



OPEN

Increased TRIM31 gene expression is positively correlated with SARS-CoV-2 associated genes TMPRSS2 and TMPRSS4 in gastrointestinal cancers

Mehmet Arda Temena & Ahmet Acar✉

Besides typical respiratory symptoms, COVID-19 patients also have gastrointestinal symptoms. Studies focusing on the gastrointestinal tumors derived from gastrointestinal tissues have raised a question whether these tumors might express higher levels of SARS-CoV-2 associated genes and therefore patients diagnosed with GI cancers may be more susceptible to the infection. In this study, we have analyzed the expression of SARS-CoV-2 associated genes and their co-expressions in gastrointestinal solid tumors, cancer cell lines and patient-derived organoids relative to their normal counterparts. Moreover, we have found increased co-expression of TMPRSS2-TMPRSS4 in gastrointestinal cancers suggesting that SARS-CoV-2 viral infection known to be mediated by this protease pair might facilitate the effects of viral infection in GI cancer patients. Further, our findings also demonstrate that TRIM31 expression is upregulated in gastrointestinal tumors, while the inhibition of TRIM31 significantly altered viral replication and viral processes associated with cellular pathways in gastrointestinal cancer samples. Taken together, these findings indicate that in addition to the co-expression of TMPRSS2-TMPRSS4 protease pair in GI cancers, TRIM31 expression is positively correlated with this pair and TRIM31 may play a role in providing an increased susceptibility in GI cancer patients to be infected with SARS-CoV-2 virus.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which can cause asymptomatic, mild to moderate and severe respiratory coronavirus disease 2019 (COVID-19) infections emerged in Wuhan (China) in December 2019^{1,2}. SARS-CoV-2 has infected approximately 400 million people to date (by February 2022) and caused more than 5,7 million deaths³. Right after its emergence, the COVID-19 pandemic has started to transform the lives of people and profoundly altered medicinal approaches in the treatment of patients. The lack of widespread access to effective vaccines worldwide and the emergence of new variants have raised new concerns regarding the transmission of the virus^{4,5}. Furthermore, different clinical responses to viral or other infections in some of the patients are known to be regulated by genetic factors or interindividual variability⁶. In terms of the transmission pathways of the virus, the evasion from natural immunity and the high infectivity of new variants are the main causes for the continuous and rapid spread of SARS-CoV-2. In order to stop the spread of the virus, it is essential to uncover the underlying molecular mechanisms of immune evasion and infectivity⁷⁻¹⁰.

Direct fusion of the cell membrane with the viral envelope or membrane fusion within the endosomes following the endocytosis allows the virus to enter the cell¹¹. SARS-CoV-2 binds to angiotensin-converting enzyme 2 (ACE2), which is widely expressed in many tissues in the human body^{12,13}. Besides this receptor, other host receptor proteins facilitating virus entry into cells have been reported¹¹. The spike glycoprotein that is responsible for the attachment of the virus to the host cell receptor, known as S protein, consists of one signal peptide, receptor-binding domain (RBD), subdomains in one subunit (S1) and heptad repeats with membrane-fusion subunit (S2)^{14,15}. After RBD-receptor interaction, the S protein undergoes proteolysis which is catalyzed through interaction of host proteins/proteases like furin; paired basic amino acid cleaving enzyme FURIN, transmembrane serine protease 2 (TMPRSS2) and cathepsin B/L (CTSB/L)¹⁶. Apart from these, there are also co-receptors that can increase viral infectivity and have a significant role to facilitate binding characteristics in cells¹⁷. For

Department of Biological Sciences, Middle East Technical University, Universiteler Mah. Dumlupınar Bulvarı 1, 06800 Çankaya, Ankara, Turkey. ✉email: acar@metu.edu.tr

example, NRP1 from the neuropilin (NRPs) family, which are highly conserved, and single-transmembrane proteins can act as a co-receptor for various molecules to facilitate the entry of SARS-CoV-2 into host cells^{18–20}. Although SR-B1 (Scavenger Receptor Class B Member 1; SCARB1) cannot directly bind to the S protein, the SCARB1 expression facilitates SARS-CoV-2 entry into cells when co-expressed with ACE2 through utilizing cholesterol recognition motifs on the subunit S1^{17,21}. Among other proteins, serine proteases are also involved in the infection process by interacting with the spike protein of the virus, one of which is TMPRSS4. Recent findings suggest that TMPRSS4 may play an important role for the coronavirus infection^{22,23}.

Even though it is widely reported that COVID-19 affects the respiratory tract, associated with alveolar damage, extrapulmonary symptoms have also been observed²⁴. Thrombotic diseases, dysfunction and rhythm disorders in the heart muscle, acute coronary syndromes, serious kidney damage, gastrointestinal symptoms, hepatocellular complications, ocular, dermatological and neurologic symptoms are among the other symptoms commonly seen in COVID-19 patients^{24–26}. Due to high expression of ACE2 in extrapulmonary tissues, there is a possibility that SARS-CoV-2 can infect these tissues. As such, endothelial damage, and thrombo-inflammation in extrapulmonary tissues of COVID-19 patients may involve ACE2-related pathways such as TMPRSS2 alongside with the dysregulation of immune response^{24,27,28}.

Although the symptoms caused by the virus are mostly seen in the respiratory tract, approximately 15% of the patients show gastrointestinal symptoms²⁹. In addition to the high expression of ACE2 in intestinal glandular cells, elevated expression of viral nucleocapsid proteins in epithelial cells of gastric, duodenal and rectal tissues and glandular enterocytes can cause susceptibility of coronavirus infection in gastrointestinal tissues^{30,31}. On the one hand, high expression of FURIN and TMPRSS2 in small intestinal epithelial cells mediates increased susceptibility to coronavirus infection in gastrointestinal tissues³². On the other hand, epithelial cells in the gut (mainly absorptive enterocytes of the colon and ileum) and esophagus also highly express ACE2 receptor^{33,34}. Moreover, glandular cells of both the duodenum and the stomach have been reported to express ACE2; thus, SARS-CoV-2 can infect intestinal epithelial cells through ACE2 receptor³³. Additionally, ACE2 receptors in the gastrointestinal tract play a crucial regulatory role in homeostasis, gut microbiome, and innate immunity. As a result, binding SARS-CoV-2 to these receptors and activating other epithelial surface proteins that can facilitate the entry of the virus may be one of the reasons for coronavirus infection in gastrointestinal tissues³⁵. Conversely, live virus and RNA material that have been detected in the stools of COVID-19 patients were reported play a role in SARS-CoV-2 entry into cells starting from the apical surface and activating epithelial cell fusion in the intestine^{36–38}. In particular, TMPRSS2 and TMPRSS4 serine proteases work together, accelerating S protein cleavage and thus increasing membrane fusion^{22,39}.

It has also been observed that the intrinsic variability of receptors in cells and different tissue factors such as structural and genetic modifications on these proteins can affect a patient's susceptibility to COVID-19⁴⁰. In one of the first reports of COVID-19 among cancer patients in China, it was reported that they experienced a serious risk of COVID-19, with other complications in addition to respiratory distress, compared to patients without cancer⁴¹. Since most of the cancer patients suffer from post treatment side effects including impaired immune system, these patients are possibly more susceptible to COVID-19^{42,43}. A recent study that investigated clinical manifestations of COVID-19, reported that cancer patients had a 3.5-fold increase in the risk of morbidity and an increased risk of severe infection compared to non-cancer patients⁴¹. In addition to this, according to one comprehensive study, authors reported that cancer patients who have encountered severe acute respiratory syndrome virus (SARS or SARS-CoV-1), Middle East Respiratory Syndrome (MERS) virus, and SARS-CoV-2 have a higher risk of severe complications and death⁴⁴. In general, patients with cancer may be much more vulnerable to SARS-CoV-2 infection⁴⁵. However, important clinical questions remain unanswered is that optimal treatment and prevention strategies for cancer patients exposed to or infected with SARS-CoV-2⁴⁶. In addition to COVID-19 risk in general cancer patient population, there have been studies demonstrating the risk of COVID-19 infection in GI cancer patients. Aznab et al., investigated COVID-19 prevalence in 279 cancer patients and reported that 72 CRC patients had COVID-19 infection while 11 patients with lung cancer, 5 with brain cancer and 12 with ovarian cancer had COVID-19 infection⁴⁷. Another study performed a meta-analysis including 6 different studies to evaluate the prevalence of COVID-19 infection in cancer patients⁴⁸. This study reported that prevalence of COVID-19 infection in a total of 205 cancer patients including GI cancer types (CRC, Esophagus and Pancreatic cancers) and they found that prevalence of COVID-19 infection in CRC patients was 20.5%, in Esophagus cancer patients was 7.6%, in Pancreatic cancer patients was 6.1%, and in lung cancer patients was 24.7%. Another meta-analysis study showed that the prevalence of COVID-19 infection in CRC patients was 45.1% by comparing 6 different studies composed of 204 different cancer patients⁴⁹. Despite all these studies show high infection risk of COVID-19 for GI cancer patients, these findings should be approached with caution due to following factors. Estimating COVID-19 infection risk in cancer patients requires more patients involved in larger studies. In addition, such studies need to take into consideration whole population beyond cancer patients with COVID-19 infection. Moreover, ethnicity bias may play role as these studies mainly rely on data obtained from few developed countries.

In this study, we performed a systematic gene expression analysis for SARS-CoV-2 genes in gastrointestinal tumor samples using publicly available datasets of solid tumors, cell lines and patient-derived organoids relative to their normal counterparts. In addition to individual gene expression changes in majority of GI cancers for *TMPRSS2*, *TMPRSS4*, *ACE2*, *CTSL*, *NRP1*, *FURIN*, and *SCARB1*, we assessed co-expression of highly expressed genes. We found that co-expressed *TMPRSS2* and *TMPRSS4* gene pair is positively correlated in GI samples. Moreover, our findings demonstrate that *TRIM31* is positively correlated with *TMPRSS2-TMPRSS4* genes in GI samples suggesting a possibility of *TRIM31* acting alongside with *TMPRSS2-TMPRSS4* protease pair in SARS-CoV-2 viral entry and promoting viral infection in GI cells. Finally, *TRIM31* knockdown data obtained from colorectal cancer cell lines resulted in changes in cellular pathways involved in viral replication and viral

Gene symbol	Gene description
<i>CTSL</i>	Cathepsin L
<i>SLC6A19</i>	Solute carrier family 6 member 19
<i>ACAT1</i>	Acetyl-CoA acetyltransferase 1
<i>FURIN</i>	Paired basic amino acid cleaving enzyme
<i>TLR7</i>	Toll like receptor 7
<i>TMPRSS4</i>	Transmembrane serine protease 4
<i>ACE2</i>	Angiotensin I converting enzyme 2
<i>TMPRSS2</i>	Transmembrane serine protease 2
<i>NRP1</i>	neuropilin 1
<i>SCARB1</i>	Scavenger receptor class B member 1

Table 1. Gene set that was found in SARS-CoV-2 entry into host cells.

processes. This supports our hypothesis whereby *TRIM31* may be linked to processes that mediate SARS-CoV-2 infection in GI cancers.

Results

Gene Set Enrichment Analysis-Molecular Signatures DataBase (GSEA-MSigDB) was utilized to obtain SARS-CoV-2 infection canonical pathway genes. The resulting genes are documented in Table 1. These genes were reported as responsible for the attachment of SARS-CoV-2 and its entry into cells. Initially we performed comparative gene expression analysis between normal gastrointestinal (GI) tissue samples. The analysis of gene expression changes between normal tissue samples revealed that *CTSL* expression in colon and esophagus mucosa, *FURIN* expression in pancreas, and *TMPRSS2* expression in stomach was found as the highest among other genes (Fig. 1a). We next investigated the gene expression changes in GI tumor samples in comparison to their normal derivatives and observed significant changes in the expression of SARS-CoV-2-associated genes in various GI tumors (Fig. 1b). When tumor tissue samples were analyzed in comparison to normal samples, colon adenocarcinoma (COAD) and rectal adenocarcinoma (READ) samples exhibited significantly elevated expression of *ACE2*, *SCARB1*, *TMPRSS2* and *TMPRSS4* genes while pancreatic adenocarcinoma (PAAD) and stomach adenocarcinoma (STAD) samples demonstrated increased expression of *ACE2*, *CTSL*, *NRP1*, *SCARB1* and *TMPRSS4* genes relative to normal samples (Fig. 1b). Further, *CTSL*, *FURIN*, *NRP1* and *SCARB1* gene expression levels were significantly higher in esophageal carcinoma (ESCA) samples in comparison to their normal matches (Fig. 1b). Differential gene expression results for each gene were given in supplementary figures in the following order: *ACE2* (Supplementary Figure S1a), *TMPRSS2* (Supplementary Figure S1b), *TMPRSS4* (Supplementary Figure S1c), *CTSL* (Supplementary Figure S1d), *FURIN* (Supplementary Figure S1e), *NRP1* (Supplementary Figure S1f), *SLC6A19* (Supplementary Figure S1g), *ACAT1* (Supplementary Figure S1h), *SCARB1* (Supplementary Figure S1i), *TLR7* (Supplementary Figure S1j).

After highly expressed genes were determined in GI solid tumor samples, a pairwise correlation analysis was performed to determine the relationship between these genes. Given some of the genes are co-expressed in tumor samples (Supplementary Figure S2), we reasoned that it may be possible to examine gene expression dependencies through their correlation analysis. After a pairwise correlation analysis of TCGA samples, *ACE2-SCARB1*, *TMPRSS2-SLC6A19* and *TMPRSS2-TMPRSS4* had a moderate positive correlation in COAD samples (Supplementary Figure S3a, p-values < 0.05 for all three; and R-values = 0.49, 0.32 and 0.4, respectively). For READ samples, in addition to positive correlation between *ACE2-SCARB1*, *TMPRSS2-SLC6A19* and *TMPRSS2-TMPRSS4*, a moderate positive correlation between *SCARB1-ACAT1* genes were also observed (Supplementary Figure S3b, p-values < 0.05 for all four; and R-values = 0.42, 0.5, 0.35 and 0.48, respectively). Moreover, we noted a moderate positive correlation between expression of *ACE2-TMPRSS2*, *TMPRSS2-TMPRSS4*, *TMPRSS2-SLC6A19*, *ACE2-SCARB1*, *CTSL-NRP1*, *NRP1-TLR7* and *SCARB1-ACAT1* in ESCA samples (Supplementary Figure S3c, p-values < 0.05 for all seven; and R-values = 0.36, 0.32, 0.38, 0.3, 0.39, 0.35 and 0.38, respectively). For PAAD samples, we observed a moderate positive correlation between *TMPRSS2-TMPRSS4*, *CTSL-TLR7*, *CTSL-NRP1* and *NRP1-TLR7* (Supplementary Figure S3d, p-values < 0.05 for all seven; and R-values = 0.45, 0.31, 0.39 and 0.62, respectively). Lastly, only *CTSL-NRP1*, *CTSL-TLR7* and *NRP1-TLR7* genes showed a moderate positive correlation in STAD samples (Supplementary Figure S3e, p-values < 0.05 for all three; and R-values = 0.51, 0.39, and 0.46, respectively). Importantly, all gastrointestinal tumor samples showed a positive correlation between *TMPRSS2* and *TMPRSS4* genes (Fig. 2). All correlation results between gene pairs in GI tumor types are presented in a supplementary file (Supplementary Figure S3). Taken together, these findings demonstrate that there is a positive correlation between the expression of a list of gene pairs specific to GI tumor types. Among these, *TMPRSS2* and *TMPRSS4* was commonly observed in all gastrointestinal tumor types.

Following a pairwise gene expression correlation analysis using GI tumor samples, the same analysis was performed using the CCLE database. By doing this, we aimed at determining whether there is a correlation between the SARS-CoV-2-associated genes similar to our observations in GI tumor types. Initial analysis was performed to eliminate cell lines exhibiting zero gene expression. Following this, the cell lines that have at least $\log_2(\text{TPM} + 1) > 0$ gene expression were included into the analysis. As a result, 63 colorectal cancer, 32 esophagus cancer, 41 pancreatic cancer and 29 stomach cancer cell lines were used (Supplementary Table S1). When we

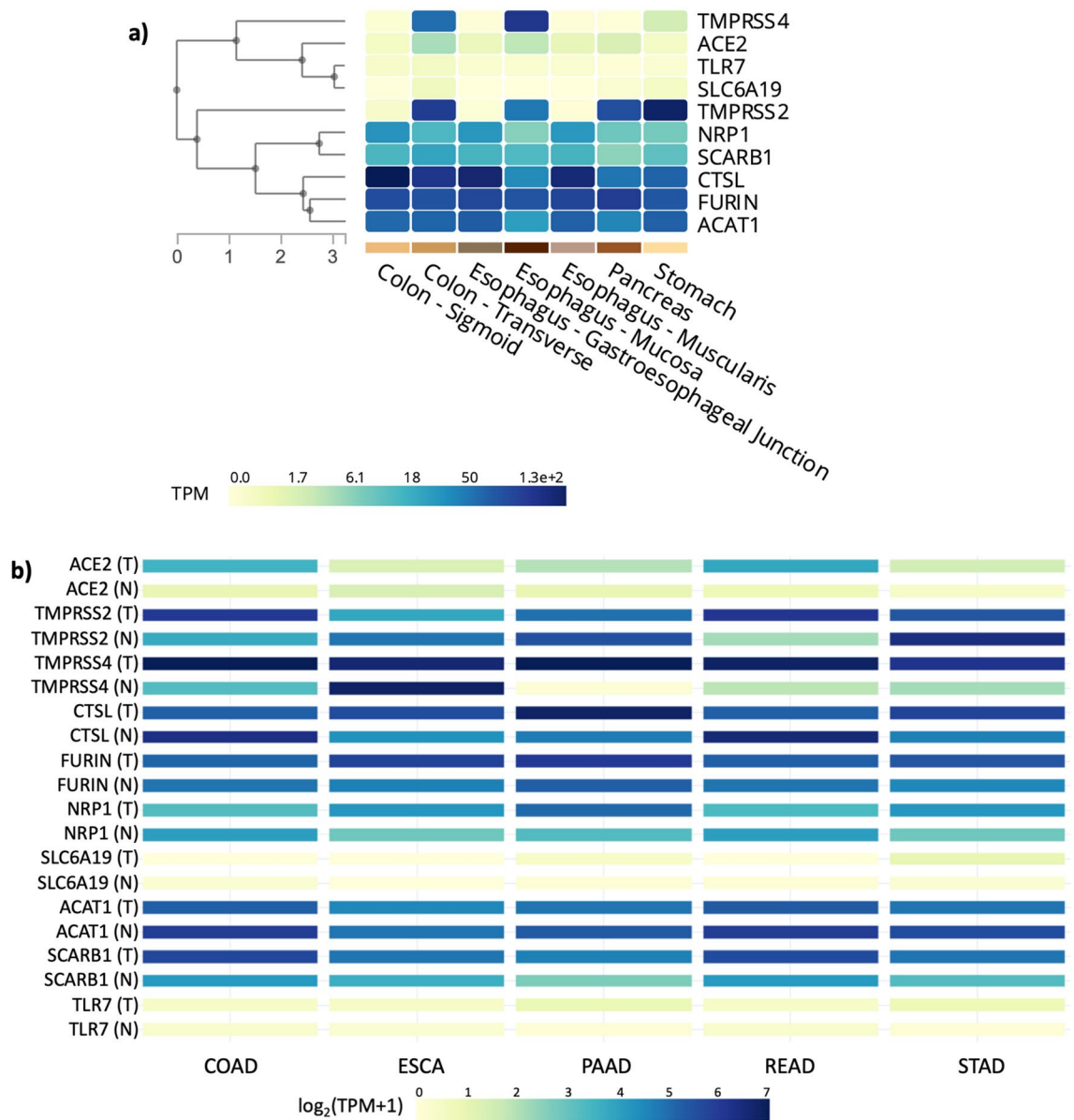


Figure 1. Heatmap of RNA-sequencing-based gene expression from the SARS-CoV-2-associated genes generated using GTEx portal for a multi-gene query in seven gastrointestinal tissues. **(a)** Comparison of the genes in different gastrointestinal normal samples. **(b)** TCGA tumor-normal samples and GTEx samples together (*T* tumor, *N* normal, *TPM* Transcripts per kilobase million-expresses RNA-sequencing reads normalized for the length of gene and sequencing depth).

performed gene expression correlation analysis for *TMPRSS2* and *TMPRSS4* genes in GI cancer cell lines, we found a moderate and statistically significant positive correlation for this gene pair in colorectal and stomach cancer cell lines (Fig. 3a, b p-values < 0.05 and R-values 0.613 and 0.559, respectively). Similarly, we observed statistically significant positive correlation for the expression of *TMPRSS2* and *TMPRSS4* genes in pancreatic and esophagus cancer cell lines (Fig. 3c, d p-values < 0.05 and R-values 0.359 and 0.521, respectively). These observations indicate that the expression of *TMPRSS2* and *TMPRSS4* gene pair are positively correlated to GI cancer cell lines. This resembles our findings in the GI tumor samples.

As *TMPRSS2* and *TMPRSS4* gene expression is positively correlated in both GI tumors and cell lines, we initially investigated whether there is an additional number of genes which are commonly associated with *TMPRSS2* and *TMPRSS4* genes in GI cancer cell lines. To address this question, we performed a correlation analysis in several GI cancer cell lines (colorectal, stomach, pancreatic, esophagus cancer cell lines) to identify a positively correlated gene(s) with *TMPRSS2* and *TMPRSS4*. We observed *TMPRSS2* and *TMPRSS4* expression to be positively correlated with seven genes which are Solute Carrier Family 44 Member 4 (*SLC44A4*), Proline Rich 15 Like (*PRR15L*), Receptor-interacting serine/threonine-protein kinase 3 (*RIPK3*), SH3 Domain Binding Glutamate Rich Protein Like 2 (*SH3BGRL2*), Transmembrane Protein 45B (*TMEM45B*), Tight Junction Protein 3 (*TJP3*) and Tripartite Motif Containing 31 (*TRIM31*) (Supplementary Table S2).

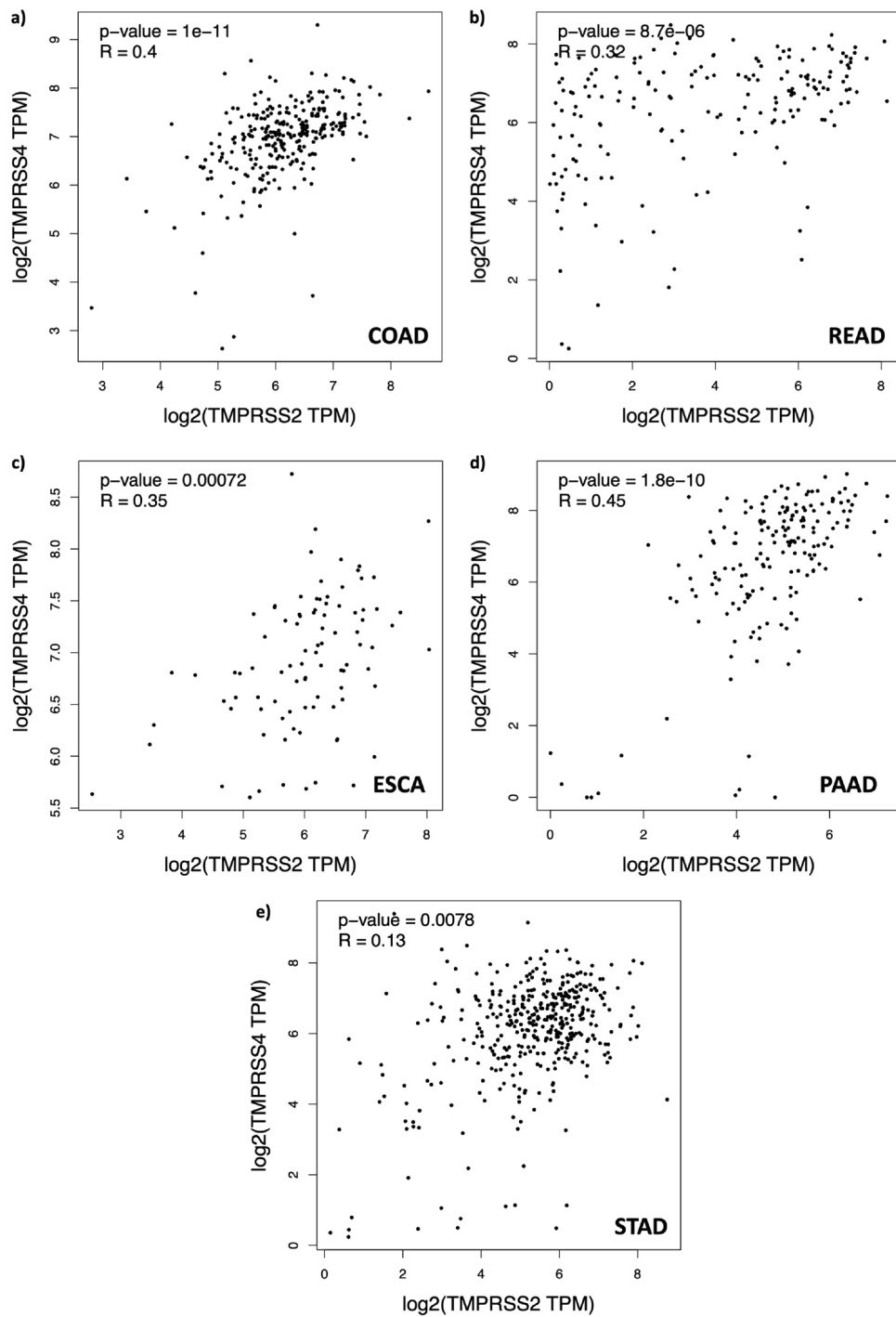


Figure 2. Correlation analysis results of *TMPRSS2* and *TMPRSS4* genes in gastrointestinal cancer tissues. Relationship of *TMPRSS2* and *TMPRSS4* with respect to their expression values shows moderate positive correlation (R -value > 0.3) in COAD, READ, ESCA and PAAD (a–d) and weak positive correlation in STAD samples (e).

We next investigated the differential gene expression of these seven genes in GI solid tumors in comparison to their matched normal tissue samples. All GI tumor samples except STAD showed significantly elevated expression of *SLC44A4* when compared to normal tissue samples (Supplementary Figure S4a). Moreover, *TMEM45B* expression was found to be significantly overexpressed in COAD, READ, STAD and PAAD except in ESCA samples (Supplementary Figure S4b). *PRR15L* showed highly increased expression in COAD and READ samples while *RIPK3* was only significantly higher in PAAD samples relative to normal (Supplementary Figure S4c, Supplementary Figure S4d). In addition, no significant change or elevation in the expression of *SH3BGRL2* was

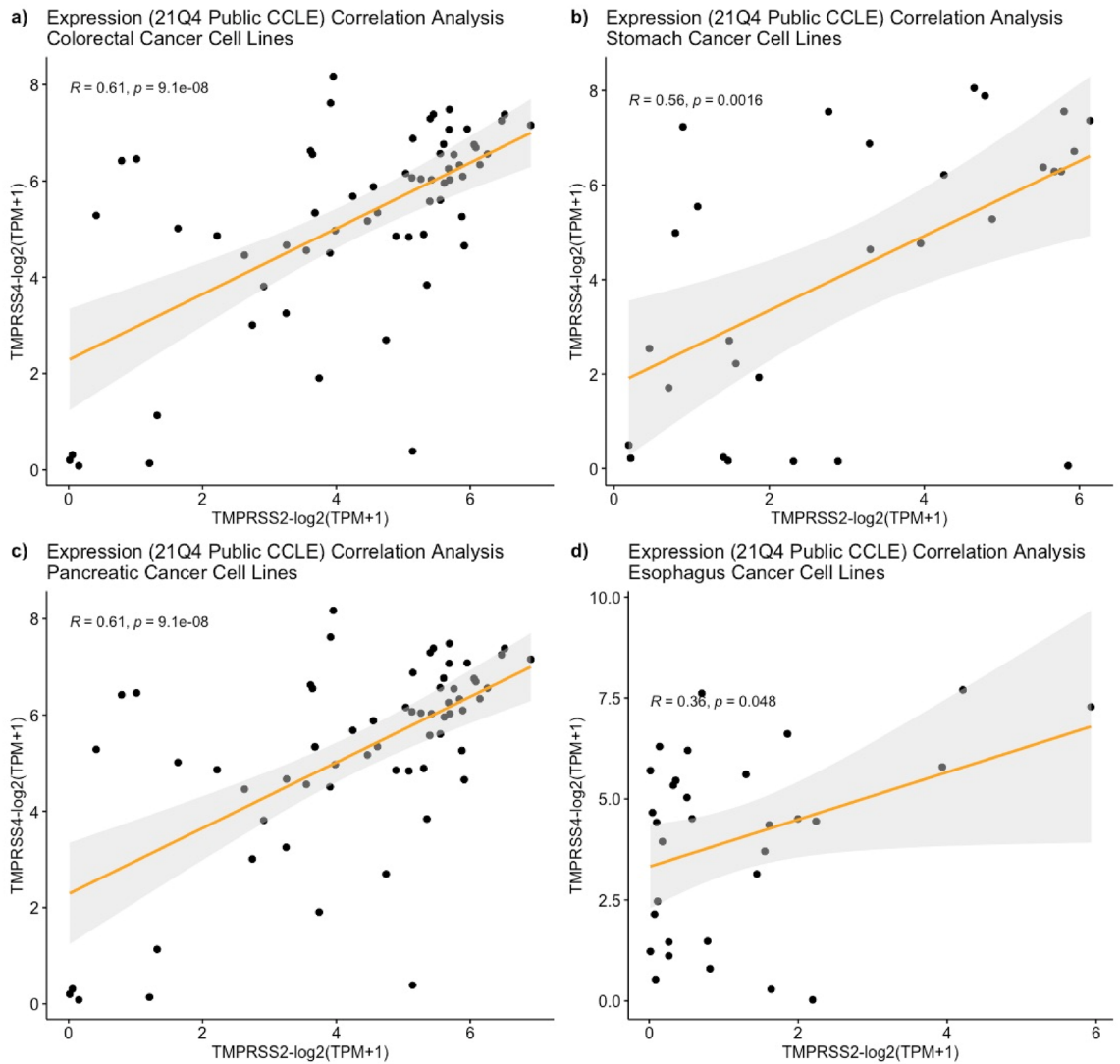


Figure 3. Positive correlation between *TMPRSS2* and *TMPRSS4* in gastrointestinal cancer cell lines from CCLE database. For colorectal (a), stomach (b) and pancreatic (c) cancer cell lines, R-value > 0.5 and for esophagus cancer cell lines (d) R-value > 0.3 while p-values for all of them < 0.05.

found in STAD and ESCA samples while it was significantly overexpressed in COAD, READ and PAAD samples (Fig. 4a). Furthermore, *TJP3* gene expression was found to be highly overexpressed in COAD, READ and PAAD samples in comparison to their matched normal samples (Fig. 4b). As a result, three genes, *SH3BGRL2*, *TJP3* and *TRIM31*, were positively correlated with *TMPRSS2* and *TMPRSS4* for all cancer cell lines and showed distinct differential expression profiles across GI solid tumors (Fig. 4). However, there was only one common gene that was both positively correlated with *TMPRSS2* and *TMPRSS4* in gastrointestinal cancer cell lines and significantly high in all gastrointestinal solid cancers, which was *TRIM31* (Fig. 4c). There was also significant positive correlation of *TRIM31* with *TMPRSS2* and *TMPRSS4* gene pair in all GI tumor samples (Supplementary Figure S5).

After *TRIM31* was found to be upregulated and positively correlated with *TMPRSS2* and *TMPRSS4* genes in both cancer cell lines and solid tumor samples, we aimed to investigate the relationship of *TRIM31* with *TMPRSS2* and *TMPRSS4* using the clinically relevant model system Patient Derived Organoids (PDOs)^{50,51}. RNA-seq expression counts data from 87 colorectal cancer PDOs with their matched tumor samples were retrieved from the Gene Expression Omnibus (GEO) database (GSE171681). Our pair-wise gene expression correlation analysis revealed that *TRIM31* gene expression in colorectal PDOs was positively correlated with *TMPRSS2* and *TMPRSS4* (Fig. 5a, b). Yet, there was no significant correlation observed between *TRIM31* to *TMPRSS2* and *TMPRSS4* in pancreatic cancer PDOs, but a positive correlation was still observed between *TMPRSS2* and *TMPRSS4* (Fig. 5c, d). Taken together, these observations indicate that *TRIM31* gene expression is positively correlated with *TMPRSS2* and *TMPRSS4* in GI cell lines, solid tumors and colorectal cancer PDOs.

This was followed by investigating the functional role of *TRIM31* and its relationship cellular pathways using a genome-wide loss of screen approach. RNAi results for *TRIM31* in colorectal cancer cell lines were obtained from the CCLE database utilizing DEMETER2^{52–56}. This model integrates large-scale pooled RNAi screening data expression changes. The relationship between expression correlation values and *TRIM31* dependency

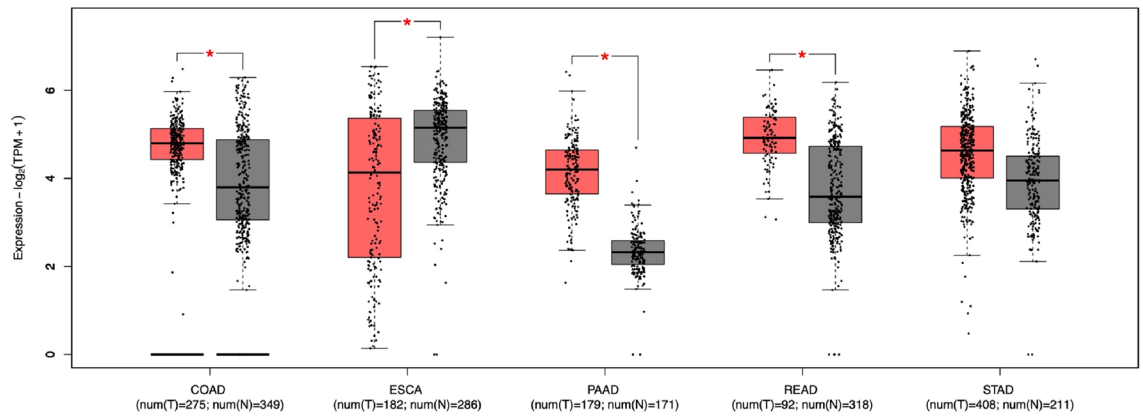
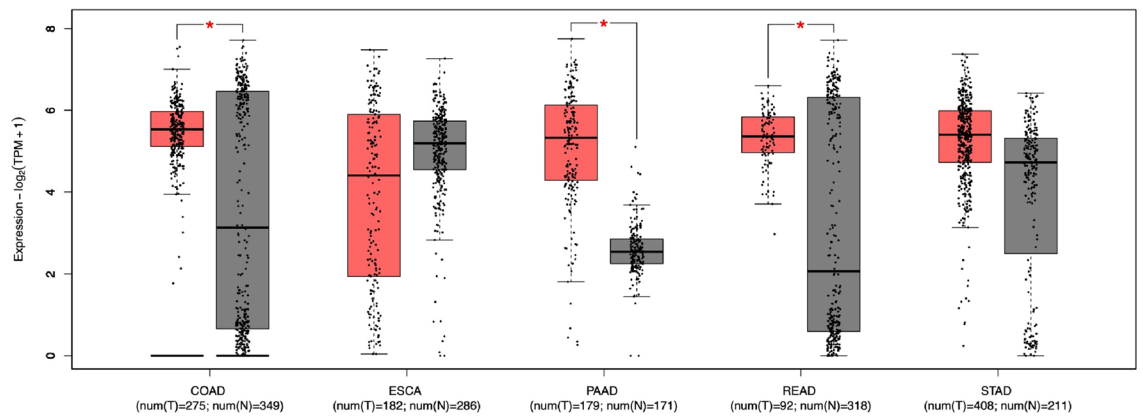
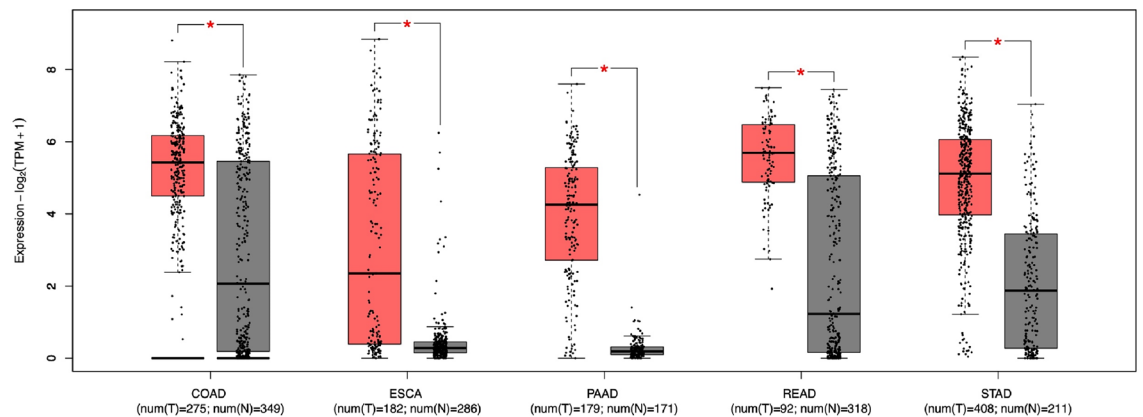
a) Expression of *SH3BGRL2* in gastrointestinal tumorsb) Expression of *TJP3* in gastrointestinal tumorsc) Expression of *TRIM31* in gastrointestinal tumors

Figure 4. Cancer type-wise expression of eight genes (a; *SH3BGRL2*, b; *TJP3*, c; *TRIM31*) in TCGA gastrointestinal solid tumor samples. Red boxes are tumor samples while gray ones are normal and boxplots with ‘*’ sign shows statistically significant differential gene expression value, p-value < 0.05. Gene Expression Profiling Interactive Analysis: colon adenocarcinoma (COAD), rectal adenocarcinoma (READ), esophagus adenocarcinoma (ESCA), pancreas adenocarcinoma (PAAD), stomach adenocarcinoma (STAD).

was determined in 20 colorectal cancer cell lines (Supplementary Table S3). Using the dataset, we performed a pathway enrichment analysis to assess *TRIM31* knockdown effect in cellular processes in colorectal cancer cell lines. Initially we have determined upregulated and downregulated genes when *TRIM31* was silenced using the RNAi approach (Fig. 6a). This analysis showed that number of genes downregulated genes (colored with blue circles) had bigger co-expression scores than upregulated genes (colored with red circles). This was then followed by the identification of related pathways when silencing of *TRIM31* caused downregulation of

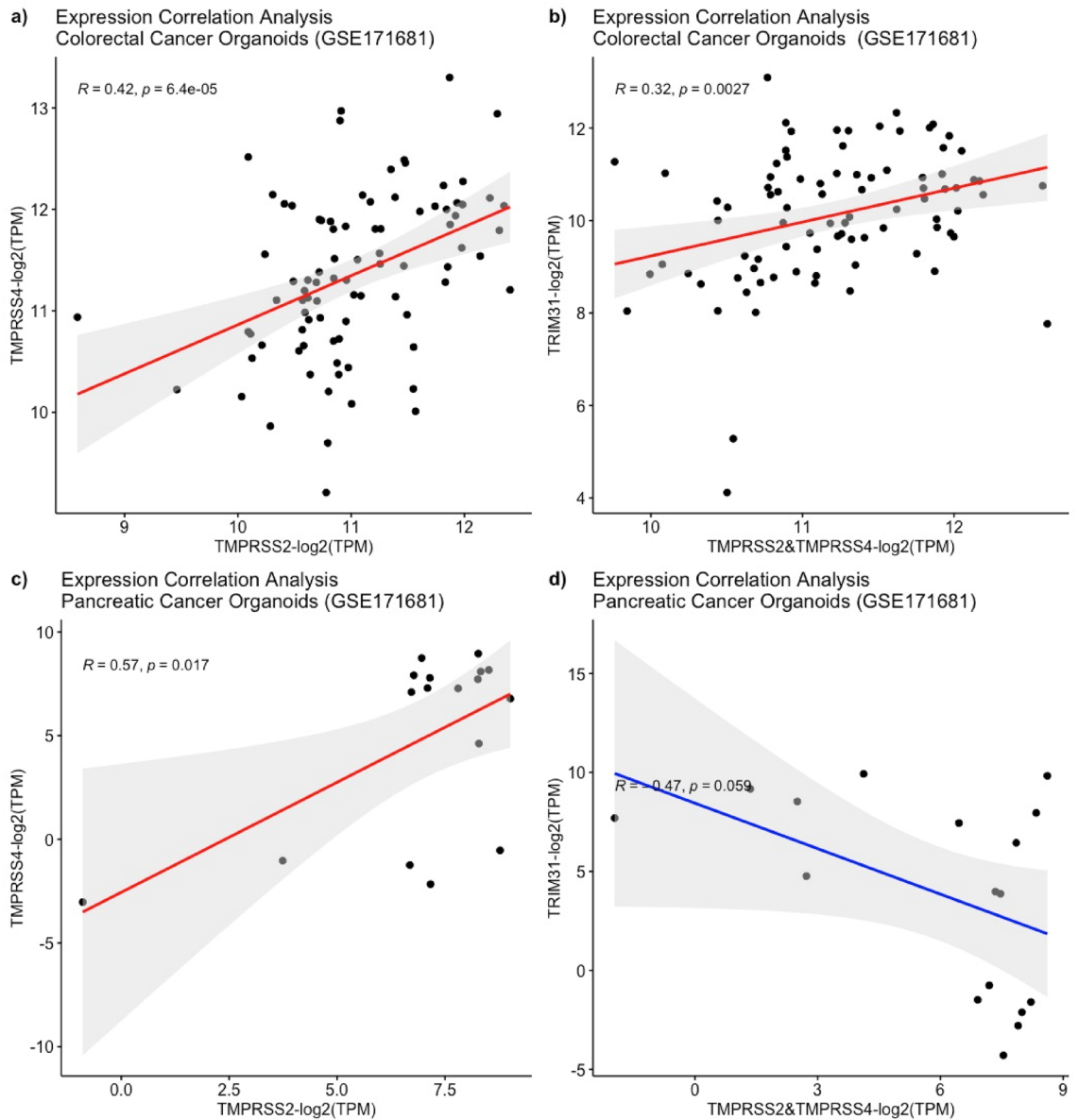


Figure 5. Correlation analysis of the *TMPRSS2*, *TMPRSS4* and *TRIM31* genes across patient derived organoids. Positive correlation between *TMPRSS2*&*TMPRSS4* and *TRIM31* was observed in colorectal cancer organoids ($n = 87$) (a,b). Only *TMPRSS2* shows a significant positive correlation with *TMPRSS4* ($n = 17$) but no significant correlation is observed between the serin protease pair and *TRIM31* in pancreatic cancer organoids (c,d).

genes marked with blue circles. This analysis demonstrated that the loss of *TRIM31* affects viral transcription (GO:0019083) (p -Value = $2.36e-17$) and viral processes (GO:0016032) (p -Value = $5.50e-08$) with a false discovery rate; p -value < 0.05 (Fig. 6b). A list of downregulated genes interacting in 18 nodes are presented (Fig. 6c). Considering most of these genes are involved in viral transcription and viral processes, *TRIM31* is likely to play a role in mediating viral transcription and viral processes in colorectal cancer. Due to the positive correlation of *TRIM31* with *TMPRSS2* and *TMPRSS4* genes, it may therefore cause higher susceptibility of viral infections, such as SARS-CoV-2, seen in cancer patients.

Discussion

With the emergence and spread of SARS-CoV-2 since December 2019, global public health concerns have reached an unprecedented level. The infection of SARS-CoV-2 in gastrointestinal tissues was soon to be discussed by a number of researchers^{26,30}. Tumors in relation to SARS-CoV-2 virus infection and their susceptibility to this viral infection are yet to be fully investigated. Given the high expression levels of ACE2 in the glandular cells of the intestine and the presence of viral nucleocapsid proteins in gastrointestinal epithelial cells and glandular enterocytes^{57,58}, we aimed to understand whether SARS-CoV-2 entry-mediated pathways are involved in gastrointestinal tumors. Our results suggest that there is a positive correlation between gene expression levels of the members of SARS-CoV-2 entry-mediated pathways, namely *TMPRSS2*, *TMPRSS4*, and *TRIM31*, which are

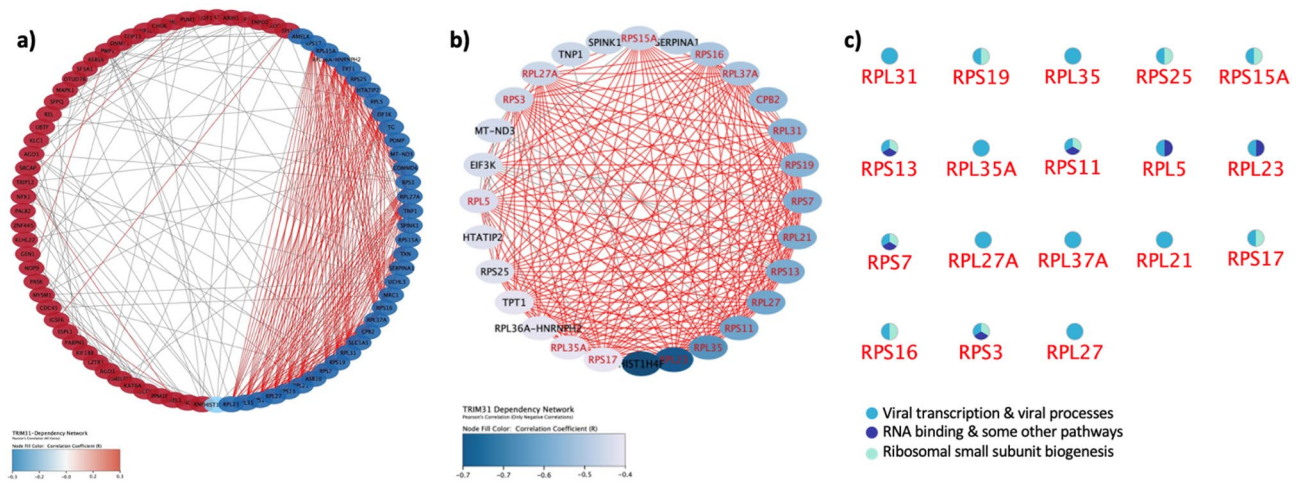


Figure 6. PPI network analysis for affected genes when *TRIM31* was silenced in 20 CRC cell lines using *rnai*. While upregulated genes with red circles exhibited weak co-expression score (less than 0.5), blue circles indicating downregulated genes showed co-expression score bigger than 0.5 (a). Blue circles exhibiting downregulated genes are highlighted with red coloured text exhibited 18 nodes indicating similar processes using gene ontology analysis (b). The 18 nodes indicating cellular processes that mostly affected by *TRIM31* silencing are presented by the names of these proteins and their cellular function such as viral transcription & viral processes, RNA binding, and ribosome small subunit biogenesis are listed (c). Majority of these cellular processes that are affected by *TRIM31* silencing in *crc* cell lines are related to viral transcription & viral processes.

involved in the viral infection of gastrointestinal tumors, cell lines and patient-derived organoids in comparison to their matched normal counterparts.

The pathophysiology of SARS-CoV-2 infection in gastrointestinal tissues may depend on many factors^{59–61}. Not only normal GI tissues but also GI tumors may easily be infected by pathogens due to their distinct mucosal membrane protein expression profiles. We therefore characterized the expression profiles of a number of genes involved in SARS-CoV-2 infection pathway and found higher *ACE2*, *TMPRSS2*, *TMPRSS4*, *SCARB1*, *CTSL* and *NRP1* expression in all GI tumor samples relative to their normal counterparts. Exceptions include *ACE2*, *TMPRSS2*, *TMPRSS4* genes in ESCA samples; *TMPRSS2* gene in STAD and PAAD samples; *CTSL* and *NRP1* genes in COAD and READ samples; and *FURIN* gene in COAD, READ, STAD and PAAD samples. Among the significantly high expression profiles of receptors and mediator proteins for SARS-CoV-2 entry into cells in GI tumors, we observed high co-expression of *TMPRSS2* and *TMPRSS4* in all GI tumors demonstrating that these mucosa specific serine proteases can facilitate the entry of the spike protein into the cell membrane. Importantly, higher co-expression of *TMPRSS2* and *TMPRSS4* relative to normal was also observed in a human gastrointestinal cancer cell line dataset and colorectal cancer organoids. The latter especially holds importance due to its clinically accepted role as a relevant model system.

Given that the co-expression analysis with *TMPRSS2-TMPRSS4* resulted in a positive correlation with *TRIM31* in all GI samples, we further investigated the functional role of *TRIM31* in GI tumors. *TRIM31* is a member of E3 ubiquitin ligase family and mediates various cellular functions in gastrointestinal tissues including antiviral immune response⁶² and tumorigenesis^{63,64}. Despite the antiviral role of mitochondrial antiviral-signaling protein (MAVS), there have been various studies, recently shown in COVID-19 infection, demonstrating that MAVS complex can be limited to facilitate viral infection⁶⁵. Amongst the role of inhibiting antiviral role of MAVS signalling, NEMO-like kinase was reported to induce the degradation of MAVS. In addition, *PLK1* was reported to phosphorylate and destroy MAVS complex⁶⁶. The dual role of MAVS complex can be context specific and especially in cancers, where immune dysregulation is commonly observed⁶⁷, MAVS can be regulated differently than its antiviral function through the activation various phosphorylation events and hence limiting the antiviral function of MAVS complex. Another study also showed that antiviral immune response mediated by MAVS complex is opposed by SARS-CoV-2 nucleocapsid protein (NP), SARS-CoV-2-NP, to facilitate NP-mediated immune evasion in on colon cancer cell line CaCo-2⁶⁸. Given the prominent consequences of COVID-19 resulting in immune deregulation and evasion⁶⁹, antiviral immune response can be impeded and therefore viral infection can be promoted in cancer cells. Furthermore, another study particularly demonstrated that MAVS expression was downregulated in human colon cancer tissue samples⁷⁰. Based on this study, we also checked MAVS expression in CRC samples and found that MAVS expression was significantly downregulated in CRC samples in comparison to normal samples (Supplementary Figure S6). Based on this finding and supporting studies^{65,66} opposing the antiviral function of MAVS signalling, we think that decreased MAVS expression can impede antiviral immune response. In the case of viral infection, attenuated levels of MAVS may result in promoting SARS-CoV-2 viral infection and consequently promoting the viral evasion. As the MAVS expression is downregulated, *TRIM31* may not perform its function as it normally mediates antiviral immune response via its interaction with the MAVS complex. *TRIM31* has been previously reported to play a role in activation of NF- κ B signalling in colorectal cancer⁷¹. In addition, NF- κ B signalling was shown to be activated by viral infection⁷² and

play a role in COVID-19 pathogenicity^{73–75}. In this process, TRIM31 may play a role in inducing the expression of NK- κ B and therefore the activation of inflammation might contribute to the viral processes. To support the role of TRIM31 in regulation viral processes, in Fig. 6, we show that inhibition of TRIM31 in colon cancer cell lines results in downregulation of viral transcription/processes related cellular pathways.

In addition, TMPRSS4 alongside with TMPRSS2 mediate SARS-CoV-2 viral entry in gut epithelial cells⁷⁶ and highly expressed in GI cancer types⁷⁷. These two proteases facilitating viral entry are co-expressed with TRIM31 which can act as promoting viral transcription and viral processes related pathways. To support this notion, we have shown that MAVS normally interacting with TRIM31 to facilitate antiviral immune response is downregulated in CRC samples (Supplementary Figure S6) and therefore TRIM31 may not function as it is normally reported in antiviral immune response. Besides the expression of *TMPRSS2*, *TMPRSS4*, and *TRIM31* in GI cancers, we have analysed the expression of these genes in all cancers, and we have found that the expression of these genes is more upregulated in majority of GI cancers than individual other cancer types (Supplementary Figure S7, Supplementary Figure S8, Supplementary Figure S9). It is important to note that we cannot rule out the possibility that other cancer types might also have increased susceptibility to COVID-19 infection. For example, other cancer types such as Lung Adenocarcinoma (LUAD) or Liver Hepatocellular Carcinoma (LIHC) may also exhibit increased susceptibility to COVID-19 infection via similar mechanisms governed by *TMPRSS2*, *TMPRSS4*, and *TRIM31*, however this remains to be further investigated.

TRIM31 RNAi expression profile obtained from 20 colorectal cancer cell lines using CCLE dataset provided evidence for *TRIM31* dependency in these cell lines. Interaction network analysis performed following the *TRIM31* knockdown in colorectal cancer cell lines highlighted a possible functional role of TRIM31 in viral replication and viral processes. This may therefore be important to develop strategies for targeting TRIM31 particularly in colorectal tumors. Studies showed that protease inhibitors such as against *TMPRSS2*, *CTSL* or other proteases may play a role in tackling SARS-CoV-2 viral entry into cells as these proteases are known to regulate viral infection^{78–80}. For example, camostat-mesylate, a *TMPRSS2* inhibitor, was found to be successful for the inhibition of viral entry into lung epithelial cells²⁷. In the light of our findings where *TRIM31* is positively correlated with *TMPRSS2-TMPRSS4* in GI tumors and knockdown of *TRIM31* mediates viral replication and viral processes, developing inhibitors targeting TRIM31 may decrease viral infection efficiency of SARS-CoV-2 in colorectal cancer patients. Taken together, our findings demonstrate that *TMPRSS2-TMPRSS4-TRIM31* genes are positively correlated together in colorectal cancer which may facilitate SARS-CoV-2 entry into cells. Hence, targeting TRIM31 may play an important role in preventing SARS-CoV-2 viral infection in colorectal cancer patients.

Methods

Data acquisition. No datasets were generated during the current study. The publicly available TCGA datasets were directly used from the TCGA Data Portal at (<https://www.cancer.gov/tcga>) for gastrointestinal solid cancers through Gene Expression Profiling Interactive Analysis portal⁸¹. For normal tissue samples, GTEx database (<https://www.gtexportal.org/home>) was used to identify expression values of selected genes in normal samples corresponding gastrointestinal tumor regions. The Cancer Cell Line Encyclopedia (CCLE) database (<https://depmap.org/portal/ccle>) was used to retrieve processed large-scale NGS expression datasets of cancer cell lines^{82–84}. Raw sequencing data of CCLE for the Ghandi M et al. 2019 publication is available through the Sequence Read Archive under accession number PRJNA523380⁸². Among whole encyclopaedia, 82 colorectal cancer, 38 esophageal cancer, 47 pancreatic cancer and 35 stomach (gastric) cancer cell lines were included in this study based on the CCLE database, the names of which are given in the supplementary file (Table S1). To further investigate the expression correlations, two independent studies and three sets of RNA-seq data from gastrointestinal tumour organoid studies (GSE171680, GSE171681⁵¹ and GSE139184⁵⁰) were used to identify gene expression correlations. Transcriptomic datasets of patient derived organoid models for only colorectal and pancreatic cancer were used since there was no available gastric and esophagus tumour organoid data and processed transcriptomic datasets available at the National Centre for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) under accession numbers: GSE171680, GSE171681 and GSE139184 (<https://www.ncbi.nlm.nih.gov/geo>).

Genes selection and differential expression. The receptor and transmembrane proteins/genes that facilitate the entry of SARS-CoV-2 into cells used in this study were selected through Molecular Signature Database of the Gene Set Enrichment (GSEA-msigdb) repository (<https://www.gsea-msigdb.org/>) query and the gene set in WP_SARSCOV2_AND_COVID19_PATHWAY was used⁸⁵. For comparative and differential expression analyses, the gene list given in Table 1 was used for the analysis of TCGA datasets for solid tumor samples with their control matches and corresponding normal tissue samples in GTEx. Gene Expression Profiling Interactive Analysis portal associated with its python package was used to compare tumor samples with normal tissue⁸¹. The expression data were first $\log_2(\text{TPM} + 1)$ transformed for differential analysis and the $\log_2\text{FC}$ was identified as median difference between tumors and normal in this portal. Genes with higher $|\log_2\text{FC}|$ values were considered differentially expressed genes. Gene expression values of GTEx samples (normal) were produced with RNA-SeqQC v1.1.9 and built-in tool in their websites were used to generate a normal tissue heatmap³³. Tissue-wise differential expression of one gene or a multi-gene signature in gastrointestinal cancer types (colon adenocarcinoma, COAD; esophagus carcinoma, ESCA; pancreatic adenocarcinoma, PAAD; rectal adenocarcinoma, READ; and stomach adenocarcinoma, STAD) was determined. For differential gene expression analysis for solid tumors of TCGA, $\log_2(\text{TPM})$ was used for logscale comparison and $|\log_2\text{FC}|$ cut-off as 1 was applied. Moreover, some of the cell lines were excluded due to zero expression or no expression data available for corresponding genes.

Gene set correlation analysis. A linear regression analysis (Pearson's correlation) was performed to explore the relationship between differentially expressed genes for each dataset. To define the relationship between differentially expressed genes, continuous variables were analyzed using Pearson's correlation analysis. For correlation analyses, the non-log scale was used for calculation and the log-scale axis was used for visualization in TCGA samples, in CCLE samples and in other datasets.

Loss-of-function screens and network analysis. Dependency analysis for cancer cell lines was performed using RNAi dependency dataset (Achilles + DRIVE + Marcotte, DEMETER2). The resulting gene expression changes demonstrating only moderate or positive relationship were further investigated. Protein–protein interaction networks functional enrichment analysis was performed using STRING (<https://string-db.org>) by all types of interactions with highest confidence interaction score (0.900). Disconnected nodes were excluded from visualization.

Statistical analysis. Pearson's correlation was calculated and concluded as strong when R value was larger than 0.7, moderate when R value was between 0.3 and 0.7 and weak when R value was between 0 and 0.3. Data were transformed into \log_2 (TPM) from raw TPM values for pancreatic cancer organoids data and filtered according to related genes by using denoted packages with RStudio. Boxplots were generated using one-way ANOVA via GEPIA2 portal using disease status (Tumor and Normal) as variables for calculating differential expression and p-value < 0.05 was considered as significant for all differential gene expressions and correlation analyses. General data handling, filtering, and plotting was performed using RStudio (<http://www.rstudio.com>) with installed packages such as dplyr, tidyverse and ggpubr^{86–88}.

Data availability

Previously published datasets including GTEx samples⁸⁹, TCGA samples⁹⁰, CCLE samples⁸², pancreatic cancer organoids⁵⁰, and colorectal cancer organoids⁵¹ were used in this study. The RNA-Seq datasets of TCGA/GTEx data are available within the GEPIA2⁸¹ and weblinks to these datasets are as following <https://portal.gdc.cancer.gov>, and <https://gtexportal.org/home>, <http://gepia2.cancer-pku.cn>, respectively. The CCLE provides public access to genomic data, which are openly available at <https://sites.broadinstitute.org/ccle/datasets>. All processed transcriptomic datasets for tumor organoid samples are available at the National Centre for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) under accession numbers of GSE171680, GSE171681 and GSE139184 (<https://www.ncbi.nlm.nih.gov/geo>).

Received: 9 February 2022; Accepted: 30 June 2022

Published online: 15 August 2022

References

1. World Health Organization. Novel Coronavirus—China. <https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON233> (2020).
2. Wu, Y. *et al.* SARS-CoV-2 is an appropriate name for the new coronavirus. *Lancet* **395**, 949–950 (2020).
3. World Health Organization. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. <https://covid19.who.int/>.
4. Pennell, N. A. *et al.* American society of clinical oncology road to recovery report: Learning from the COVID-19 experience to improve clinical research and cancer care. *J. Clin. Oncol.* **39**, 155–169 (2020).
5. Charles, G. & Anderson-Nathe, B. The way ahead past COVID: Worse than before. *Child Youth Serv.* **42**, 1–2 (2021).
6. Casanova, J.-L. & Abel, L. Lethal infectious diseases as inborn errors of immunity: Toward a synthesis of the germ and genetic theories. *Annu. Rev. Pathol.* **16**, 23–50 (2021).
7. World Health Organization. Update on Omicron. <https://www.who.int/news/item/28-11-2021-update-on-omicron>.
8. Fontanet, A. *et al.* SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet* **397**, 952–954 (2021).
9. Korber, B. *et al.* Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827.e19 (2020).
10. Shang, J. *et al.* Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci.* **117**, 11727–11734 (2020).
11. Zhang, Q. *et al.* Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal Transduct. Target. Ther.* **6**, 233 (2021).
12. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
13. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
14. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
15. Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221–224 (2020).
16. 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* **78**, 779–784.e5 (2020).
17. Wei, C. *et al.* SARS-CoV-2 manipulates the SR-B1-mediated HDL uptake pathway for its entry. *bioRxiv* (2020).
18. Cantuti-Castelvetri, L. *et al.* Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **1979**(370), 856–860 (2020).
19. Daly, J. L. *et al.* Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science* **1979**(370), 861–865 (2020).
20. Kyrou, I., Randevara, H. S., Spandidos, D. A. & Karteris, E. Not only ACE2—the quest for additional host cell mediators of SARS-CoV-2 infection: Neuropilin-1 (NRP1) as a novel SARS-CoV-2 host cell entry mediator implicated in COVID-19. *Signal Transduct. Target. Ther.* **6**, 21 (2021).
21. Wei, C. *et al.* HDL-scavenger receptor B type 1 facilitates SARS-CoV-2 entry. *Nat. Metab.* **2**, 1391–1400 (2020).
22. Zang, R. *et al.* TMPRSS2 and TMPRSS4 mediate SARS-CoV-2 infection of human small intestinal enterocytes. *BioRxiv* <https://doi.org/10.1101/2020.04.21.054015> (2020).
23. Wruck, W. & Adjaye, J. SARS-CoV-2 receptor ACE2 is co-expressed with genes related to transmembrane serine proteases, viral entry, immunity and cellular stress. *Sci. Rep.* **10**, 21415 (2020).
24. Gupta, A. *et al.* Extrapulmonary manifestations of COVID-19. *Nat. Med.* **26**, 1017–1032 (2020).
25. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).

26. Guan, W. *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
27. Hoffmann, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280 (2020).
28. Li, F., Li, W., Farzan, M. & Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **1979**(309), 1864–1868 (2005).
29. Yang, L. & Tu, L. Implications of gastrointestinal manifestations of COVID-19. *Lancet Gastroenterol. Hepatol.* **5**, 629–630 (2020).
30. Xiao, F. *et al.* Evidence for gastrointestinal infection of SARS-CoV-2. *Gastroenterology* **158**, 1831–1833 (2020).
31. Lamers, M. M. *et al.* SARS-CoV-2 productively infects human gut enterocytes. *Science* **1979**(369), 50–54 (2020).
32. Ding, S. & Liang, T. J. Is SARS-CoV-2 also an enteric pathogen with potential fecal-oral transmission? A COVID-19 virological and clinical review. *Gastroenterology* **159**, 53–61 (2020).
33. Zhang, H. *et al.* The digestive system is a potential route of 2019-nCoV infection: a bioinformatics analysis based on single-cell transcriptomes. *BioRxiv* (2020).
34. Xu, H. *et al.* High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *Int J Oral Sci* **12**, 8 (2020).
35. Groff, A. *et al.* Gastrointestinal manifestations of COVID-19: A review of what we know. *Ochsner J.* **21**, 177–180 (2021).
36. Darvish-Damavandi, M. *et al.* An analysis of SARS-CoV-2 cell entry genes identifies the intestine and colorectal cancer as susceptible tissues. *Br. J. Surg.* **107**, e452–e454 (2020).
37. Wang, W. *et al.* Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA* **323**, 1843–1844 (2020).
38. Wu, Y. *et al.* Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *Lancet Gastroenterol. Hepatol.* **5**, 434–435 (2020).
39. Kumar, A. *et al.* Relevance of SARS-CoV-2 related factors ACE2 and TMPRSS2 expressions in gastrointestinal tissue with pathogenesis of digestive symptoms, diabetes-associated mortality, and disease recurrence in COVID-19 patients. *Med. Hypotheses* **144**, 110271 (2020).
40. Shi, Y. *et al.* Host susceptibility to severe COVID-19 and establishment of a host risk score: Findings of 487 cases outside Wuhan. *Crit. Care* **24**, 1–4 (2020).
41. Liang, W. *et al.* Cancer patients in SARS-CoV-2 infection: A nationwide analysis in China. *Lancet Oncol.* **21**, 335–337 (2020).
42. Edge, R. *et al.* Cancer care disruption and reorganisation during the COVID-19 pandemic in Australia: A patient, carer and healthcare worker perspective. *PLoS ONE* **16**, e0257420 (2021).
43. Ng, W. H. *et al.* Comorbidities in SARS-CoV-2 patients: a systematic review and meta-analysis. *MBio* **12**, e03647–e3720 (2021).
44. Knisely, A. *et al.* Coronavirus spectrum infections (COVID-19, MERS, SARS) in cancer patients: A systematic review of the literature. *Cancer Invest.* **38**, 436–444 (2020).
45. Dai, M. *et al.* Patients with cancer appear more vulnerable to SARS-CoV-2: A multicenter study during the COVID-19 outbreak. *Cancer Discov.* **10**, 783–791 (2020).
46. Kuderer, N. M. & Lyman, G. H. COVID-19, cancer, and consequences: Where are we now?. *Cancer Invest.* **38**, 431–435 (2020).
47. Aznab, M. Evaluation of COVID 19 infection in 279 cancer patients treated during a 90-day period in 2020 pandemic. *Int. J. Clin. Oncol.* **25**, 2 (2020).
48. Wang, B. & Huang, Y. Which type of cancer patients are more susceptible to the SARS-COV-2: Evidence from a meta-analysis and bioinformatics analysis. *Crit. Rev. Oncol./Hematol.* **153**, 2 (2020).
49. Antikchi, M. H. *et al.* The risk and prevalence of COVID-19 infection in colorectal cancer patients: A systematic review and meta-analysis. *J. Gastrointest. Cancer* **52**, 2 (2021).
50. Osuna de la Peña, D. *et al.* Bioengineered 3D models of human pancreatic cancer recapitulate in vivo tumour biology. *Nat. Commun.* **12**, 5623 (2021).
51. Cho, E. J. *et al.* Immuno-genomic classification of colorectal cancer organoids reveals cancer cells with intrinsic immunogenic properties associated with patient survival. *J. Exp. Clin. Cancer Res.* **40**, 230 (2021).
52. Tsherniak, A. *et al.* Defining a cancer dependency map. *Cell* **170**, 564–576 (2017).
53. Marcotte, R. *et al.* Functional genomic landscape of human breast cancer drivers, vulnerabilities, and resistance. *Cell* **164**, 293–309 (2016).
54. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci.* **105**, 20380–20385 (2008).
55. Cowley, G. S. *et al.* Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. *Sci. Data* **1**, 1–12 (2014).
56. McDonald, E. R. III. *et al.* Project DRIVE: A compendium of cancer dependencies and synthetic lethal relationships uncovered by large-scale, deep RNAi screening. *Cell* **170**, 577–592 (2017).
57. Maa, L. M. *et al.* SARS-CoV-2 productively infects human gut enterocytes. *Science* **369**, 50–54 (2020).
58. Zhang, H. *et al.* Digestive system is a potential route of COVID-19: an analysis of single-cell coexpression pattern of key proteins in viral entry process. *Gut* **69**, 1010–1018 (2020).
59. Zang, R. *et al.* TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **5**, 3582 (2020).
60. Ruochen, Z. *et al.* TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **5**, 3582 (2020).
61. Mahmoud, I. S. & Jarrar, Y. B. Targeting the intestinal TMPRSS2 protease to prevent SARS-CoV-2 entry into enterocytes—prospects and challenges. *Mol. Biol. Rep.* **48**, 4667–4675 (2021).
62. Song, H. *et al.* The E3 ubiquitin ligase TRIM31 attenuates NLRP3 inflammasome activation by promoting proteasomal degradation of NLRP3. *Nat. Commun.* **7**, 2 (2016).
63. Wang, H., Yao, L., Gong, Y. & Zhang, B. TRIM31 regulates chronic inflammation via NF- κ B signal pathway to promote invasion and metastasis in colorectal cancer. *Am. J. Transl. Res.* **10**, 1247–1259 (2018).
64. Venuto, S. & Merla, G. E3 ubiquitin ligase TRIM proteins. *Cell Cycle Mitosis. Cells* **8**, 2 (2019).
65. Li, S. Z. *et al.* Phosphorylation of MAVS/VISA by Nemo-like kinase (NLK) for degradation regulates the antiviral innate immune response. *Nat. Commun.* **10**, 2 (2019).
66. Vitour, D. *et al.* Polo-like kinase 1 (PLK1) regulates interferon (IFN) induction by MAVS. *J. Biol. Chem.* **284**, 2 (2009).
67. Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 2 (2010).
68. Wang, S. *et al.* Targeting liquid–liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity. *Nat. Cell Biol.* **23**, 52 (2021).
69. Taefehshokr, N., Taefehshokr, S., Hemmat, N. & Heit, B. Covid-19: Perspectives on innate immune evasion. *Front. Immunol.* **11**, 48 (2020).
70. Zhang, W. *et al.* The mitochondrial protein MAVS stabilizes p53 to suppress tumorigenesis. *Cell Rep.* **30**, 97 (2020).
71. Wang, H., Yao, L., Gong, Y. & Zhang, B. Trim31 regulates chronic inflammation via nf-kb signal pathway to promote invasion and metastasis in colorectal cancer. *Am. J. Transl. Res.* **10**, 151 (2018).
72. Santoro, M. G. NEW EMBO MEMBER'S REVIEW: NF-kappaB and virus infection: who controls whom. *EMBO J.* **22**, 6 (2003).
73. Kandasamy, M. NF- κ B signalling as a pharmacological target in COVID-19: potential roles for IKK β inhibitors. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **394**, 85 (2021).
74. Robles, J. P. *et al.* The spike protein of SARS-CoV-2 induces endothelial inflammation through integrin α 5 β 1 and NF- κ B signaling. *J. Biol. Chem.* **298**, 95 (2022).
75. Calado, M. B., da Silva Santana, C. E. & Crovella, S. Do inflammasome impact COVID-19 severity?. *Virus Dis.* **32**, 68 (2021).

76. Zang, R. *et al.* TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **5**, 25 (2020).
77. Katopodis, P. *et al.* COVID-19 and SARS-CoV-2 host cell entry mediators: Expression profiling of TMRSS4 in health and disease. *Int. J. Mol. Med.* **47**, 68 (2021).
78. Iketani, S. *et al.* Lead compounds for the development of SARS-CoV-2 3CL protease inhibitors. *Nat. Commun.* **12**, 37 (2021).
79. Coelho, C., Gallo, G., Campos, C. B., Hardy, L. & Wurtele, M. Biochemical screening for SARS-CoV-2 main protease inhibitors. *PLoS ONE* **15**, 5 (2020).
80. Banerjee, R., Perera, L. & Tillekeratne, L. M. V. Potential SARS-CoV-2 main protease inhibitors. *Drug Discov. Today* **26**, 2 (2021).
81. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560 (2019).
82. Ghandi, M. *et al.* Next-generation characterization of the cancer cell line encyclopedia. *Nature* **569**, 503–508 (2019).
83. Stransky, N. *et al.* Pharmacogenomic agreement between two cancer cell line data sets. *Nature* **528**, 84–87 (2015).
84. Barretina, J. *et al.* The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
85. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
86. RStudio | Open source & professional software for data science teams. (2021).
87. dplyr: A Grammar of Data Manipulation. (2021).
88. Wickham, H. *et al.* Welcome to the tidyverse. *J. Open Sour. Softw.* **4**, 1686 (2019).
89. Lonsdale, J. *et al.* The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
90. Network, C. G. A. R. *et al.* The cancer genome Atlas pan-cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).

Acknowledgements

We thank The Scientific and Technological Research Council of Turkey (TUBITAK) (Grant No. 118C197). Ahmet Acar is currently supported by the International Fellowship for Outstanding Researchers Program administrated by TUBITAK.

Author contributions

M.A.T. performed data analysis. A.A. and M.A.T. interpreted the results. A.A. and M.A.T. wrote the manuscript. All authors contributed to manuscript preparation and editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-15911-2>.

Correspondence and requests for materials should be addressed to A.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022