PLOS PATHOGENS

RESEARCH ARTICLE

The endogenous cellular protease inhibitor SPINT2 controls SARS-CoV-2 viral infection and is associated to disease severity

Carlos Ramirez Alvarez[1☯], Carmon Kee[2,3☯], Ashwini Kumar Sharma[1], Leonie Thomas[1], Florian I. Schmidt[4], Megan L. Stanifer[5], Steeve Boulant[2,3]*, Carl Herrmann[1]*

1 Health Data Science Unit, Medical Faculty Heidelberg and BioQuant, Heidelberg, Germany, **2** Department of Infectious Diseases, Virology, Heidelberg University, Heidelberg, Germany, **3** Research Group "Cellular Polarity and Viral Infection", German Cancer Research Center (DKFZ), Heidelberg, Germany, **4** Institute of Innate Immunity, Medical Faculty, University of Bonn, Bonn, Germany, **5** Department of Infectious Diseases, Molecular Virology, Heidelberg University, Heidelberg, Germany

☯ These authors contributed equally to this work.
* steeve.boulant@med.uni-heidelberg.de (SB); carl.herrmann@bioquant.uni-heidelberg.de (CH)

Citation: Ramirez Alvarez C, Kee C, Sharma AK, Thomas L, Schmidt FI, Stanifer ML, et al. (2021) The endogenous cellular protease inhibitor SPINT2 controls SARS-CoV-2 viral infection and is associated to disease severity. PLoS Pathog 17(6): e1009687. https://doi.org/10.1371/journal.ppat.1009687

Editor: Benhur Lee, Icahn School of Medicine at Mount Sinai, UNITED STATES

Received: March 31, 2021

Accepted: June 2, 2021

Published: June 28, 2021

Data Availability Statement: For this work we used the following datasets available in public repositories: scRNA-seq profiles of Calu-3 and H1299 cell lines (10.1016/j.isci.2021.102151); scRNA-seq from ileum derived organoids (10.15252/msb.202110232); translatome and proteome quantifications of Caco-2 cell line (10.1038/s41586-020-2332-7); bulk RNA-seq profiles of lung samples of COVID-19 deceased patient autopsies (10.1038/s41467-020-20139-7); Peripheral Blood Mononuclear Cells (10.1038/

Abstract

COVID-19 outbreak is the biggest threat to human health in recent history. Currently, there are over 1.5 million related deaths and 75 million people infected around the world (as of 22/12/2020). The identification of virulence factors which determine disease susceptibility and severity in different cell types remains an essential challenge. The serine protease *TMPRSS2* has been shown to be important for S protein priming and viral entry, however, little is known about its regulation. *SPINT2* is a member of the family of Kunitz type serine protease inhibitors and has been shown to inhibit *TMPRSS2*. Here, we explored the existence of a co-regulation between *SPINT2*/*TMPRSS2* and found a tightly regulated protease/inhibitor expression balance across tissues. We found that *SPINT2* negatively correlates with SARS-CoV-2 expression in Calu-3 and Caco-2 cell lines and was down-regulated in secretory cells from COVID-19 patients. We validated our findings using Calu-3 cell lines and observed a strong increase in viral load after *SPINT2* knockdown, while over-expression lead to a drastic reduction of the viral load. Additionally, we evaluated the expression of *SPINT2* in datasets from comorbid diseases using bulk and scRNA-seq data. We observed its down-regulation in colon, kidney and liver tumors as well as in alpha pancreatic islets cells from diabetes Type 2 patients, which could have implications for the observed comorbidities in COVID-19 patients suffering from chronic diseases.

Author summary

Since early 2020, the world is facing a unique pandemic situation triggered by the appearance of a new coronavirus, Sars-CoV-2. While it shares similarities with previous coronaviruses, it shows some unique features, which requires in-depth investigations to understand its mechanisms of action, in particular the type of cells which can be infected,

s41591-020-0944-y); Nasopharynx samples from COVID-19 patients (10.1038/s41587-020-0602-4); Human Cell Landscape (10.1038/s41586-020-2157-4); The Cancer Genome Atlas. Tumor scRNA-seq datasets: Human Hepatocellular Carcinoma (10.2139/ssrn.3624436), renal Clear Cell Carcinoma (10.1126/science.aat1699), Lung Adenocarcinoma (10.1038/s41591-019-0750-6), Colon Adenocarcinoma (10.1038/ng.3818), pancreatic cells from Diabetes Type 2 patients (0.1016/j.cmet.2016.08.020); ATAC-seq data from Human intestinal organoids (10.1152/physiolgenomics.00113.2019). All the codes for the data processing and analysis are provided in the following GitHub repository: https://github.com/hdsu-bioquant/covid19-comorbidity.

Funding: This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) project numbers 415089553 (Heisenberg program), 240245660 (SFB1129), and 278001972 (TRR186) to SB, and project 416072091 to MS. SB also receveived support from the state of Baden Wuerttemberg (AZ: 33.7533.-6-21/5/1), the Bundesministerium Bildung und Forschung (BMBF) (01KI20198A) and the German Federal Ministry of Education and Research, Network University Medicine - Organo-Strat COVID-19. We also acknowledge funding from the Helmholtz International Graduate School for Cancer Research to CK. CRA, SB, CH acknowledge funding through the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 272983813 (TRR 179). SB and CH acknowledge funding through the DFG program "Identification of the molecular origins of comorbidity in COVID-19 patients" (Project-ID 458633366). This study was partly supported by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation): SFB1403-414786233 (F.I.S.), TRR237-369799452 (F.I.S.), and the Emmy Noether Programme 322568668 (F.I.S.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

and the mechanism of cellular entry. We describe a new important actor which plays a decisive role in the entry of the virus into the cell. The serine protease inhibitor *SPINT2* inhibits the well described serine protease *TMPRSS2*, and we show that inhibiting or over-expressing this gene leads to an increase, respectively a strong decrease of the number of infected cells. Intriguingly, this gene has been described in the context of diseases which are known to be comorbid with COVD19, for example Type 2 diabetes and several malignancies. Using published datasets, we show that in these comorbid diseases, SPINT2 is downregulated, which might lead to an increase of the infectivity of cells in disease-tissues such as endocrine cells of diabetic patients or cells from tumor tissues.

Introduction

SARS-CoV-2 entry requires a two-step process: first, the envelope protein spike (S) binds to the viral cellular receptor Angiotensin-converting enzyme 2 (*ACE2*) membrane protein [1] and is then proteolytically activated by cellular serine proteases (SPs) like *TMPRSS2*, *TMPRSS4* and *Furin* [2–4]. A wider spectrum of serine proteases might also contribute to viral priming since *ACE2* and *TMPRSS2* co-expression at the gene expression level is mainly restricted to specific cell types like ciliated, alveolar type 2 and secretory cells in lungs and enterocytes in colon and ileum [5]. A possible role of serine proteases and their inhibitors in SARS-CoV-2 viral infection has been examined in recent reviews [6, 7]. *TMPRSS2* has been proposed as a putative drug target [3, 8, 9] and as a biomarker for COVID19 disease severity [10, 11]. Despite its central role, the regulation of *TMPRSS2* is poorly understood, although its activation by androgen response elements has been documented in normal and tumor prostate tissues [12].

SPINT2, a member of the Kunitz-type serine proteases inhibitors [13] has been shown to inhibit TMPRSS2 protease activity [14, 15] which could have implications for COVID-19 disease. A down-regulation of *SPINT1*, *SPINT2* and *SERPINA1* has been reported in colorectal Caco-2 cells infected with SARS-CoV-2 [16], however, a clear association between *SPINT2* activation and viral permissivity has not been confirmed. To fill this gap, we evaluated the existence of coregulation between *SPINT2*/*TMPRSS2* and found common transcription factors (TF) associated with genomic loci for both genes which was also in line with a consistent correlation of the genes across cell types. This coregulation suggested a modulation of SARS-CoV-2 infection by *SPINT2*. We could corroborate a negative correlation between *SPINT2* gene expression and SARS-CoV-2 viral load in Calu-3 and Caco-2 cell lines. To validate our findings we knocked-down *SPINT2* in Calu-3 and A549 cell lines and observed an increase in the number of SARS-CoV-2 infected cells. We hypothesized that *SPINT2* levels would be lower in SARS-CoV-2 target cells from COVID-19 patients with severe symptoms which we could indeed observe in secretory cells from nasopharynx samples. This suggests that *SPINT2* can be used as a biomarker for disease susceptibility. Finally, it is known that *SPINT2* is down-regulated among different types of tumors [17] and we were able to corroborate this by systematically evaluating bulk- and scRNA-seq datasets which suggests a possible association to the COVID-19 comorbidity observed in cancer patients.

Results and discussions

SPINT2 and *TMPRSS2* are coregulated across tissues

TMPRSS2 proteolytic activity inhibition by *SPINT2* has been previously reported [14, 15]. We investigated a coregulation between *SPINT2* and *TMPRSS2*, as a similar shared regulation through the transcription factor (TF) *CDX2* has been described for *SPINT1* and *ST14*

(Matriptase) previously in enterocytes [18]. Since *SPINT2* is also able to regulate *ST14* activity in small and large intestines [19] we decided to use enterocytes as a model to test this hypothesis. In order to find common TFs regulators of *SPINT2*/*TMPRSS2* we performed two independent analysis: *i*) A footprinting analysis of chromatin open regions using ATAC-seq data from Human Intestinal Organoids [20] identifying potential TF binding sites and *ii*) Using scRNA-seq data from ileum derived organoids [21] we calculated the activity of transcription factors based on the gene expression of their targets using the SCENIC algorithm [22]. TF activities were then correlated to *SPINT2* and *TMPRSS2* gene expression. We identified common TFs inferred to be bound to the open chromatin sites in the *SPINT2* and *TMPRSS2* genomic loci (**Fig 1A**, top and bottom, respectively) and those with TF activities positively correlated to both *SPINT2* and *TMPRSS2* gene expression (**Fig 1B**, top right quadrant). Comparing these two sets of TFs, we found ten shared regulators: *ELF3*, *FOS*, *FOSL1*, *FOXC1*, *IRF1*, *IRF7*, *JUND*, *JUNB*, *ONECUT3* and *KLF4*. Interestingly, many of these regulators play a role in immune response upon infection, suggesting a possible feedback mechanism. SARS-CoV-2 infection has been shown to upregulate *FOS* expression in Huh7.5 and A549 cell lines [23]. *IRF1* and *IRF7* are interferon regulatory factors which regulate infection responses, and have been observed to be upregulated in COVID-19 patients [24]. *JUNB* has been found in SARS-CoV-2 infection gene expression signatures in Calu-3 and Caco-2 cell lines [25]. *ELF3* is an

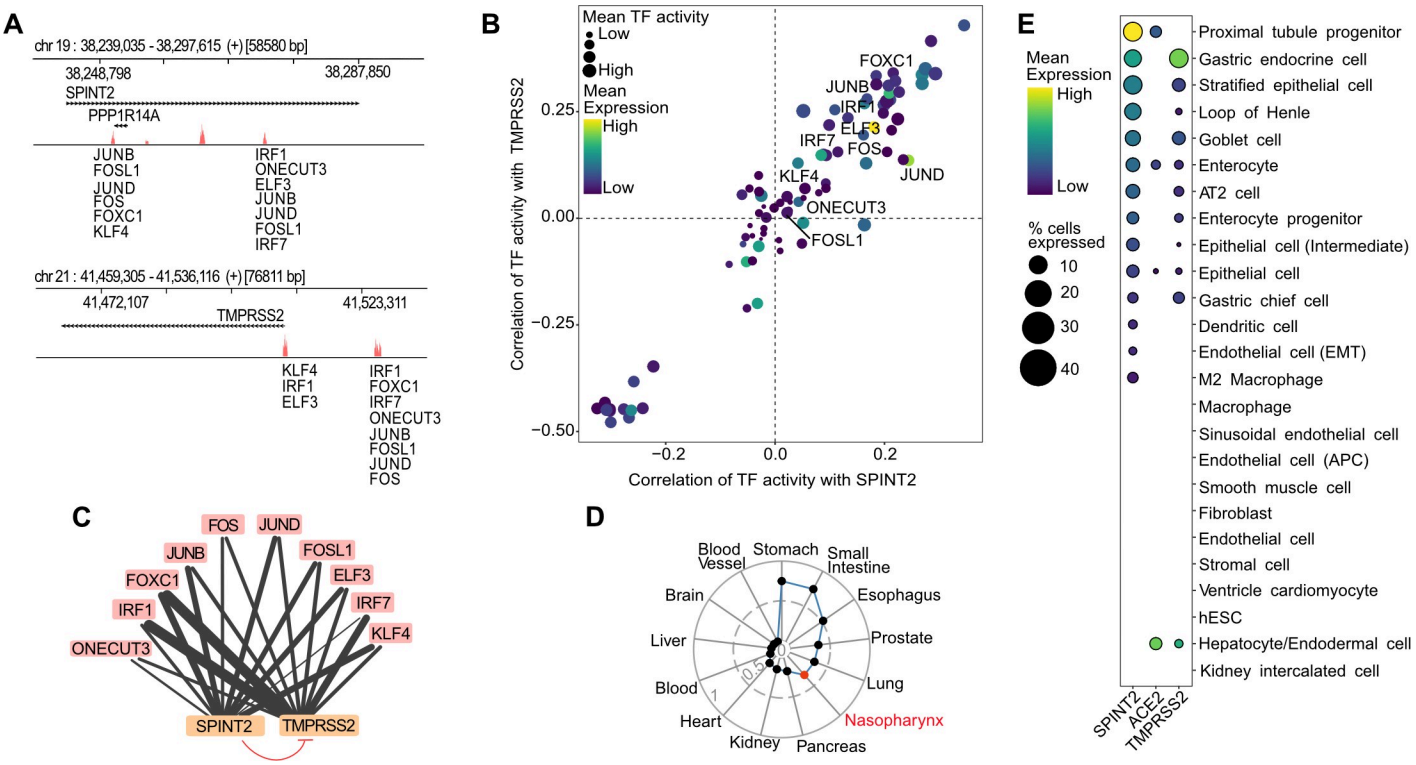

Fig 1. Coregulation of *SPINT2* and *TMPRSS2*. **A**. Peaks associated to open chromatin sites in regions containing the *SPINT2* and *TMPRSS2* genes (+/- 5kb of the coding regions) using ATAC-seq data from Human Intestinal Organoids. TF Binding Sites (TFBS) are scored based on the height and width of the peaks and the similarity of the sequences in the peaks to TF motifs. **B**. Correlation values of TF activities to *SPINT2* and *TMPRSS2* gene expression levels using scRNA-seq data from ileum derived organoids. Labeled TFs in (*A*) and (*B*) correspond to the intersection of inferred TFs bound to TFBS in *(A)* and those with positive correlation of TF activities to both *SPINT2* and *TMPRSS2* in *(B)*, top right quadrant. **C**. Model for *SPINT2* and *TMPRSS2* coregulation. Edges from TFs (blue) to *SPINT2* and *TMPRSS2* (salmon) represent inferred binding shown in *(A)*, *(B)* and the edge width reflects TF binding scores. The blunt red arrow represents the inhibition of *TMPRSS2* enzymatic activity by *SPINT2*. **D**. Correlation of *SPINT2* and *TMPRSS2* across tissues using bulk RNA-seq from GTEx. The additional red point labeled as nasopharynx represents the correlation of an independent dataset from nasopharynx tissue of COVID-19 patients. **E**. Gene expression of *ACE2*, *SPINT2* and *TMPRSS2* across cell types using HCL scRNA-seq data. EMT: Endothelial to Mesenchymal Transition. APC: Antigen Presenting Cell. hESC: human Embryonic Stem Cells.

https://doi.org/10.1371/journal.ppat.1009687.g001

important factor controlling the development of epithelium tissues [26] including intestinal epithelia [27]. Based on our analysis, we depict the regulatory model in which *SPINT2* and *TMPRSS2* gene expression are coregulated by common TFs in ileum enterocytes, hence maintaining the protease/inhibitor balance (**Fig 1C**). On the other hand, down-regulation of TMPRSS2 enzymatic activity by SPINT2 possibly maintains viral load at a low level. In line with this, we observed that the expression of *SPINT2* and *TMPRSS2* are positively correlated across normal human tissues (**Fig 1D**). Interestingly, tissues which have been shown to be targets for the virus display the highest correlation between *SPINT2* and *TMPRSS2* expression. Also, at single cell resolution, we observed that both genes are specifically co-expressed in many cell types [28] (**Fig 1E**) which corroborates the inferred coregulation.

Deriving a SARS-CoV-2 permissivity signature

Given the previous observed association between *SPINT2* and *TMPRSS2*, a major SARS-CoV-2 virulence factor [3], we asked whether *SPINT2* could account for differences in viral permissivity. Calu-3 and H1299 cells have been previously reported as SARS-CoV-2 permissive and non-permissive cell lines, respectively [25]. To test our hypothesis, we inferred a SARS-CoV-2 permissivity signature using the following approach: we calculated Differentially Expressed Genes (DEGs) between non-infected Calu-3 and H1299 cells (set *z* in **Fig 2A**). Because we were interested in *a priori* permissivity factors we excluded from this list any viral-induced

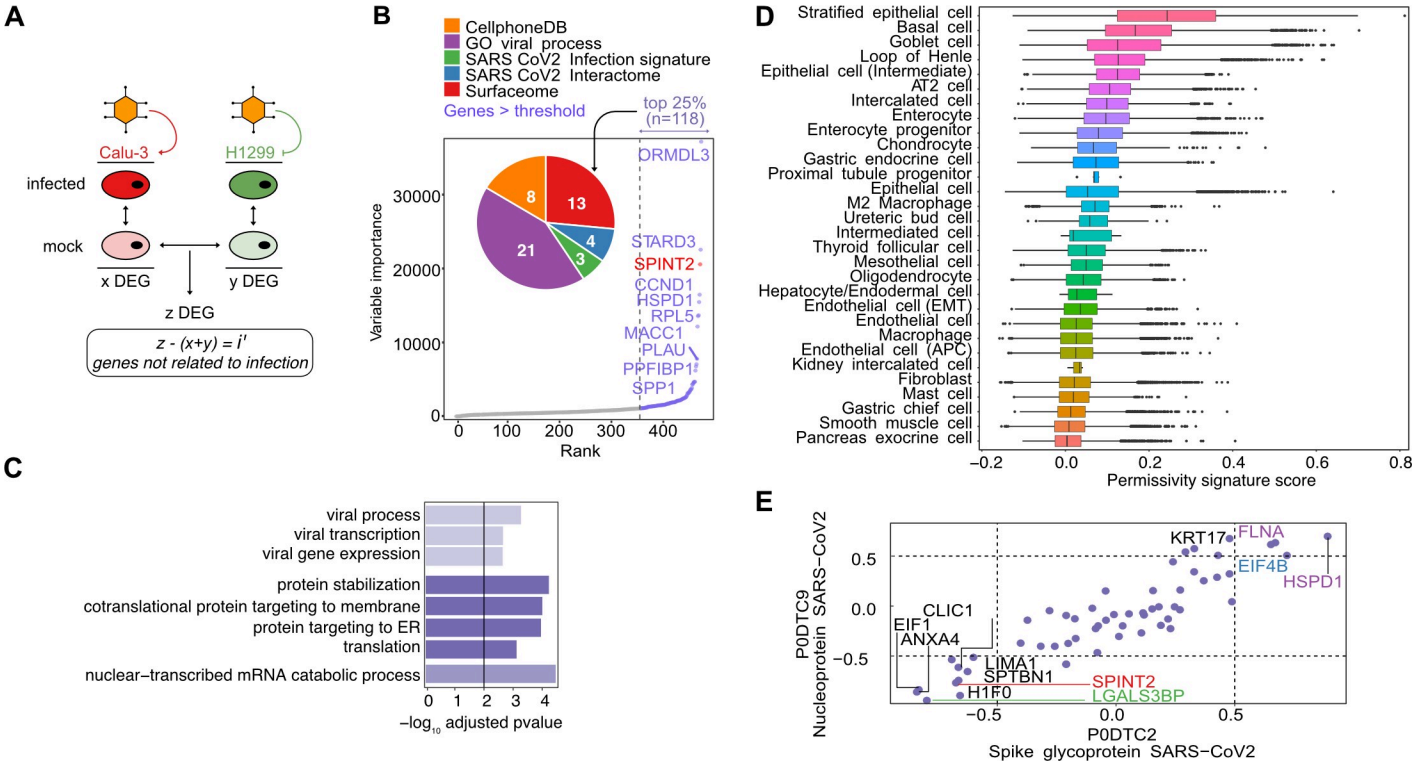

Fig 2. Permissivity genes are associated to viral related processes and correlated to viral expression. A. Permissivity signature genes (z) were found using differential gene expression analysis (DEA) between Calu-3 and H1299 permissive and non-permissive cell lines. Similarly, infection signatures for both cell lines were calculated independently (x + y) and filtered out from the permissive signature (*z*) resulting in (*i'*). **B.** Permissive genes ranked by importance in predicting the cumulative sum of gene expression of viral genes. Only the top 10 ranked genes are labeled with names. Top 25% genes are shown in color and *SPINT2* is highlighted in red. The inset plot show the number of genes in the intersection with each geneset. See the main text for the description of each gene set. **C.** Pathway Enrichment Analysis of the top 25% ranked genes. **D.** Ranking of cells in HCL using the permissivity scores based on the same genes as in *(B)* and *(C)*. **E.** Correlation to viral proteins (S and N genes) in the translatome data of Caco-2 cells. Gene names are coloured as in the pie chart in *(B)*. APC: Antigen Presenting Cells. AT2: Alveolar Type 2 cells.

https://doi.org/10.1371/journal.ppat.1009687.g002

genes. In order to do so, we calculated DEGs between the infected *vs* mock-infected cells from both Calu-3 (x) and H1299 cell lines (y) and then filtered out these genes from z to obtain non-viral inducible genes *(i')* (**Fig 2A**). Filtered genes ($x + y$) shown in the **S1A Fig** and listed in **S1 Table** are enriched in pathways related to RNA and protein synthesis and viral processes (**S1B Fig**) as has been also previously reported [29]. Using this approach, we identified a set of 480 candidate genes which might contribute to infection permissivity (**S2 Table**). We used normalized expression values for these 480 genes as input to train a Random Forest (RF) model for predicting the cumulative sum of viral gene expression in Calu-3 infected cells (**Fig 2B**). We then used the top 25% ranked genes for further analysis. *SPINT2* was found among the top ranked genes (**Fig 2B**). Of the top ranked genes, 21 corresponded to genes with functional annotations related to viral infection (**Fig 2B** inset pie chart and **S2 Table**), four have been reported to participate and/or interact directly with SARS-CoV-2 [30], eight corresponded to curated receptors [31] or ligands in the CellPhone Data Base [32] and 13 are cell membrane surface proteins. Importantly, by intersecting the set of genes in the permissivity signature with a reported SARS-CoV-2 viral infection transcriptional signature [33], we only found 3 hits (*CXCL5*, *LGALS3BP* and *EHF*), confirming that most of the identified genes in our permissivity signature are not viral-induced, but likely represent *a priori* susceptibility factors. We found the following Heat Shock Proteins (HSP): *HSPB1*, *HSPA8* and *HSPD1* to be differentially expressed. In a previous study a different HSP, *HSP90*, was observed to correlate to SARS-CoV-2 viral load in Calu-3 cells and its inhibition reduced viral infection [25]. We also found several ribosomal proteins (*RPL9*, *RPL23*, *RPL26*, *RPL28*, *RPL38*, *RPS7*, *RPS12* and *RPS27A*) and elongation factors (*EIF3A*, *EIF4A2* and *EIF4B*) which could be related to viral protein translation and ER stress response [34]. In order to confirm that the permissivity signature are not just reflecting tissue specific or immune signatures, a Pathway Enrichment Analysis (PEA) was performed using the top ranked genes. Interestingly, we found an enrichment of host-viral interactions processes, protein stabilization and Endoplasmic Reticulum (ER) trafficking pathways (**Fig 2C**).

Next, we investigated if these susceptibility genes, identified from lung-derived cell lines, are also expressed in other cell types. Therefore, we ranked the cell types in the Human Cell Landscape dataset (HCL) [28] based on the permissivity score derived from our top ranked genes in the permissivity signature (**Fig 2D**) and found that stratified epithelial, basal, AT2 lung cells and enterocytes were among the top-ranked cell types which correspond to cell types known to be infected by the virus [35, 36].

In order to further refine our permissivity signature by going beyond transcriptional levels, we used protein expression levels of a previously released proteomic dataset from SARS-CoV-2 infected Caco-2 cells [16]. We determined the Spearman correlations of the translation rates for the top ranked genes to that of the N and S viral proteins (**Fig 2E**). Some of the highly correlated genes (both negative and positive) have been previously reported to participate in viral infection processes. For example, *LGALS3BP* is a glycoprotein secreted molecule with antiviral properties observed in HIV and Hantavirus infection [37, 38] and in the regulation of LPS induced endotoxin shock in murine models [39]. *CLIC1* has been previously identified as a virulence factor of Merkel Cell Polyomavirus (MCPyV) which is upregulated during infection and promotes the development of Merkel Cell Carcinoma [40]. *HSPD1* has been shown to promote viral infection of HIV, HBV and Influenza viruses [41]. Interestingly, *SPINT2* was consistently correlated to viral translation (**Fig 2E**). Furthermore, the correlation of *SPINT2* with viral gene expression is negative and this trend is consistent in both Caco-2 and Calu-3 cell lines, indicating a repressive role on SARS-CoV-2 infection (**S1C and S1D Fig**). Hence, these findings suggest that *SPINT2* represents a permissivity factor that negatively correlates with SARS-CoV-2 infection.

SPINT2 knockdown increases SARS-CoV-2 infection in a *TMPRSS2*-dependent manner in cell lines

To experimentally validate the negative correlation of *SPINT2* expression with SARS-CoV-2 viral gene expression, we hypothesized that this gene could have a direct influence on SARS-CoV-2 infection by impairing early steps of viral entry. Hence, to test our hypothesis, we knocked-down *SPINT2* using small-hairpins RNA in the human lung carcinoma derived line Calu-3 cells. *SPINT2* expression was readily detectable in wild-type (WT) Calu-3 cells **(Fig 3A and 3B).** When Calu-3 cells were transduced with a specific shRNA directed against *SPINT2*, *SPINT2* levels were significantly decreased compared to WT cells or cells transduced with a scrambled shRNA at both the transcript (Fig 3A) and protein level (Fig 3B). To address the impact of *SPINT2* knocked-down on the permissivity of Calu-3 cells to SARS-CoV-2, WT, scrambled and *SPINT2* knocked-down cells were infected with SARS-CoV-2 using the same multiplicity of infection (MOI) and, at 24 hours post-infection (hpi), the impact of *SPINT2* silencing on SARS-CoV-2 infection was addressed by immunofluorescence using an antibody directed against the nucleocapsid protein of SARS-CoV-2. Knock-down of *SPINT2* resulted in more than two-fold increase in the number of cells positive for SARS-CoV-2 at 24 hpi (Fig 3E and 3F). Concommittently, knockdown of *SPINT2* was also associated with an increase in SARS-CoV-2 replication as monitored by reverse transcription quantitative PCR against the viral genome at 24 hpi (Fig 3C). Evaluation of the number of infectious virus particles released by cells revealed no significant increase in virus production and release upon knockdown of *SPINT2* (Fig 3D). Interestingly, even when using higher MOI, loss of *SPINT2* always resulted in an increased infection of Calu-3 cells as shown by the increase of the number of SARS-CoV-2 infected cells (S2A and S2B Fig) and viral genome replication (S2C Fig**).** Of note, at the highest MOI tested, we noticed a significant drop in the SARS-CoV-2 replication (S2C Fig) due to the fact that at high MOI of infection, most of the cells are already dead at 24 hpi (S2A Fig). To validate that the observed increase of SARS-CoV-2 infection upon *SPINT2* knockdown was not limited to Calu-3 cells, the human lung carcinoma derived line A549 cells overexpressing the SARS-CoV-2 receptor *ACE2* were exploited. Similar to our findings with Calu-3 cells, although only a partial knock-down of *SPINT2* was achieved in A549 cells (S3A Fig), we observed an increase of the percentage of SARS-CoV-2 infected cells at 24 hpi (S3B and S3C Fig). This increase of the number of infected cells was associated with an significant increase of viral replication when comparing *SPINT2* KD to the scrambled control (S3D Fig).

In order to test the hypothesis whether *SPINT2* modulation of viral load is dependent on *TMPRSS2*, we monitored its fold change expression. Interestingly, *TMPRSS2* gene expression was found to be higher in *SPINT2* knocked-down cells when compared to WT or scramble cells both in Calu-3 cells and in A549 cells (**Figs** 3G and S4A). Finally, to control that the increase in SARS-CoV-2 infection observed upon *SPINT2* silencing was dependent on the *TMPRSS2*-mediated activation of the virus, we employed a pharmacological approach to inhibit the activity of the *TMPRSS2* protease. Treatment of Calu-3 cells with the *TMPRSS2* inhibitor (Camostat mesylate) resulted in an almost complete inhibition of viral infection as monitored by quantifying the number of SARS-CoV-2 infected cells at 24 hpi (Fig 3H) and by quantifying replication using quantitative RT-PCR against the SARS-CoV-2 genome (S4B Fig). As described above, knockdown of *SPINT2* resulted in an increase of SARS-CoV-2 expression compared to cells treated with scrambled control shRNA and this increase was abrogated upon treatment of cells with the inhibitor of *TMPRSS2* (**Figs** 3H **and** S4B).

SPINT2 overexpression decreases SARS-CoV-2 infection

As knockdown of *SPINT2* resulted in an increase of SARS-CoV-2 infection, we hypothesized that a greater amount of *SPINT2* in cells will result in an inhibition of *TMPRSS2* which will in

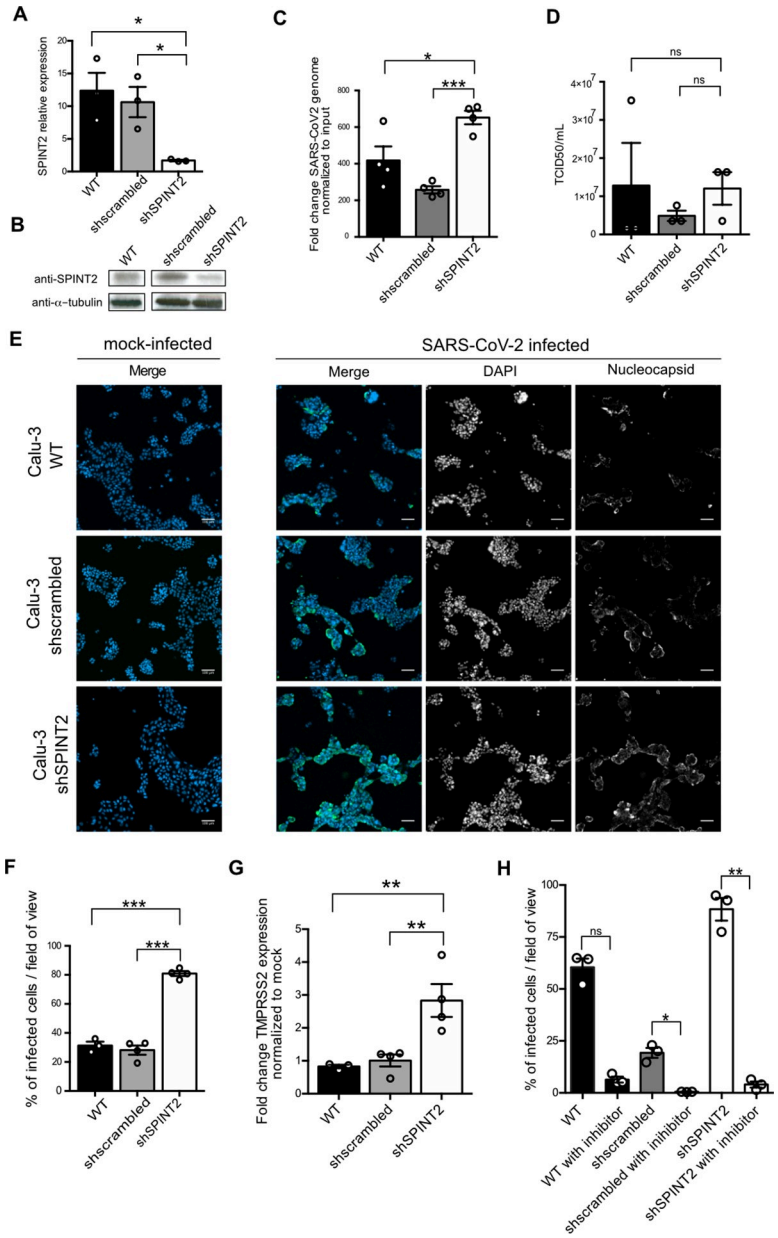

Fig 3. *SPINT2* knockdown in Calu-3 cells promotes SARS-CoV-2 infection. A. Relative expression of *SPINT2* normalized to the housekeeping gene TBP in wild type Calu-3, scrambled-treated and *SPINT2* KD Calu-3 cells. **B.** Western blot of *SPINT2* and α-tubulin as control in wild type, sh-scrambled and *SPINT2* KD Calu-3 cells. **C.** Quantification of SARS-CoV-2 replication. RNA was harvested at 24hpi, and q-RT-PCR was used to evaluate the copy number of the SARS-CoV-2 genome. Data are normalized to inoculum used for infection. **D**. Quantification of *de-novo* virus particles released by WT, sh-scrambled and *SPINT2* KD cells. **E.** Representative immunofluorescence images of Calu-3 wild type (WT) cells and cells treated with scrambled control shRNA (sh-scrambled) and shRNA targeting *SPINT2* (shSPINT2) for mock-infected and SARS-CoV-2 infected conditions. Infection was detected by indirect immunofluorescence against the nucleocapsid N and Nuclei were stained with DAPI**. F**. Quantification of the percentage of SARS-CoV-2 infected cells from C. **G**. Quantification of *TMPRSS2* expression upon *SPINT2* silencing. q-RT-PCR was used to evaluate the expression level of *TMPRSS2* in WT, sh-scrambled and *SPINT2 KD* Calu-3 cells. Relative expression level of *TMPRSS2* in SARS-CoV-2 infected cells normalized to that of the mock-infected cells. **H.** Quantification of the percentage of SARS-CoV-2 infected cells for WT, sh-scrambled and *SPINT2 KD* Calu-3 cells mock-treated or treated with the TMPRSS2 inhibitor camostat mesylate. Infection was detected by indirect immunofluorescence using an anti-nucleocapsid antibody at 24hpi. Data are normalized to inoculum used for infection. Error bars indicate standard deviation. n = 3 biological replicates. $P<0.05$ *, $P<0.01$ **, $P<0.001$ ***, $P<0.0001$ ****. Analysis was done by a two-tailed unpaired t-test with Welch's correlation for the respective cell lines.

https://doi.org/10.1371/journal.ppat.1009687.g003

turn prevent infection. In order to test this hypothesis we monitored *TMPRSS2* gene expression in Calu-3 cells overexpressing *SPINT2*. First, *SPINT2* was overexpressed in Calu-3 cells (**Fig 4A**) and both WT and cells overexpressing *SPINT2* were infected by SARS-CoV-2. At 24 hpi, cells were immunostained for the nucleocapsid N and quantification of the number and percentage of infected cells revealed that overexpression of *SPINT2* negatively impacted SARS-CoV-2 infection (**Fig 4B and 4C**). We observed a decrease in viral replication by measuring SARS-CoV-2 genome levels using RT-PCR (**Fig 4D**). We also monitored gene expression of *TMPRSS2* under *SPINT2* overexpression and observed a downregulation of *TMPRSS2* (**Fig 4E**) consistent with the *SPINT2* knock-down results. Together, these data strongly suggest that SARS-CoV-2 infection negatively correlates with *SPINT2* expression levels which is in full agreement with our observations using previously reported data in cell lines (**Figs 2E and S1C and S1D**).

SPINT2 modulates infection by controlling viral entry

To address which step of the SARS-CoV-2 lifecycle is favored upon *SPINT2* knocked-down, we quantified the number of SARS-CoV-2 infected Calu-3 cells overtime (**Fig 5A**). SARS-CoV-2 infected cells (immunostained against the nucleocapsid N) were readily detectable as early as 4 hpi (**Fig 5A and 5B**) which was consistent with previous reports. Interestingly, at any time points post-infection, cells knocked-down for *SPINT2* were always found more infected compared to cells treated with a control scrambled shRNA as monitored by quantifying the number of infected cells (**Fig 5B**), viral replication (**Fig 5C**) or production of *de-novo* infectious virus particles (**Fig 5D**). Together, these results suggest that knock-down of *SPINT2* favors early steps of the virus lifecycle, most probably the entry step given the function of *SPINT2* in controlling the *TMPRSS2* activity. To directly address whether loss of *SPINT2* promotes SARS-CoV-2 entry in cells, we thought of uncoupling the SARS-CoV-2 entry step to the rest of the viral lifecycle. For this, we exploited the Vesicular Stomatitis Virus pseudotyped with the SARS-CoV-2 spike in A549 [42]. Entry of this engineered virus relies on the SARS-CoV-2 spike protein while the rest of its lifecycle corresponds to the VSV replication/assembly cycle. Infection of A549 cells overexpressing *ACE2* with the spike pseudotyped VSV encoding the green fluorescent protein (GFP) was readily detectable as early as 4 hpi. Interestingly, upon *SPINT2* knocked-down, we observed an increase in the number and percentage of spike-pseudotyped VSV infected cells compared to cells treated with an scrambled control shRNA (**Fig 5E and 5F**). All together, these data strongly suggest that loss of *SPINT2* leads to greater infection by SARS-CoV-2 by promoting entry into the target cells.

SPINT2 is negatively correlated to viral load and is down-regulated in severe COVID-19 cases

Given the observed negative correlation between *SPINT2* expression and SARS-CoV-2 infection in cell lines (**Figs 2E and S1C and S1D**) we next investigated if *SPINT2* expression is associated with disease severity in COVID-19 patients. We used a publicly available scRNA-seq dataset on nasopharynx swabs samples from patients with severe and mild symptoms [43]. We correlated a list of serine proteases and inhibitors (SPRGs, **S3 Table**) to the viral RNA reads and found that *SPINT2* was the second most negatively correlated gene (**Fig 6A**). Then, we selected the cell cluster with the highest expression of *SPINT2*, which correspond to secretory cells (**S5A Fig**) and among these cells, observed a lower *SPINT2* gene expression in cells from critical COVID19 cases compared to moderate cases (**Fig 6B**). This finding is particularly relevant since secretory cells are primary targets of viral infection [44]. We also evaluated data on Peripheral Blood Mononuclear Cells (PBMC) from severe COVID-19 patients [45]. In this

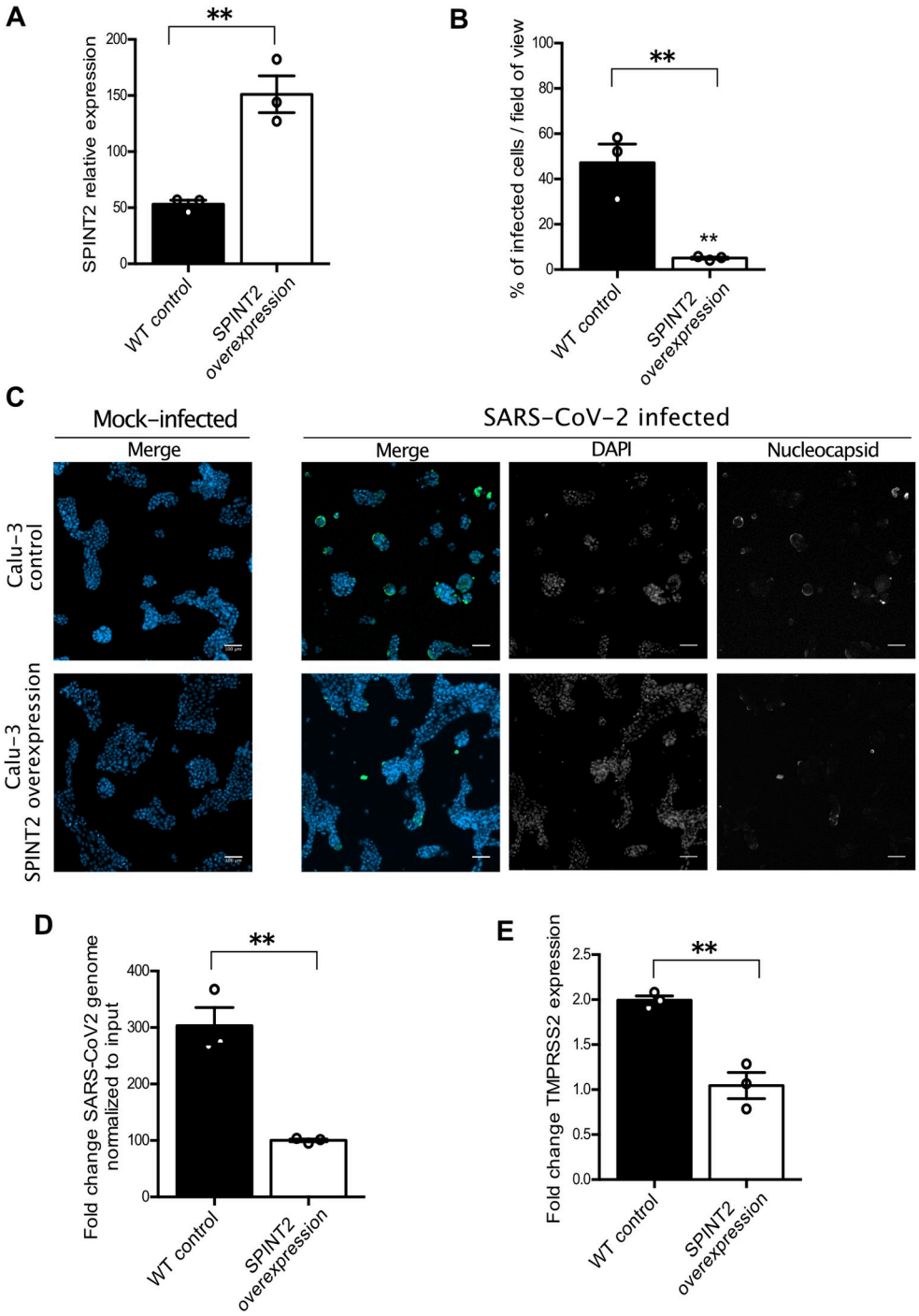

Fig 4. *SPINT2* overexpression in Calu-3 cell lines negatively impacts SARS-CoV-2 infection. A. Relative expression of *SPINT2* normalized to the housekeeping gene TBP in empty vector control and *SPINT2* overexpressing Calu-3 cells. **B.** Quantification of the percentage of infected cells in Calu-3 cell lines transfected with an empty vector or a vector overexpressing *SPINT2* for mock-infected and SARS-CoV-2 infected conditions. **C.** Representative immunofluorescence images of the infected cells quantified in B. Infection was detected by indirect immunofluorescence against the nucleocapsid N (green) and Nuclei were stained with DAP (blue). **D.** Quantification of SARS-CoV-2 replication. RNA was harvested at 24 hpi, and q-RT-PCR was used to evaluate the copy number of the SARS-CoV-2 genome. Data are normalized to inoculum used for infection. **E.** Relative expression of *TMPRSS2* normalized to that of the respective mock-treated cells in empty vector control and *SPINT2* overexpression Calu-3 cells. Error bars indicate standard deviation. n = 3 biological replicates. $P<0.05$ *, $P<0.01$ **, $P<0.001$ ***, $P<0.0001$ ****. Analysis was done by a two-tailed unpaired t-test with Welch's correlation.

https://doi.org/10.1371/journal.ppat.1009687.g004

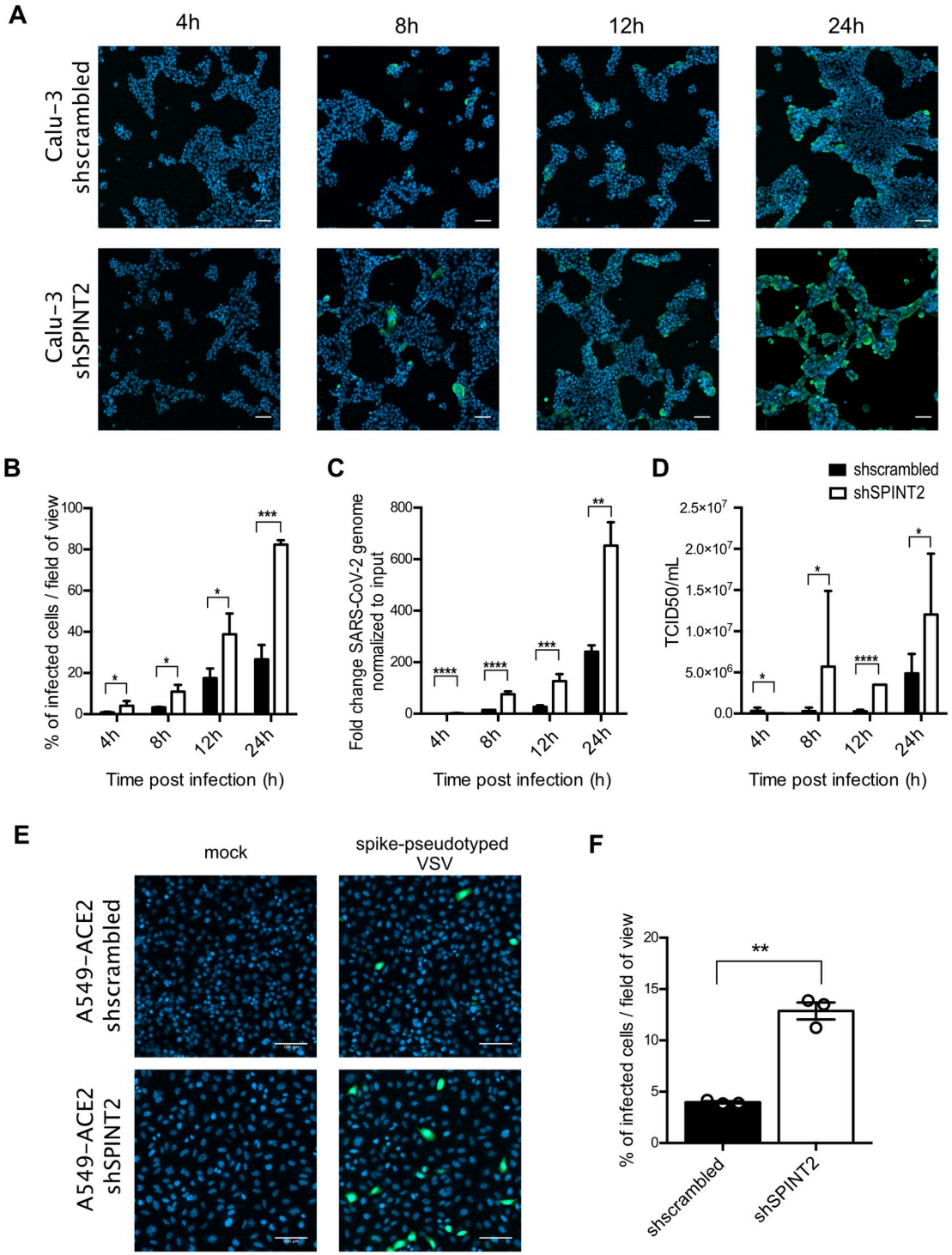

Fig 5. *SPINT2* silencing impacts an early step of the SARS-CoV-2 life cycle. A. Representative immunofluorescence images of Calu-3 cells treated with scrambled control shRNA (sh-scrambled) and shRNA targeting *SPINT2* (shSPINT2) for mock-infected and SARS-CoV-2 infected conditions at indicated time point post-infection. **B.** Quantification of the percentage of SARS-CoV-2 infected cells from A. **C.** Quantification of SARS-CoV-2 replication. RNA was harvested at the indicated time points, and q-RT-PCR was used to evaluate the copy number of the SARS-CoV-2 genome. Data are normalized to inoculum used for infection. **D.** Quantification of *de-novo* virus particles released by sh-scrambled and *SPINT2* KD cells at the indicated time points post-infection. **E.** Representative fluorescence images of sh-scrambled and *SPINT2* KD A549 cells expressing *ACE2* for mock-infected and VSV-GFP pseudotyped with SARS-CoV-2 spike infected cells. **F.** Infection was detected by fluorescence imaging of GFP and

the percentage of infected cells was quantified. Nuclei were stained with DAPI (blue). Error bars indicate standard deviation. n = 3 biological replicates. P<0.05 *, P<0.01 **, P<0.001 ***, P<0.0001 ****. Analysis was done by a two-tailed unpaired t-test with Welch's correlation for the respective time point for B-D. Two-tailed unpaired t-test with Welch's correlation was performed for E.

https://doi.org/10.1371/journal.ppat.1009687.g005

dataset, *SPINT2* was found to be strongly expressed in Dendritic Cells (DC), plasmacytoid DC (pDC), and stem cells (SC) and eosinophils (**S5B Fig**). Among these cells, again, we observed lower *SPINT2* expression in patients from Intensive Care Units (ICU) (**Fig 6C**). Additionally, we could also corroborate the negative correlation of *SPINT2* and viral load using bulk RNA-seq data from lung autopsies of COVID-19 deceased patients [46]. We calculated the correlations of gene expression between *SPINT2*, *ACE2* and *TMPRSS2* to E, M, N and S viral genes and observed the similar negative correlation (**S5C Fig**). Collectively, this evidence suggests that *SPINT2* expression level could be associated to COVID-19 disease severity.

SPINT2 is down-regulated in multiple tumor types and pancreatic cells from T2D patients

COVID-19 patients with previous records of chronic diseases like cancer or diabetes are considered at higher risk [47–51]. Also, *SPINT2* gene silencing by promoter hypermethylation has

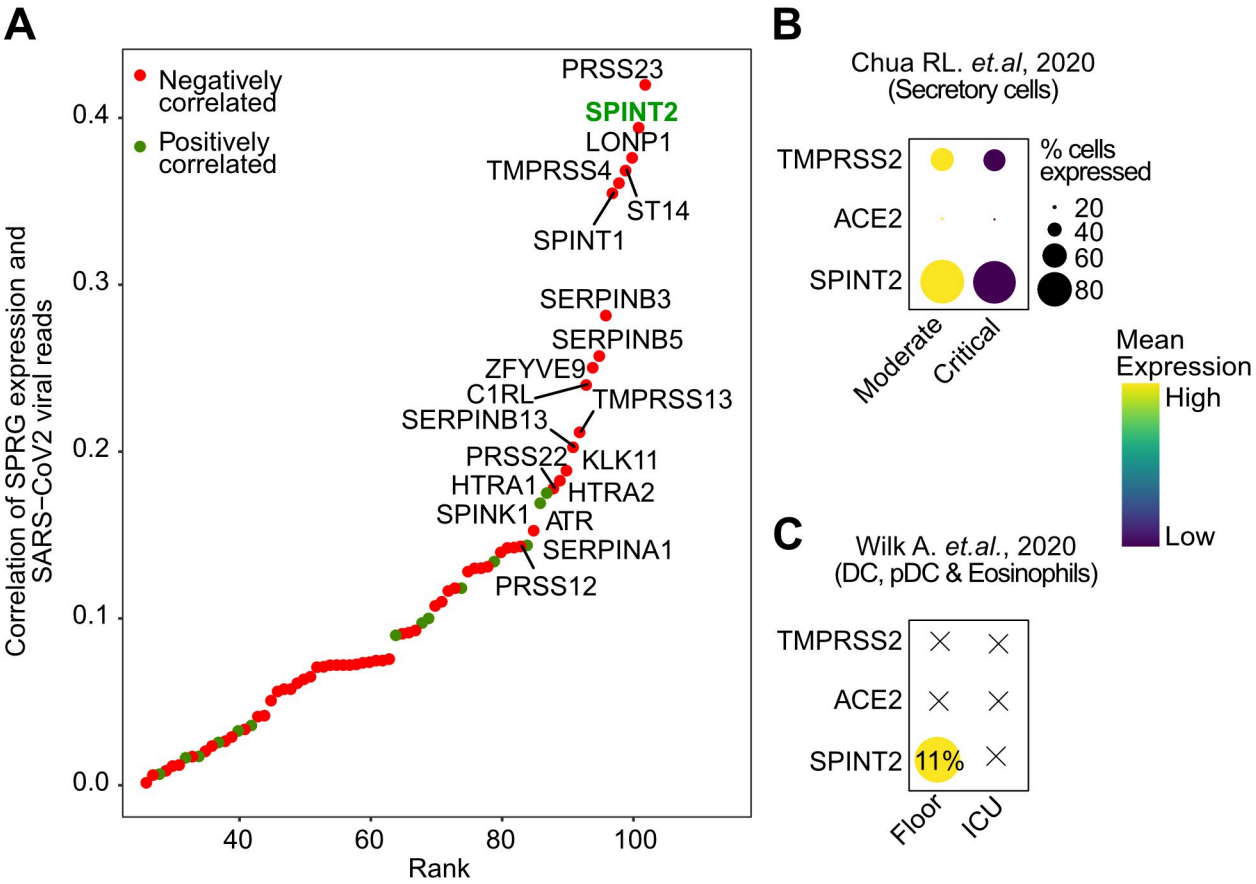

Fig 6. *SPINT2* expression is associated with disease severity. A. Spearman correlation values of SPRGs to SARS-CoV-2 viral reads in scRNA-seq of nasopharynx samples from COVID-19 patients. Top 20 most correlated SPRGs are labeled and *SPINT2* is highlighted in green. **B**. *SPINT2*, *ACE2* and *TMPRSS2* gene expression in severe and mild cases COVID-19 patients using datasets from Chua RL et al, 2020 **C.** Same as in (*B*) but using the Wilk A et al, 2020 data on Peripheral Blood Mononuclear Cells derived from COVID-19 patients

https://doi.org/10.1371/journal.ppat.1009687.g006

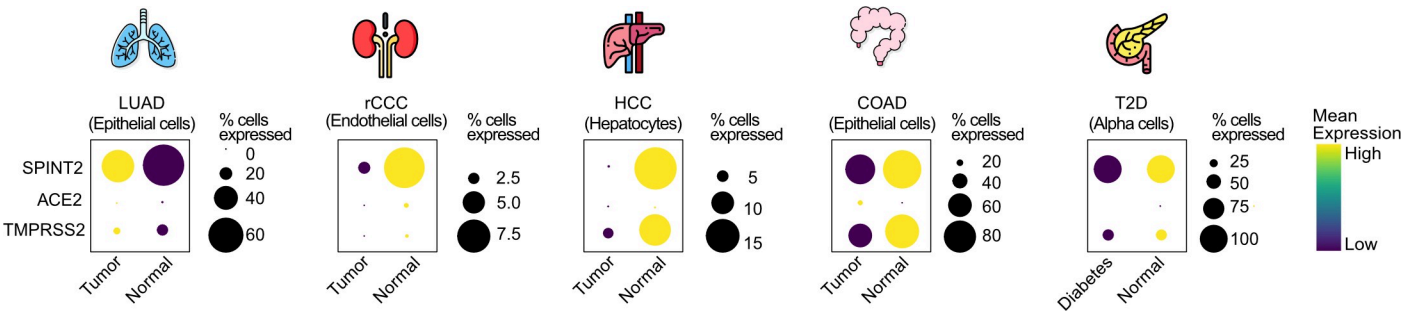

Fig 7. *SPINT2* is down-regulated in tumors and diabetic pancreatic cells. Dot plots showing gene expression and percentage of cells expressing *SPINT2*, *ACE2* and *TMPRSS2* in tumors and pancreatic alpha cells. The selected cell type for each tumor type is indicated in each case. Lung Adenocarcinoma (LUAD), renal Clear Cell Carcinoma (rCCC), Hepatocyte Cell Carcinoma (HCC), Colon Adenocarcinoma (COAD), Type 2 diabetes (T2D).

https://doi.org/10.1371/journal.ppat.1009687.g007

been reported in multiple tumor types which promotes tumor progression [17, 52–54]. For this reason, we hypothesized that *SPINT2* down-regulation in tumor cells would increase viral infection permissivity which among others, could be one of the mechanisms behind the comorbidity observed in COVID-19 patients. We screened lung, colon, liver and hepatic tumor datasets to evaluate the differences in *SPINT2* gene expression between tumor and paired normal samples. We found statistically significant down-regulation of *SPINT2* in the kidneys and liver tumors (**S6A Fig**). Similarly, using comparable tumor scRNA-seq datasets [55–59] we observed a down-regulation of *SPINT2* in colon adenocarcinoma (epithelial cells), renal clear cell carcinoma (endothelial cells) and hepatocellular carcinoma (hepatocytes) (**Fig 7**). Interestingly, we were able to detect *SPINT2* down-regulation in colorectal tumor epithelial cells at single cell level but not in bulk RNA-seq data suggesting that *SPINT2* expression might be modulated in specific cell subtypes (**Figs 7** and **S6A**). In lung adenocarcinomas, we found *SPINT2* upregulation in tumors both in TCGA and scRNA-seq data which might reflect the existence of different determinants for comorbidity in lung tissues independent of *SPINT2* modulation. We also looked at the expression of *SPINT2* in pancreatic cells from diabetes type 2 (DT2) patients [59]. Islet cells have high *SPINT2* expression when compared to other cell types like endothelial cells (**S6B Fig**). We observed a strong down-regulation of *SPINT2* in alpha-cells of DT2 patients, which have been shown to be primary targets of the SARS-CoV-2 virus. This down-regulation of the virulence associated factor *SPINT2* might contribute to the comorbidity between COVID19 and DT2.

Discussion

In this study, we describe a tight protease-inhibitor/protease balance at the gene expression level between *SPINT2* and *TMPRSS2*, a major co-receptor of SARS-CoV-2. We found Transcription Factor Binding Sites (TFBS) for ten regulators including *IRF1*, *IRF3*, *JUNB*, *JUND* and *ELF3* whose TF activities were found to be correlated to both *SPINT2* and *TMPRSS2* gene expression which suggests their possible role as common regulators of both genes. Interestingly, *ELF3* and *IRF7* TF activity has been found to be modulated in SARS-CoV-2 infected vs bystander enterocytes from ileum [21], which could point to viral load modulation mediated by *TMPRSS2* and *SPINT2* through these TFs. We show that *SPINT2* and *TMPRSS2* gene expression levels are correlated across cell types and tissues. Interestingly, known SARS-CoV-2 target tissues have high correlation values and co-expression for both genes which suggest that *SPINT2* could play a role in SARS-CoV-2 viral entry.

Currently, it is unclear what the molecular signatures are that determine viral permissivity and how they are related to disease severity. We inferred a SARS-CoV-2 permissivity signature, using differentially expressed genes between permissive and non-permissive cell lines from which we removed viral induced genes. We were able to find *SPINT2* in this permissivity signature and observed a negative correlation to SARS-CoV-2 viral load in Calu-3 cells. We also corroborated this trend at the protein level in Caco-2 cells. During the preparation of these manuscript, a study from Bojkova D et al, 2020 was published suggesting a possible role of *SPINT1*, *SPINT2* and *SERPINA1* in viral infection by observing the down-regulation of their protein levels in infected cells and also by evaluating the effect of Aprotinin a non-specific SP inhibitor on viral load [9]. In contrast, we could not observe SPINT2 as a viral-induced gene. Such discrepancy might be attributed to a difference in the level of information used between protein and mRNA levels.

However, here for the first time by knocking down and overexpressing *SPINT2*, we provide direct causal evidence that *SPINT2* is indeed able to modulate SARS-CoV-2 infection. This modulation is dependent on *TMPRSS2* and we provide evidence that *SPINT2* knock-down impacts SARS-CoV-2 infection by affecting viral entry. *SPINT2* inhibits *TMPRSS2* enzymatic activity through its KD1 and KD2 domains [15]. Interestingly, beyond direct *TMPRSS2* enzymatic activity inhibition we could observe an up-regulation of *TMPRSS2* mRNA expression in the *SPINT2* knocked-down Calu-3 cells. Further on, we also observed the same trend in the overexpression experiments, that is, *TMPRSS2* was down regulated after *SPINT2* overexpression. Further investigation is needed to explore this regulation at the gene expression level. It has been previously reported that *SPINT2* can regulate gene expression through different mechanisms apart from direct inhibition. For example, *SPINT2* can modulate the serine protease *ST14* protein activity by regulating its shedding from the cell membrane of mouse intestinal epithelial cells [19] *SPINT2* has been reported to regulate transcription of certain genes like *CDK1A* via histone methylation [60]. Then, we speculate that *SPINT2* levels could also regulate *TMPRSS2* transcriptionaly, independently of direct enzymatic activity inhibition. Our findings clearly show that *SPINT2* regulates SARS-CoV-2 viral infection through the inhibition of *TMPRSS2*, since a drug induced inhibition of TMPRSS2 abrogates the infection-promoting effect of *SPINT2*.

We found a lower expression of *SPINT2* in secretory cells from COVID-19 patients with severe symptoms [43]. This could have implications for COVID-19 disease severity since secretory cells have been shown to be the target of SARS-CoV viral infection using organotypic human airway epithelial cultures [44]. We found *SPINT2* in the permissivity signature from which we filtered out viral induced genes, suggesting that this gene could be used as a marker for predicting COVID-19 disease susceptibility *prior* to infection, however this needs to be further evaluated.

Serine proteases (SPs) have been reported to be abnormally regulated in diverse chronic diseases [17, 61–63]. For example, during carcinogenic development SPs influence metastasis and cancer progression [64, 65], while in the context of diabetes they control fibrinolysis, coagulation and inflammation which in turn affects disease severity [62]. This led us to hypothesize that shared molecular mechanisms between some chronic diseases and COVID-19 could be explained in part by the regulation of *SPINT2*. We observed *SPINT2* down-regulation in Hepatocellular Carcinoma (HCC), Colon Adenocarcinoma (COAD) and renal Clear Cell Carcinoma (rCCC) tumor cells. *SPINT2* down-regulation in liver has been reported to contribute to the development of HCC by the binding and inhibition of the serine protease *HGFA* which transforms Hepatocyte Growth Factor (*HGF*) into its active form which in turn promotes metastasis, cell growth and angiogenesis [17, 66] and the same mechanism has been suggested for rCCC [12]. A marked down-regulation of *SPINT2* can be observed in alpha islets

pancreatic cells from diabetes patients. It has been reported that islet cells can be infected by SARS-CoV-2 which could contribute to the onset of acute diabetes [67]. Hence, these results suggests that kidney, colon and liver tumor types as well as pancreatic islets cells from diabetic patients could be more permissive and susceptible to SARS-CoV-2 viral infection due to an imbalance of *SPINT2* gene expression, which could lead to the disruption of the protease-inhibitor/protease balance [68].

In conclusion, we showed for the first time that *SPINT2* is a permissivity factor that modulates SARS-CoV-2 infection. This modulation could be explained by the balance of *TMPRSS2*/*SPINT2* (serine protease/inhibitor) that we observed at the gene expression level across several tissues. We also found lower *SPINT2* gene expression in samples from COVID-19 patients with severe symptoms, hence, this gene might represent a biomarker for predicting disease severity. We also found *SPINT2* down-regulation in tumor types which could have implications for the observed comorbidities in COVID-19 patients with cancer.

Methods

Cell line and viruses

Human lung adenocarcinoma cell lines Calu-3 (ATCC HTB-55), human lung carcinoma cell lines A549 (ATCC CCL-185) stably expressing ACE2 (kindly provided by Dr. Ralf Bartenschlager) and Vero E6 cells (ATCC CRL 1586) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with Glutamax (Gibco), 10% fetal bovine serum and 1% penicillin/streptomycin. Calu-3 cell lines stably expressing the *SPINT2* knockdown and *SPINT2* overexpression were generated by lentiviral transduction.

Production of lentiviral constructs expressing shRNA against *SPINT2* and lentiviral constructs overexpressing *SPINT2*

Oligonucleotides encoding the sequence for *SPINT2* knockdown were designed from the TRC library based on Genetic Perturbation Platform (GPP) Web Portal, cloneID: TRCN0000073581 (Box 1) [69] Annealed oligonucleotides were ligated with the AgeI-HF and EcoRI-HF digested pLKO.1 puro vector (Add gene #8453) using the T4 DNA Ligase (New England Biolabs). The pENTR223 vector encoding the ORF of human SPINT2 (Genbank ID: CV023579) was obtained from the DKFZ Genomics & Proteomics Core Facility and then subsequently cloned by recombinational Gateway Cloning into an expression vector pWPI. Both resulting plasmids were transformed into E. coli DH5α-competent cells.

Amplified plasmid DNA was purified using the NucleoBondR PC 100 kit by Marchery-Nagel following the manufacturer's instructions.

Box 1: Oligonucleotides for shRNA expression. Bold characters mark the respective target sequence

Name	Sequence
shSPINT2 fw	CCGG**CAGCTGGTGAAGAACACATAT**CTCGAGATATGTGTTCTTCACCAGCTGTTTTTG
shSPINT2 rev	AATTCAAAAA**CAGCTGGTGAAGAACACATAT**CTCGAGATATGTGTTCTTCACCAGCTG
shScrambled fw	CAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTT
shScrambled rev	AAAACAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTG

https://doi.org/10.1371/journal.ppat.1009687.t001

Lentivirus production and selection of stable cell lines

HEK293T cells (ATCC CRL-3216) were seeded on 10 cm^2 dishes and allowed to adhere for 36 hours. The cells were transfected with 4 μg of pMD2.G (Addgene #12259), 4 μg of psPAX2 (Addgene #12260) and 8 μg of purified pLKO.1 plasmid containing the shRNA constructs upon reaching 70% confluency. Cell supernatant containing lentivirus was harvested 72 h post-transfection, filtered through a 45 μM Millex HA-filter (Merck Millipore) and purified by ultracentrifugation at 27,000x g for 90min. $2x10^5$ Calu-3 cells were seeded onto collagen coated 6-well plates 24 h prior to transduction. Cell medium was replaced with 3 mL medium containing 20 μL of the purified lentivirus and 3μl polybrene transfection reagent (Merck Millipore). Medium was supplemented with 10 μg/mL puromycin for selection of successfully transduced cells two to three days after transduction.

SARS-CoV-2 viral infection

The SARS-CoV-2 isolate used in the experiments was obtained from the swab of a SARS-CoV-2 positive patient from the Heidelberg University Hospital. The virus was isolated and propagated in Vero E6 cells. All SARS-CoV-2 infections were performed with a multiplicity of infection of 0.04 as determined in Vero E6 cells. Prior to infection, culture media was removed and virus was added to cells and incubated for 1 hour at 37˚C. Fresh media was added back to the cells upon virus removal.

SARS-CoV-2 spike protein pseudotyped VSV assay

A549 cells expressing ACE2 were seeded at 10000 cells/ well in a 96-well plate 24h prior to infection. Spike pseudotyped VSV was added to the wells and the infection was allowed to proceed 8 hpi, media was removed, samples were washed 1X with PBS and fixed in 4% paraformaldehyde (PFA) for 20 mins at room temperature (RT). Cells were washed in 1X PBS, permeabilized in 0.5% Triton-X for 15 mins at RT and then incubated with DAPI for 30 mins at RT. Cells were washed in 1X PBS three times and maintained in PBS. Cells were imaged on a Zeiss Cell Discoverer 7 to quantify the number of infected cells relative to the number of nuclei. For details about the VSV SARS-CoV-2 S Δ18 eGFP, refer to [42].

RNA isolation, cDNA synthesis and qPCR

Cells were harvested 24 hours post infection for RNA isolation using RNAeasy RNA extraction kit (Qiagen) as per manufacturer's instructions. Complementary DNA was synthesized using iSCRIPT reverse transcriptase (BioRad) from 250 ng of total RNA per 20μL reaction according to the manufacturer's instructions. Quantitative RT-PCR assay was performed using iTaq SYBR green (BioRad) as per manufacturer's instructions. The expression of target genes was normalized to endogenous control *TBP*. Primer sequences are indicated in Box 2.

Box 2: Primer sequences for qPCR

Target gene	Species name	Forward sequence	Reverse sequence
TBP	Human	ccactcacagactctcacaac	ctgcggtacaatcccagaact
COV21	SARS-CoV-2	gcctcttctgttcctcatcac	agacagcatcaccgccattg
SPINT2	Human	attgcctcttttctcatcac	gactcccataaagaaagcac
TMPRSS2	Human	acctgatcacaccagccatg	cttcgaagtgaccagaggcc

https://doi.org/10.1371/journal.ppat.1009687.t002

The fold change in SARS-CoV-2 genome copy number was calculated using input as a reference. Input samples were harvested directly post-infection and accounted for the basal viral genome copy number detected due to viruses attaching to the cell membrane.

In-cell Western (TCID50)

Vero E6 cells were seeded at 20,000 cells/ well into a 96-well plate 24h prior to infection. 100uL of supernatant was added to the first wells and seven 1:10 serial dilutions were made. The cells were incubated for 24h and then fixed with 2% PFA for 20mins at RT. Cells were washed twice with 1X PBS upon PFA removal and then permeabilized for 15mins with 0.5% Triton-X in PBS. Blocking was carried out with 1:2 dilution of Li-Cor blocking buffer (Li-Cor) in PBS for 30mins at RT. Cells were then incubated with primary antibody against dsRNA, J2 (Scicons: 10010500, 1:1000) for 1h at RT. Cells were washed three times with PBS containing 0.1% Tween 20. Cells were then incubated with secondary antibody (anti-mouse CW800) and DNA dye Draq5 (Abcam) diluted 1:10.000 in blocking buffer for 1h at RT. Cells were again washed three times with PBS containing 0.1% Tween 20. The plate was then imaged on a LICOR (Li-Cor) imager.

Indirect immunofluorescence assay

Cells were seeded on a 48-well plate at 50.000 cells/well. Cells were fixed in 4% paraformaldehyde (PFA) for 20 mins at RT 24 hours post infection. Cells were washed in 1X PBS and permeabilized in 0.5% Triton-X for 15 mins at RT. 30 minutes of blocking were carried out using 3% BSA-PBS at RT. Mouse monoclonal antibody against SARS-CoV-2 Nucleocapsid (NC) protein (Sino biologicals MM05) as primary antibody was diluted in 1% BSA- phosphate-buffered saline (PBS) and incubated for 1h at RT. Cells were washed with 1X PBS three times and incubated with secondary antibodies conjugated with AF488 (Molecular Probes) and DAPI for 30–45 mins at RT. Cells were washed in 1X PBS three times and maintained in PBS. Cells were imaged on a Zeiss Cell Discoverer 7microscope to quantify the number of infected cells relative to the number of nuclei.

Western blot

Cells were rinsed once with 1X PBS and lysed with 1X RIPA (150 mM sodium chloride, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), 50 mM Tris, pH 8.0 with phosphatase and protease inhibitors (Sigma-Aldrich)) for 5 mins at room temperature (RT). Lysates were collected and equal protein amounts were separated by SDS-PAGE and blotted onto a nitrocellulose membrane by wet-blotting (Bio-Rad). Membranes were blocked with 5% BSA in TBS containing 0.1% Tween 20 (TBS-T) for two hours at RT. Primary antibodies against *SPINT2* (Sigma Aldrich HPA011101, 1:500) and α-tubulin (Sigma Aldrich T9026, 1:1000) were diluted in blocking buffer and incubated overnight at 4˚C. Membranes were washed 3X in TBS-T for 15 mins at RT. Secondary antibodies were diluted in blocking buffer and incubated at RT for 1 hour with rocking. Membranes were washed 3X in TBS-T for 15 mins at RT. HRP detection reagent (GE Healthcare) was mixed 1:1 and incubated at RT for 5 mins. Membranes were exposed to film and developed.

TMPRSS2 inhibition assay

Calu-3 cells were seeded onto a 48-well plate 24 hours prior to treatment. Cells were incubated with 5 μM of Camostat mesylate (Sigma Aldrich, SML0057) for 30 mins prior to virus infection and throughout the 1 hour virus infection. After infection, fresh media containing

Camostat mesylate was added to the cells and incubated for another 24h. Cells were then harvested 24 hpi.

Statistics and computational analyses and statistics. In order to quantify infected cells from indirect immunofluorescent stained samples, ilastik 1.2.0 was used on DAPI images to generate a mask representing each nucleus as an individual object. These masks were used on CellProfiler 3.1.9 to measure the intensity of the conjugated secondary antibodies in each nucleus. A threshold was set based on the basal fluorescence of non infected samples, and all nuclei with a higher fluorescence were considered infected cells.

Calu-3 and H1299 cells preprocessing

For Calu-3 cells we filter out cells with an extremely high number of detected genes (>50,000) which probably corresponds to doublets. In H1299, since few cells were detected to be infected, because this line is non-permissive, in order to obtain DEGs we defined infected cells as those with cumulative sum of viral genes expression >0.

*Assessing non-viral induced permissivity signatures*As we wanted to differentiate between permissivity and infection signatures, we first looked for differentially expressed genes in SARS-CoV-2 permissive *vs* non-permissive cell lines and then we removed all the genes which were up- or down-regulated during infection (**Fig 2A**). We performed Differential Expression Analysis using Seurat [70] (DEA) between Calu-3 and H1299 cells in non-infected mock cells at 4 hours of culture (*z*). Then, we obtained DEGs of Calu-3 infected *vs* mock at 12 hours post infection (*x*); we did the same with H1299 infected cells *vs* non-infected cells (H1299 infected cells were defined as explained above). In all DEA we set a Log Fold Change (FC) = 0.25 threshold. Finally, we removed these infection signatures from the DEGs of Calu-3 *vs* H1299 to obtain the permissivity gene signature (*i'*). A gene set enrichment analysis was performed to evaluate the composition of removed genes using enrichR.

Ranking genes using RF and pathway enrichment analysis

A Random Forest (RF) regression analysis was performed using the normalized gene expression of the permissivity signature to predict the cumulative sum of the expression of viral genes in Calu-3 cells at 12 hpi. We trained the RF using a random subsample of 75% and tested the results with the remaining set. Next, we estimated the feature importance for each of the permissivity signature genes and performed enrichment analysis on the top 25% ranked genes.

Scoring permissivity signatures

For the scoring of cells based on the permissivity signature among cell types in the HCL dataset, we used the top 25% RF ranked genes and applied the AddModuleScore function of Seurat setting *nbin* = 100.

SPINT2 expression correlation to viral gene expression

For the translatome correlation analysis, the summed intensity normalized values were used as provided in the study [16]. In order to compute the correlations of SPRGs (**S3 Table**) to the viral reads in the scRNA-seq data from Chua RL et al, 2020 the raw count matrices were extracted from the Seurat object provided by the authors, splitted by sample and then imputed using scimpute [71] with the following parameters: drop_thr = 0.5 and Kcluster equal to the number of annotated cell types in each matrix. The imputed matrices were then merged and log2 normalized. Finally, correlations were performed restricted to infected cells (viral read counts>0). In the bulk RNA-seq data from deceased COVID-19 patients log2 RPM of

normalized counts are used. In both cases correlation to viral genes were carried out using spearman coefficients.

scRNA-seq data preprocessing

In order to have a standardized workflow for the processing of scRNA-seq data we used SCT normalization using the Seurat workflow for every dataset except for Human Cell Landscape data where log2 normalization and scaling were performed since this dataset is large and using SCT was unpractical. HCC data were downloaded from GEO (GSE149614) and reprocessed. We used the Louvain method implemented in Seurat for community detection and clusters were identified by using tissue markers. We used the markers used to characterize cell types from an independent scRNA-seq human liver atlas [72] and using these markers identified clusters of epithelial, endothelial, hepatocytes, Kupffer and NK cells. For kidney, colon, prostate tumors, pancreatic cells from T2D, PBMC and Airways epithelium from SARS-CoV-2 patients's datasets and the annotations were used as provided in the corresponding publications (see *Data and script availability* section).

SPINT2 expression in TCGA tumors data

TPM normalized counts from tumor samples in TCGA were downloaded (https://www.cancer.gov/tcga). Our analysis was restricted to tissue and sample matched tumor and normal samples only. Difference in average expression was estimated using Wilcoxon Test with Holm correction.

Inference of transcription binding sites

A footprinting analysis was carried out using the TOBIAS pipeline [73] with a default parameters setting of MACS–nomodel–shift -100 –extsize 200 –broad. Then, we extracted the inferred Transcription Factor Binding Sites (TFBS) for those TF with activities found to be positively correlated to both *SPINT2* and *TMPRSS2* using the single cell RNA-seq data. TFBS were visualised using the PlotTracks TOBIAS function and the network was built in Cytoscape [74]. Edges in the network represent TF binding scores.

Supporting information

S1 Fig. SPINT2 is negatively correlated to SARS-CoV-2 viral expression. A. Upset plot showing the intersections of Calu-3 vs H1299 DEGs with Calu-3 and H1299 viral-induced genes signatures (see Methods). Removed genes from the permissivity signature are shown in green. **B**. Gene set enrichment of filtered out genes (green intersections in A). **C**. Correlation of *SPINT2* to the cumulative sum of normalized expression values of SARS-CoV-2 genes in Calu-3 cells. **D**. Correlation of *SPINT2* to S viral protein translation rates in Caco-2 cells. (TIFF)

S2 Fig. *SPINT2* knockdown in Calu-3 cells promotes SARS-CoV-2 infection. A. Representative immunofluorescence images of Calu-3 cells treated with scrambled control shRNA (sh-scrambled) and shRNA targeting *SPINT2* (shSPINT2) infected with SARS-CoV-2 at indicated MOI (as determined in Vero cells). Infection was detected by indirect immunofluorescence against the nucleocapsid N and Nuclei were stained with DAPI. MOI 0.16 and MOI 0.24 are shown. **B.** Quantification of the percentage of SARS-CoV-2 infected cells at the indicated MOI for cells treated with scrambled control shRNA (sh-scrambled) and shRNA targeting *SPINT2* (shSPINT2). **C.** Quantification of SARS-CoV-2 replication as in D. RNA was harvested at 24 hpi, and q-RT-PCR was used to evaluate the copy number of the SARS-CoV-2 genome. Data

are normalized to inoculum used for infection for each MOI. Error bars indicate standard deviation. n = 3 biological replicates. P<0.05 *, P<0.01 **, P<0.001 ***, P<0.0001 ****. Analysis was done by a two-tailed unpaired t-test with Welch's correlation for the respective MOI.
(TIFF)

S3 Fig. *SPINT2* knockdown in A549-ACE2 cells promotes SARS-CoV-2 infection. A. Relative expression of *SPINT2* normalized to the housekeeping gene TBP in wild type, sh-scrambled and *SPINT2* KD A549-ACE2 cells. **B.** Percentage of infected cells in A549-ACE2 wild type (WT) cells and cells treated with scrambled control shRNA (sh-scrambled) and shRNA targeting *SPINT2* (shSPINT2) for mock-infected and SARS-CoV-2 infected conditions. Infection was detected by indirect immunofluorescence against the nucleocapsid N and Nuclei were stained with DAPI. **C.** Representative immunofluorescence images of the cells shown in B. **D.** Quantification of SARS-CoV-2 replication. RNA was harvested at 24 hpi, and q-RT-PCR was used to evaluate the copy number of the SARS-CoV-2 genome. Data are normalized to inoculum used for infection. Error bars indicate standard deviation. n = 3 biological replicates. P<0.05 *, P<0.01 **, P<0.001 ***, P<0.0001 ****, Analysis was done by a two-tailed unpaired t-test with Welch's correlation using *SPINT2* KD A549 cells as reference.
(TIFF)

S4 Fig. *SPINT2* knockdown modulates viral infection via *TMPRSS2*. A. Quantification of *TMPRSS2* expression upon *SPINT2* silencing. q-RT-PCR was used to evaluate the expression level of *TMPRSS2* in WT, sh-scrambled and *SPINT2 KD* in A549 cells transduced with *ACE2*. Relative expression level of *TMPRSS2* in SARS-CoV-2 infected cells normalized to that of the mock infected cells. **B.** Quantification of SARS-CoV-2 replication upon *SPINT2* KD and *TMPRSS2* inhibition conditions in Calu-3 cells. RNA was harvested at 24 hpi and q-RT-PCR was used to evaluate the copy number of the SARS-CoV-2 genome. Data are normalized to inoculum used for infection. Error bars indicate standard deviation. n = 3 biological replicates. P<0.05 *, P<0.01 **, P<0.001 ***, P<0.0001 ****. Analysis was done by a two-tailed unpaired t-test with Welch's correlation for the respective cell lines.
(TIFF)

S5 Fig. *SPINT2* expression in patients's datasets. *ACE2*, *SPINT2* and *TMPRSS2* expression in Chua RL et al, 2020 (**A**) and Blish C et al, 2020 (**B**) scRNA-seq datasets. *SPINT2* expressing cell types shown in the main text are highlighted in red for both datasets. **C.** Correlation of *ACE2*, *SPINT2* and *TMPRSS2* to viral proteins using bulk RNA-seq from COVID-19 deceased patients in Desai N et al, 2020.
(TIFF)

S6 Fig. Analysis of SARS-CoV-2 associated comorbidity. A. Normalized gene expression (TPM) of *SPINT2* in different cancers. P-values from Wilcoxon Test. **B**. *SPINT2* gene expression in pancreatic cells from diabetic patients. Hepatocytes and pancreatic alpha cells clusters highlighted in red represent cells for which gene expression profiles are shown in *Fig 5*.
(TIFF)

S1 Table. List of DEG between infected and mock-infected cells in Calu-3 and H1299 cell lines. The 344 genes are likely to represent response genes upon infection, which we filter out of our permissivity signature.
(XLSX)

S2 Table. List of genes in the permissivity signature. The table lists the statistical parameters of the differential expression analysis between the mock-infected Calu-3 and H1299 cell lines.
(XLSX)

S3 Table. List of serine proteases related genes (SPRGs).
(XLSX)

Acknowledgments

We would like to thank Dr. Vibor Laket and Dr. Sylvia Olberg for their support through the Infectious Diseases Imaging Platform (IDIP).

Author Contributions

Conceptualization: Carlos Ramirez Alvarez, Carl Herrmann.

Formal analysis: Carlos Ramirez Alvarez, Ashwini Kumar Sharma, Leonie Thomas.

Investigation: Carmon Kee, Megan L. Stanifer.

Methodology: Carlos Ramirez Alvarez, Carmon Kee, Florian I. Schmidt, Megan L. Stanifer, Carl Herrmann.

Project administration: Steeve Boulant, Carl Herrmann.

Supervision: Steeve Boulant, Carl Herrmann.

Validation: Florian I. Schmidt.

Visualization: Ashwini Kumar Sharma.

Writing – original draft: Carlos Ramirez Alvarez, Carl Herrmann.

References

1. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020; 367: 1260–1263. https://doi.org/10.1126/science.abb2507 PMID: 32075877
2. Bestle D, Heindl MR, Limburg H, Van Lam van T, Pilgram O, Moulton H, et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. Life Sci Alliance. 2020; 3: e202000786. https://doi.org/10.26508/lsa.202000786 PMID: 32703818
3. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, 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.e8. https://doi.org/10.1016/j.cell.2020.02.052 PMID: 32142651
4. Zang R, Gomez Castro MF, McCune BT, Zeng Q, Rothlauf PW, Sonnek NM, et al. TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. Sci Immunol. 2020; 5: eabc3582. https://doi.org/10.1126/sciimmunol.abc3582 PMID: 32404436
5. Sungnak W, Huang N, Bécavin C, Berg M, Queen R, Litvinukova M, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. Nat Med. 2020; 26: 681–687. https://doi.org/10.1038/s41591-020-0868-6 PMID: 32327758
6. Seth S, Batra J, Srinivasan S. COVID-19: Targeting Proteases in Viral Invasion and Host Immune Response. Front Mol Biosci. 2020; 7: 215. https://doi.org/10.3389/fmolb.2020.00215 PMID: 33195400
7. Fuentes-Prior P. Priming of SARS-CoV-2 S protein by several membrane-bound serine proteinases could explain enhanced viral infectivity and systemic COVID-19 infection. J Biol Chem. 2020. https://doi.org/10.1074/jbc.REV120.015980 PMID: 33268377
8. Hoffmann M, Schroeder S, Kleine-Weber H, Müller MA, Drosten C, Pöhlmann S. Nafamostat mesylate blocks activation of SARS-CoV-2: New treatment option for COVID-19. Antimicrobial agents and chemotherapy. American Society for Microbiology; 2020. https://doi.org/10.1128/AAC.00754-20 PMID: 32312781
9. Bojkova D, Bechtel M, McLaughlin K-M, McGreig JE, Klann K, Bellinghausen C, et al. Aprotinin inhibits SARS-CoV-2 replication. Cells. 2020; 9: 2377. https://doi.org/10.3390/cells9112377 PMID: 33143316

10. Asselta R, Paraboschi EM, Mantovani A, Duga S. ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. Aging (Albany NY). 2020; 12: 10087–10098. https://doi.org/10.18632/aging.103415 PMID: 32501810

11. Hou Y, Zhao J, Martin W, Kallianpur A, Chung MK, Jehi L, et al. New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. BMC Med. 2020; 18: 216. https://doi.org/10.1186/s12916-020-01673-z PMID: 32664879

12. Lucas JM, Heinlein C, Kim T, Hernandez SA, Malik MS, True LD, et al. The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. Cancer Discov. 2014; 4: 1310–1325. https://doi.org/10.1158/2159-8290.CD-13-1010 PMID: 25122198

13. Steinmetzer T, Hardes K. The antiviral potential of host protease inhibitors. Activation of Viruses by Host Proteases. Cham: Springer International Publishing; 2018. pp. 279–325.

14. Straus MR, Kinder JT, Segall M, Dutch RE, Whittaker GR. SPINT2 inhibits proteases involved in activation of both influenza viruses and metapneumoviruses. Virology. 2020; 543: 43–53. https://doi.org/10.1016/j.virol.2020.01.004 PMID: 32056846

15. Ko C-J, Hsu T-W, Wu S-R, Lan S-W, Hsiao T-F, Lin H-Y, et al. Inhibition of TMPRSS2 by HAI-2 reduces prostate cancer cell invasion and metastasis. Oncogene. 2020; 39: 5950–5963. https://doi.org/10.1038/s41388-020-01413-w PMID: 32778768

16. Bojkova D, Klann K, Koch B, Widera M, Krause D, Ciesek S, et al. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. Nature. 2020; 583: 469–472. https://doi.org/10.1038/s41586-020-2332-7 PMID: 32408336

17. Roversi FM, Olalla Saad ST, Machado-Neto JA. Serine peptidase inhibitor Kunitz type 2 (SPINT2) in cancer development and progression. Biomed Pharmacother. 2018; 101: 278–286. https://doi.org/10.1016/j.biopha.2018.02.100 PMID: 29499401

18. Danielsen ET, Olsen AK, Coskun M, Nonboe AW, Larsen S, Dahlgaard K, et al. Intestinal regulation of suppression of tumorigenicity 14 (ST14) and serine peptidase inhibitor, Kunitz type -1 (SPINT1) by transcription factor CDX2. Sci Rep. 2018; 8. https://doi.org/10.1038/s41598-018-30216-z PMID: 30087389

19. Friis S, Sales KU, Schafer JM, Vogel LK, Kataoka H, Bugge TH. The protease inhibitor HAI-2, but not HAI-1, regulates matriptase activation and shedding through prostasin. J Biol Chem. 2014; 289: 22319–22332. https://doi.org/10.1074/jbc.M114.574400 PMID: 24962579

20. Yin S, Ray G, Kerschner JL, Hao S, Perez A, Drumm ML, et al. Functional genomics analysis of human colon organoids identifies key transcription factors. Physiol Genomics. 2020; 52: 234–244. https://doi.org/10.1152/physiolgenomics.00113.2019 PMID: 32390556

21. Triana S, Metz Zumaran C, Ramirez C, Kee C, Doldan P, Shahraz M, et al. Single-cell analyses reveal SARS-CoV-2 interference with intrinsic immune response in the human gut. bioRxiv. bioRxiv; 2020. https://doi.org/10.1101/2020.10.21.348854

22. Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc. 2020; 15: 2247–2276. https://doi.org/10.1038/s41596-020-0336-2 PMID: 32561888

23. Patra T, Meyer K, Geerling L, Isbell TS, Hoft DF, Brien J, et al. 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. https://doi.org/10.1371/journal.ppat.1009128 PMID: 33284859

24. Karki R, Sharma BR, Tuladhar S, Williams EP, Zalduondo L, Samir P, et al. COVID-19 cytokines and the hyperactive immune response: Synergism of TNF-α and IFN-γ in triggering inflammation, tissue damage, and death. bioRxiv. 2020. https://doi.org/10.1101/2020.10.29.361048 PMID: 33140051

25. Emanuel W, Kirstin M, Vedran F, Asija D, Theresa GL, Roberto A, et al. Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. bioRxiv. bioRxiv; 2020. https://doi.org/10.1101/2020.05.05.079194

26. Ng AY-N, Waring P, Ristevski S, Wang C, Wilson T, Pritchard M, et al. Inactivation of the transcription factor Elf3 in mice results in dysmorphogenesis and altered differentiation of intestinal epithelium. Gastroenterology. 2002; 122: 1455–1466. https://doi.org/10.1053/gast.2002.32990 PMID: 11984530

27. Sengez B, Aygün I, Shehwana H, Toyran N, Tercan Avci S, Konu O, et al. The Transcription Factor Elf3 Is Essential for a Successful Mesenchymal to Epithelial Transition. Cells. 2019; 8. https://doi.org/10.3390/cells8080858 PMID: 31404945

28. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, et al. Construction of a human cell landscape at single-cell level. Nature. 2020; 581: 303–309. https://doi.org/10.1038/s41586-020-2157-4 PMID: 32214235

29. Wyler E, Mösbauer K, Franke V, Diag A, Gottula LT, Arsiè R, et al. Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy. iScience. 2021; 24: 102151. https://doi.org/10.1016/j.isci.2021.102151 PMID: 33585804

30. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, O'Meara MJ, et al. A SARS-CoV-2-human protein-protein interaction map reveals drug targets and potential drug-repurposing. bioRxivorg. 2020. https://doi.org/10.1101/2020.03.22.002386 PMID: 32511329

31. Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, et al. A mass spectrometric-derived cell surface protein atlas. PLoS One. 2015; 10: e0121314. https://doi.org/10.1371/journal.pone.0121314 PMID: 25894527

32. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc. 2020; 15: 1484–1506. https://doi.org/10.1038/s41596-020-0292-x PMID: 32103204

33. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, Møller R, Panis M, Sachs D, et al. SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. bioRxiv. bioRxiv; 2020. https://doi.org/10.1101/2020.03.24.004655

34. Meller A, Coombs D, Shalgi R. Fluorescent polysome profiling reveals stress-mediated regulation of HSPA14-ribosome interactions. Cold Spring Harbor Laboratory. 2020. p. 860833. https://doi.org/10.1101/860833

35. Lamers MM, Beumer J, van der Vaart J, Knoops K, Puschhof J, Breugem TI, et al. SARS-CoV-2 productively infects human gut enterocytes. Science. 2020; 369: 50–54. https://doi.org/10.1126/science.abc1669 PMID: 32358202

36. Mulay A, Konda B, Garcia G, Yao C, Beil S, Sen C, et al. SARS-CoV-2 infection of primary human lung epithelium for COVID-19 modeling and drug discovery. bioRxivorg. 2020. https://doi.org/10.1101/2020.06.29.174623 PMID: 32637946

37. Lodermeyer V, Ssebyatika G, Passos V, Ponnurangam A, Malassa A, Ewald E, et al. The antiviral activity of the cellular glycoprotein LGALS3BP/90K is species specific. J Virol. 2018; 92. https://doi.org/10.1128/JVI.00226-18 PMID: 29743357

38. Hepojoki J, Strandin T, Hetzel U, Sironen T, Klingström J, Sane J, et al. Acute hantavirus infection induces galectin-3-binding protein. J Gen Virol. 2014; 95: 2356–2364. https://doi.org/10.1099/vir.0.066837-0 PMID: 25013204

39. Hong C-S, Park M-R, Sun E-G, Choi W, Hwang J-E, Bae W-K, et al. Gal-3BP Negatively Regulates NF-κB Signaling by Inhibiting the Activation of TAK1. Front Immunol. 2019; 10: 1760. https://doi.org/10.3389/fimmu.2019.01760 PMID: 31402917

40. Stakaitytė G, Nwogu N, Lippiat JD, Blair GE, Poterlowicz K, Boyne JR, et al. The cellular chloride channels CLIC1 and CLIC4 contribute to virus-mediated cell motility. J Biol Chem. 2018; 293: 4582–4590. https://doi.org/10.1074/jbc.RA117.001343 PMID: 29462791

41. Wyżewski Z, Gregorczyk KP, Szczepanowska J, Szulc-Dąbrowska L. Functional role of Hsp60 as a positive regulator of human viral infection progression. Acta Virol. 2018; 62: 33–40. https://doi.org/10.4149/av_2018_104 PMID: 29521101

42. Koenig P-A, Das H, Liu H, Kümmerer BM, Gohr FN, Jenster L-M, et al. Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. Science. 2021; 371. https://doi.org/10.1126/science.abe6230 PMID: 33436526

43. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. Nat Biotechnol. 2020; 38: 970–979. https://doi.org/10.1038/s41587-020-0602-4 PMID: 32591762

44. Zhu N, Wang W, Liu Z, Liang C, Wang W, Ye F, et al. Morphogenesis and cytopathic effect of SARS-CoV-2 infection in human airway epithelial cells. Nat Commun. 2020; 11: 3910. https://doi.org/10.1038/s41467-020-17796-z PMID: 32764693

45. Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martínez-Colón GJ, McKechnie JL, et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat Med. 2020; 26: 1070–1076. https://doi.org/10.1038/s41591-020-0944-y PMID: 32514174

46. Desai N, Neyaz A, Szabolcs A, Shih AR, Chen JH, Thapar V, et al. Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. medRxiv. 2020. https://doi.org/10.1101/2020.07.30.20165241 PMID: 32766600

47. Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. Kidney Int. 2020; 97: 829–838. https://doi.org/10.1016/j.kint.2020.03.005 PMID: 32247631

48. Jothimani D, Venugopal R, Abedin MF, Kaliamoorthy I, Rela M. COVID-19 and the liver. J Hepatol. 2020; 73: 1231–1240. https://doi.org/10.1016/j.jhep.2020.06.006 PMID: 32553666

49. Sanyaolu A, Okorie C, Marinkovic A, Patidar R, Younis K, Desai P, et al. Comorbidity and its Impact on Patients with COVID-19. SN Compr Clin Med. 2020; 1–8. https://doi.org/10.1007/s42399-020-00363-4 PMID: 32838147

50. Guan W-J, Liang W-H, Zhao Y, Liang H-R, Chen Z-S, Li Y-M, et al. Comorbidity and its impact on 1590 patients with Covid-19 in China: A Nationwide Analysis. Eur Respir J. 2020 [cited 18 Dec 2020]. https://doi.org/10.1183/13993003.00547-2020 PMID: 32217650

51. Guan W-J, Liang W-H, He J-X, Zhong N-S. Cardiovascular comorbidity and its impact on patients with COVID-19. Eur Respir J. 2020; 55. https://doi.org/10.1183/13993003.01227-2020 PMID: 32341104

52. Fukushima T, Kawaguchi M, Yamamoto K, Yamashita F, Izumi A, Kaieda T, et al. Aberrant methylation and silencing of the SPINT2 gene in high-grade gliomas. Cancer Sci. 2018; 109: 2970–2979. https://doi.org/10.1111/cas.13732 PMID: 29987920

53. Dong W, Chen X, Xie J, Sun P, Wu Y. Epigenetic inactivation and tumor suppressor activity of HAI-2/SPINT2 in gastric cancer. Int J Cancer. 2010; 127: 1526–1534. https://doi.org/10.1002/ijc.25161 PMID: 20063316

54. Qu Y, Dang S, Hou P. Gene methylation in gastric cancer. Clin Chim Acta. 2013; 424: 53–65. https://doi.org/10.1016/j.cca.2013.05.002 PMID: 23669186

55. Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. Science. 2018; 361: 594–599. https://doi.org/10.1126/science.aat1699 PMID: 30093597

56. Lu Y, Yang A, Quan C, Zhang H, Pan Y, Li Y, et al. A Single-Cell Atlas of the Multicellular Ecosystem of Primary and Metastatic Hepatocellular Carcinoma. 2020. https://doi.org/10.2139/ssrn.3624436

57. Li H, Courtois ET, Sengupta D, Tan Y, Chen KH, Goh JJL, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat Genet. 2017; 49: 708–718. https://doi.org/10.1038/ng.3818 PMID: 28319088

58. Laughney AM, Hu J, Campbell NR, Bakhoum SF, Setty M, Lavallée V-P, et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. Nat Med. 2020; 26: 259–269. https://doi.org/10.1038/s41591-019-0750-6 PMID: 32042191

59. Segerstolpe Å, Palasantza A, Eliasson P, Andersson E-M, Andréasson A-C, Sun X, et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. Cell Metab. 2016; 24: 593–607. https://doi.org/10.1016/j.cmet.2016.08.020 PMID: 27667667

60. Seget-Trzensiok K, Bernhard SV, Kuffer C, Krastev DB, Theis M, Keuper K, et al. USP28 and SPINT2 mediate cell cycle arrest after whole genome doubling. bioRxiv; 2020. https://doi.org/10.1101/2020.09.18.303834

61. Poddar NK, Maurya SK, Saxena V. Role of Serine Proteases and Inhibitors in Cancer. Proteases in Physiology and Pathology. Singapore: Springer Singapore; 2017. pp. 257–287.

62. Finotti P. The role played by serine proteases in the development and worsening of vascular complications in type 1 diabetes mellitus. Curr Diabetes Rev. 2006; 2: 295–305. https://doi.org/10.2174/157339906777950624 PMID: 18220634

63. Sharony R, Yu P-J, Park J, Galloway AC, Mignatti P, Pintucci G. Protein targets of inflammatory serine proteases and cardiovascular disease. J Inflamm (Lond). 2010; 7: 45. https://doi.org/10.1186/1476-9255-7-45 PMID: 20804552

64. Martin CE, List K. Cell surface–anchored serine proteases in cancer progression and metastasis. Cancer Metastasis Rev. 2019; 38: 357–387. https://doi.org/10.1007/s10555-019-09811-7 PMID: 31529338

65. Goel A, Chauhan SS. Role of proteases in tumor invasion and metastasis. Indian J Exp Biol. 1997; 35: 553–564. PMID: 9357157

66. Parr C, Sanders AJ, Jiang WG. Hepatocyte growth factor activation inhibitors—therapeutic potential in cancer. Anticancer Agents Med Chem. 2010; 10: 47–57. https://doi.org/10.2174/1871520611009010047 PMID: 20015004

67. Yang J-K, Lin S-S, Ji X-J, Guo L-M. Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes. Acta Diabetol. 2010; 47: 193–199. https://doi.org/10.1007/s00592-009-0109-4 PMID: 19333547

68. Mason SD, Joyce JA. Proteolytic networks in cancer. Trends Cell Biol. 2011; 21: 228–237. https://doi.org/10.1016/j.tcb.2010.12.002 PMID: 21232958

69. Yang X, Boehm JS, Yang X, Salehi-Ashtiani K, Hao T, Shen Y, et al. A public genome-scale lentiviral expression library of human ORFs. Nat Methods. 2011; 8: 659–661. https://doi.org/10.1038/nmeth.1638 PMID: 21706014

70. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cold Spring Harbor Laboratory. 2020. p. 2020.10.12.335331. https://doi.org/10.1101/2020.10.12.335331

71. Li WV, Li JJ. An accurate and robust imputation method scImpute for single-cell RNA-seq data. Nat Commun. 2018; 9: 997. https://doi.org/10.1038/s41467-018-03405-7 PMID: 29520097

72. Aizarani N, Saviano A, Sagar, Mailly L, Durand S, Herman JS, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature. 2019; 572: 199–204. https://doi.org/10.1038/s41586-019-1373-2 PMID: 31292543

73. Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, et al. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat Commun. 2020; 11: 4267. https://doi.org/10.1038/s41467-020-18035-1 PMID: 32848148

74. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13: 2498–2504. https://doi.org/10.1101/gr.1239303 PMID: 14597658