

## Article

# Identification of Transcription Factors Regulating SARS-CoV-2 Tropism Factor Expression by Inferring Cell-Type-Specific Transcriptional Regulatory Networks in Human Lungs

Haonan Tong <sup>1,†</sup>, Hao Chen <sup>2,3,\*,†</sup> and Cranos M. Williams <sup>1,\*</sup>

<sup>1</sup> Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695, USA; htong@ncsu.edu

<sup>2</sup> Program in Genetics, Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC 27695, USA

<sup>3</sup> College of Forestry, Shandong Agricultural University, Tai'an 271018, China

\* Correspondence: hchen22@ncsu.edu (H.C.); cmwilli5@ncsu.edu (C.M.W.)

† These authors contributed equally to this work.

**Abstract:** Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus that caused the coronavirus disease 2019 (COVID-19) pandemic. Though previous studies have suggested that SARS-CoV-2 cellular tropism depends on the host-cell-expressed proteins, whether transcriptional regulation controls SARS-CoV-2 tropism factors in human lung cells remains unclear. In this study, we used computational approaches to identify transcription factors (TFs) regulating SARS-CoV-2 tropism for different types of lung cells. We constructed transcriptional regulatory networks (TRNs) controlling SARS-CoV-2 tropism factors for healthy donors and COVID-19 patients using lung single-cell RNA-sequencing (scRNA-seq) data. Through differential network analysis, we found that the altered regulatory role of TFs in the same cell types of healthy and SARS-CoV-2-infected networks may be partially responsible for differential tropism factor expression. In addition, we identified the TFs with high centralities from each cell type and proposed currently available drugs that target these TFs as potential candidates for the treatment of SARS-CoV-2 infection. Altogether, our work provides valuable cell-type-specific TRN models for understanding the transcriptional regulation and gene expression of SARS-CoV-2 tropism factors.

**Keywords:** SARS-CoV-2; single-cell RNA-seq; differential network analysis



**Citation:** Tong, H.; Chen, H.; Williams, C.M. Identification of Transcription Factors Regulating SARS-CoV-2 Tropism Factor Expression by Inferring Cell-Type-Specific Transcriptional Regulatory Networks in Human Lungs. *Viruses* **2022**, *14*, 837. <https://doi.org/10.3390/v14040837>

Academic Editor: Petri Susi

Received: 3 March 2022

Accepted: 5 April 2022

Published: 17 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Coronavirus disease 2019 (COVID-19), caused by SARS-CoV-2, has led to a global public health crisis. Since the first case was reported in late December 2019, SARS-CoV-2 has spread to 215 countries, infected more than 437 million humans, and has led to more than six million deaths, primarily among the elderly [1]. COVID-19 has a high mortality, which is largely due to acute respiratory distress syndrome in the lungs [2]. Since SARS-CoV-2 is mainly transmitted via respiratory droplets, the lung is the preliminary target organ for SARS-CoV-2 infection [3]. Pathological investigations, including postmortem biopsies, have confirmed major pulmonary damage in the lungs as the most likely cause of death, and lifelong damage has been seen in the cases examined [4–6]. Understanding how SARS-CoV-2 infects human lung cells is important in order to identify effective treatment options for COVID-19.

SARS-CoV-2 infection in cells (SARS-CoV-2 cellular tropism) is caused by the binding of the viral spike (S) proteins to cellular receptors (e.g., angiotensin-converting enzyme 2; ACE2) and S protein priming by host cell proteases (e.g., transmembrane protease serine protease 2; TMPRSS2) [7–9]. The co-expression of ACE2 and TMPRSS2 is widely used as a marker for identifying cells that have the potential to be infected by SARS-CoV-2 [7]. Therefore, many studies have revealed a subset of tissues and cell types that

are potentially susceptible to SARS-CoV-2 using single-cell RNA-seq (scRNA-seq) and immunocytochemistry for profiling the co-expression of ACE2 and TMPRSS2 across healthy human tissues [10]. However, ACE2 is not the only functional receptor for SARS-CoV-2. The cell-surface receptors neuropilin-1 NRP1 and CD209 also facilitate the entry of SARS-CoV-2 to specific cells [11–14]. Additionally, recent studies have tested whether other coronavirus (SARS-CoV, MERS-CoV, and HCoV) receptors work as SARS-CoV-2 receptors. The results have shown that MERS receptor DPP4, SARS-CoV receptors CLEC4G/M, and HCoV-229E receptor CD13 cannot be SARS-CoV-2 receptors because of their inability to bind with SARS-CoV-2 [15–17]. The SARS-CoV-2 S protein has been found to be cleaved by several host cell proteases, including CTSL [18], FURIN [19,20], TMPRSS4 [21], and TMPRSS11A [22]. Additionally, in response to virus entry, cells usually express interferon-induced transmembrane proteins (IFITMs) that block the fusion capacity of many enveloped RNA viruses [23]. SARS-CoV-2, however, can hijack IFITM1, IFITM2, and IFITM3 to enhance the infection of the viruses [24,25]; thus, IFITMs serve as factors helping the entry of SARS-CoV-2 into cells. In summary, SARS-CoV-2 cellular tropism depends on the host-cell-expressed proteins (cellular tropism factors), which include receptors (ACE2, NRP1, CD209), proteases (TMPRSS2, 4, 11A, CTSL, and FURIN), and IFITMs (IFITM 1, 2, and 3).

There is a lack of experimental data that identify the transcriptional factors (TFs) that regulate the cellular tropism factors in lung cells. The spatial–temporal expression profiles of these SARS-CoV-2 cellular tropism factor proteins are largely determined by their upstream TFs, whose regulatory roles can differ depending on the cell types involved. The construction of transcriptional regulatory networks (TRNs) can identify potential upstream TFs controlling SARS-CoV-2 tropism factors and elucidate how regulatory relationships are established between TFs and targets [26,27]. The susceptibility of each type of lung cell to SARS-CoV-2 infection can differ widely [18,28,29] and the expression of tropism factors in each cell is heterogeneous [30–32], suggesting that there is heterogeneity within each lung cell type in terms of regulating the expression of cellular tropism factors. Therefore, it is essential to construct TRNs in a cell-type-specific way in order to identify the potential regulators controlling SARS-CoV-2 tropism factors in different types of lung cells.

Cell type-specific TRNs that are constructed before and after SARS-CoV-2 infection could be used to compare and identify changes in gene regulation between a healthy and a diseased condition using differential network analysis [33,34]. The differential analysis of TRNs can identify topological differences between networks, which can be quantified using network centrality measures [35]. The most commonly used centrality measures for TRN include degree, betweenness, closeness, pagerank, and eigenvalue [36]. Using these measures, the nodes (TFs) with an important role in regulating target genes can be identified in each network [37,38]. The centrality-based analysis of the TFs in the TRNs of COVID-19-infected donors and healthy donors can not only provide insight into the transcriptional regulation alteration of SARS-CoV-2 tropism after virus infection but can also help identify the candidate nodes (TFs) that may serve as targets for existing drugs to inhibit the virus entry into the human lungs.

Here, we applied a network approach to investigate the regulators of SARS-CoV-2 tropism factors using human lung scRNA-seq data from healthy and COVID-19-infected donors. To study the heterogeneity of the SARS-CoV-2 tropism factors between cell types, we identified cell-type-specific tropism-factor-associated gene modules and potential regulators of these modules specific to each cell type. We then refined the connections within cell-type-specific networks using motif information including ChIP-Seq. This allowed for the identification of candidate TFs that directly regulated the SARS-CoV-2 tropism factors within each cell type. This type of network was constructed for each type of lung cell from healthy donors and COVID-19 patients. Centrality-based differential network analyses were performed for the same type of lung cell across healthy donors and COVID-19 patients and were used to identify how SARS-CoV-2 affected the regulatory signatures of specific cell types. Finally, we applied a centrality-based approach to identify the master TFs from each cell-type-specific TRN that regulate tropism factors. The existing drugs targeting these

master TFs were identified as candidates that could be repurposed to mitigate SARS-CoV-2 infection.

## 2. Materials and Methods

### 2.1. Acquisition and Precession of Public Datasets

Single-cell RNA sequence (scRNA-seq) data were retrieved from NCBI, EBI, and the European Genome-Phenome Archive (EGA), including bronchoalveolar lavage fluid (BALF) from 16 samples taken from COVID-19 patients and 16 samples taken from healthy donors. For the donors of COVID-19 patients, 9 samples came from Liao et al. [39] (GSE145926: C141, C142, C143, C144, C145, C146, C148, C149, C152), 2 samples came from Chua et al. [40] (EGAS00001004481: BIHCoV01\_BL\_1 and BIHCoV04\_BL\_1), and 5 samples came from Grant et al. [41] (GSE155249: GSM4698176, GSM4698177, GSM4698178, GSM4698180, GSM4698182). For healthy donors, 3 samples came from Liao et al. [39] (GSE145926: C51, C52, C100), 1 sample came from Morse et al. [42] (GSE128033), 4 human bronchial epithelial cell samples came from Lukassen et al. [43] (EGAD00001006185), and 8 samples came from Reyfman et al. [44] (GSE122960: GSM3489182, GSM3489185, GSM3489187, GSM3489189, GSM3489191, GSM3489193, GSM3489195, GSM3489197). The clinical characteristics of the scRNA-seq samples are summarized in Table S1. All raw data and processed data used for analysis are available upon request.

### 2.2. Assembly of Multiple Distinct scRNA-Seq Datasets

The Seurat framework [45] was used for the comprehensive integration of single-cell data. Quality control was performed for each gene-cell-barcode matrix by examining the mitochondrial contamination (mitochondrial gene percentage < 0.1), the total number of detected molecules (unique molecular identifier count > 1000), and the number of unique genes in a cell (filtering cells that had unique feature counts of less than 200). Feature expression measurements were normalized for each cell by dividing the total counts and multiplying by a scale factor of 10,000, followed by log-transformation (natural logarithm of the given value plus one). For each dataset, the top 2000 variable genes with the highest standardized variance were determined using variance stabilizing transformation [46]. These features were then ranked by the number of datasets they were deemed to be variable in, breaking ties using the median variable feature rank across datasets. The top 2000 genes were selected. The cell pairs were identified by searching for mutual nearest neighbors (MNNs) [47] in a low-dimensional representation (90 dimensions to use) of both datasets from canonical correlation analysis (CCA) [48]. The identified anchors were used to compute all pairwise distances between datasets (total number of cells in the smaller dataset divided by the total number of anchors between the two datasets). The resulting distance matrix was clustered to determine a guide tree. The datasets were then iteratively merged by calculating a transformation matrix based on corresponding anchors [45] for dimensionality reduction and clustering.

### 2.3. Unsupervised Dimensionality Reduction and Clustering

We performed unsupervised dimensionality reduction and clustering according to the method of [45]. A linear transformation was applied by shifting and scaling the expression of each gene to be zero mean and unit variance across cells. Then, a principal component analysis (PCA) was performed on the scaled integrated data. The detected principal components (PCs) were ranked based on their percentage of variance and it was observed that 50 PCs captured the majority of the true signal. Among density-based clustering algorithms that can be used to partition gene expression into arbitrary-sized clusters [49], we used a graph-based clustering analysis, as its clustering results can reflect the cell types involved with a high accuracy [50]. A K-nearest neighbor (KNN) graph of cells was constructed based on the Euclidean distance in PCA space. A shared nearest neighbor (SNN) graph was constructed by calculating the neighborhood overlap (Jaccard index) between every cell and its 20 nearest neighbors in the KNN graph. The Louvain algorithm

was used as the modularity optimization technique for extracting communities from the constructed SNN graph [51,52]. The graph-based clustering results were visualized using Uniform Manifold Approximation and Projection (UMAP) based on the identified top 50 PCs.

#### 2.4. Marker Gene Selection and the Classification of Cell Types

Marker genes from each cluster were identified via performing Wilcoxon rank sum test on uncorrected gene expressions, which showed an at least 0.25-fold difference (log-scale) between the two groups of cells. We manually annotated cell clusters with broad cell-type classes using the top 2 cluster marker genes compared to the canonical lung cell-type markers annotated in [53]. Clusters without clear markers for distinguishing types were excluded from further analysis.

#### 2.5. Module Generation and Direct TF–Target Regulon Identification

Potential regulators were identified for each cell type as TFs whose expression was detected in 0.5% of cells. GRNBoost2 [54] was then used to predict cell-type-specific weighted regulatory interactions, which were then organized into ranking TF–target regulon lists for each of the cell types. A module (i.e., a set of co-expressed TFs and a predicted target gene) was generated based on the top 50 regulators for a gene from the ranking list. Multiple modules were created for each tropism factor gene and their predicted direct regulators. To prune indirect regulatory relationships, the proposed gene modules' connections validated by the TF–target binding information (with the levels A: high confidence; B: likely confidence; C: medium confidence; D: low confidence taken from [55] used as the TF–target scoring criteria) were considered as direct regulation and kept for further analysis.

#### 2.6. Defining the Regulatory Direction of Each TF–Target Regulon

The regulatory direction of our identified direct TF–target regulon was defined based on TF–target data sources [56–58]. We extracted the overexpression effect data of TFs from [56], which included transcriptome analyses of cell lines overexpressing each of the 514 human TFs. The data processing was followed by [56] in order to determine the activation or repression effect of each TF on their downstream genes. We extracted the knock out and knock down effect data of TFs from the KnockTF database [57]. Following the database instructions, differentially expressed genes (DEGs) were identified for each TF knock out/knock down using an FDR <0.05. For each TF perturbation experiment, the TF activity for its downstream DEGs was assigned as activation if the knock out/down resulted in a decrease in gene expression, and vice versa. For TF–target regulons not covered by the above analyses, the TRRUST v2 reference database [58] was used to query their regulatory direction. The regulatory functions of 9396 regulatory interactions of 800 human transcription factors were downloaded from the TRRUST v2 database to annotate predicted TRNs. If there was any conflict regarding the assignment of the type of regulation, we prioritized the signs as follows: overexpression > knock out > knock down > TRRUST v2. For the regulations that were not annotated from these databases, the Pearson product moment correlation between the expression of the TF and its targets was used to determine the type of transcriptional regulation taking place [59]. TF–target pairs that were positively correlated (Pearson correlation  $\leq 0.03$ ) were assigned as activating regulation, whereas TF–target pairs that were negatively correlated (Pearson correlation  $\leq -0.03$ ) were assigned as repressing regulation.

#### 2.7. Network Organization and Visualization

For each cell type, the TF–target regulons were interconnected into a predicted multi-level gene network that was organized and visualized using Cytoscape 3.8.0.

### 2.8. Differential Expression Analysis

A two-sided Wilcoxon rank sum test was performed to identify the differentially expressed genes of each cell type in COVID-19 patients and healthy donors. Bonferroni correction was performed to adjust the  $p$ -value based on the total number of genes. Genes were determined to be differentially expressed in a cell type if the adjusted  $p$ -value was less than 0.05.

### 2.9. Differential Network Analysis

Tools from graph theory were used for pinpointing important genes in cell-type-specific function and drivers linked to the disease [36]. In each cell-type-specific gene network, various measures of centrality—namely, degree, betweenness, closeness, eigenvalues, and pagerank—were used to calculate the importance of the TFs. All these centralities were calculated using the package NetworkX 2.4. Single TFs that simultaneously decreased or increased the five centralities from healthy individuals to COVID-19 patients were identified as potential important TFs whose topological features significantly differed between healthy individuals and COVID-19 patients. A cell-type-specific hub TF was identified with maximum difference for any centrality measurement.

Degree centrality is defined as the number of edges that are connected to a point. Our inferred networks are directed. Therefore, the degree of a gene is the sum of the indegree (number of incoming edges) and outdegree (number of outgoing edges) of the gene. Betweenness centrality is calculated from the number of times each node falls in the shortest path of a network. Genes with a higher betweenness will have more control over the network by acting as bridges in the transcriptional cascades. The closeness centrality of a node is measured as the inverse distance between the node and all others in the network. Genes with a high level of closeness have the shortest distance to other genes and thus are able to spread transcriptional information through the network efficiently. In this study, both betweenness and closeness centrality were normalized to compare the network property between networks with different numbers of genes. The eigenvalue centrality of a node is the corresponding element of the principal eigenvector paralleling the largest eigenvalue of the graph adjacency matrix. Genes with a high eigenvalue are likely to be influential in the gene network, since they are connected to other influential genes. Pagerank, a variation of eigenvalue centrality, measures the importance of a node recursively based on the number of incoming edges (indegree) and the quality (i.e., pagerank) of the nodes linking to the node. Different from eigenvalue centrality, pagerank centrality models the probability that the information will continue flowing in the network, as represented by a damping factor (typically 0.85). Pagerank has been used to prioritize TFs in gene regulatory networks with the underlying assumption that genes with a high pagerank are likely to be transcriptionally regulated by [60].

### 2.10. Hybrid Centrality Measures

Three centrality measures, degree, betweenness, and closeness, were combined into one measure that determined the importance of genes in a network [61]. For each gene in a cell-type-specific TRN, the three centrality values were ranked and the consensus measure was calculated as the mean of the centrality ranks.

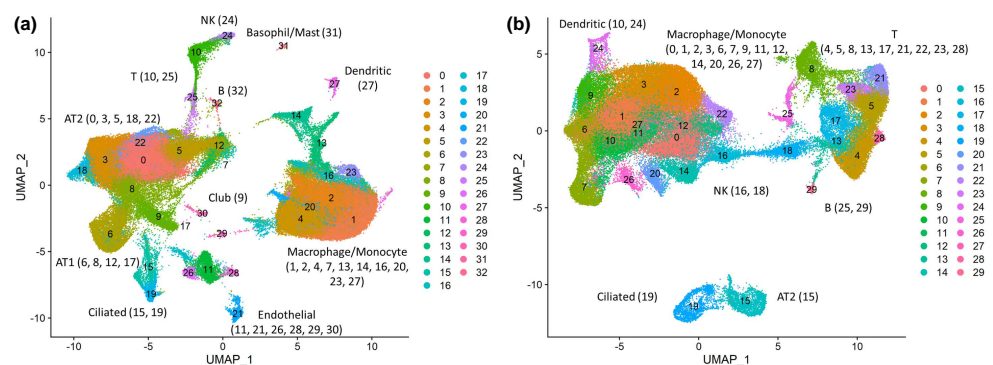
### 2.11. Mapping Drug Targets

The drugs and their annotations for their mode of action, gene target, and undergone clinical investigations were imported from The Drug Repurposing Hub [62] from the ConnectivityMap (CMap) database [63]. Information on chemical compounds and their corresponding gene targets was extracted from The Probes & Drugs portal [64].

### 3. Results

#### 3.1. Expression Landscape of the SARS-CoV-2 Tropism Factors Are Largely Different in Each Cell Type

Previous studies identified receptors (ACE2, NRP1, CD209), proteases (TMPRSS2, 4, 11A, CTSL, and FURIN), and IFITMs (IFITM 1, 2, and 3) as SARS-CoV-2 tropism factors. To investigate the expression specificity of these tropism factors in various cell types within healthy lungs or within SARS-CoV-2-infected lungs, we retrieved public scRNA-seq data from bronchoalveolar lavage fluid (BALF) samples obtained from 16 healthy donors [39,42–44] and 16 COVID-19 patients [39–41] (Table S1). We applied a graph-based clustering analysis of the scRNA-seq data from healthy donors and identified 33 distinct cell clusters. These cell clusters were visualized in two-dimensional spaces using the Uniform Manifold Approximation and Projection (UMAP) manifold learning technique (Figure 1a). UMAP is used to construct the fuzzy simplicial set, which is associated with a low-dimensional data representation from manifold approximation [65] and has been used to visualize cell types based on data from scRNA-seq [66]. We classified these clusters into 11 cell types based on their marker genes [53], including club cells, ciliated cells, alveolar epithelial type 1 cells (AT1), alveolar epithelial type 2 cells (AT2), B cells, T cells, natural killer cells (NK), basophil/mast cells, macrophage/monocyte cells, dendritic cells, and endothelial cells (Figure 1a). For scRNA-seq data from COVID-19 patients, we identified 30 clusters that were classified into 7 cell types, including ciliated cells, alveolar epithelial type 2 cells (AT2), B cells, T cells, natural killer cells (NK), macrophage/monocyte cells, and dendritic cells (Figure 1b). The BALFs of COVID-19 patients have higher proportions of immune cells (macrophage/monocyte cell, T cell, B cell, and NK cell) than those obtained from healthy donors and lower proportions of some epithelial cells (ciliated cells and AT2 cells) (Table S2). We quantified the expression of each gene in the specific cell types based on their percentage of gene-expressing cells and their average expression (Figure 2a,b; Tables S3 and S4).

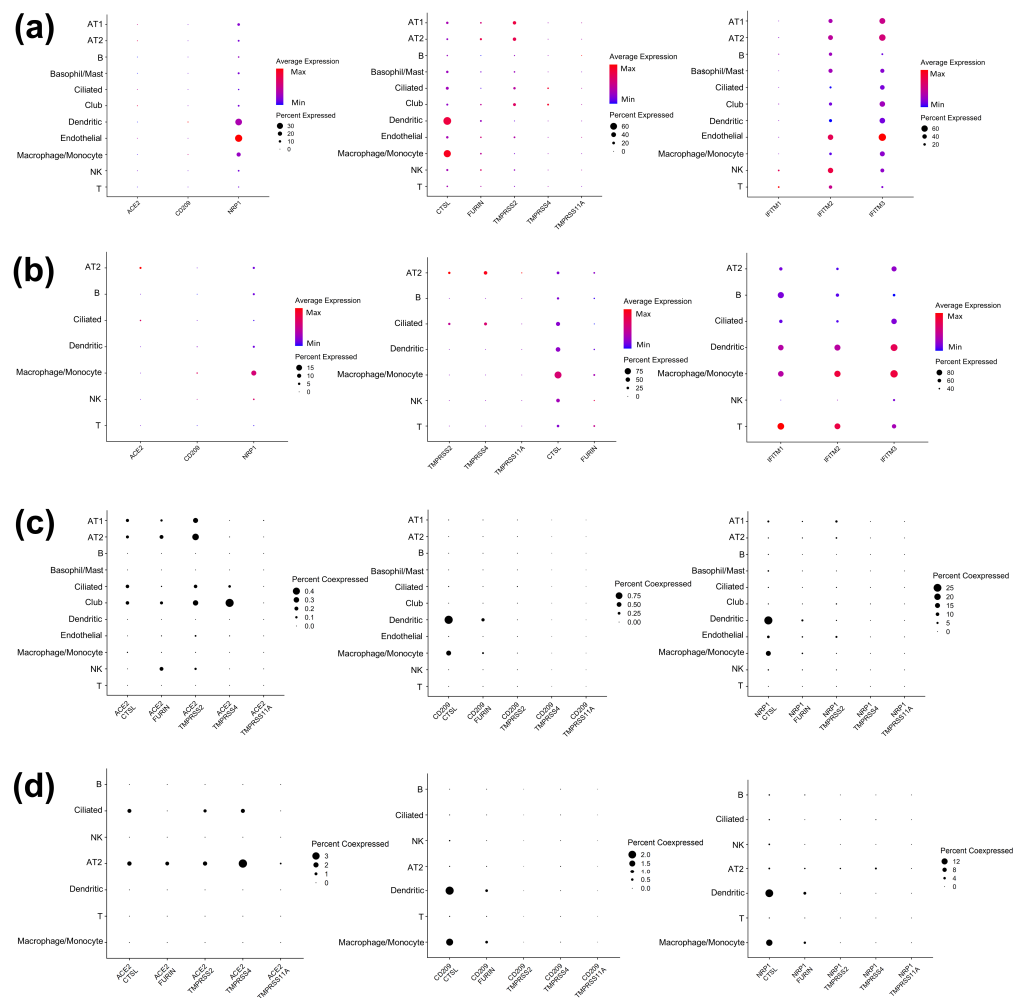


**Figure 1.** Bronchoalveolar landscapes in healthy donors and COVID-19 patients. (a) Uniform Manifold Approximation and Projection (UMAP) presentation of 11 major cell types and associated clusters in BALFs from healthy donors. Healthy cell clusters are annotated with a numeric number and assigned to a cell type. These cell types include club cells, ciliated cells, alveolar epithelial type 1 cells (AT1), alveolar epithelial type 2 cells (AT2), B cells, T cells, natural killer cells (NK), basophil/mast cells, macrophages/monocytes, dendritic cells, and endothelial cells. (b) UMAP presentation of 7 major cell types and their associated clusters in BALFs from COVID-19 patients. Infected cell clusters are annotated with a numeric number and assigned to a cell type. Identified cell types include ciliated cells, alveolar epithelial type 2 cells (AT2), B cells, T cells, natural killer cells (NK), macrophage/monocyte cells, and dendritic cells.

The expression of experimentally validated SARS-CoV-2 tropism factors (ACE2; NRP1; CD209; TMPRSS2, 4, and 11A; CTSL; FURIN; IFITM 1, 2, and 3) was profiled in specific cell types from healthy lungs. The highest proportion of ACE2-expressing cell types was epithelial cell types (club cells, ciliated cells, AT1, and AT2), and most of the CD209-

expressed cells were clustered into immune cells (macrophage/monocyte and dendritic cells) (Figure 2a; Table S3). In contrast with ACE2 and CD209, all healthy lung cell types had a proportion of cells expressing NRP1 (Figure 2a; Table S3). Our analysis further examined the co-expression of these receptors and their associated proteases in healthy lungs. ACE2 was preferentially co-expressed with TMPRSS2, 4, FURIN, and CTSL in multiple types of epithelial cells (Figure 2; Table S3) compared to immune cells, consistent with previous studies [21,67,68]. CD209 was mainly co-expressed with CTSL and FURIN in dendritic cells and macrophage/monocyte cells (Figure 2c; Table S4). Likewise, NRP1 was preferentially co-expressed with CTSL within dendritic cells and macrophage/monocyte cells (Figure 2c; Table S4). Additionally, we found that IFITM2 and 3 were highly expressed in most lung cell types, while IFITM1 was only preferentially expressed in natural killer cells and T cells (Figure 2a). Next, we analyzed the expression profiles of the tropism factors in the lung cells of COVID-19 patients.

To investigate the effects of virus infection on the expression of tropism factors, the lung cellular expression landscapes of COVID-19 patients were screened on a cell-type-specific level (Figure 2b; Table S3). With the infection of SARS-CoV-2, most of the ACE2-expressing cells were clustered into AT2 and ciliated cells, and the proportion of cells expressing ACE2 was increased in these two cell types seen in infected lungs compared with in healthy lungs (Figure 2a,b; Table S3). Likewise, the proportion of CD209 in macrophage/monocyte cells and dendritic cells was increased in the lungs of COVID-19 patients compared to in healthy lungs (Figure 2a,b). By contrast, the proportion of NRP1-expressing cells was decreased in most lung cell types in COVID-19 patients (ciliated cells, B cells, T cells, natural killer cells, macrophages/monocytes, and dendritic cells) compared to the healthy donors (Figure 2a,b; Table S3). Considering that the entry of SARS-CoV-2 into cells requires the expression of both receptor and proteases, we next investigated the co-expression of receptors (ACE2, CD209, and NRP1) with proteases. Consistent with data from healthy donors, we observed the co-expression of ACE2 with TMPRSS2, 4, FURIN, and CTSL in AT2 and ciliated cells. However, we observed an increased proportion of cells expressing these ACE2–protease pairs in each of the cell types in COVID-19 patients' lungs (Figure 2d; Table S4). Compared to the healthy donors, the proportion of macrophage/monocyte cells expressing CD209–CTSL and FURIN was increased, whereas the proportion of dendritic cells expressing CD209–CTSL was unchanged in COVID-19 patients' lungs (Figure 2d; Table S4). In contrast, in macrophage/monocyte and dendritic cells, the proportion co-expressing NRP1 with CTSL and FURIN was reduced in COVID-19 patients' lungs (Figure 2d; Table S4). Taken together, the differential co-expression pattern of receptor–protease pairs in lung cells between healthy donors and COVID-19 patients suggested an enhanced SARS-CoV-2 tropism in AT2 and ciliated cells via ACE2 and macrophage/monocyte cells via CD209 and a reduced SARS-CoV-2 tropism in macrophages/monocytes and dendritic cells via NRP1 after SARS-CoV-2 infection. For the IFITMs, their expression was increased overall in SARS-CoV-2-affected lung cells compared to cells from healthy lungs (Figure 2a,b, Table S3). In summary, the expression level of the tropism factors (Figure 2a,b) and the proportion of specific cell types co-expressing each pair of receptors and proteases (Figure 2c,d) differed between the lung cells of healthy donors and those of COVID-19 patients, suggesting that SARS-CoV-2 altered the TRNs regulating tropism factor expression. These results, combined with the differential expression of these tropism factors in each cell type, prompted us to construct TRNs to reveal the regulation of these tropism factors at the cell-type-specific level. We then compared the differential topology of the cell-type-specific TRNs between healthy lung cells and lung cells infected with SARS-CoV-2.



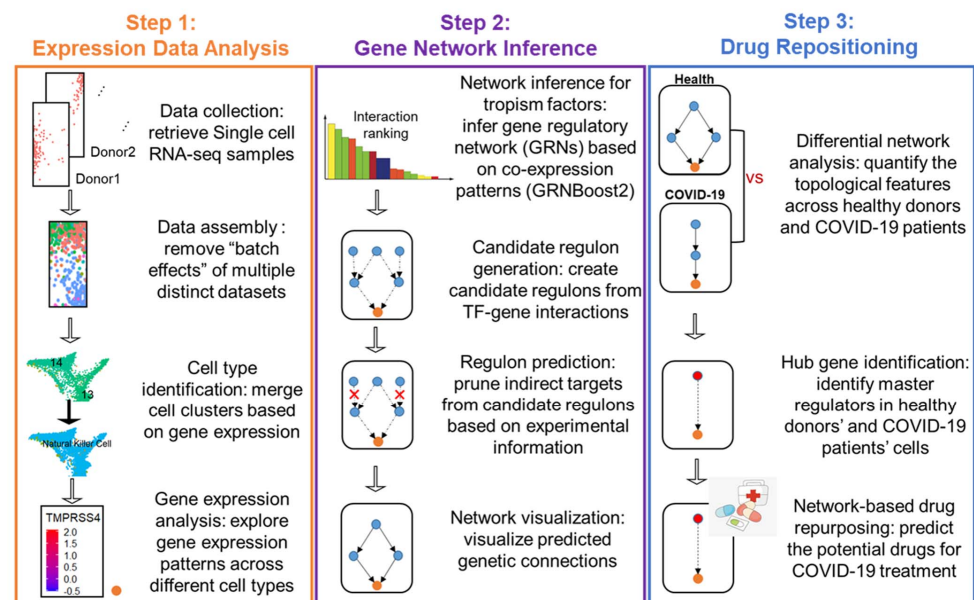
**Figure 2.** Dot plots showing the expression of SARS-CoV-2 tropism factors in 11 cell types from the BALFs of healthy donors (a) and 7 cell types from the BALFs of COVID-19 patients (b). These factors include entry receptors (ACE2, CD209, and NRP1; left), entry proteases (CTSL, Furin, TMPRSS2, 4, and 11A; middle), and interferons (IFITM1, 2, and 3; right). The values of raw expressions were normalized and log-transformed. The size of the dots represents the proportion of cells in each cell type with the expression of each gene, while the color indicates the mean expression of each gene. Dot plots illustrate the co-expression of each receptor–protase pair in 11 cell types from the BALFs of healthy donors (c) and 7 cell types from the BALFs of COVID-19 patients (d). These receptor–protase pairs include ACE2-CTSL, Furin, TMPRSS2, 4, and 11A (left); CD209-CTSL, Furin, TMPRSS2, 4, and 11A (middle); and NRP1-CTSL, Furin, TMPRSS2, 4, and 11A (right). The size of the dots represents the proportion of cells in each type of co-expressed gene of each receptor–protase pair.

### 3.2. Cell-Type-Specific Trns Affecting Tropism Factors

Three key steps were used for the construction of cell-type-specific TRNs for healthy lung cells and lung cells from COVID-19 patients (Figure 3). In Step 1 (single-cell gene expression analysis), we used the integrated scRNA-seq of lung samples obtained from 16 healthy donors [39,42–44] and 16 COVID-19 patients [39–41], then used Seurat’s scRNA-seq toolbox to categorize the healthy lung cells into 11 cell types and the COVID-19 patient cells into 7 cell types [45,48]. In Step 2 (TRN inference), we used 1639 human TFs [69] as predefined regulators and 11 experimentally validated SARS-CoV-2 tropism factors as targets for inferring TF–target regulatory networks via the regression tool GRNboost2 [54]. GRNboost2 infers TF–target regulatory networks based on the gene co-expression pattern for each cell type. GRNBoost2 adopts the GENIE3 algorithm, which was the winner of the DREAM5 network challenge, and uses gradient boosting to reduce the time needed to



infer a TRN [54]. The direct regulators of the tropism factors are identified by incorporating each target and its 50 top co-expressed TFs as bottom-layered regulators. The network is then expanded by identifying the 50 top co-expressed TFs with bottom-layered regulators as the top-layered regulators. We filter out the indirect TF–target interaction from the candidate regulon based on the CHIP-Seq data and TF-motif binding data [55] (Table S5). The regulatory direction (activation or repression) of each TF–target is defined based on the public experimental data sources for TF overexpression, TF knock out, and TF knock down, with the data ordered as follows: overexpression > knock out > knock down [56–58] (Table S5). For the TF–target pairs for which there is no experimental evidence, we used in silico correlation analysis to elucidate their regulatory direction. Finally, 18 two-layered TRNs were constructed for the 11 major lung cell types of the healthy donors and the 7 major lung cell types of the COVID-19 patients (Figures S1–S26). In Step 3 (network-based drug repositioning), we quantified the topological features of cell-type-specific TRN across healthy donors and COVID-19 patients via differential network analysis, then used these features to identify the hub TFs that were important for regulating tropism factors in the cells of healthy individuals and COVID-19 patients. The drugs affecting the expression and activity of the identified hub TFs were used as candidates for mitigating the tropism of SARS-CoV-2.

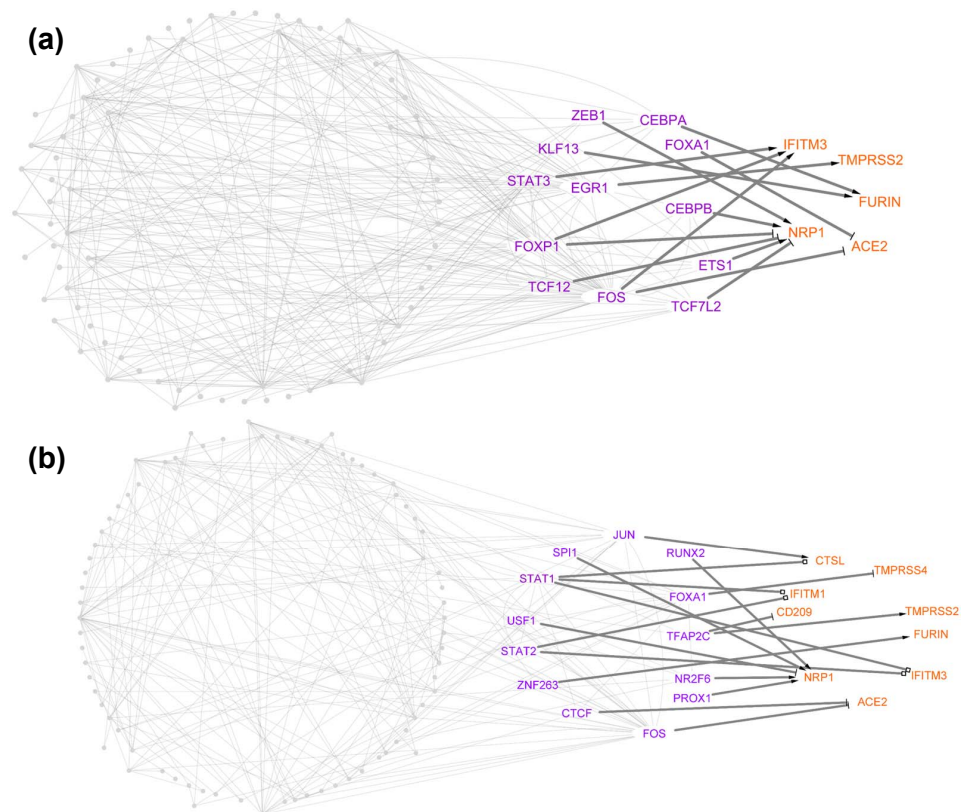


**Figure 3.** Methodology flowchart. In Step 1, the scRNA-seq of the BALF samples obtained from healthy donors and COVID-19 patients was integrated to identify each cell's identity and analyze the gene expression within each cell type. In Step 2, the cell-type-specific gene expression was used to reconstruct the transcriptional regulatory network. In Step 3, the roles of TFs were evaluated in the TRNs of each cell type across healthy individuals and SARS-CoV-2 patients, thus providing a clue for repurposing drugs for use in COVID-19 treatment.

### 3.3. Identification of TFs Directly Regulating Tropism Factors within Each Cell Type

The TRNs from healthy cell types were analyzed and compared to determine how TFs regulate the tropism factors expressed in each healthy cell type. In total, we identified 144 TF–target relationships, including 49 unique TFs directly targeting 11 tropism factors from the networks across all healthy cell types (Table S5; Figure 4a; Figures S1a, S2a, S3a, S4a, S5a, S6a, S7, S8, S9, and S10). We first checked whether the canonical receptor–protease pair ACE2 and TMPRSS2 were directly regulated by a single TF in each specific cell type. Network analyses showed that no TFs had this ability, which suggested that ACE2- and TMPRSS2-mediated SARS-CoV-2 cellular tropism may depend on cooperative regulation by multiple TFs. As such, TMPRSS2 was directly activated by EGR1 and ACE2 was directly

activated by c-FOS and repressed by FOXA1 in AT2 cells (Figure 4a). In addition, ACE2 was directly regulated by STAT1 and TMPRSS2 was directly activated by FOXA1 in ciliated cells (Figure S2a). Of these four TFs (EGR1, c-Fos, FOXA1, and STAT1), STAT1 and FOXA1 have been shown to regulate ACE2 or TMPRSS2 [70–72]. STAT1 is required for SARS-CoV-1 infection [73] and enhances ACE2 expression in mammary tissue during pregnancy [71]. FOXA1 is driven by the androgen receptors and regulates the expression of TMPRSS2 and ACE2 in prostate cancer cells [74,75]. Together with previous studies showing that SARS-CoV-2 virus infection induces EGR1 expression [76] and activates c-FOS via MAPKs using its spike protein [77], we hypothesized that SARS-CoV-2 infection may enhance the tropism of AT2 cells through regulatory cascades of EGR1 and c-Fos and the tropism factors ACE2 and TMPRSS2.



**Figure 4.** Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in AT2. Inferred regulation for tropism factors for AT2 in healthy donors (a) and in COVID-19 patients (b) based on an integrated analysis of in silico and experimental data. The orange nodes represent tropism factors, the purple nodes represent TFs that directly regulate tropism factors, and the gray nodes represent the TFs that indirectly regulate tropism factors. Direct regulations for tropism factors are differentiated between transcriptional activation (delta shape arrows), repression (T shape arrows), and undetermined (square shape arrows).

We then asked the question of whether any of the TFs were able to directly regulate any receptor–protease pairs in the individual healthy cell types. Three TFs—CEBPA in monocytes/macrophages (Figure S4a), c-Fos in endothelial cells (Figure S10), and STAT1 in monocytes/macrophages (Figure S4a) were identified to have this ability. Of endothelial cells co-expressing any receptor–protease pairs, almost 50% of these cells co-express NRP1 and CTSL, which are directly targeted by c-Fos (Table S4). Of monocyte/macrophage cells co-expressing any receptor–protease pairs, about 79% of these co-express NRP1 and CTSL (STAT1 target), and only 4% of these cells co-express NRP1 and TMPRSS2 (CEBPA target). Our results suggested that the regulons c-Fos-NRP1-CTSL and STAT1-NRP1-TMPRSS2 have the potential to regulate tropism in monocyte/macrophage cells (Figure S4a) and

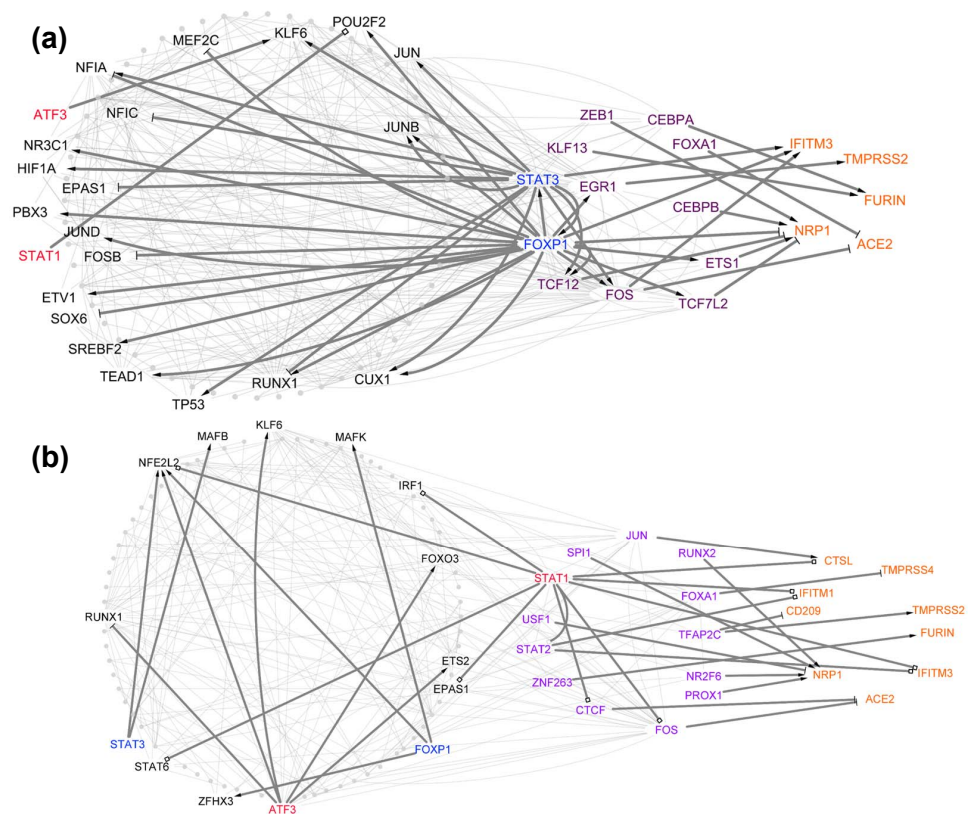
endothelial cells (Figure S10a), respectively. Additionally, NRP1-CTSL was also identified as the major co-expressed receptor–protease pair in AT1 cells (Figure S8), B cells (Figure S1a), basophil/mast cells (Figure S7), club cells (Figure S9), dendritic cells (Figure S3a), NK cells (Figure S5a), and T cells (Figure S6a). Instead of regulating the NRP1-CTSL pair using a single TF, our networks showed that multiple TFs coordinate to regulate NRP1 and CTSL in these healthy cell types, suggesting the robustness of these TRNs in cell types controlling the cellular tropism of SARS-CoV-2.

To further understand how each type of lung cell transcriptionally regulates its cellular tropism after SARS-CoV-2 infection, we analyzed and compared the TRNs of different cell types of COVID-19 patients. From the TRNs, we could identify 35 unique TFs which directly regulate 10 tropism factors through 119 TF–target relationships across patient lung cells (Table S5; Figure 4b; Figures S1b, S2b, S3b, S4b, S5b, and S6b). Of these 35 TFs, 10 TFs (SPI1, USF1, NR2F6, KLF6, HNF4G, MAFB, IRF1, NR1H2, ZNF274, and ELF3) were not determined to regulate tropism factors in the TRNs of healthy cell types. We then identified what tropism factors were directly regulated by these 10 TFs in the lung cell types of COVID-19 patients. Our networks showed that SPI1 regulates NRP1 and FURIN, HNF4G regulates TMPRSS2, ZNF274 regulates TMPRSS4, and the last seven TFs only target NRP1 in either epithelial cells (AT2, ciliated cells) or immune cells (dendritic, monocytes/macrophages, NK cells, and T cells) (Table S5). These TFs may be responsible for the dynamic NRP1 mRNA expression change in the lungs of severely ill COVID-19 patients [78]. However, none of these 10 COVID-19 patients TRNs' TFs regulate receptor–protease pairs in the same cell type from COVID-19 patients. We then queried whether the individual TFs could directly regulate receptor–protease pairs in each of these cell-type-specific TRNs of COVID-19 patients. As a result, five TFs were identified with this regulatory role, including HIF1A, which inhibits ACE2 and FURIN in ciliated cells and monocyte/macrophage cells (Figure S4b); c-Jun, which regulates NRP1 and CTSL in dendritic and monocyte/macrophage cells (Figure S4b) and CD209 and CTSL in T cells (Figure S6b); c-Fos, which regulates NRP1 and CTSL in monocyte/macrophage cells (Figure S4b); STAT1, which regulates NRP1 and TMPRSS2 in T cells (Figure S6b); and TFAP2C, which regulates CD209 and TMPRSS2 in AT2 cells (Figure 4b). It is noteworthy that c-Jun and c-Fos form an AP-1 early response transcription factor that could be activated by the SARS-CoV-1 nucleocapsid protein [79]. Given the similarity of the sequence identity between SARS-CoV-2 and SARS-CoV-1 in nucleocapsid protein [80], SARS-CoV-2 nucleocapsid protein might also activate the AP-1 early response transcription factor, as SARS-CoV-1 nucleocapsid protein does. With the direct activation of AP-1 (c-Jun and c-Fos) in NRP1–CTSL pairs, monocyte/macrophage cells may be more susceptible to SARS-CoV-2 infection in COVID-19 patients compared to healthy donors. To further understand the transcriptional regulation of the SARS-CoV-2 tropism between lung cell types in healthy and infected states, we quantitatively measured the differential role of TFs in healthy-donor TRNs versus COVID-19-patient TRNs.

### 3.4. Centrality Analysis for Predicting TFs That Are Important in Host Cell Tropism

The regulatory importance of each TF (node) in each network can be evaluated through its centrality (degree, betweenness, closeness, pagerank, and eigenvalues). Knowing the centrality differences between the cell-type-specific TRNs of healthy donors and COVID-19 patients would be helpful for screening out the TFs whose regulatory role is mostly influenced by SARS-CoV-2 infection [36,61]. First, we calculated the centralities of each common TF in the TRNs of healthy individuals and COVID-19 patients for each cell type. Next, we subtracted all the centralities of each common TF in the COVID-19 network and healthy network for the same cell types (AT2, B, ciliate, dendritic, macrophage/monocyte, T, and NK); the TFs showing increases or decreases in the five centralities were used for further analysis (Figure 5; Figures S21, S22, S23, S24, S25, and S26). We identified 53 TFs with positive rewiring (increased centrality change) which were more central in COVID-19 networks (Table S6) and suggested that these TFs might play more important roles in regulating the tropism factors after infection with SARS-CoV-2. We identified 48 TFs

with decreased centralities which were more important in the networks of healthy lung cells. To test whether our centrality-identified TFs were functionally associated with SARS-CoV-2 infection, we selected the TFs displaying any of the largest centrality differences between the networks of healthy individuals and COVID-19 patients and assessed their characteristics using the existing literature. A total of 16 unique TFs (FOXP1, STAT3, ATF3, STAT1, SMAD3, ETS1, ETV6, BHLHE40, SPI1, NFKB1, IRF1, RUNX1, MEF2C, RARA, FOS, CREB5) were characterized for further analysis (Table S7). Of these 16 TFs, 9 TFs (FOXP1, STAT1, STAT3, ATF3, SMAD3, NFKB1, RARA, FOS, IRF1) have been identified as potential drug targets for preventing SARS-CoV-2 infection or inhibiting the activity of SARS-CoV-2 [15,72,81–86], suggesting that centrality analysis could be a useful tool for identifying TFs from networks regulating tropism factors.



**Figure 5.** Differential network analysis for ciliated cells. Red (blue) nodes represent TFs that show a significant increase (decrease) in network centralities from the TRN of healthy donors to the TRN of COVID-19 patients. These red (blue) nodes are marked in the TRN of healthy ciliated cells (a) and the TRN of ciliated cells from COVID-19 patients (b). The orange nodes represent tropism factors, the purple nodes represent TFs that directly regulate the tropism factors, and the gray nodes represent the TFs that indirectly regulate the tropism factors. Direct regulations of tropism factors are differentiated between transcriptional activation (delta-shaped arrows), repression (T-shaped arrows), and undetermined (square-shaped arrows).

#### 4. Discussion

The expression of host tropism factors is used as a molecular indicator for the susceptibility of lung cells to SARS-CoV-2 [87]. Targeting the transcriptional regulation of tropism factors has been proposed as a viable treatment strategy to prevent SARS-CoV-2 infection [75]. The single-cell sequence of each lung cell type represents a distinct transcriptional landscape [53]. Therefore, it is necessary to elucidate the transcriptional regulatory roadmap in each lung cell type, thereby identifying the TFs regulating tropism factors with the potential to be used as therapeutic targets. We performed a computational analysis of multiple existing public single-cell sequencing datasets of healthy and SARS-CoV-2

infected donors' lung cells and obtained the count matrix depicting the gene abundances for all single cells. Based on these single-cell data, co-expression modules associated with SARS-CoV-2 host cell tropism factors were created for each lung cell type for healthy donors and COVID-19 patients using GRNboost2. Next, the indirect targets were pruned from these modules based on an available collection of experimentally validated TF–target interactions. Lastly, the activity and direction of these regulons were determined based on the effect of TF perturbation on the regulon's target genes from the public databases. With these three steps, we established the multiplier hierarchical TRNs targeting tropism factors in a lung-cell-specific manner, which could help us to answer the questions underlying the transcriptional regulatory program for host cell tropism.

#### *4.1. Case Study 1: Stat1 Has a Different Regulatory Role for Host Cell Tropism Factors in Healthy vs. SARS-CoV-2-Infected Donors*

SARS-CoV-2 N protein has been reported to suppress phosphorylation and the nuclear trafficking of STAT1, thus inhibiting the regulatory activity of STAT1 [88]. Based on our cell-type-specific TRNs, we also identified STAT1 as a direct regulator of tropism factors in some same lung cell types (ciliated, dendritic, and macrophage/monocyte cells) across healthy and infected donors. We then checked whether the regulatory effect of STAT1 on tropism factors was partially diminished in same cell type TRN in healthy and infected states using the STAT1 regulon, which was scored as “A: high confidence” or “B: likely confidence”. For healthy donor TRNs, STAT1 directly targets ACE2 in ciliated cells, CTSL in dendritic cells, and NRP1 and CTSL in macrophage/monocyte cells. In contrast, the direct regulatory activity of STAT1 towards ACE2, NRP2, and CTSL could not be detected in the same cell types' TRNs from SARS-CoV-2 patients. The results suggest that our TRNs could be used to identify SARS-CoV-2 infection-affected TFs that target tropism factors, providing additional insights for subsequent experiments that extend our knowledge of the virus–host interactions of SARS-CoV-2.

#### *4.2. Case Study 2: Differential Centrality Analysis Can Be Used to Identify Important TFs Whose Expression Is Not Differentially Changed between Healthy and SARS-CoV-2-Infected Donors*

We applied the specific centrality measures for each TF in the TRNs and identified the TFs whose regulatory role in cell tropism became more pronounced after SARS-CoV-2 infection based on the centrality differences of each TF between healthy donors and COVID-19 patients. Traditionally, differential gene expression analysis between different conditions can help to uncover important regulatory genes [89]. Here, we query whether TFs with an increased/decreased network centrality between healthy and COVID-19 cell-type-specific networks have a corresponding increased/decreased transcriptional expression in each cell type from healthy to SARS-CoV-2-infected lungs. To answer this question, we took the 16 TFs displaying the largest centrality differences between the healthy and COVID-19 networks of the same cell types (FOXP1, STAT3, ATF3, STAT1, SMAD3, ETS1, ETV6, BHLHE40, SPI1, NFKB1, IRF1, RUNX1, MEF2C, RARA, FOS, CREB5) for further analysis. We found that the centrality and transcriptional expression of most TFs in each cell type were consistently increased/decreased from healthy to COVID-19 patients (Tables S7 and S8), suggesting that the identification of relative changes in the centrality of single-cell regulatory networks can be used as an alternative to differential expression analysis to classify the biological importance of genes. However, the centrality of eight unique TFs (three of five TFs for B, one of four TFs for ciliated, one of four TFs for macrophage/monocyte, three of four TFs for NK, and one of four TFs for T cells) was not changed in the same direction as expression (centrality decreased with gene expression not decreased; centrality increased with gene expression not increased) (Tables S7 and S8). We hypothesized that this inconsistent change between the centrality and expression of these TFs may be attributed to the effect of SARS-CoV-2 virus infection on these TFs at other regulatory levels, such as protein modification. A literature-curated analysis was undertaken to test whether these eight TFs were modified at the protein level. Six TFs (SMAD3, ETS1, RARA, CREB5, FOS, ETV6) could be phosphorylated [90–95], and two TFs (BHLHE40 and FOXP1) could be

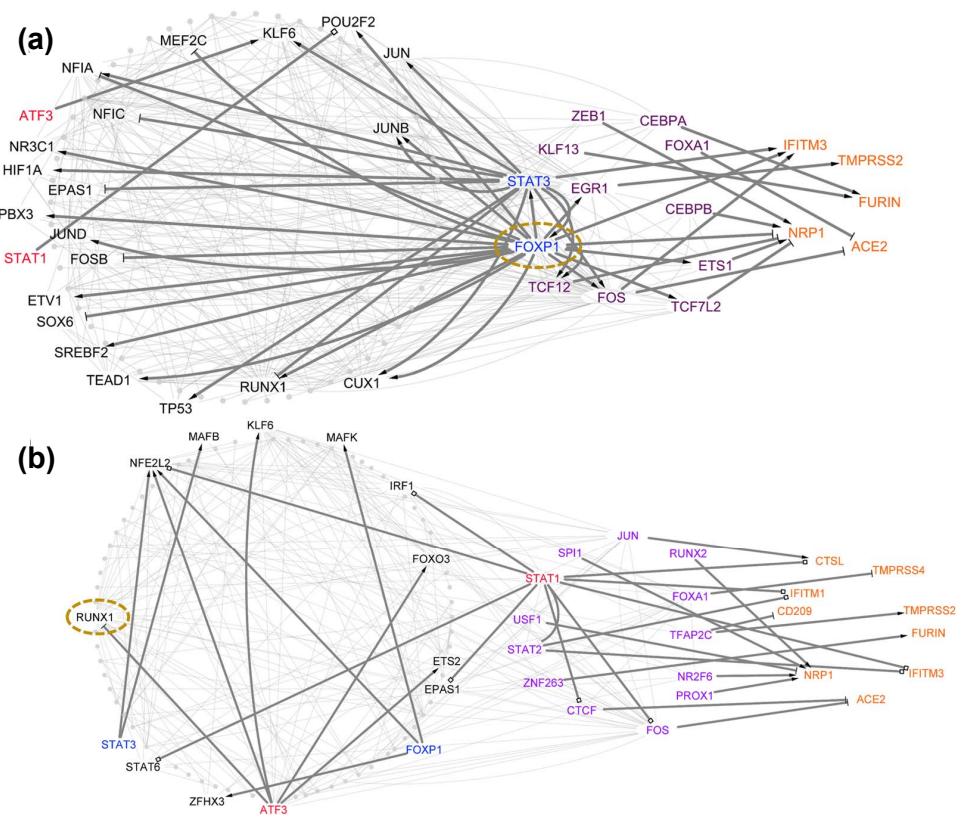
sumoylated [96,97]. It is noteworthy that the phosphorylation of SMAD3 and FOS could be induced by hepatitis virus infection [90,94], as our results supported. In summary, our differential analysis suggested that the transcriptional regulation of SARS-CoV-2 tropism differs between the lungs of healthy individuals and COVID-19 patients within the same cell types. Our results highlight important TFs present within the cells of healthy individuals and COVID-19 patients and further highlight known candidate drugs that target these central TFs.

#### *4.3. Case Study 3: Hub Tfs Identified by Network Centrality Analysis Could Be Used as Targets for Drug Repositioning to Prevent SARS-CoV-2 Infection and Transmission*

We first ranked each TF within each TRN using its cumulative centrality, which is a combination of degree, betweenness, and closeness (Materials and Methods). The TFs with the top cumulative centrality were then selected as the master regulators in each of the TRNs (Table S9). From 11 healthy cell-type TRNs and 8 COVID-19 cell-type TRNs, we identified 8 unique central TFs, FOXP1, RUNX1, FOS, SMAD3, JUN, NFKB1, PPARG, and STAT3. Only the top TF, FOXP1, was specifically identified in the TRNs of healthy lung cell types, which includes AT1 (Figure S11), AT2 (Figure 6a), club cells (Figure S12), NK cells (Figure S13a), and T cells (Figure S14a). In contrast, many TFs were identified specifically as the top TFs in COVID-19 patients' lung cell types, including SMAD3 in ciliated cells (Figure S15b), NFKB1 and JUN in dendritic cells (Figure S16b), PPARG in macrophage/monocyte cells (Figure S17b), and STAT3 in T cells (Figure S14b). In addition, we also identified two TFs serving as the top TFs in the lung cell type TRNs from both healthy individuals and COVID-19 patients. The two TFs were RUNX1 in the TRNs of healthy B cells (Figure S18a), basophil/mast cells (Figure S19), dendritic cells (Figure S16a), macrophage/monocyte cells (Figure S17a), NK cells (Figure S13a) and COVID-19 patients' AT2 cells (Figure 6b) and B cells (Figure 3b) and FOS in the TRNs of healthy ciliated (Figure S15a) and endothelial cells (Figure S20) and COVID-19 patients' ciliated cells (Figure S15b) and NK cells (Figure S13b). Only FOS in ciliated cells (Figure S15) and RUNX1 in B cells (Figure S18) remained as the top TFs in the TRNs of the same cell type before and after SARS-CoV-2 infection.

We then screened candidate drugs targeting these eight top TFs based on the Drug Repurposing Hub [62]. A total of 67 drugs in clinical and launch stages (Table S10) could be identified for use in 6 of these 8 TFs (FOS, SMAD3, JUN, NFKB1, PPARG, and STAT3), and no drugs were identified for FOXP1 and RUNX1. We then queried the Probes & Drugs portal [64] to identify the compounds targeting FOXP1 and RUNX1. No compound was identified for FOXP1, and eight compounds targeting RUNX1 were identified (Table S10). RUNX1 was the only gene targeted by CBF $\beta$  Inhibitor. CBF $\beta$  Inhibitor has been shown to alter the ability of RUNX1 to bind to target genes and alter their expression, thereby inhibiting the growth of leukemia and breast cancer cell lines [98]. Given that RUNX1 is the top TF for the TRNs of many cell-types, including AT2 cells, the repurposing of the CBF $\beta$  Inhibitor to affect RUNX1 may be beneficial for preventing the SARS-CoV-2 infection and transmission. In addition to AT2, ciliated cells are also the primary cells infected by SARS-CoV-2 [99,100]. FOS, as the top TF from healthy ciliated cell TRN, is targeted by the launched drug ephedrine (Table S10). Ephedrine and its functional analog pseudoephedrine have been used for treating patients with severe COVID-19 pneumonia [101,102]. However, except for FOS, ephedrine regulates 19 other genes; thus, more extensive studies are needed to elucidate the role of FOS and the effect of ephedrine on the treatment of COVID-19 patients. Regarding the TRN of COVID-19 patients' ciliated cells, its top TF SMAD3 was only identified to be targeted by SIS3. SIS3 inhibits the activity of SMAD3, and it has been shown to have a protective effect on SARS-CoV-2-induced death [103]. We then investigated whether the rest of these 67 drugs specifically target one of the TFs without having any effect on the other genes. We identified T-5224 specifically targeting c-Jun, 15 drugs specifically targeting PPARG, and 8 drugs specifically targeting STAT3 (Table S10). Additionally, we found that bardoxolone-methyl could affect PPARG and STAT3 together without affecting other genes. Given that Jun, PPARG, and STAT3 are the top TFs from COVID-19 patients'

TRNs, these identified drugs may be candidates for reducing SARS-CoV-2 infection for these cell types.



**Figure 6.** Identification of master TFs in AT2. Dotted circled nodes represent the identified master TFs with the highest hybrid-network centrality. (a) FOXP1 is the predicted master TF for healthy donors. (b) RUNX1 is the predicted master TF for COVID-19 patients.

#### 4.4. Limitations of the Present Approach

Firstly, the lung cell types are not completely the same in healthy and SARS-CoV-2-infected donors. Given that SARS-CoV-2 infection can result in the death of lung cells, particularly for several epithelial cell types, such as AT1 cells, club cells, and endothelial cells [104], we could not construct TRNs for all lung cell types and compare the TRNs between healthy and infected donors in our analyses. Secondly, our network construction stochastic algorithm could not guarantee achieving exactly the same results for independent runs on the same input data due to the stochastic nature of the inference approach. The inference approach, the GRNBoost2 algorithm (Step 2 of the Results Section: “Cell-type-specific TRN affecting tropism factors”), is a tree-based stochastic algorithm that uses repeated random samplings of genes to model the expression of target genes for the sets of TFs [105]. Therefore, each regulon composed of TF and target genes might slightly differ between each independent run. We generally expected to achieve consistent results from run to run by averaging the predictions of single regression trees using the gradient boosting machine (GBM) [106]. Thirdly, we employed the information of the TF–target interactions for 1541 human TFs to annotate the direct regulation and pruning of false positives in our predicted single-cell TRNs [55]. This was achieved by sacrificing the predictive power of the TRN models if SARS-CoV-2 could induce a TF–target interaction which was not included in these known TF–target interactions, in which case these regulons would be classified as false negatives in the TRN construction. Fourthly, our pipeline could not infer interactions affected by post-transcriptional regulation. Our GRN boost 2 network inference was based on the principle that if a TF regulates target genes, a change in TF expression will affect the expression of its target genes. However, if the regulatory activity of one TF is not controlled

based on its expression but relies on post-translational modifications, such as dimerization with another TF or protein modifications such as phosphorylation, our pipeline could not characterize this type of regulon.

## 5. Conclusions

Overall, this study tried to enrich the knowledge of the transcriptional regulation of SARS-CoV-2 tropism in human lungs by constructing the cell-type-specific TRNs of healthy donors and COVID-19 patients. Our studies comprehensively investigated the expression of 11 tropism factors for 11 healthy lung cell types and 7 cell types from COVID-19 patients. We then integrated co-expression analysis and experimentally validated TF–target information to construct the TRNs regulating these 11 tropism factors in each lung cell type. A total of 11 healthy cell type TRNs and 7 COVID patients' TRNs were established to visualize the transcriptional regulatory cascades used for activating or repressing the expression of tropism factors. Our differential network analysis elucidated the altered regulatory role of TFs in the same cell types in healthy and SARS-CoV-2-infected states which are responsible for the expression of differential tropism factors. We screened out the key regulators from each TRN and repurposed the drugs targeting these regulators for the potential treatment of SARS-CoV-2 infection. The pipeline described here can also be used to understand the effects of other virus perturbation on cell regulatory networks.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v14040837/s1>, Figure S1: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in B cells; Figure S2: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in ciliated cells; Figure S3: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in dendritic cells; Figure S4: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in macrophage/monocyte cells; Figure S5: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in NK cells; Figure S6: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in T cells; Figure S7: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in healthy basophil/mast cells; Figure S8: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in healthy AT1 cells; Figure S9: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in healthy club cells; Figure S10: Transcriptional regulatory networks controlling SARS-CoV-2 tropism factors in healthy endothelial cells; Figure S11: Identification of master TFs in healthy AT1 cells; Figure S12: Identification of master TFs in healthy club cells; Figure S13: Identification of master TFs in NK cells; Figure S14: Identification of master TFs in T cells; Figure S15: Identification of master TFs in ciliated cells; Figure S16: Identification of master TFs in dendritic cells; Figure S17: Identification of master TFs in macrophage/monocyte cells; Figure S18: Identification of master TFs in B cells; Figure S19: Identification of master TFs in healthy basophil/mast cells; Figure S20: Identification of master TFs in healthy endothelial cells; Figure S21: Differential network analysis for B cells; Figure S22: Differential network analysis for ciliated cells; Figure S23: Differential network analysis for dendritic cells; Figure S24: Differential network analysis for macrophage/monocyte cells; Figure S25: Differential network analysis for NK cells; Figure S26: Differential network analysis for T cells; Table S1: Clinical characteristics (a) for 16 healthy donors including sex, age, and ethnicity (b) and for 16 COVID-19 including pneumonia, ARDS, timing from symptom onset, etc.; Table S2: The number of cells and the number of read count for each cell type from healthy donors and COVID-19 patients; Table S3: Gene expression analysis for the detected cell types in healthy donors and COVID-19 patients; Table S4: Co-expression analysis for the detected cell types in healthy donors and COVID-19 patients; Table S5: The binding and regulatory direction data of each TF–target from all cell-type-specific TRNs; Table S6: The centralities changes of TFs whose five centralities simultaneously alter in TRNs from healthy state to SARS-CoV-2 affected state; Table S7: The list of TFs whose centralities are simultaneously changed and displaying any of maximum centrality divergence between healthy and COVID-19 network; Table S8: Differential expressed genes between the healthy donors and COVID-19 patients; Table S9: Hybrid centrality rankings for the TFs regulating tropism factors; Table S10: Drugs targeting master regulators repurposed from drug databases.



**Author Contributions:** H.T.: conceptualization, methodology, investigation, formal analysis, writing—original draft, and writing—review and editing; H.C.: conceptualization, formal analysis, writing—original draft, and writing—review and editing; C.M.W.: conceptualization, methodology, writing—review and editing, supervision, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** H.T. and H.C. were funded by the National Science Foundation (IOS-1444561).

**Institutional Review Board Statement:** Not applicable, since we used publicly available data.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data were uploaded as Supplementary Materials.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hopkins, H. COVID-19 Case Tracker. Follow Global Cases and Trends. 2020. Available online: <https://coronavirus.jhu.edu/map.html> (accessed on 2 March 2022).
2. Hu, B.; Guo, H.; Zhou, P.; Shi, Z.L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **2021**, *19*, 141–154. [[CrossRef](#)] [[PubMed](#)]
3. Guan, W.J.; Ni, Z.Y.; Hu, Y.; Liang, W.H.; Ou, C.Q.; He, J.X.; Liu, L.; Shan, H.; Lei, C.L.; Hui, D.S.; et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [[CrossRef](#)] [[PubMed](#)]
4. Bradley, B.T.; Maioli, H.; Johnston, R.; Chaudhry, I.; Fink, S.L.; Xu, H.; Najafian, B.; Deutsch, G.; Lacy, J.M.; Williams, T.; et al. Histopathology and ultrastructural findings of fatal COVID-19 infections in Washington State: A case series. *Lancet* **2020**, *396*, 320–332. [[CrossRef](#)]
5. Huang, Y.; Zhao, N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. *Psychiatry Res.* **2020**, *288*, 112954. [[CrossRef](#)]
6. Xu, J.; Xu, X.; Jiang, L.; Dua, K.; Hansbro, P.M.; Liu, G. SARS-CoV-2 induces transcriptional signatures in human lung epithelial cells that promote lung fibrosis. *Respir. Res.* **2020**, *21*, 1–12. [[CrossRef](#)]
7. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T.S.; Herrler, G.; Wu, N.H.; Nitsche, A.; et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, *181*, 271–280. [[CrossRef](#)]
8. Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J.; et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **2020**, *11*, 1–12. [[CrossRef](#)]
9. Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 11727–11734. [[CrossRef](#)]
10. Singh, M.; Bansal, V.; Feschotte, C. A single-cell RNA expression map of human coronavirus entry factors. *Cell Rep.* **2020**, *32*, 108175. [[CrossRef](#)]
11. Vankadari, N.; Wilce, J.A. Emerging COVID-19 coronavirus: Glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26. *Emerg. Microbes Infect.* **2020**, *9*, 601–604. [[CrossRef](#)]
12. Wang, P.G.; Tang, D.J.; Hua, Z.; Wang, Z.; An, J. Sunitinib reduces the infection of SARS-CoV, MERS-CoV and SARS-CoV-2 partially by inhibiting AP2M1 phosphorylation. *Cell Discov.* **2020**, *6*, 1–5. [[CrossRef](#)] [[PubMed](#)]
13. Amraei, R.; Yin, W.; Napoleon, M.A.; Suder, E.L.; Berrigan, J.; Zhao, Q.; Olejnik, J.; Chandler, K.B.; Xia, C.; Feldman, J.; et al. CD209L/L-SIGN and CD209/DC-SIGN act as receptors for SARS-CoV-2. *ACS Cent. Sci.* **2021**, *7*, 1156–1165. [[CrossRef](#)] [[PubMed](#)]
14. Cantuti-Castelvetri, L.; Ojha, R.; Pedro, L.D.; Djannatian, M.; Franz, J.; Kuivanen, S.; van der Meer, F.; Kallio, K.; Kaya, T.; Anastasina, M.; et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **2020**, *370*, 856–860. [[CrossRef](#)] [[PubMed](#)]
15. Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov.* **2020**, *6*, 1–18. [[CrossRef](#)] [[PubMed](#)]
16. Xi, C.R.; Di Fazio, A.; Nadvi, N.A.; Patel, K.; Xiang, M.S.W.; Zhang, H.E.; Deshpande, C.; Low, J.K.; Wang, X.T.; Chen, Y.; et al. A novel purification procedure for active recombinant human DPP4 and the inability of DPP4 to bind SARS-CoV-2. *Molecules* **2020**, *25*, 5392. [[CrossRef](#)] [[PubMed](#)]
17. Tsou, E.P.S.; Sule, G.; Rubio, M.G.; Amin, M.A.; Zuo, Y.; Knight, J.; Kanthi, Y.; Fox, D. Identification of CD13 as a Potential Cause for SARS-CoV-2-triggered Hyperinflammation and Thrombosis. *Arthritis Rheumatol.* **2020**, *72*.
18. Yang, L.; Han, Y.; Nilsson-Payant, B.E.; Gupta, V.; Wang, P.; Duan, X.; Tang, X.; Zhu, J.; Zhao, Z.; Jaffré, F.; et al. A human pluripotent stem cell-based platform to study SARS-CoV-2 tropism and model virus infection in human cells and organoids. *Cell Stem Cell* **2020**, *27*, 125–136. [[CrossRef](#)]
19. Hoffmann, M.; Kleine-Weber, H.; Pöhlmann, S. A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* **2020**, *78*, 779–784. [[CrossRef](#)]

20. Johnson, B.A.; Xie, X.; Kalveram, B.; Lokugamage, K.G.; Muruato, A.; Zou, J.; Zhang, X.; Juelich, T.; Smith, J.K.; Zhang, L.; et al. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* **2021**, *591*, 293–299. [[CrossRef](#)]
21. Zang, R.; Castro, M.F.G.; McCune, B.T.; Zeng, Q.; Rothlauf, P.W.; Sonnek, N.M.; Liu, Z.; Brulois, K.F.; Wang, X.; Greenberg, H.B.; et al. TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **2020**, *5*, eabc3582. [[CrossRef](#)]
22. Zhang, C.; Zhang, Y.; Zhang, S.; Wang, Z.; Sun, S.; Liu, M.; Chen, Y.; Dong, N.; Wu, Q. Intracellular autoactivation of TMPRSS11A, an airway epithelial transmembrane serine protease. *J. Biol. Chem.* **2020**, *295*, 12686–12696. [[CrossRef](#)] [[PubMed](#)]
23. Bailey, C.C.; Zhong, G.; Huang, I.C.; Farzan, M. IFITM-family proteins: The cell's first line of antiviral defense. *Annu. Rev. Virol.* **2014**, *1*, 261–283. [[CrossRef](#)] [[PubMed](#)]
24. Bozzo, C.P.; Nchioua, R.; Volcic, M.; Wettstein, L.; Weil, T.; Krüger, J.; Heller, S.; Conzelmann, C.; Müller, J.; Gross, R.; et al. IFITM proteins promote SARS-CoV-2 infection and are targets for virus inhibition in vitro. *Nat. Commun.* **2021**, *12*, 4584. [[CrossRef](#)] [[PubMed](#)]
25. Shi, G.; Kenney, A.D.; Kudryashova, E.; Zani, A.; Zhang, L.; Lai, K.K.; Hall-Stoodley, L.; Robinson, R.T.; Kudryashov, D.S.; Compton, A.A.; et al. Opposing activities of IFITM proteins in SARS-CoV-2 infection. *EMBO J.* **2021**, *40*, e106501. [[CrossRef](#)]
26. Barabási, A.L.; Gulbahce, N.; Loscalzo, J. Network medicine: A network-based approach to human disease. *Nat. Rev. Genet.* **2011**, *12*, 56–68. [[CrossRef](#)]
27. Hu, J.X.; Thomas, C.E.; Brunak, S. Network biology concepts in complex disease comorbidities. *Nat. Rev. Genet.* **2016**, *17*, 615–629. [[CrossRef](#)]
28. Cagno, V. SARS-CoV-2 cellular tropism. *Lancet Microbe* **2020**, *1*, e2. [[CrossRef](#)]
29. Hou, Y.J.; Okuda, K.; Edwards, C.E.; Martinez, D.R.; Asakura, T.; Dinnon III, K.H.; Kato, T.; Lee, R.E.; Yount, B.L.; Mascenik, T.M.; et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **2020**, *182*, 429–446. [[CrossRef](#)]
30. Hikmet, F.; Méar, L.; Edvinsson, Å.; Micke, P.; Uhlén, M.; Lindskog, C. The protein expression profile of ACE2 in human tissues. *Mol. Syst. Biol.* **2020**, *16*, e9610. [[CrossRef](#)]
31. Qi, F.; Qian, S.; Zhang, S.; Zhang, Z. Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochem. Biophys. Res. Commun.* **2020**, *526*, 135–140. [[CrossRef](#)]
32. Murgolo, N.; Therien, A.G.; Howell, B.; Klein, D.; Koeplinger, K.; Lieberman, L.A.; Adam, G.C.; Flynn, J.; McKenna, P.; Swaminathan, G.; et al. SARS-CoV-2 tropism, entry, replication, and propagation: Considerations for drug discovery and development. *PLoS Pathog.* **2021**, *17*, e1009225. [[CrossRef](#)] [[PubMed](#)]
33. de la Fuente, A. From 'differential expression' to 'differential networking'—Identification of dysfunctional regulatory networks in diseases. *Trends Genet.* **2010**, *26*, 326–333. [[CrossRef](#)] [[PubMed](#)]
34. Del Sol, A.; Balling, R.; Hood, L.; Galas, D. Diseases as network perturbations. *Curr. Opin. Biotechnol.* **2010**, *21*, 566–571. [[CrossRef](#)] [[PubMed](#)]
35. Jalili, M.; Salehzadeh-Yazdi, A.; Gupta, S.; Wolkenhauer, O.; Yaghmaie, M.; Resendis-Antonio, O.; Alimoghaddam, K. Evolution of centrality measurements for the detection of essential proteins in biological networks. *Front. Physiol.* **2016**, *7*, 375. [[CrossRef](#)]
36. Iacono, G.; Massoni-Badosa, R.; Heyn, H. Single-cell transcriptomics unveils gene regulatory network plasticity. *Genome Biol.* **2019**, *20*, 1–20. [[CrossRef](#)]
37. Resendis-Antonio, O.; Freyre-Gonzalez, J.A.; Menchaca-Mendez, R.; Gutiérrez-Ríos, R.M.; Martínez-Antonio, A.; Avila-Sanchez, C.; Collado-Vides, J. Modular analysis of the transcriptional regulatory network of *E. coli*. *Trends Genet.* **2005**, *21*, 16–20. [[CrossRef](#)]
38. Resendis-Antonio, O.; Hernández, M.; Mora, Y.; Encarnación, S. Functional modules, structural topology, and optimal activity in metabolic networks. *PLoS Comput. Biol.* **2012**, *8*, e1002720. [[CrossRef](#)]
39. Liao, M.; Liu, Y.; Yuan, J.; Wen, Y.; Xu, G.; Zhao, J.; Cheng, L.; Li, J.; Wang, X.; Wang, F.; et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **2020**, *26*, 842–844. [[CrossRef](#)]
40. Chua, R.L.; Lukassen, S.; Trump, S.; Hennig, B.P.; Wendisch, D.; Pott, F.; Debnath, O.; Thürmann, L.; Kurth, F.; Völker, M.T.; et al. COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. *Nat. Biotechnol.* **2020**, *38*, 970–979. [[CrossRef](#)]
41. Grant, R.A.; Morales-Nebreda, L.; Markov, N.S.; Swaminathan, S.; Querrey, M.; Guzman, E.R.; Abbott, D.A.; Donnelly, H.K.; Donayre, A.; Goldberg, I.A.; et al. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature* **2021**, *590*, 635–641. [[CrossRef](#)]
42. Morse, C.; Tabib, T.; Sembrat, J.; Buschur, K.L.; Bittar, H.T.; Valenzi, E.; Jiang, Y.; Kass, D.J.; Gibson, K.; Chen, W.; et al. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur. Respir. J.* **2019**, *54*. [[CrossRef](#)] [[PubMed](#)]
43. Lukassen, S.; Chua, R.L.; Trefzer, T.; Kahn, N.C.; Schneider, M.A.; Muley, T.; Winter, H.; Meister, M.; Veith, C.; Boots, A.W.; et al. SARS-CoV-2 receptor ACE 2 and TMPRSS 2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* **2020**, *39*, e105114. [[CrossRef](#)] [[PubMed](#)]
44. Reyfman, P.A.; Walter, J.M.; Joshi, N.; Anekalla, K.R.; McQuattie-Pimentel, A.C.; Chiu, S.; Fernandez, R.; Akbarpour, M.; Chen, C.I.; Ren, Z.; et al. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **2019**, *199*, 1517–1536. [[CrossRef](#)] [[PubMed](#)]
45. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck III, W.M.; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive integration of single-cell data. *Cell* **2019**, *177*, 1888–1902. [[CrossRef](#)]

46. Hafemeister, C.; Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **2019**, *20*, 1–15. [[CrossRef](#)]
47. Haghverdi, L.; Lun, A.T.; Morgan, M.D.; Marioni, J.C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **2018**, *36*, 421–427. [[CrossRef](#)]
48. Butler, A.; Hoffman, P.; Smibert, P.; Papalexi, E.; Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **2018**, *36*, 411–420. [[CrossRef](#)]
49. Madhulatha, T.S. An overview on clustering methods. *arXiv* **2012**, arXiv:1205.1117.
50. Xu, C.; Su, Z. Identification of cell types from single-cell transcriptomes using a novel clustering method. *Bioinformatics* **2015**, *31*, 1974–1980. [[CrossRef](#)]
51. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [[CrossRef](#)]
52. Waltman, L.; Van Eck, N.J. A smart local moving algorithm for large-scale modularity-based community detection. *Eur. Phys. J. B* **2013**, *86*, 1–14. [[CrossRef](#)]
53. Travaglini, K.J.; Nabhan, A.N.; Penland, L.; Sinha, R.; Gillich, A.; Sit, R.V.; Chang, S.; Conley, S.D.; Mori, Y.; Seita, J.; et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **2020**, *587*, 619–625. [[CrossRef](#)] [[PubMed](#)]
54. Moerman, T.; Aibar Santos, S.; Bravo González-Blas, C.; Simm, J.; Moreau, Y.; Aerts, J.; Aerts, S. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **2019**, *35*, 2159–2161. [[CrossRef](#)] [[PubMed](#)]
55. Garcia-Alonso, L.; Holland, C.H.; Ibrahim, M.M.; Turei, D.; Saez-Rodriguez, J. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res.* **2019**, *29*, 1363–1375. [[CrossRef](#)] [[PubMed](#)]
56. Nakatake, Y.; Ko, S.B.; Sharov, A.A.; Wakabayashi, S.; Murakami, M.; Sakota, M.; Chikazawa, N.; Ookura, C.; Sato, S.; Ito, N.; et al. Generation and profiling of 2,135 human ESC lines for the systematic analyses of cell states perturbed by inducing single transcription factors. *Cell Rep.* **2020**, *31*, 107655. [[CrossRef](#)] [[PubMed](#)]
57. Feng, C.; Song, C.; Liu, Y.; Qian, F.; Gao, Y.; Ning, Z.; Wang, Q.; Jiang, Y.; Li, Y.; Li, M.; et al. KnockTF: A comprehensive human gene expression profile database with knockdown/knockout of transcription factors. *Nucleic Acids Res.* **2020**, *48*, D93–D100. [[CrossRef](#)] [[PubMed](#)]
58. Han, H.; Cho, J.W.; Lee, S.; Yun, A.; Kim, H.; Bae, D.; Yang, S.; Kim, C.Y.; Lee, M.; Kim, E.; et al. TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **2018**, *46*, D380–D386. [[CrossRef](#)]
59. Van de Sande, B.; Flerin, C.; Davie, K.; De Waegeneer, M.; Hulselmans, G.; Aibar, S.; Seurinck, R.; Saelens, W.; Cannoodt, R.; Rouchon, Q.; et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nature Protoc.* **2020**, *15*, 2247–2276. [[CrossRef](#)]
60. Ding, H.; Yang, Y.; Xue, Y.; Seninge, L.; Gong, H.; Safavi, R.; Califano, A.; Stuart, J.M. Prioritizing transcriptional factors in gene regulatory networks with PageRank. *IScience* **2021**, *24*, 102017. [[CrossRef](#)]
61. Koschützki, D.; Schreiber, F. Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene Regul. Syst. Biol.* **2008**, *2*, GRSB–S702. [[CrossRef](#)]
62. Corsello, S.M.; Bittker, J.A.; Liu, Z.; Gould, J.; McCarren, P.; Hirschman, J.E.; Johnston, S.E.; Vrcic, A.; Wong, B.; Khan, M.; et al. The Drug Repurposing Hub: A next-generation drug library and information resource. *Nat. Med.* **2017**, *23*, 405–408. [[CrossRef](#)] [[PubMed](#)]
63. Subramanian, A.; Narayan, R.; Corsello, S.M.; Peck, D.D.; Natoli, T.E.; Lu, X.; Gould, J.; Davis, J.F.; Tubelli, A.A.; Asiedu, J.K.; et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **2017**, *171*, 1437–1452. [[CrossRef](#)] [[PubMed](#)]
64. Skuta, C.; Popr, M.; Muller, T.; Jindrich, J.; Kahle, M.; Sedlak, D.; Svozil, D.; Bartunek, P. Probes & Drugs portal: An interactive, open data resource for chemical biology. *Nat. Methods* **2017**, *14*, 759–760. [[PubMed](#)]
65. McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv* **2018**, arXiv:1802.03426.
66. Dorrity, M.W.; Saunders, L.M.; Queitsch, C.; Fields, S.; Trapnell, C. Dimensionality reduction by UMAP to visualize physical and genetic interactions. *Nature Commun.* **2020**, *11*, 1–6. [[CrossRef](#)] [[PubMed](#)]
67. Ziegler, C.G.; Allon, S.J.; Nyquist, S.K.; Mbano, I.M.; Miao, V.N.; Tzouanas, C.N.; Cao, Y.; Yousif, A.S.; Bals, J.; Hauser, B.M.; et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* **2020**, *181*, 1016–1035. [[CrossRef](#)] [[PubMed](#)]
68. Zhao, M.M.; Yang, W.L.; Yang, F.Y.; Zhang, L.; Huang, W.J.; Hou, W.; Fan, C.F.; Jin, R.H.; Feng, Y.M.; Wang, Y.C.; et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct. Target. Ther.* **2021**, *6*, 1–12. [[CrossRef](#)]
69. Lambert, S.A.; Jolma, A.; Campitelli, L.F.; Das, P.K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T.R.; Weirauch, M.T. The human transcription factors. *Cell* **2018**, *172*, 650–665. [[CrossRef](#)]
70. Pedersen, K.B.; Chodavarapu, H.; Lazartigues, E. Forkhead box transcription factors of the FOXA class are required for basal transcription of angiotensin-converting enzyme 2. *J. Endocr. Soc.* **2017**, *1*, 370–384. [[CrossRef](#)]
71. Hennighausen, L.; Lee, H.K. Activation of the SARS-CoV-2 receptor Ace2 through JAK/STAT-dependent enhancers during pregnancy. *Cell Rep.* **2020**, *32*, 108199. [[CrossRef](#)]

72. Matsuyama, T.; Kubli, S.P.; Yoshinaga, S.K.; Pfeffer, K.; Mak, T.W. An aberrant STAT pathway is central to COVID-19. *Cell Death Differ.* **2020**, *27*, 3209–3225. [CrossRef] [PubMed]
73. Hogan, R.J.; Gao, G.; Rowe, T.; Bell, P.; Flieder, D.; Paragas, J.; Kobinger, G.P.; Wivel, N.A.; Crystal, R.G.; Boyer, J.; et al. Resolution of primary severe acute respiratory syndrome-associated coronavirus infection requires Stat1. *J. Virol.* **2004**, *78*, 11416–11421. [CrossRef] [PubMed]
74. Parolia, A.; Cieslik, M.; Chu, S.C.; Xiao, L.; Ouchi, T.; Zhang, Y.; Wang, X.; Vats, P.; Cao, X.; Pitchiaya, S.; et al. Distinct structural classes of activating FOXA1 alterations in advanced prostate cancer. *Nature* **2019**, *571*, 413–418. [CrossRef] [PubMed]
75. Qiao, Y.; Wang, X.M.; Mannan, R.; Pitchiaya, S.; Zhang, Y.; Wotring, J.W.; Xiao, L.; Robinson, D.R.; Wu, Y.M.; Tien, J.C.Y.; et al. Targeting transcriptional regulation of SARS-CoV-2 entry factors ACE2 and TMPRSS2. *Proc. Natl. Acad. Sci. USA* **2021**, *118*. [CrossRef] [PubMed]
76. Chilakamarti, R.; Bikul, D. Regulation of early growth response-1 (Egr-1) gene expression by Stat1-independent type I interferon signaling and respiratory viruses. *Comput. Math. Biophys.* **2021**, *1*, 289–303.
77. Patra, T.; Meyer, K.; Geerling, L.; Isbell, T.S.; Hoft, D.F.; Brien, J.; Pinto, A.K.; Ray, R.B.; Ray, R. SARS-CoV-2 spike protein promotes IL-6 trans-signaling by activation of angiotensin II receptor signaling in epithelial cells. *PLoS Pathog.* **2020**, *16*, e1009128. [CrossRef] [PubMed]
78. Mayi, B.S.; Leibowitz, J.A.; Woods, A.T.; Ammon, K.A.; Liu, A.E.; Raja, A. The role of Neuropilin-1 in COVID-19. *PLoS Pathog.* **2021**, *17*, e1009153. [CrossRef]
79. He, R.; Leeson, A.; Andonov, A.; Li, Y.; Bastien, N.; Cao, J.; Osioy, C.; Dobie, F.; Cutts, T.; Ballantine, M.; et al. Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein. *Biochem. Biophys. Res. Commun.* **2003**, *311*, 870–876. [CrossRef]
80. Wang, H.; Li, X.; Li, T.; Zhang, S.; Wang, L.; Wu, X.; Liu, J. The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur. J. Clin. Microbiol. Infect. Dis.* **2020**, *39*, 1629–1635. [CrossRef]
81. Bagheri, A.; Moezzi, S.M.I.; Mosaddeghi, P.; Parashkouhi, S.N.; Hoseini, S.M.F.; Badakhshan, F.; Negahdaripour, M. Interferon-inducer antivirals: Potential candidates to combat COVID-19. *Int. Immunopharmacol.* **2021**, *91*, 107245. [CrossRef]
82. Chakrabarty, B.; Das, D.; Bulusu, G.; Roy, A. Network-based analysis of fatal comorbidities of COVID-19 and potential therapeutics. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2021**, *18*, 1271–1280 [CrossRef] [PubMed]
83. Khan, A.T.A.; Khalid, Z.; Zahid, H.; Yousaf, M.A.; Shakoori, A.R. A computational and bioinformatic analysis of ACE2: An elucidation of its dual role in COVID-19 pathology and finding its associated partners as potential therapeutic targets. *J. Biomol. Struct. Dyn.* **2020**, *40*, 1813–1829 [CrossRef] [PubMed]
84. Mamoor, S. The Activating Transcription Factor ATF3 is Differentially Expressed and Transcriptionally Induced in Multiple Models of Coronavirus Infection. 2020. Available online: <https://osf.io/dhnpm> (accessed on 2 March 2022).
85. Shamir, I.; Abutbul-Amitai, M.; Abbas-Egbariya, H.; Pasmanik-Chor, M.; Paret, G.; Nevo-Caspi, Y. STAT3 isoforms differentially affect ACE2 expression: A potential target for COVID-19 therapy. *J. Cell. Mol. Med.* **2020**, *24*, 12864–12868. [CrossRef] [PubMed]
86. Yuan, S.; Yin, X.; Meng, X.; Chan, J.F.W.; Ye, Z.W.; Riva, L.; Pache, L.; Chan, C.C.Y.; Lai, P.M.; Chan, C.C.S.; et al. Clofazimine broadly inhibits coronaviruses including SARS-CoV-2. *Nature* **2021**, *593*, 418–423. [CrossRef] [PubMed]
87. Liu, J.; Li, Y.; Liu, Q.; Yao, Q.; Wang, X.; Zhang, H.; Chen, R.; Ren, L.; Min, J.; Deng, F.; et al. SARS-CoV-2 cell tropism and multiorgan infection. *Cell Discov.* **2021**, *7*, 1–4. [CrossRef]
88. Mu, J.; Fang, Y.; Yang, Q.; Shu, T.; Wang, A.; Huang, M.; Jin, L.; Deng, F.; Qiu, Y.; Zhou, X. SARS-CoV-2 N protein antagonizes type I interferon signaling by suppressing phosphorylation and nuclear translocation of STAT1 and STAT2. *Cell Discov.* **2020**, *6*, 1–4. [CrossRef]
89. Myers, J.S.; von Lersner, A.K.; Robbins, C.J.; Sang, Q.X.A. Differentially expressed genes and signature pathways of human prostate cancer. *PLoS ONE* **2015**, *10*, e0145322. [CrossRef]
90. Murata, M.; Yoshida, K.; Yamaguchi, T.; Matsuzaki, K. Linker phosphorylation of Smad3 promotes fibro-carcinogenesis in chronic viral hepatitis of hepatocellular carcinoma. *World J. Gastroenterol. WJG* **2014**, *20*, 15018. [CrossRef]
91. Cowley, D.O.; Graves, B.J. Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. *Genes Dev.* **2000**, *14*, 366–376. [CrossRef]
92. Nasr, R.; Guillemin, M.C.; Ferhi, O.; Soilihi, H.; Peres, L.; Berthier, C.; Rousselot, P.; Robledo-Sarmiento, M.; Lallemand-Breitenbach, V.; Gourmel, B.; et al. Eradication of acute promyelocytic leukemia-initiating cells through PML-RARA degradation. *Nat. Med.* **2008**, *14*, 1333–1342. [CrossRef]
93. Olsen, J.V.; Vermeulen, M.; Santamaria, A.; Kumar, C.; Miller, M.L.; Jensen, L.J.; Gnad, F.; Cox, J.; Jensen, T.S.; Nigg, E.A.; et al. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* **2010**, *3*, ra3. [CrossRef] [PubMed]
94. Watanabe, T.; Hiasa, Y.; Tokumoto, Y.; Hirooka, M.; Abe, M.; Ikeda, Y.; Matsuura, B.; Chung, R.T.; Onji, M. Protein kinase R modulates c-Fos and c-Jun signaling to promote proliferation of hepatocellular carcinoma with hepatitis C virus infection. *PLoS ONE* **2013**, *8*, e67750. [CrossRef] [PubMed]
95. Hock, H.; Shimamura, A. ETV6 in hematopoiesis and leukemia predisposition. In *Seminars in Hematology*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 54, pp. 98–104.
96. Rocca, D.L.; Wilkinson, K.A.; Henley, J.M. SUMOylation of FOXP1 regulates transcriptional repression via CtBP1 to drive dendritic morphogenesis. *Sci. Rep.* **2017**, *7*, 1–12. [CrossRef] [PubMed]

97. Kunz, K.; Wagner, K.; Mendler, L.; Hölper, S.; Dehne, N.; Müller, S. SUMO signaling by hypoxic inactivation of SUMO-specific isopeptidases. *Cell Rep.* **2016**, *16*, 3075–3086. [[CrossRef](#)]
98. Illendula, A.; Gilmour, J.; Grembecka, J.; Tirumala, V.S.S.; Boulton, A.; Kuntimaddi, A.; Schmidt, C.; Wang, L.; Pulikkan, J.A.; Zong, H.; et al. Small molecule inhibitor of CBF $\beta$ -RUNX binding for RUNX transcription factor driven cancers. *EBioMedicine* **2016**, *8*, 117–131. [[CrossRef](#)]
99. Mason, R.J. Pathogenesis of COVID-19 from a Cell Biology Perspective. *Eur. Respir. J.* **2020**, *55*, 2000607. [[CrossRef](#)]
100. Mulay, A.; Konda, B.; Garcia, G., Jr.; Yao, C.; Beil, S.; Villalba, J.M.; Koziol, C.; Sen, C.; Purkayastha, A.; Kolls, J.K.; et al. SARS-CoV-2 infection of primary human lung epithelium for COVID-19 modeling and drug discovery. *Cell Rep.* **2021**, *35*, 109055. [[CrossRef](#)]
101. An, X.; Zhang, Y.; Duan, L.; Jin, D.; Zhao, S.; Zhou, R.; Duan, Y.; Lian, F.; Tong, X. The Direct Evidence and Mechanism of Traditional Chinese Medicine Treatment of COVID-19. *Biomed. Pharmacother.* **2021**, *137*, 111267. [[CrossRef](#)]
102. Zhang, L.; Li, J.; Zhou, M.; Chen, Z. Summary of 20 tracheal intubation by anesthesiologists for patients with severe COVID-19 pneumonia: Retrospective case series. *J. Anesth.* **2020**, *34*, 599–606. [[CrossRef](#)]
103. Wei, J.; Alfajaro, M.M.; DeWeirdt, P.C.; Hanna, R.E.; Lu-Culligan, W.J.; Cai, W.L.; Strine, M.S.; Zhang, S.M.; Graziano, V.R.; Schmitz, C.O.; et al. Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cell* **2021**, *184*, 76–91. [[CrossRef](#)]
104. Li, S.; Zhang, Y.; Guan, Z.; Li, H.; Ye, M.; Chen, X.; Shen, J.; Zhou, Y.; Shi, Z.L.; Zhou, P.; et al. SARS-CoV-2 triggers inflammatory responses and cell death through caspase-8 activation. *Signal Transduct. Target. Ther.* **2020**, *5*, 1–10. [[CrossRef](#)] [[PubMed](#)]
105. Huynh-Thu, V.A.; Irrthum, A.; Wehenkel, L.; Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* **2010**, *5*, e12776. [[CrossRef](#)] [[PubMed](#)]
106. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]