us to discover 33 such CSPs.

Dynamic Gene Regulation Modeling and Comparative Transcriptional Analyses Elucidate Core Transcriptional Circuitry and Regulatory Signatures Pertinent to SARS-CoV-2 Infection

To further understand the biological characteristics, regulatory relationships, and molecular events associated with the nodes in CSI, we incorporated transcriptome data of COVID-19 patients derived from bronchoalveolar lavage fluid (BALF) cells and peripheral blood mononuclear cells (PBMC) with our CSI data

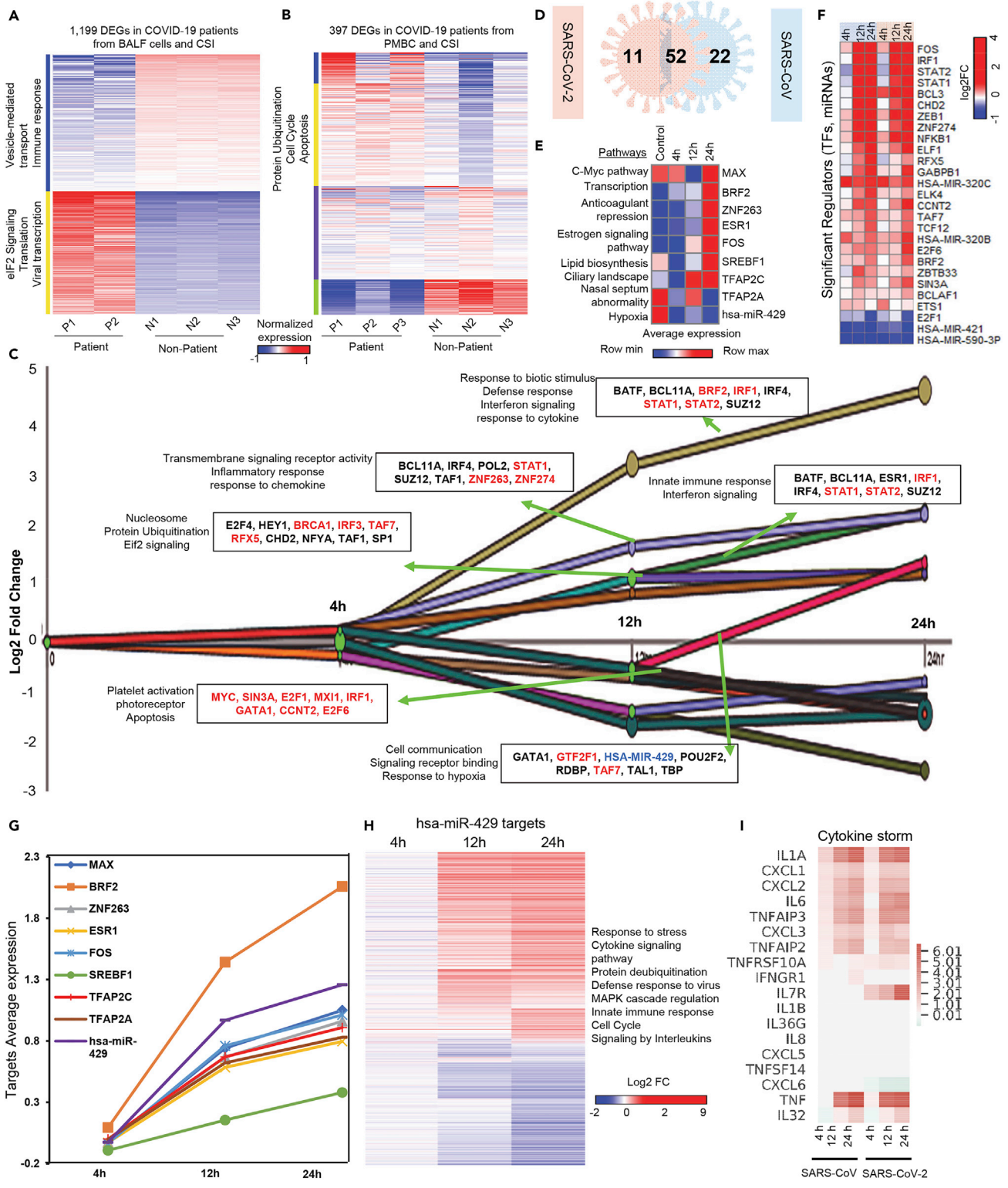


Figure 4. Dynamic Gene Regulation Modeling of Transcriptional Signatures Pertinent to SARS-CoV-2 Infection

(A) Heatmap of 1,199 differentially expressed genes (DEGs) in COVID-19 patients derived bronchoalveolar lavage fluid (BALF) cells transcriptome and Calu-3-specific human-SARS-CoV-2 interactome (CSI). The heatmap was clustered based on *k*-mean with upregulated cluster (570 DEGs) with maximum genes

Figure 4. *Continued*

enriched in eIF2 signaling/translation pathways and viral transcription. Downregulated cluster (629 DEGs) is enriched in dampening the immune response and vesicle-mediated transport. P and N denote patients and controls, respectively.

(B) Heatmap of 397 DEGs common between transcriptome of peripheral blood mononuclear cells (PMBC) derived from P (patients) and N (controls) and Calu-3-specific human-SARS-CoV-2 interactome (CSI). The heatmap was clustered based on *k*-mean with upregulated cluster maximum genes (47 and 155) enriched in protein ubiquitination, cell cycle, and apoptosis.

(D) Venn diagrams illustrating common and unique significant regulators SARS-CoV-2 and SARS-CoV.

(C) Dynamic regulatory event mining of 5,495 cumulative DEGs in SARS-CoV-2 across 24 h of infection reconstructed by incorporating static protein-DNA interaction, microRNA (miRNA)/target interaction data with time series proteomics, and miRNAs expression (GSE148729). The regulators only differentially expressed in Calu-3 SARS-CoV-2 temporal transcriptomes are highlighted. Significant regulators (transcription factors; TFs/miRNAs) control the regulation dynamics ($P < 0.001$ for TFs, $P < 0.05$ for miRNAs). Major highlighted pathways are cytokine storm, interferon signaling, eIF2 signaling/protein ubiquitination, and apoptosis, and hypoxia response occurs during 24 h with a total of 63 regulators (TFs/miRNAs) involved in dynamic modulation.

(E and F) Shared and unique significant regulators (TFs, miRNAs) identified in SARS-CoV-2- and SARS-CoV-infected Calu-3 cells. Unique significant differentially expressed regulator (MAX, BRF2, ZNF263, ESR1, FOS, SREBF1, TFAP2C, TFAP2A, and hsa-miR-429) in SARS-CoV-2-infected Calu-3 cells involved in C-Myc pathway, transcription, anticoagulation, estrogen signaling, lipid biosynthesis, ciliary landscape, nasal septum abnormality, and hypoxia (E) shared significant regulators with differential transcriptional amplitude in response to SARS-CoV-2 compared with SARS-CoV over time (F).

(G) The average expression of target genes of SARS-CoV-2 unique regulator across 24 h of infection in Calu-3 cells.

(H) The expression patterns of hsa-miR-429 target genes in response to SARS-CoV-2 infection at the indicated time points. The enriched pathways are listed.

(I) Heatmap illustrating comparative expression analysis of cytokine-storm-related genes in response to SARS-CoV and SARS-CoV-2 infection at the indicated time.

"See also Figures S2, S3, and S4 and Table S4."

(Xiong et al., 2020). Overall, SARS-CoV-2 infection exhibited largely different transcriptional signatures for BALF cells and PBMC (Xiong et al., 2020). We identified a set of 1,199 and 397 differentially expressed genes (DEGs) in BALF cells and PBMC, respectively ($FDR \leq 0.05$, Figures 4A and B, and Table S4). Thus, CSI constitutes over 50% of transcriptomes pertaining to both BALF cells and PBMC. Intriguingly, in BALF cells, we observed that the upregulated cluster A is enriched with eIF2 signaling/translation and viral transcription pathway, whereas the downregulated clusters are enriched in vesicle-mediated transport and immune response pathway (Figure 4A). Conversely, one major cluster that is significantly upregulated in PBMC is enriched in the cell cycle, T cell receptor regulation of apoptosis, and protein ubiquitination pathway (Figure 4B). Moreover, the mRNAs encoding 33 CSPs exhibit a strong positive correlation between SARS-CoV and SARS-CoV-2 infection at both 12 h ($r^2 = 0.973$) and 24 h ($r^2 = 0.975$) (Figure S1G). Taken together, these data further support the notion that significantly enriched protein modules in CSI are involved in the SARS-CoV-2 pathogenesis.

To reveal the regulatory circuitry and molecular events pertinent to SARS-CoV-2 infection, we performed probabilistic modeling using the iDREM (interactive Dynamic Regulatory Events Miner) framework that incorporates protein-DNA interactions, protein-protein interactions, and miRNA-target interactions with temporal proteomics, small RNA sequencing, and transcriptomics (Ding et al., 2018). Given that iDREM requires time course transcriptional profiling data, we used recently reported temporal (three-time points; 4 h, 12 h, and 24 h) SARS-CoV-2 and SARS-CoV transcriptomics (both RNA- and miR-seq) as well as SARS-CoV-2 proteomics dataset in Calu-3 cell lines (Bojkova et al., 2020; Emanuel et al., 2020). Moreover, we only focused on those upstream transcriptional regulators (transcription factors; TFs and microRNAs; miRNAs) and downstream target genes that were differentially expressed in SARS-CoV-2- and SARS-CoV-mediated dynamic regulatory networks. This dynamic regulatory modeling identified several bifurcation points, where a set of TFs/miRNAs regulates their potential co-expressed and downstream target genes in both viral infections (Figures 4C, S2A, and Table S4). Based on the expression trajectories and path expression patterns, we identified a total of 63 and 74 significant regulators that were expressed in SARS-CoV-2- and SARS-CoV-infected Calu-3 cells, respectively ($P < 0.001$ for TFs and 0.05 for miRNAs, Table S4). By comparing these significant regulators of SARS-CoV-2 with the transcriptome of BALF cells and PBMC COVID-19 patients, we identified 14 significant regulators (MEF2A, IRF1, IRF3, HEY1, ELK4, BRCA1, ZEB1, MEF2C, TCF12, KAT2A, BCL11A, E2F1, RFX5, and SREBF1) (Table S4). Similarly, 19 significant regulators (MEF2A, MAFF, USF1, TRIM28, GATA3, RXRA, EGR1, IRF1, IRF3, HEY1, ELK4, BRCA1, ZEB1, MEF2C, TCF12, KAT2A, BCL11A, E2F1, and RFX5) are shared between the transcriptomes that originated from SARS-CoV and BALF/PBMC COVID-19 patients (Table S4). Taken together, we demonstrated that the significant regulators discovered in SARS-CoV-2-infected Calu-3 cell line resemble the transcriptomes of COVID-19 patients.

During our dynamic regulatory event mining, we observed the first wave of differential regulation and activation of regulators at 12h post-SARS-CoV-2 infection. At this bifurcative transcriptional event, we found a

set of six TFs (STAT1, STAT2, BCL3, CHD2, ZEB1, and NFKB) and two miRNAs (hsa-miR-421 and hsa-miR-429), which were also expressed in the Calu-3 cell's transcriptome. We also found that four TFs (HEY1, ZEB1, RFX5, and E2F1) are differently expressed in BALF and PBMC COVID-19 patients. The next bifurcation occurred at 24 h postinfection, comprising four TFs (STAT1, STAT2, IRF1, and MXI1) and three miRNAs (hsa-miR-320b, hsa-miR-320c, and hsa-miR-590-3p) expressed in SARS-CoV-2-infected Calu-3 cells (Table S4). Overall, we reported 52 significant regulators (TFs and miRNAs) that participate during both 24 h post-SARS-CoV-2 and -SARS-CoV infection, whereas 11 and 22 regulators are involved specifically in SARS-CoV-2 and SARS-CoV infection, respectively (Figures 4D–4F). It is important to note that nine of these eleven significant regulators (MAX, BRF2, ZNF263, ESR1, FOS, SREBF1, TFAP2C, TFAP2A, and hsa-miR-429) are highly differentially regulated in SARS-CoV-2-infected Calu-3 cells (Figure 4E), further substantiating the importance of our discovery. These nine regulators are known to be involved in the C-Myc pathway, transcription, anticoagulation, estrogen signaling, lipid biosynthesis, ciliary landscape, nasal septum abnormality, and hypoxia (Figure 4E). Among the common significant regulators for SARS-CoV-2 and SARS-CoV, we observed differences in transcriptional amplitudes for numerous genes, thereby distinguishing the progression of disease upon infection with these two coronaviruses (Figure 4F). Furthermore, to understand the behavior of SARS-CoV-2 unique regulators, we analyzed the average expression of their target genes. Intriguingly, we found that the average expression of the target genes of these regulators is increasing during SARS-CoV-2 infection (Figure 4G). Here, we also found that hsa-miR-429, which targets over 750 genes involved in response to stress, cytokine signaling pathway, protein de-ubiquitination, defense response to the virus, MAPK cascade regulation, innate immune response, cell cycle, and signaling by interleukins during SARS-CoV-2 infection, is one of the significantly downregulated regulators (Figure 4H). Although we found similar sets of target genes regulated by diverse sets of TFs at different stages of infection, we also discovered multiple combinations of TFs regulating similar sets of downstream genes (Figure 4C and S2A). This reflects the intricate nature of dynamic regulatory relationships between TFs/miRNAs and their targets.

Relationship between SARS-CoV-2 Transcriptional Circuitry and CSI

Next, we primarily focused on the relationship between SARS-CoV-2 transcriptional circuitry and five major pathways/signaling events enriched in CSI, i.e. cytokine storm with interferon signaling, eIF2 signaling/translation, protein ubiquitination pathway, T cell receptor regulation of apoptosis, and response to hypoxia. In the first example of cytokine storm and interferon signaling, we identified a total of 13 TFs (BATF, BCL11A, BRF2, IRF1, IRF4, STAT1, STAT2, SUZ12, POL2, TAF1, ZNF263, ZNF274, and ESR1) that act as significant regulators (Figure 4C, Table S4). Predominantly, we found that six TFs (ZNF263, IRF1, ESR1, STAT2, SUZ12, and STAT1) and one master regulator (IRF1) are early transcriptional players activated at 12 h and 24 h postinfection. It is important to note that the majority of these genes/TFs are related to inflammatory/immune regulatory processes and display dissimilar expression in response to SARS-CoV-2 and SARS-CoV (Figure 4C and S2A). Furthermore, we found eight cytokine storm-related genes (IFNGR1, TNFAIP3, IL32, IL6, TNFRSF10A, TNF, IL7R, and CXCL1) that are a part of CSI indicating that members of CSI participate in a cytokine storm (Figure 4C, Table S4). Moreover, we also observed significant transcriptional amplitude differences for these cytokine storm genes, particularly IL7R and TNFRSF10A in response to SARS-CoV-2 and SARS-CoV (Figure 4I). Also, we found that 40 interferon signaling genes are a part of CSI (Figure S2B and Table S4). Collectively, we showed that CSI nodes also participate in both cytokine storm and interferon signaling.

Similarly, for eIF2 signaling/translation and protein ubiquitination pathways, we identified 10 significant TFs (E2F4, HEY1, BRCA1, IRF3, TAF7, RFX5, CHD2, NFYA, TAF1, and SP1) regulating these pathways over 24-h time course. Interestingly, we report that six TFs (CHD2, BRCA1, RFX5, IRF3, TAF7, and SP1) are a part of the CSI network and four of those TFs (BRCA1, IRF3, TAF7, and RFX5) are significantly expressed in Calu-3 SARS-CoV-2 transcriptome (Figure 4C and Table S4). Although IRF3 is a direct target of SARS-CoV-2 protein, BRCA1 possesses 47 interacting proteins, indicating the significance of network centrality analysis in our study. Similarly, in the gene subset related to T cell receptor regulation of apoptosis, we identified a total of eight TFs (MYC, SIN3A, E2F1, MXI1, IRF1, GATA1, CCNT2, and E2F6) that are differentially expressed in Calu-3 SARS-CoV-2 infection. We report that six of these significant regulators (E2F1, SIN3A, IRF1, CCNT2, MYC, and E2F6) are a part of CSI. Specifically, MYC and SIN3A have 35 and 24 interactions, respectively, and E2F6 is the first neighbor of SARS-CoV-2 target protein (Table S4) in the CSI network. Finally, for the hypoxia response pathway, eight significant regulators (GATA1, GTF2F1, hsa-miR-429, POU2F2, RDBP, TAF7, TAL1, and TBP) are regulating the response to hypoxia along with cell

communication and signaling receptor binding process (Table S4). It is important to note that significant oxygen deprivation is recorded in severe COVID-19 patients (Geier and Geier, 2020). Interestingly, three regulators (GTF2F1, TAF7, and hsa-miR-429) are differentially expressed in Calu-3 SARS-CoV-2 temporal transcriptome and two of them (TAF7 and GTF2F1) are also a part of CSI network (Table S4). Furthermore, using the miRbase target prediction tool combined with the SARS-CoV-2 expression profile, we showed that hsa-miR-429 can potentially regulate over 750 genes involved in SARS-CoV-2 infection (Table S4). Intriguingly, 573 of these 752 hsa-miR-429 target genes are a part of CSI, further substantiating the importance of CSI involvement in the SARS-CoV-2 pathogenesis.

In addition to comparative transcriptional output and dynamic regulatory circuitry analyses between SARS-CoV-2 and SARS-CoV, we also compared the transcriptional amplitude differences between SARS-CoV-2 and other respiratory viruses in different cell types (Figures S2–S4). By comparing the DEGs of transcriptomes originated from A549, Calu-3, and NHBE cells that were infected with the human respiratory syncytial virus (RSV), influenza A/Puerto Rico/8/1934 (H1N1) virus (IAV), human parainfluenza virus 3 (HPIV3), and SARS-CoV-2 (Blanco-Melo et al., 2020), we revealed the significantly altered pathways in SARS-CoV-2 (Figure S2C). We found several immune-related signaling pathways including numerous interleukins signaling cascades significantly enriched in SARS-CoV-2-infected Calu-3 cells (Figure S2C). Subsequently, we revealed the differential regulation of CSP genes, blood coagulation abnormalities, hepatic necrosis, cytokine storm, mitochondrial inheritance, respiratory failure, and SARS-CoV-2 significant regulators in SARS-CoV-2-infected Calu-3 cells (Figures S3 and S4).

Overall, we demonstrated the dynamic transcription patterns of CSI genes and CSPs that participate in the cytokine storm, interferon signaling, eIF2 signaling/translation, protein ubiquitination pathway, and T cell receptor regulation of apoptosis and hypoxia (Table S4). We also identified a set of TFs that potentially regulate the aforementioned pathways' genes at various stages of the SARS-CoV-2 infection in COVID-19 patients. Finally, we showed distinct patterns of key pathway genes in SARS-CoV-2 infection when compared with SARS-CoV and other respiratory viruses.

Translational Implications of This Study

Based on our interactome data as well as published investigations, SARS-CoV-2 proteins interact with human proteins resulting in significant alterations to signaling pathways, which provide conducive micro-environment for replication of SARS-CoV-2 (a purine biosynthesis enzyme IMPDH2 is one such example), activation of multiple protein degradation pathways (Cathepsin B/L, UBE2I etc.), and modulation in the human tissues' metabolic regulation (mTOR pathway and its direct/indirect regulators, e.g., AMPK) (Gordon et al., 2020). These alterations are accompanied by the intense immune reactions manifested as concerted activation of multiple pro-inflammatory signals (epigenetic modifiers including bromodomain proteins, HDACs, ELF1, STATs, JUN, and SIN3A etc.), which could be the underpinning of the cytokine storm. Alterations in interferon response factors (IRFs), as observed here, could be of importance in regard to teasing out the complex underlying mechanism of SARS-CoV-2 pathogenesis. For example, changes in IRF1/IRF3/IRF4 TFs may be related to the observed alterations in the expression of STATs and dependent signaling. Some of these pathways are also key in type I interferon induction (Hadjadj et al., 2020). Accordingly, we believe that effective and prompt therapeutic strategies for blocking SARS-CoV-2 infection are not straight forward. Notably, keeping in view the long asymptomatic latency of virus infection, single-agent treatment could at the most be marginally or partially effective, depending on the stage of the disease progression. However, combinatorial drug treatment approaches developed logically based on the potential inhibition of multiple molecular targets identified in our interactome analysis and other published data may prove highly efficacious in dampening the aggressive and lethal progression of the disease. A cocktail of three or four drugs is also used for the effective management of HIV-1 infection in humans (Pemovska et al., 2018; Pirrone et al., 2011). We propose in our formula that the cocktail of therapeutic agents should include an antiviral drug combined with a protease inhibitor and a metabolic immune modulator/anti-inflammatory agent (NSAID). Because multiple proteases and protein degradative signaling pathways may be involved in the pathogenesis of SARS-CoV-2 infection, it is hard to select one effective drug in this category. We observed in our analysis that calcium channel blockers may be more effective, as calcium is required for the activation of proteases and anti-hypertensive calcium channel blocker drugs may limit the cellular availability of free calcium. Thus, we propose that a cocktail of FDA-approved drugs (for other indications) such as Remdesivir, Verapamil, and Rapamycin may be one such potential combination. Other similar combinations of the FDA-approved drugs may also be contemplated. Interestingly, Massachusetts

General Hospital has initiated a clinical trial of nitric oxide (NO; NCT04305457) in these patient populations. Similar trials are initiated at other locations. These trials are based on the earlier success of NO in SARS-CoV-2 patients. Likewise, an iron chelator and antioxidant deferoxamine (phase I/phase II): (NCT04333550) is reported to be tested in these patients in Iran. Indeed, genes encoding proteins implicated in antioxidants and NO-related processes are also enriched in CSI (Figure S1C). A number of clinical trials have recently been initiated employing anti-inflammatory drugs, naproxen (phase III): (NCT04325633), and hydroxychloroquine (Phase III): (NCT04329611). Therefore, it is likely that depending on the stage of disease progression, these drugs may manifest some advantage and drug combinations may also be modulated accordingly. Similarly, ACE2 inhibition is still considered of importance in the view of a recently initiated clinical trial of its inhibitor losartan (phase I): (NCT04335123) in the COVID-19 patients. In addition, a class of phytochemical agents known as rocaglamides (Iwasaki et al., 2019), which are potent inhibitors of PHBs could be a new class of agents effective against SARS-CoV-2, particularly in combination with an antiviral. Finally, we observed that CSI proteins are related to signaling pathways that regulate metabolism. Thus, changes in cellular metabolism could play an important role in causing morbidity and mortality in these patients. Because caloric restriction and exercise are regulators of these pathways, careful and disease-stage-specific implementation of proper diet and exercise may show some complementary advantages.

DISCUSSION

In the last two decades, intra- and inter-species interactomes have been generated in a number of prokaryotes and eukaryotes including human, mouse, worm, and plant models (Luck et al., 2017; McCormack et al., 2016; Mishra et al., 2019; Vidal et al., 2011). Investigating such interactomes has indicated that diverse cellular networks are governed by universal laws and led to the discovery of shared and distinct molecular components and signaling pathways implicated in viral pathogenicity. In the present study, we constructed a Calu-3-specific human-SARS-CoV-2 interactome (CSI) by integrating the lung-epithelial-cells-specific co-expression network with the human interactome. We determined that CSI displayed features of scale-freeness and was enriched in different centrality measures. The identification of structural modules demonstrated the relationships with a set of functional pathways in CSI. In-depth network analyses revealed 33 most influential nodes. Additional noteworthy findings pertain to SARS-CoV-2 transcriptional signatures, regulatory relationships among diverse pathways in CSI, and overall SARS-CoV-2 pathogenesis including the cytokine storm and interferon signaling.

We constructed a comprehensive and robust CSI, a human-viral interactome that displayed scale-free properties (r^2 = 0.92; Figure 1D). We also showed that the SARS-CoV-2-interacting proteins (SIPs) exhibit increased average centrality indices compared with the remaining proteins in the network (Figure 2, and Figures S1A and S1B). Numerous human-viral interactomes have previously been generated to uncover global principles of viral entry, infection, and disease progression. These include human T cell lymphotropic viruses, Epstein-Barr virus, hepatitis C virus, influenza virus, human papillomavirus, dengue virus, Ebola virus, HIV-1, and SARS-CoV (Ackerman et al., 2018; Becerra et al., 2017; Bosl et al., 2019), and all of these interactomes exhibited a power-law distribution. Another significant tenet of interactomes is the existence of modular units or modules, defined as sets of densely connected clusters within a network that exhibit heightened connectivity among nodes within a module. Such nodes within a module have previously been deemed to possess similar biological functions or belong to the same functional pathways (Vlaic et al., 2018). Since nodes in CSI form protein complexes to coronavirus infection, we extracted several functional modules from our network (Figures 2E–L). Nuclear transport machinery is an attractive target of a wide spectrum of viruses to facilitate the transport of viral RNAs by interfering with host nucleocytoplasmic trafficking and thereby averting the host immune responses (Yarbrough et al., 2014). In our viral infection/mRNA transport module, SARS-CoV-2 was shown to target two nucleoporins, NUP98 and NUP214 (Figure 2H). A question still to be addressed is how SARS-CoV-2 mechanistically targets these nucleoporins to enhance its virulence. Recently, the global posttranslational modification landscape in response to SARS-CoV-2 infection indicated an increase in the phosphorylation of NUP98, a significant CSP in our study (Bouhaddou et al., 2020) (Figure 3E). Given that ORF6 of both SARS-CoV and SARS-CoV-2 interact with NUP98 and SARS-CoV blocks the transport of STAT1 to the nucleus (Frieman et al., 2007), it is highly likely that SARS-CoV-2-mediated NUP98 posttranslational modification would lead to the inhibition of host mRNAs export including STAT1 translocation through the nuclear pore. Another significant CSP in this pathway is UBE2I (Figure 3E), ubiquitin-conjugating enzyme E2I that constitutes the core component in the cell's sumoylation pathway. Because the sumoylation of the nucleoporins is necessary for their functions

(Texari and Stutz, 2015), we hypothesized that interferences in the host nucleocytoplasmic trafficking of mRNAs/proteins are partially dependent on the SARS-CoV-2 interaction with UBE2I.

The ubiquitin proteasome system (UPS) constitutes the major protein degradation system of eukaryotic cells that participates in a wide range of cellular processes and is another critical target of diverse viruses (Tang et al., 2018). UPS plays an indispensable role in fine-tuning the regulation of inflammatory responses. For instance, proteasome-mediated activation of NF-κB regulates the expression of proinflammatory cytokines including tumor necrosis factor alpha (TNF-α), interleukin 1 β (IL-1β), and IL-8. Similarly, UPS is essential for the regulation of leukocyte proliferation (Tang et al., 2018). The UPS is generally considered a double-edged sword in viral pathogenesis. For example, UPS is a powerhouse that eliminates viral proteins to control viral infection, but at the same time viruses hijack UPS machinery for their propagation (Tang et al., 2018). In case of herpes simplex virus type 1, varicella-zoster virus, and simian varicella virus, induction of nuclear factor κB (NF-κB)-mediated host innate immunity is suppressed by the manipulation of UPS components (Ye et al., 2017). Moreover, it was revealed that UPS plays crucial roles at multiple stages of coronavirus infection (Raaben et al., 2010). In our study, the ubiquitin proteasome module was composed of several members of 26S proteasome ATPase or non-ATPase regulatory subunits, which includes PSMA2 (Figure 3E). It still needs to be determined whether PSMA2 plays important roles in the expression of proinflammatory cytokines and is potentially involved in the cytokine storm. While the mechanistic evaluation of SARS-CoV-2 interaction with these two high-value targets needs to be explored, both the mRNA and protein expression corresponding to PSMD8 were recently shown to be decreased by up to 30% in aged keratinocytes (Ishii et al., 2018). Because reduced proteasome activity results in aggregation of aberrant proteins that perturb cellular functions, we hypothesized that SARS-CoV-2 targets these proteins to interfere with endoplasmic reticulum (ER)-mediated cellular responses. Indeed, the aged population is more prone to severe disease pathogenesis (Davies et al., 2020). AP2M1, the μ subunit of clathrin adaptor protein complex 2 (AP-2), is another significant CSP in this pathway (Figure 3E). Interestingly, AP2M1 was previously shown to be essential for the assembly of HCV (Neveu et al., 2012). However, it remains to be determined how SARS-CoV-2 mechanistically targets AP2M1 for its assembly and/or disruption of intracellular trafficking.

Another hallmark of viral pathogenesis is the dysregulation of the host cell cycle in order to create a conducive environment for their replication (Fan et al., 2018). For instance, some viruses suppress the functions of T or B lymphocytes to subvert the immune system, whereas carcinogenic viruses are known to promote cell cycle progression by antagonizing with cell cycle checkpoints (Fan et al., 2018; Tutuncuoglu et al., 2020). In the case of SARS-CoV-2, we identified three centriole assembly proteins, CEP350, CEP135, and CDK5RAP2, interacting with viral factors (Figure 3E), indicating the importance of cell cycle progression in the pathogenesis of this deadly virus. The highly connected module pertains to eIF2 signaling and is comprised of protein translation-related proteins such as RPS and RPLs. Indeed, these ribosomal proteins have been shown to interact with viral RNA for viral proteins biosynthesis and are subsequently required for viral replication in the host cells (Li, 2019). Noteworthy, two ribosomal proteins, RPL36 and RPS20, were found to interact with several SARS-CoV-2 viral factors. Moreover, both of these proteins are also CSPs that harbor increased centrality measures (Figure 3E). Intriguingly, RPS20 has been demonstrated to operate as an immune factor that activates a TLR3-mediated antiviral (Li, 2019). It remains to be addressed whether RPS20 is a "double whammy" target of SARS-CoV-2 for (1) hijacking this important factor for viral translation and replication and (2) suppressing a critical immune signaling pathway. Regardless, ribosomal proteins are critical targets of numerous viruses and play equally essential roles in developing antiviral therapeutics (Li, 2019). Another noteworthy module is the T cell receptor regulation of apoptosis. Indeed, it was recently reported that SARS-CoV-2 infection may cause lymphocyte apoptosis demonstrated by overall cell count and transcriptional signatures in PBMC of COVID-19 patients (Chen et al., 2020; Xiong et al., 2020). Taken together, our module-based functional analyses identified several novel molecular components, structural and functional modules, and overall provided insights into the pathogenesis of SARS-CoV-2.

Our network topology analyses discovered 33 CSPs that have been implicated in several above-described modules and pathways (Figure 3E). To provide a system-wide perspective of the importance of these CSPs in COVID-19, we categorized these CSPs into three groups based on their possible functionality. Group-1 includes CSPs that are potentially relevant to modifying host response following SARS-CoV-2 infection. These include SPECC1, DCTN2, BCL2L1, ATP6V1A, KPNA2, KPNA4, AP2M1, UBE2I, CCDC86, ETFA, RAB1A, RAB2A, RAB5C, and RAB7A (Figure 3E). We hypothesized that such CSPs are important in creating

a protective environment in the host tissue following the viral infection. For example, SPECC1 under the predisposing conditions of oxidative stress regulates TGF-β signaling and autophagy, whereas DCTN2 specifically affects autophagy (Zhang et al., 2020). Moreover, BCL2L1 regulates apoptosis by blocking mitochondrial VDAC channels and is considered important in controlling ROS production and autophagy regulation (Naghdi and Hajnoczky, 2016). Another mitochondrial ribosomal protein, MRPS25, is also targeted by SARS-CoV-2, and if altered it could lead to mitochondrial dysfunction such as those underlying encephalomyopathy. ATP6V1A is considered important in receptor-mediated endocytosis and protein sorting and could also be involved in early autophagy regulation (Merkulova et al., 2015). Similarly, KPNA2 and KPNA4 under hypoxic conditions control the cytoplasmic-nuclear transport mechanism. Interestingly, influenza polymerase nuclear trafficking is known to be regulated by these proteins (Christiansen and Dyrskjot, 2013; Resa-Infante et al., 2019). SARS-CoV-2 could also utilize these or similar groups of proteins for its effective shuttling between various subcellular compartments. Although AP2M1 and UBE2I were discussed earlier, CCDC86, an important cytokine-induced protein that regulates various immune responses, was shown to be expressed in the hippocampus of adult macaque monkeys, indicating the importance of this CSP in neurological disorders during COVID-19 (Shishkov et al., 2013). RAB and RHO group of ras proteins may be involved in augmenting inflammatory signaling pathways (Prashar et al., 2017). Group-2 CSPs that we identified are likely to be hijacked by SARS-CoV-2 for its entry, proliferation, and survival in the host tissue. In this category, one of the most important CSPs is prohibitin (PHB; Figure 3E). PHB is an important protein shown to be a receptor for dengue and chikungunya viruses (Kuadkitkan et al., 2010; Wintachai et al., 2012). Although it has been shown that ACE2 serves as the main receptor for SARS-CoV-2 entry into the cells (Shang et al., 2020), it is quite interesting that pathogenesis of the viral infection is not significantly different between the populations of hypertensive patients who receive or do not receive ACE2 inhibitors (Kuster et al., 2020; Li et al., 2020; Meng et al., 2020; Vaduganathan et al., 2020). Therefore, it is plausible that under certain physiological conditions when SARS-CoV-2 does not engage with the ACE2 receptor for its entry into the cell, PHB or other yet to be discovered proteins serve as an alternative receptors or co-receptors. It also remains to be demonstrated whether viral entry through different receptors is context dependent. Another CSP Integrin β1 encoded by ITGB1 was recently shown to be required for the entry of rabies virus (Shuai et al., 2020). Whether ITGB1 could also promote the entry of SARS-CoV-2 is another question that needs to be addressed. Similarly, PPP1CA was shown to regulate HIV-1 transcription by modulating CDK9 phosphorylation (Tyagi et al., 2015) and thus is potentially involved in the gene regulation of SARS-CoV-2. Another significant CSP, CAV1, regulates aquaporin 4 and is involved in water/glycerin transport and metabolic reprogramming. Dysregulation of this protein is possibly related to neuropathic pain (Yang et al., 2015). Indeed, a number of neurologic impairments were reported to be associated with COVID-19. As discussed earlier, RPL20, RPL36, and NUP98 are other three CSPs that are targeted by SARS-CoV-2 for its entry and proliferation (Figure 3E). Similarly, NUP98 can also be utilized for viral entry into the nucleus. Additional three CSPs in this category, RPL36 and RPS20 (Li, 2019; Lv et al., 2017), could be employed for viral transcription and protein synthesis (Figure 3E). Finally, Group-3 CSPs are proteins that SARS-CoV-2 may utilize both to facilitate its proliferation as well as to induce a conducive environment in the host tissue for its sustenance and pathogenesis (Figure 3E). These CSPs include DDX21, DDX5, UPF1, EXOSC, STOM, and members of Group-1 with overlapping functions. The dead box proteins (DDX21 and DDX5), which have helicase activity, are important in regulating RNA biogenesis, RNA transport, and protein translation, as well as other processes (Cheng et al., 2018). It is important to note that SARS-CoV-2 not only regulates the expression of these genes but also affects many other genes related to protein translation, such as eIF2A. These proteins are known to interact with Jun, a part of the AP-1 transcription factor complex and a significant CSP in our study, which is altered by SARS-CoV-2 infection and plays key roles in regulating oxidative and inflammatory signaling pathways (Gazon et al., 2017). COVID-19 patients are known to manifest upregulated inflammatory and oxidative signaling pathways. In addition, UPF1 is another RNA helicase that is involved in the RNA decay pathway (Kim and Maquat, 2019) and is a target of SARS-CoV-2. It is interesting to note that a subset of COVID-19 patients manifests pulmonary thrombosis and microvascular thrombosis leading to severe morbidity (Ackermann et al., 2020). An important CSP pertinent to this phenomenon in our study is STOM, as its loss of function can cause a type of hemolytic anemia (Genetet et al., 2017). Moreover, Group-3 and Group-1 share CSPs that exhibit overlapping functions. Among these CSPs, EEF1A1, PPIA, RAB1A, RAB2A, RAB5C, RAB7A, and UBE2I are identified as the ones that are potentially associated with the pathogenesis of the cytokine storm as observed in some severely affected patient populations. Intriguingly, EEF1A1, a target of several viruses, is known to be activated upon inflammation (Maruyama et al., 2007). This CSP is independently identified as one of the major regulators in human-SARS-CoV-2-predicted interactome

(Guzzi et al., 2020). The CSPs, which regulate protein folding and translation, could be utilized by SARS-CoV-2 to halt host protein translation, folding, and protein quality control. In addition, we also identified these translation-related proteins as the first neighbors of some of these CSPs. These CSP complexes play key roles in promoting cell death, causing inflammation, and acting enzymatically as viral integrases. Collectively, these CSPs and their first neighbors could, directly and indirectly, perform intricate pathophysiological functions but those mentioned here could be the key effects of COVID-19 on host tissue dysregulation. This classification is also crucial for the design of effective therapeutic interventions against COVID-19. Finally, we presented transcriptional modeling of CSI genes including CSPs that participate in the cytokine storm, interferon signaling, eIF2 signaling/translation, protein ubiquitination pathway, hypoxia, and T cell receptor regulation of apoptosis. Thus, these signaling pathways and TFs discovered through our analyses could provide important clues about effective drug targets and their combinations that can be administered at different stages of COVID-19.

In conclusion, we generated a human-SARS-CoV-2 interactome, integrated virus-related transcriptome to the host interactome, discovered COVID-19 pertinent structural and functional modules, identified high-value viral targets, and performed dynamic transcriptional modeling. Thus, our integrative network-biology-based framework led us to uncover the underlying molecular mechanisms and pathways of SARS-CoV-2 pathogenesis.

Limitations of the Study

The PPI network used in this study is composed of ~19,000 nodes. Thus, the interactome does not represent the full scope of cellular activities and in turn, may lack some targets of SARS-CoV-2. In addition, Gordon et al. (Gordon et al., 2020) viral target dataset is generated in an ex-vivo experiment using individual SARS-CoV-2 proteins expression in HEK-293T/17 cells followed by affinity-purification–mass spectrometry (AP-MS)-based proteomics. We expect both false-positive and false-negative targets in such a study. For instance, Spike protein interaction with ACE-2 was not recapitulated with high confidence. In addition, there is no significant comparable overlap between BALF and PBMC cell transcriptome of COVID-19 patients. This could have facilitated in better understanding of transcriptome-wide differences. Also, a detailed temporal dataset of SARS-CoV-2 infection to death/recovery host transcriptome is lacking, which may provide immense clarity in activated regulators during different stages of infection and death/recovery. Finally, we may expect additional drug-protein interactions besides proposed in this study. Likewise, additional drug combinations may prove to be more effective than proposed.

Resource Availability

Lead Contact

Requests for materials and communications with the journal should be addressed to M.S.M. (smukhtar@uab.edu).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

All datasets used for this study are accessible through Table S files. Standard DESeq method is used for RNASeq Analysis. The Cystoscope tools and NetworkX (2.4) is used for network analysis.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101526.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (IOS-1557796) to M.S.M., and U54 ES 030246 from NIH/NIEHS to M.A. and M.S.M. The authors thank Dr. Karolina Mukhtar for editing and critically reading the manuscript.

AUTHOR CONTRIBUTIONS

M.S.M., N.K., and M.A. conceived the project. N.K., B.M., and A.M. performed network-based and statistical analyses. M.S.M. and M.A. wrote the first draft of the manuscript except for the Methods section. N.K. and B.M. wrote the Methods section. All the authors discussed the results, critically reviewed the manuscript, and provided valuable comments/edits.

DECLARATION OF INTERESTS

The authors declare no competing interests. The authors also declare no financial interests.

Received: April 15, 2020
Revised: July 30, 2020
Accepted: August 31, 2020
Published: September 25, 2020

REFERENCES

Abreu, A.L., Souza, R.P., Gimenes, F., and Consolaro, M.E. (2012). A review of methods for detect human Papillomavirus infection. Virol. J. *9*, 262.

Ackerman, E.E., Kawakami, E., Katoh, M., Watanabe, T., Watanabe, S., Tomita, Y., Lopes, T.J., Matsuoka, Y., Kitano, H., Shoemaker, J.E., et al. (2018). Network-guided discovery of influenza virus replication host factors. mBio *9*, e02002–e02018.

Ackermann, M., Verleden, S.E., Kuehnel, M., Haverich, A., Welte, T., Laenger, F., Vanstapel, A., Werlein, C., Stark, H., Tzankov, A., et al. (2020). Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. N. Engl. J. Med. *383*, 120–128.

Ahmed, H., Howton, T.C., Sun, Y., Weinberger, N., Belkhadir, Y., and Mukhtar, M.S. (2018). Network biology discovers pathogen contact points in host protein-protein interactomes. Nat. Commun. *9*, 2312.

Arabidopsis Interactome Mapping, C. (2011). Evidence for network evolution in an Arabidopsis interactome map. Science *333*, 601–607.

Becerra, A., Bucheli, V.A., and Moreno, P.A. (2017). Prediction of virus-host protein-protein interactions mediated by short linear motifs. BMC Bioinformatics *18*, 163.

Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.C., Uhl, S., Hoagland, D., Moller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., et al. (2020). Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell *181*, 1036–1045.9.

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., and Munch, C. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. Nature *583*, 469–472.

Bosl, K., Ianevski, A., Than, T.T., Andersen, P.I., Kuivanen, S., Teppor, M., Zusinaite, E., Dumpis, U., Vitkauskiene, A., Cox, R.J., et al. (2019). Common nodes of virus-host interaction revealed through an integrated network analysis. Front. Immunol. *10*, 2186.

Bouhaddou, M., Memon, D., Meyer, B., White, K.M., Rezelj, V.V., Correa Marrero, M., Polacco, B.J., Melnyk, J.E., Ulferts, S., Kaake, R.M., et al. (2020). The global phosphorylation landscape of SARS-CoV-2 infection. Cell *182*, 685–712.e19.

Calderwood, M.A., Venkatesan, K., Xing, L., Chase, M.R., Vazquez, A., Holthaus, A.M., Ewence, A.E., Li, N., Hirozane-Kishikawa, T., Hill, D.E., et al. (2007). Epstein-Barr virus and virus human protein interaction maps. Proc. Natl. Acad. Sci. U S A. *104*, 7606–7611.

Cascella, M., Rajnik, M., Cuomo, A., Dulebohn, S.C., and Di Napoli, R. (2020). Features, Evaluation and Treatment Coronavirus (COVID-19) (StatPearls).

Center for Disease Control (2020). Coronavirus (COVID-19). https://www.cdc.gov/coronavirus/2019-ncov/index.html.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet *395*, 507–513.

Cheng, W., Chen, G., Jia, H., He, X., and Jing, Z. (2018). DDX5 RNA helicases: emerging roles in viral infection. Int. J. Mol. Sci. *19*, 1122.

Christiansen, A., and Dyrskjot, L. (2013). The functional role of the novel biomarker karyopherin alpha 2 (KPNA2) in cancer. Cancer Lett. *331*, 18–23.

Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., group, C.C.-w., and Eggo, R.M. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat. Med. *26*, 1205–1211.

de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M.S., Aublin-Gex, A., Agaugue, S., Meiffren, G., Pradezynski, F., Faria, B.F., Chantier, T., et al. (2008). Hepatitis C virus infection protein network. Mol. Syst. Biol. *4*, 230.

Devkota, P., Danzi, M.C., and Wuchty, S. (2018). Beyond degree and betweenness centrality: alternative topological measures to predict viral targets. PLoS One *13*, e0197595.

Ding, J., Hagood, J.S., Ambalavanan, N., Kaminski, N., and Bar-Joseph, Z. (2018). iDREM: interactive visualization of dynamic regulatory networks. PLoS Comput. Biol. *14*, e1006019.

Emanuel, W., Kirstin, M., Vedran, F., Asija, D., Theresa, G.L., Roberto, A., Filippos, K., David, K., Salah, A., Christopher, B., et al. (2020). Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. bioRxiv. https://doi.org/10.1101/2020.05.05.079194.

Fan, Y., Sanyal, S., and Bruzzone, R. (2018). Breaking bad: how viruses subvert the cell cycle. Front. Cell. Infect. Microbiol. *8*, 396.

Frieman, M., Yount, B., Heise, M., Kopecky-Bromberg, S.A., Palese, P., and Baric, R.S. (2007). Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. J. Virol. *81*, 9812–9824.

Garbutt, C.C., Bangalore, P.V., Kannar, P., and Mukhtar, M.S. (2014). Getting to the edge: protein dynamical networks as a new frontier in plant-microbe interactions. Front. Plant Sci. *5*, 312.

Gazon, H., Barbeau, B., Mesnard, J.M., and Peloponese, J.M., Jr. (2017). Hijacking of the AP-1 signaling pathway during development of ATL. Front. Microbiol. *8*, 2686.

Geier, M.R., and Geier, D.A. (2020). Respiratory conditions in coronavirus disease 2019 (COVID-19): important considerations regarding novel treatment strategies to reduce mortality. Med. Hypotheses *140*, 109760.

Genetet, S., Desrames, A., Chouali, Y., Ripoche, P., Lopez, C., and Mouro-Chanteloup, I. (2017). Stomatin modulates the activity of the anion exchanger 1 (AE1, SLC4A1). Sci. Rep. *7*, 46170.

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature *583*, 459–468.

Gulbahce, N., Yan, H., Dricot, A., Padi, M., Byrdsong, D., Franchi, R., Lee, D.S., Rozenblatt-Rosen, O., Mar, J.C., Calderwood, M.A., et al.

(2012). Viral perturbations of host networks reflect disease etiology. Plos Comput. Biol. *8*, e1002531.

Guzzi, P.H., Mercatelli, D., Ceraolo, C., and Giorgi, F.M. (2020). Master regulator analysis of the SARS-CoV-2/human interactome. J. Clin. Med. *9*, 982.

Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Pere, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al. (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science *369*, 718–724.

Hsu, L.Y., Chia, P.Y., and Lim, J.F. (2020). The novel coronavirus (SARS-CoV-2) epidemic. Ann. Acad. Med. Singapore *49*, 1–3.

Huttlin, E.L., Bruckner, R.J., Paulo, J.A., Cannon, J.R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M.P., Parzen, H., et al. (2017). Architecture of the human interactome defines protein communities and disease networks. Nature *545*, 505–509.

Ishii, M.A., Miyachi, K.J., Cheng, B., and Sun, B.K. (2018). Aging-associated decline of epidermal PSMD8 contributes to impaired skin function. J. Invest. Dermatol. *138*, 976–978.

Iwasaki, S., Iwasaki, W., Takahashi, M., Sakamoto, A., Watanabe, C., Shichino, Y., Floor, S.N., Fujiwara, K., Mito, M., Dodo, K., et al. (2019). The translation inhibitor rocaglamide targets a bimolecular cavity between eIF4A and polypurine RNA. Mol. Cell *73*, 738–748 e739.

Kim, Y.K., and Maquat, L.E. (2019). UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. RNA *25*, 407–422.

Kuadkitkan, A., Wikan, N., Fongsaran, C., and Smith, D.R. (2010). Identification and characterization of prohibitin as a receptor protein mediating DENV-2 entry into insect cells. Virology *406*, 149–161.

Kuster, G.M., Pfister, O., Burkard, T., Zhou, Q., Twerenbold, R., Haaf, P., Widmer, A.F., and Osswald, S. (2020). SARS-CoV2: should inhibitors of the renin-angiotensin system be withdrawn in patients with COVID-19? Eur. Heart J. *41*, 1801–1803.

Li, G., Hu, R., and Zhang, X. (2020). Antihypertensive treatment with ACEI/ARB of patients with COVID-19 complicated by hypertension. Hypertens. Res. *43*, 588–590.

Li, S. (2019). Regulation of ribosomal proteins on viral infection. Cells *8*, 508.

Luck, K., Sheynkman, G.M., Zhang, I., and Vidal, M. (2017). Proteome-scale human interactomics. Trends Biochem. Sci. *42*, 342–354.

Lv, H., Dong, W., Qian, G., Wang, J., Li, X., Cao, Z., Lv, Q., Wang, C., Guo, K., and Zhang, Y. (2017). uS10, a novel Npro-interacting protein, inhibits classical swine fever virus replication. J. Gen. Virol. *98*, 1679–1692.

Maruyama, T., Nara, K., Yoshikawa, H., and Suzuki, N. (2007). Txk, a member of the non-receptor tyrosine kinase of the Tec family, forms a complex with poly(ADP-ribose) polymerase 1 and elongation factor 1alpha and regulates interferon-gamma gene transcription in Th1 cells. Clin. Exp. Immunol. *147*, 164–175.

McCormack, M.E., Lopez, J.A., Crocker, T.H., and Mukhtar, M.S. (2016). Making the right connections: network biology and plant immune system dynamics. Curr. Plant Biol. *5*, 2–12.

Meng, J., Xiao, G., Zhang, J., He, X., Ou, M., Bi, J., Yang, R., Di, W., Wang, Z., Li, Z., et al. (2020). Renin-angiotensin system inhibitors improve the clinical outcomes of COVID-19 patients with hypertension. Emerg. Microbes Infect. *9*, 757–760.

Merkulova, M., Paunescu, T.G., Azroyan, A., Marshansky, V., Breton, S., and Brown, D. (2015). Mapping the H(+) (V)-ATPase interactome: identification of proteins involved in trafficking, folding, assembly and phosphorylation. Sci. Rep. *5*, 14827.

Mishra, B., Kumar, N., and Mukhtar, M.S. (2019). Systems biology and machine learning in plant-pathogen interactions. Mol. Plant Microbe Interact. *32*, 45–55.

Mishra, B., Sun, Y., Ahmed, H., Liu, X., and Mukhtar, M.S. (2017). Global temporal dynamic landscape of pathogen-mediated subversion of Arabidopsis innate immunity. Sci. Rep. *7*, 7849.

Mukhtar, M.S., Carvunis, A.R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., Tasan, M., Galli, M., Hao, T., Nishimura, M.T., et al. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. Science *333*, 596–601.

Nabirotchkin, S., Peluffo, A.E., Rinaudo, P., Yu, J., Hajj, R., and Cohen, D. (2020). Next-generation drug repurposing using human genetics and network biology. Curr. Opin. Pharmacol. *S1471-4892*, 30123–30127.

Naghdi, S., and Hajnoczky, G. (2016). VDAC2-specific cellular functions and the underlying structure. Biochim. Biophys. Acta *1863*, 2503–2514.

Neveu, G., Barouch-Bentov, R., Ziv-Av, A., Gerber, D., Jacob, Y., and Einav, S. (2012). Identification and targeting of an interaction between a tyrosine motif within hepatitis C virus core protein and AP2M1 essential for viral assembly. PLoS Pathog. *8*, e1002845.

Pan, A., Lahiri, C., Rajendiran, A., and Shanmugham, B. (2016). Computational analysis of protein interaction networks for infectious diseases. Brief. Bioinform. *17*, 517–526.

Pemovska, T., Bigenzahn, J.W., and Superti-Furga, G. (2018). Recent advances in combinatorial drug screening and synergy scoring. Curr. Opin. Pharmacol. *42*, 102–110.

Pfefferle, S., Schopf, J., Kogl, M., Friedel, C.C., Muller, M.A., Carbajo-Lozoya, J., Stellberger, T., von Dall'Armi, E., Herzog, P., Kallies, S., et al. (2011). The SARS-coronavirus-host interactome: identification of cyclophilins as target for pan-coronavirus inhibitors. Plos Pathog. *7*, e1002331.

Pirrone, V., Thakkar, N., Jacobson, J.M., Wigdahl, B., and Krebs, F.C. (2011). Combinatorial approaches to the prevention and treatment of HIV-1 infection. Antimicrob. Agents Chemother. *55*, 1831–1842.

Prashar, A., Schnettger, L., Bernard, E.M., and Gutierrez, M.G. (2017). Rab GTPases in immunity and inflammation. Front. Cell. Infect. Microbiol. *7*, 435.

Raaben, M., Posthuma, C.C., Verheije, M.H., te Lintelo, E.G., Kikkert, M., Drijfhout, J.W., Snijder, E.J., Rottier, P.J., and de Haan, C.A. (2010). The ubiquitin-proteasome system plays an important role during various stages of the coronavirus infection cycle. J. Virol. *84*, 7869–7879.

Resa-Infante, P., Bonet, J., Thiele, S., Alawi, M., Stanelle-Bertram, S., Tuku, B., Beck, S., Oliva, B., and Gabriel, G. (2019). Alternative interaction sites in the influenza A virus nucleoprotein mediate viral escape from the importin-alpha7 mediated nuclear import pathway. FEBS J. *286*, 3374–3388.

Roohvand, F., Maillard, P., Lavergne, J.P., Boulant, S., Walic, M., Andreo, U., Goueslain, L., Helle, F., Mallet, A., McLauchlan, J., et al. (2009). Initiation of hepatitis C virus infection requires the dynamic microtubule network: role of the viral nucleocapsid protein. J. Biol. Chem. *284*, 13778–13791.

Rozenblatt-Rosen, O., Deo, R.C., Padi, M., Adelmant, G., Calderwood, M.A., Rolland, T., Grace, M., Dricot, A., Askenazi, M., Tavares, M., et al. (2012). Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. Nature *487*, 491–495.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., and Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. Nature *581*, 221–224.

Shapira, S.D., Gat-Viks, I., Shum, B.O., Dricot, A., de Grace, M.M., Wu, L., Gupta, P.B., Hao, T., Silver, S.J., Root, D.E., et al. (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell *139*, 1255–1267.

Shishkov, R., Chervenkov, T., Yamashima, T., and Tonchev, A.B. (2013). Expression of Cyclon/CCDC86, a novel nuclear protein, in the hippocampus of adult non-human primates. J. Neuroimmunol *258*, 96–99.

Shuai, L., Wang, J., Zhao, D., Wen, Z., Ge, J., He, X., Wang, X., and Bu, Z. (2020). Integrin beta1 promotes peripheral entry by Rabies virus. J. Virol. *94*, e01819–19.

Simonis, N., Rual, J.F., Lemmens, I., Boxus, M., Hirozane-Kishikawa, T., Gatot, J.S., Dricot, A., Hao, T., Vertommen, D., Legros, S., et al. (2012). Host-pathogen interactome mapping for HTLV-1 and -2 retroviruses. Retrovirology *9*, 26.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. *43*, D447–D452.

Tang, Q., Wu, P., Chen, H., and Li, G. (2018). Pleiotropic roles of the ubiquitin-proteasome system during viral propagation. Life Sci. *207*, 350–354.

Texari, L., and Stutz, F. (2015). Sumoylation and transcription regulation at nuclear pores. Chromosoma *124*, 45–56.

Tutuncuoglu, B., Cakir, M., Batra, J., Bouhaddou, M., Eckhardt, M., Gordon, D.E., and Krogan, N.J. (2020). The landscape of human cancer proteins targeted by SARS-CoV-2. Cancer Discov. *10*, 916–921.

Tyagi, M., Iordanskiy, S., Ammosova, T., Kumari, N., Smith, K., Breuer, D., Ilatovskiy, A.V., Kont, Y.S., Ivanov, A., Uren, A., et al. (2015). Reactivation of latent HIV-1 provirus via targeting protein phosphatase-1. Retrovirology *12*, 63.

Vaduganathan, M., Vardeny, O., Michel, T., McMurray, J.J.V., Pfeffer, M.A., and Solomon, S.D. (2020). Renin-Angiotensin-aldosterone system inhibitors in patients with Covid-19. N. Engl. J. Med. *382*, 1653–1659.

Vidal, M., Cusick, M.E., and Barabasi, A.L. (2011). Interactome networks and human disease. Cell *144*, 986–998.

Vitali, F., Cohen, L.D., Demartini, A., Amato, A., Eterno, V., Zambelli, A., and Bellazzi, R. (2016). A network-based data integration approach to support drug repurposing and multi-target therapies in triple negative breast cancer. PLoS One *11*, e0162407.

Vlaic, S., Conrad, T., Tokarski-Schnelle, C., Gustafsson, M., Dahmen, U., Guthke, R., and Schuster, S. (2018). ModuleDiscoverer: identification of regulatory modules in protein-protein interaction networks. Sci. Rep. *8*, 433.

Wessling, R., Epple, P., Altmann, S., He, Y., Yang, L., Henz, S.R., McDonald, N., Wiley, K., Bader, K.C., Glasser, C., et al. (2014). Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. Cell Host Microbe *16*, 364–375.

Wintachai, P., Wikan, N., Kuadkitkan, A., Jaimipuk, T., Ubol, S., Pulmanausahakul, R., Auewarakul, P., Kasinrerk, W., Weng, W.Y., Panyasrivanit, M., et al. (2012). Identification of prohibitin as a Chikungunya virus receptor protein. J. Med. Virol. *84*, 1757–1770.

Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., et al. (2020). A new coronavirus associated with human respiratory disease in China. Nature *579*, 265–269.

Xiong, Y., Liu, Y., Cao, L., Wang, D., Guo, M., Jiang, A., Guo, D., Hu, W., Yang, J., Tang, Z., et al. (2020). Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. Emerg. Microbes Infect. *9*, 761–770.

Yang, J.X., Hua, L., Li, Y.Q., Jiang, Y.Y., Han, D., Liu, H., Tang, Q.Q., Yang, X.N., Yin, C., Hao, L.Y., et al. (2015). Caveolin-1 in the anterior cingulate cortex modulates chronic neuropathic pain via regulation of NMDA receptor 2B subunit. J. Neurosci. *35*, 36–52.

Yarbrough, M.L., Mata, M.A., Sakthivel, R., and Fontoura, B.M. (2014). Viral subversion of nucleocytoplasmic trafficking. Traffic *15*, 127–140.

Ye, R., Su, C., Xu, H., and Zheng, C. (2017). Herpes simplex virus 1 ubiquitin-specific protease UL36 abrogates NF-kappaB activation in DNA sensing signal pathway. J. Virol. *91*, e02417–16.

Zhang, B., Liu, Z., Cao, K., Shan, W., Liu, J., Wen, Q., and Wang, R. (2020). Circ-SPECC1 modulates TGFbeta2 and autophagy under oxidative stress by sponging miR-33a to promote hepatocellular carcinoma tumorigenesis. Cancer Med. *9*, 5999–6008.

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

iScience, Volume *23*

Supplemental Information

Integrative Network Biology Framework Elucidates Molecular Mechanisms of SARS-CoV-2 Pathogenesis

Nilesh Kumar, Bharat Mishra, Adeel Mehmood, Mohammad Athar,M Shahid Mukhtar

Supplemental Information

Transparent Methods

Human-SARS-CoV-2 interactome data acquisition

To build human interactome, we assembled a comprehensive protein-protein interactions (PPIs) set comprising experimentally validated PPIs from STRING database (Szklarczyk et al., 2015) and four additional proteomes-scale interactome studies *i.e.* Human Interactome I and II, BioPlex, QUBIC, and CoFrac (reviewed in (Luck et al., 2017). The resulting human interactome had 18,906 nodes (proteins) with 444,633 edges (interactions). Our human interactome contained 1,200 more proteins and 93,189 interactions that were not included in the previous study (Ding et al., 2018). We collected a total of 394 SARS-CoV-2 interacting proteins (SIPs) from two recent studies, encompassing 332 proteins of SARS-CoV-2-human interactions (Szklarczyk et al., 2015), and 62 proteins of SARS-CoV- and MERS-CoV- human interactions (Ding et al., 2018) (Table S1).

Differential gene expression analysis on SARS-CoV, and SARS-CoV-2 datasets

We obtained RNA-seq data for GSE147507 and GSE148729 from the GEO database (Edgar et al., 2002) and used iDEP; an interactive web tool to calculate differential gene expression between infection and mock treatments at their respective time points. Briefly, iDEP utilizes DESeq2 R package. DESeq2 is an R package for the analysis of gene expression RNA-seq data (Love et al., 2014). Specifically, it uses the linear model for analyzing designed experiments and the assessment of differential expression. A threshold of 2 log fold change and FDR ≤ 0.05 was set for differential expression analysis

of all microarray experiments. For comparative study of SARS-CoV-2 expression pattern, we downloaded an expression dataset of RNAs isolated from the bronchoalveolar lavage fluid (BALF) cells and peripheral blood mononuclear cells (PBMC) of COVID-19 patients (Xiong et al., 2020). The criteria for filtering out significant genes were kept as adjusted p-value < 0.05.

Network Integration and Topology Analysis

To extract the Calu-3-specific human-SARS-CoV-2 Interactome (CSI), we integrated the cumulative DEGs during SARS-CoV-2 and SARS-CoV infection in Calu-3 cell and SARS-CoV-2–Human Interactions (12,852 nodes with 84,100 edges) including 373 SIPs. The resulting CSI network contains 4,176 nodes with 18,630 edges including 181 SIPs with all possible interactions, including their first and second neighbors (Table S1). Network topology analyses were performed using NetworkX (version 2.4)(Hagberg et al., 2008). Python (version 3.7.6) package was used except weighted *k*-shell-decomposition, for which we downloaded w*k*-shell-decomposition Cystoscope App (version 1.0) and MCODE (Bader and Hogue, 2003) for cluster extraction. Cytoscape (Version 7.3.2) was used to visualize all the networks (Shannon et al., 2003).

Gene Ontology Functional Enrichment Analysis

The functional enrichment analysis was done by Ingenuity Pathway Analysis (IPA), WikiPathways, GO biological process, ClueGO, and enricher for human phenotype ontology and rare diseases term with their statistically significant parameters (Kuleshov et al., 2016). Heatmaps were made using iDEP and using the Seaborn Python package (Ge et al., 2018).

Reconstructing SARS-CoV-2 Responsive Dynamic Regulatory Events and Comparing the Transcriptional Response of Significant Regulators of SARS-CoV-2 with other Respiratory Viruses

Interactive visualization of dynamic regulatory networks (iDREM) is a method that incorporates static and time series proteomics expression (Bojkova et al., 2020), gene and miRNA expression (GSE148729) post SARS-CoV-2/SARS-CoV-infection datasets to reconstruct condition-specific reaction network in an unsupervised manner (Emanuel et al., 2020). Additionally, the regulatory model identifies specific stimulated pathways and genes, which uses statistical analysis to recognize regulators (TFs/miRNAs) that vary in activity among models. We implemented iDREM on 5,495 and 4,544 cumulative differentially expressed genes across 4, 12 and 24 hours of SARS-CoV-2 and SARS-CoV infection, respectively, along with log2 normalization and all human 954,377 TFs/targets collections from encode database (Consortium, 2012). Additionally, we utilized 108 miRNAs expressed across 4, 12 and 24 h of SARS-CoV-2 and SARS-CoV infection with all 1,097,066 reported miRNA/target interactions collection. Furthermore, we included a recent time course proteomics study (2h, 6h, 12h, and 24h) on SARS-CoV-2 for a third layer of expression where 6h time was considered as 4h for similar time scale to transcriptomics (Bojkova et al., 2020). The dynamic activated pathways regulated by TFs/miRNAs were generated by EBI human gene ontology function. To compare the transcriptional differences between SARS-CoV-2 and other respiratory viruses, we utilized GSE147507 (series 3, 4, 5, 7, 8 and 9) (Blanco-Melo et al., 2020). Specifically, we focused on A549, Calu-3 and NHBE cells that were infected with human respiratory syncytial virus (RSV), influenza A/Puerto Rico/8/1934 (H1N1) virus (IAV), human

parainfluenza virus 3 (HPIV3) and SARS-CoV-2 as previously reported (Blanco-Melo et al., 2020).

Statistical analyses

Hypergeometric test, linear regression (r^2), and Student *t*-test were performed using R version 3.3.1 as well as online Stat Trek tool.

Supplemental Figures

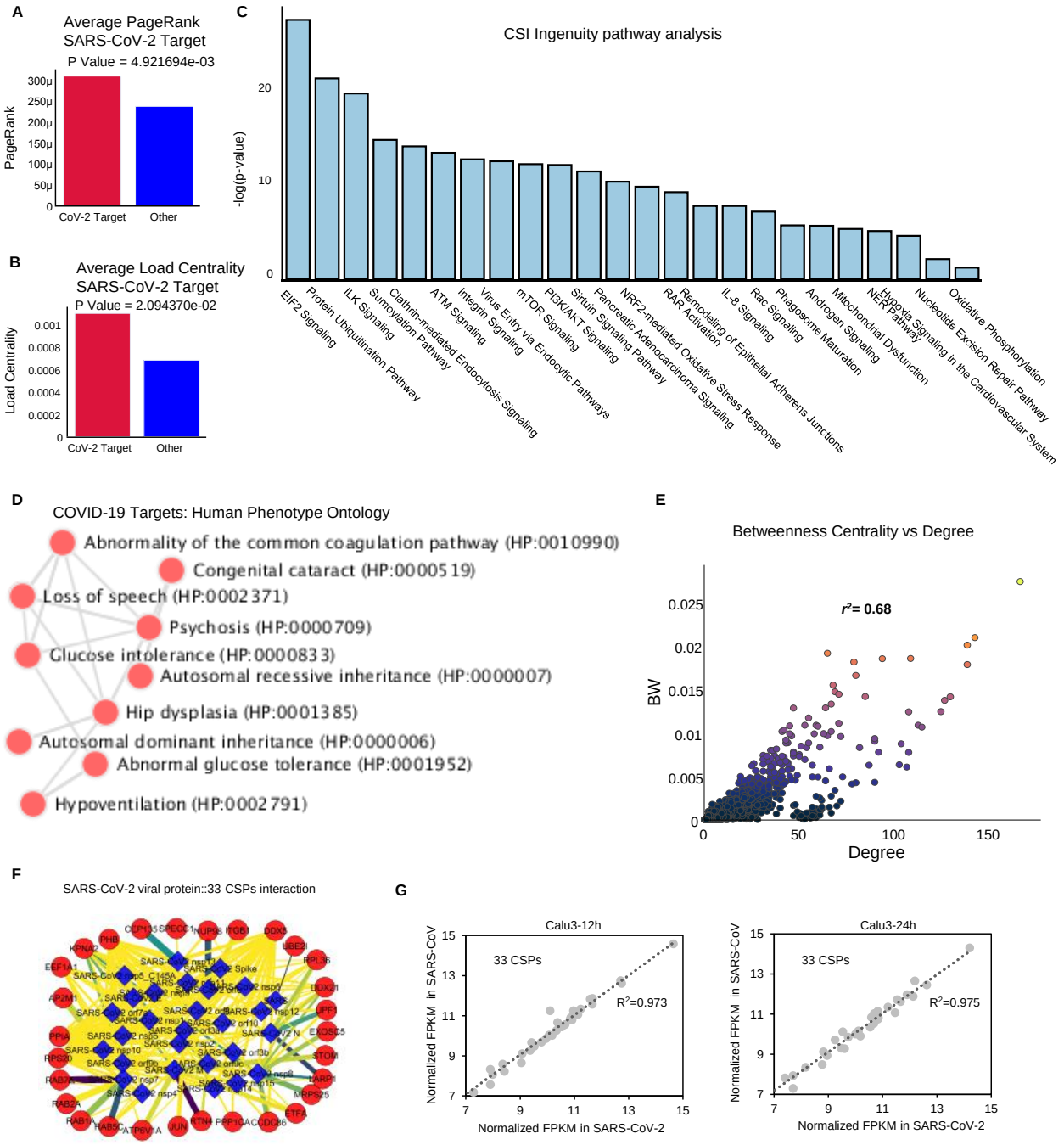


Figure S1: SARS-CoV-2 Interacting Proteins (SIPs) additional structural and functional properties are enriched in Calu-3-specific human-SARS-CoV-2 Interactome (CSI). Related to Figure 2 and 3. (A) Average PageRank of SIPs (300e-6) is significantly higher than that other interacting proteins (240e-06) in CSI network (*t*- test,

P < 4.92e-3). **(B)** SIPs display significantly increased average load centrality of (0.0011) compared to other interacting proteins (0.000672) in CSI network (*t*- test, *P* < 2.09e-02). **(C)** Ingenuity Pathway Analysis (IPA) of identified significantly enriched canonical pathways in CSI proteins (-log(*P*-value) ≥1.2). **(D)** Significantly enriched human phenotype ontology is identified using Enrichr. Abnormality of the common coagulation pathway, developmental cataract, loss of speech, psychosis, abnormal glucose tolerance, glucose intolerance, autosomal recessive inheritance, hip dysplasia, autosomal dominant inheritance, glucose intolerance and hypoventilation terms are enriched. **(E)** Correlation between betweenness and degree (r^2 = 0.68). **(F)** Network representation of SARS-CoV-2 viral protein interaction with 33 CSPs (Nodes: Red= viral proteins, Blue= CSPs significantly targeted by viral protein, grey= CSPs with insignificant viral protein interaction; edge width= MIST score, edge color= AvgSpec). **(G)** Correlation coefficient of the normalized expression levels of 33 CSPs in response to SARS-CoV and SARS-CoV-2 in Calu-3 cells at 12 h (hour; left) and 24h (right) with indicated R^2 values.

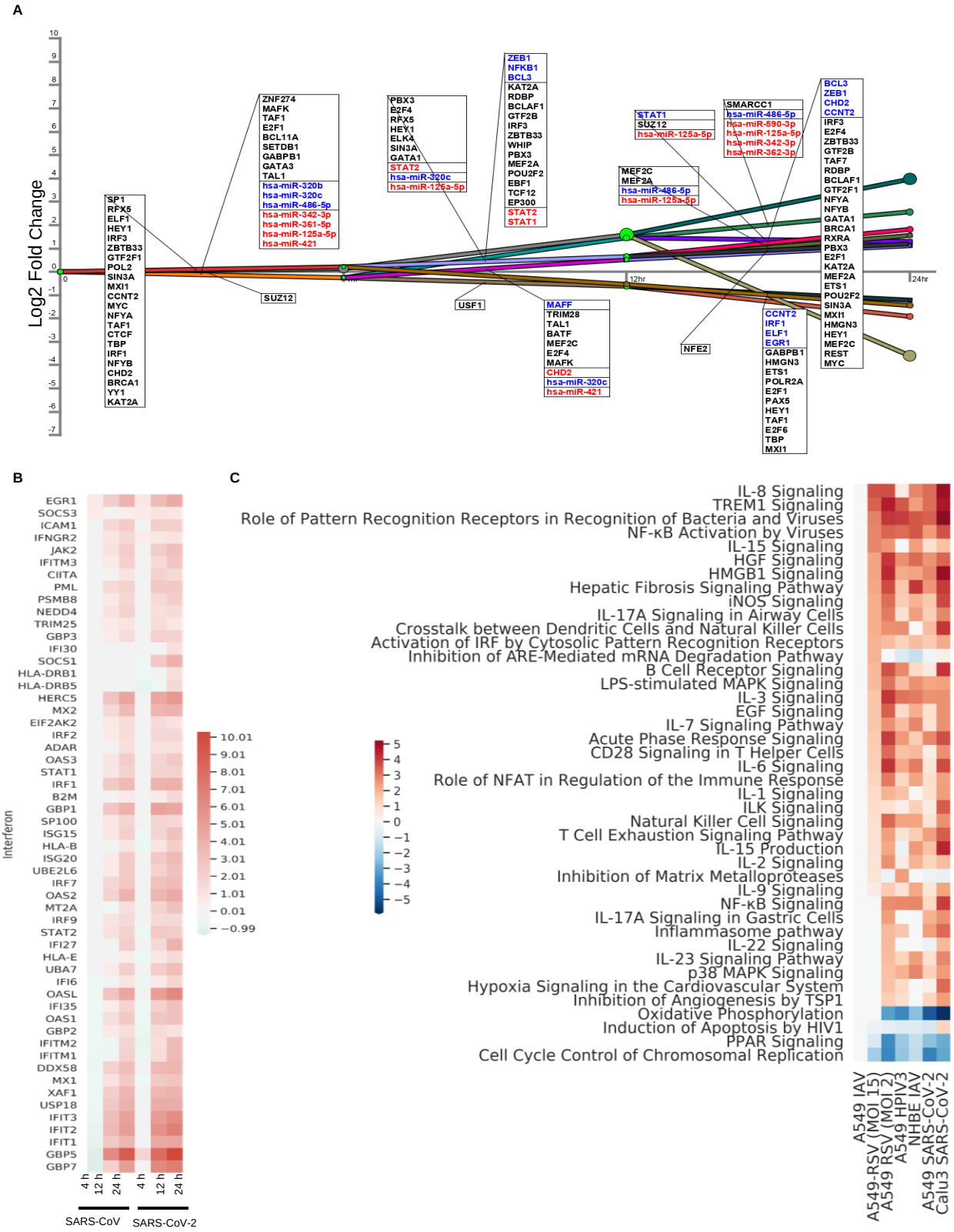


Figure S2: comparative differentially expressed genes (DEGs) analysis between SARS-CoV-2 with other Respiratory Viruses. Related to Figure 4. (A) Dynamic

regulatory event mining of 4,544 cumulative DEGs in SARS-CoV across 24 hours (h) of infection. Significant regulators (Transcription Factors; TFs/miRNAs) control the regulation dynamics (P<0.001 for TFs, P< 0.05 for miRNAs) are highlighted.**(B)** Heatmap depicting DEGs of interferon-related genes in response to SARS-CoV and SARS-CoV-2 infection at the indicated time points. **(C)** Ingenuity Pathway Analysis (IPA) of DEGs derived from A549, Calu-3 and NHBE cells that were infected with human respiratory syncytial virus (RSV), influenza A/Puerto Rico/8/1934 (H1N1) virus (IAV), human parainfluenza virus 3 (HPIV3) and SARS-CoV-2. The level of significant is displayed using heatmap with indicated scale.

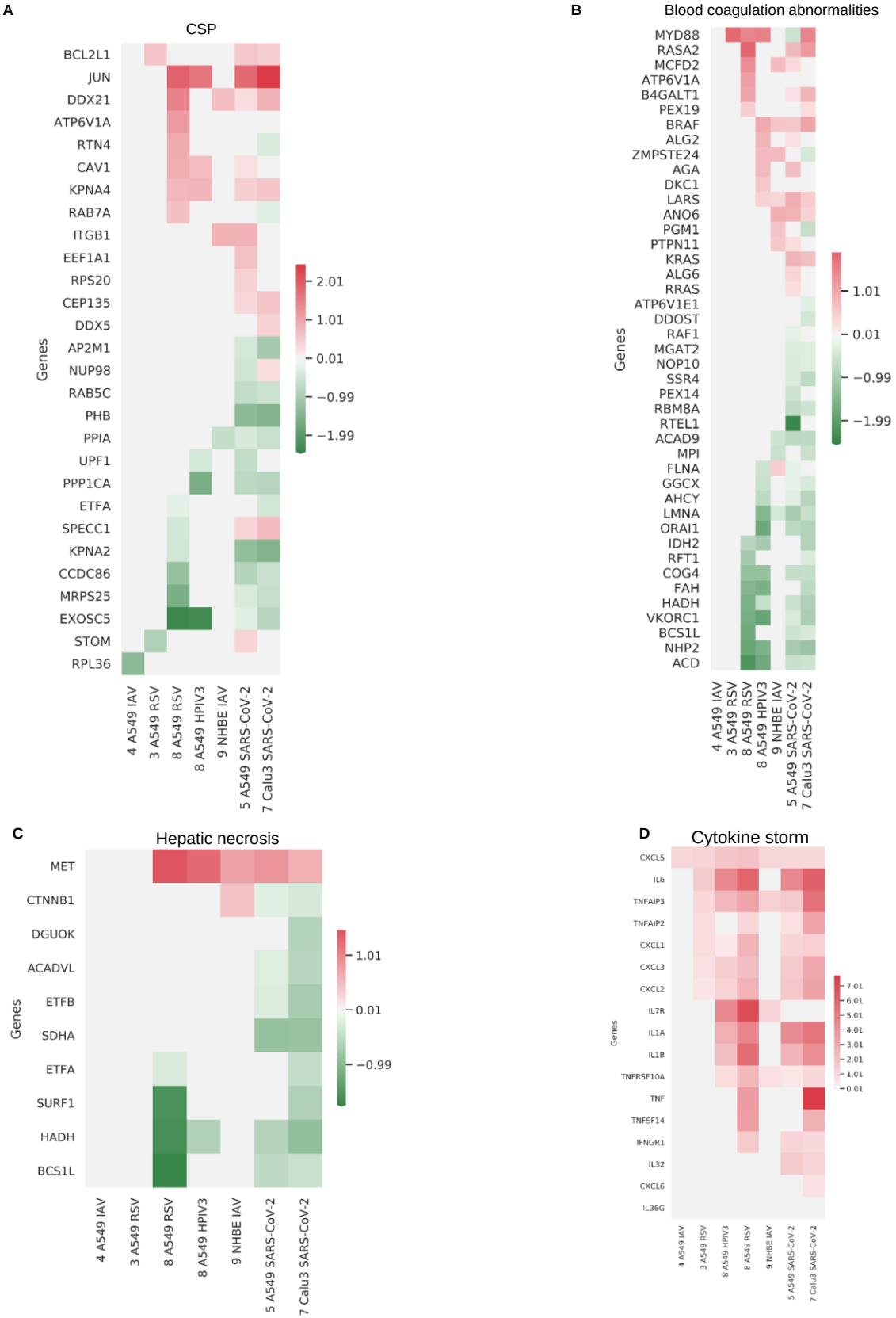

A
CSP
BCL2L1
JUN
DDX21
ATP6V1A
RTN4
CAV1
KPNA4
RAB7A
ITGB1
EEF1A1
RPS20
CEP135
DDX5
AP2M1
NUP98
RAB5C
PHB
PPIA
UPF1
PPP1CA
ETFA
SPECC1
KPNA2
CCDC86
MRPS25
EXOSC5
STOM
RPL36
Genes
2.01
1.01
0.01
−0.99
−1.99
4 A549 IAV
3 A549 RSV
8 A549 RSV
8 A549 HPIV3
9 NHBE IAV
5 A549 SARS-CoV-2
7 Calu3 SARS-CoV-2
B
Blood coagulation abnormalities
MYD88
RASA2
MCFD2
ATP6V1A
B4GALT1
PEX19
BRAF
ALG2
ZMPSTE24
AGA
DKC1
LARS
ANO6
PGM1
PTPN11
KRAS
ALG6
RRAS
ATP6V1E1
DDOST
RAF1
MGAT2
NOP10
SSR4
PEX14
RBM8A
RTEL1
ACAD9
MPI
FLNA
GGCX
AHCY
LMNA
ORAI1
IDH2
RFT1
COG4
FAH
HADH
VKORC1
BCS1L
NHP2
ACD
Genes
1.01
0.01
−0.99
−1.99
4 A549 IAV
3 A549 RSV
8 A549 RSV
8 A549 HPIV3
9 NHBE IAV
5 A549 SARS-CoV-2
7 Calu3 SARS-CoV-2
C
Hepatic necrosis
MET
CTNNB1
DGUOK
ACADVL
ETFB
SDHA
ETFA
SURF1
HADH
BCS1L
Genes
1.01
0.01
−0.99
4 A549 IAV
3 A549 RSV
8 A549 RSV
8 A549 HPIV3
9 NHBE IAV
5 A549 SARS-CoV-2
7 Calu3 SARS-CoV-2
D
Cytokine storm
CXCL5
IL6
TNFAIP3
TNFAIP2
CXCL1
CXCL3
CXCL2
IL7R
IL1A
IL1B
TNFRSF10A
TNF
TNFSF14
IFNGR1
IL32
CXCL6
IL36G
Genes
7.01
6.01
5.01
4.01
3.01
2.01
1.01
0.01
4 A549 IAV
3 A549 RSV
8 A549 HPIV3
8 A549 RSV
9 NHBE IAV
5 A549 SARS-CoV-2
7 Calu3 SARS-CoV-2

Figure S3: Comparative differentially expressed genes (DEGs) analysis among diverse respiratory viruses. Related to Figure 3, Figure 4, and Figure S1. DEGs pertinent to CSI Significant Proteins (33 CSPs) **(A)**, blood coagulation abnormalities **(B)**, hepatic necrosis **(C)** and cytokine storm **(D)** derived from A549, Calu-3 and NHBE cells that were infected with human respiratory syncytial virus (RSV), influenza A/Puerto Rico/8/1934 (H1N1) virus (IAV), human parainfluenza virus 3 (HPIV3) and SARS-CoV-2. The level of significant is displayed using heatmap with indicated scale. Numeral next to the cell and virus types indicate different series from GSE147507 dataset.

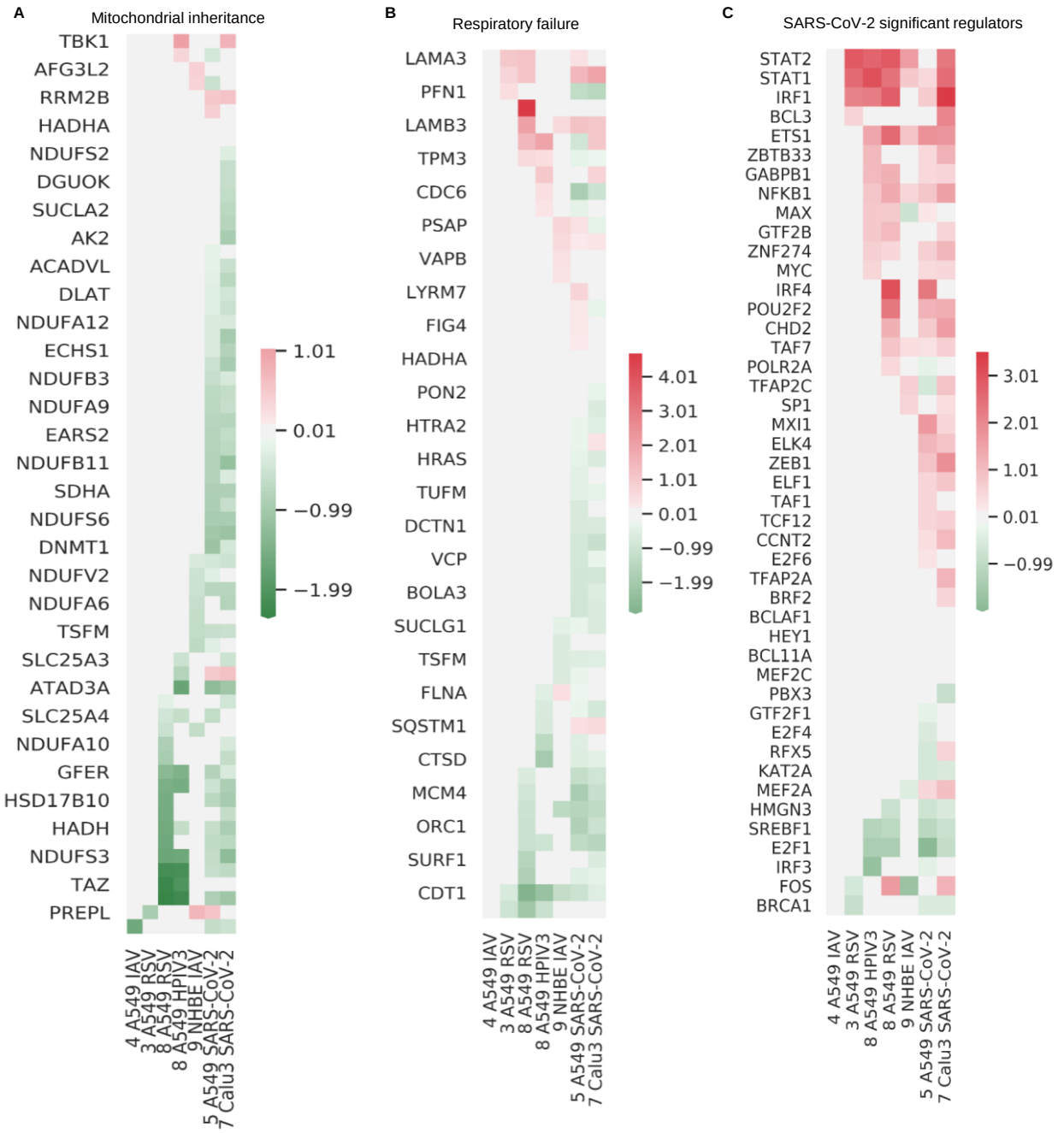


Figure S4: The expression patterns of important pathways-related genes in diverse respiratory viruses. Related to Figure 1, Figure 4, and Figure S2. Transcriptional levels of genes pertinent to mitochondrial inheritance (CSPs) **(A)**, respiratory failure **(B)** and SARS-CoV-2 significant regulators from iDREM analysis **(C)** in the form of heatmap has been displayed. Cell and virus types are indicated. Diverse viruses are RSV :(human respiratory syncytial virus), IAV (influenza A/Puerto Rico/8/1934 (H1N1) virus), HPIV3 (human parainfluenza virus 3) and SARS-CoV-2. Numeral next to the cell and virus types indicate different series from GSE147507 dataset.

Supplemental References:

Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics *4*, 2.

Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.C., Uhl, S., Hoagland, D., Moller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D.*, et al.* (2020). Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. Cell *181*, 1036-1045 e1039.

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

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Ding, J., Hagood, J.S., Ambalavanan, N., Kaminski, N., and Bar-Joseph, Z. (2018). iDREM: Interactive visualization of dynamic regulatory networks. PLoS Comput Biol *14*, e1006019.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res *30*, 207-210.

Emanuel, W., Kirstin, M., Vedran, F., Asija, D., Theresa, G.L., Roberto, A., Filippos, K., David, K., Salah, A., Christopher, B.*, et al.* (2020). Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. bioRxiv.

Ge, S.X., Son, E.W., and Yao, R. (2018). iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. BMC Bioinformatics *19*, 534.

Hagberg, A., Swart, P., and S Chult, D. (2008). Exploring network structure, dynamics, and function using NetworkX Los Alamos National Lab (No. LA-UR-08-05495; LA-UR-05408-05495).

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A.*, et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res *44*, W90-97.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Luck, K., Sheynkman, G.M., Zhang, I., and Vidal, M. (2017). Proteome-Scale Human Interactomics. Trends Biochem Sci *42*, 342-354.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P.*, et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res *43*, D447-452.

Xiong, Y., Liu, Y., Cao, L., Wang, D., Guo, M., Jiang, A., Guo, D., Hu, W., Yang, J., Tang, Z.*, et al.* (2020). Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. Emerg Microbes Infect *9*, 761-770.