



OPEN

## An RNA-immunoprecipitation via CRISPR/dCas13 reveals an interaction between the SARS-CoV-2 5'UTR RNA and the process of human lipid metabolism

Yurika Shimizu<sup>1,2</sup>, Srinivas Bandaru<sup>1,6</sup>, Mari Hara<sup>1</sup>, Sonny Young<sup>7</sup>, Toshikazu Sano<sup>3</sup>, Kaya Usami<sup>4</sup>, Yuta Kurano<sup>5</sup>, Suni Lee<sup>1</sup>, Naoko Kumagai-Takei<sup>1</sup>, Shogo Takashiba<sup>2</sup>, Shunji Sano<sup>3</sup> & Tatsuo Ito<sup>1</sup>✉

We herein elucidate the function of SARS-CoV-2 derived 5'UTR in the human cells. 5'UTR bound host cellular RNAs were immunoprecipitated by gRNA-dCas13 (targeting luciferase RNA fused to SARS-CoV-2 5'UTR) in HEK293T and A549 cells. The 5'UTR bound RNA extractions were predominantly enriched for regulating lipid metabolism. Overexpression of SARS-CoV-2 5'UTR RNA altered the expression of factors involved in the process of the human Mevalonate pathway. In addition, we found that HMG-CoA reductase inhibitors were shown to suppress SARS-CoV-2 5'UTR-mediated translation activities. In conclusion, we deduce the array of host RNAs interacting with SARS-CoV-2 5'UTR that drives SARS-CoV-2 translation and influences host metabolic pathways.

The outbreak of a novel strain of coronavirus—SARS-CoV-2 causing severe respiratory illness designated as COVID-19, surged as a large-scale global pandemic claiming more than 1 million lives while requiring intensive care hospitalizations for 2–10% of the infected cases worldwide<sup>1</sup>. Accumulated data corroborated since a year of the pandemic has shown COVID-19 patients to present diverse symptoms with severe immune dysregulation of unknown incentives<sup>2–5</sup>.

It is noteworthy that a recent meta-analysis establishes a positive link between COVID-19 vulnerability with obesity<sup>6–10</sup>. Hospital-based studies showed that people with obesity were 113% more likely to contract SARS-CoV-2 and 74% of them required intensive care admissions<sup>11</sup>. Furthermore, patients in the excessive BMI with metabolic-related fatty liver disease demonstrated an enhanced risk of severe COVID-19 disease<sup>12</sup>. These studies explicitly indicate obesity as a potential comorbid condition for COVID-19 affecting the overall course of the disease by influencing viral load, tissue damage, and mortality.

Although the rate of mutation in SARS-CoV-2 is slow (two single nucleotide substitutions per month)<sup>13</sup>, nevertheless, more than 12,000 mutations have been documented so far ([www.gisaid.org](http://www.gisaid.org)). Among these, only a few of them result in the virus's protein variations which make it competent for human-to-human transmission and was responsible for severe illness, hospitalization and mortality rate especially associated with Beta and Delta variant<sup>4,5</sup>.

Viruses like SARS-CoV-2 exploit the complexity of eukaryotic initiation to gain access to the host machinery for protein production<sup>14</sup>. The genome of SARS-CoV-2 consists of approximately 30 kb of linear, non-segmented,

<sup>1</sup>Department of Hygiene, Kawasaki Medical School, 577 Matsushima, Kurashiki, Okayama 701-0192, Japan. <sup>2</sup>Department of Pathophysiology - Periodontal Science, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Okayama 700-8525, Japan. <sup>3</sup>Department of Surgery, Division of Pediatric Cardiothoracic Surgery, University of California San Francisco, San Francisco, CA, USA. <sup>4</sup>Okayama University Medical School, Okayama 700-8558, Japan. <sup>5</sup>Kawasaki Medical School, Kurashiki, Okayama 701-0192, Japan. <sup>6</sup>Koneru Lakshmaiah Educational Foundation, Green Fields, Vaddeswaram, Andhra Pradesh 522302, India. <sup>7</sup>Stanford University, Stanford, CA 94305, USA. ✉email: [tataito@med.kawasaki-m.ac.jp](mailto:tataito@med.kawasaki-m.ac.jp)

5'-capped, polycistronic, polyadenylated RNA molecules, comprising a total of 14 open reading frames (ORFs)<sup>15</sup>. SARS-CoV-2 genomic RNA is capped at the 5' end, and translation initiation efficiency is regulated by a stable hairpin structure near the cap structure<sup>16,17</sup>. The 5' UTR of many RNA viruses has a unique stem-loop structure and is known to recruit important host translation regulatory factors including proteins and RNA, but studies on the physical binding of such factors to the SARS-CoV-2 RNA remains to be elucidated. Understanding the interaction between viral RNA and host factors will be useful in seeking to develop effective therapeutic targets for this disease<sup>18,19</sup>.

In the present study, employing CRISPR/dCas13 aided targeting and immunoprecipitation of SARS-CoV-2 5'UTR bound host RNAs, we herein provide the possible model of regulation of SARS-CoV-2 in human cells and the way it affects the host cellular metabolism.

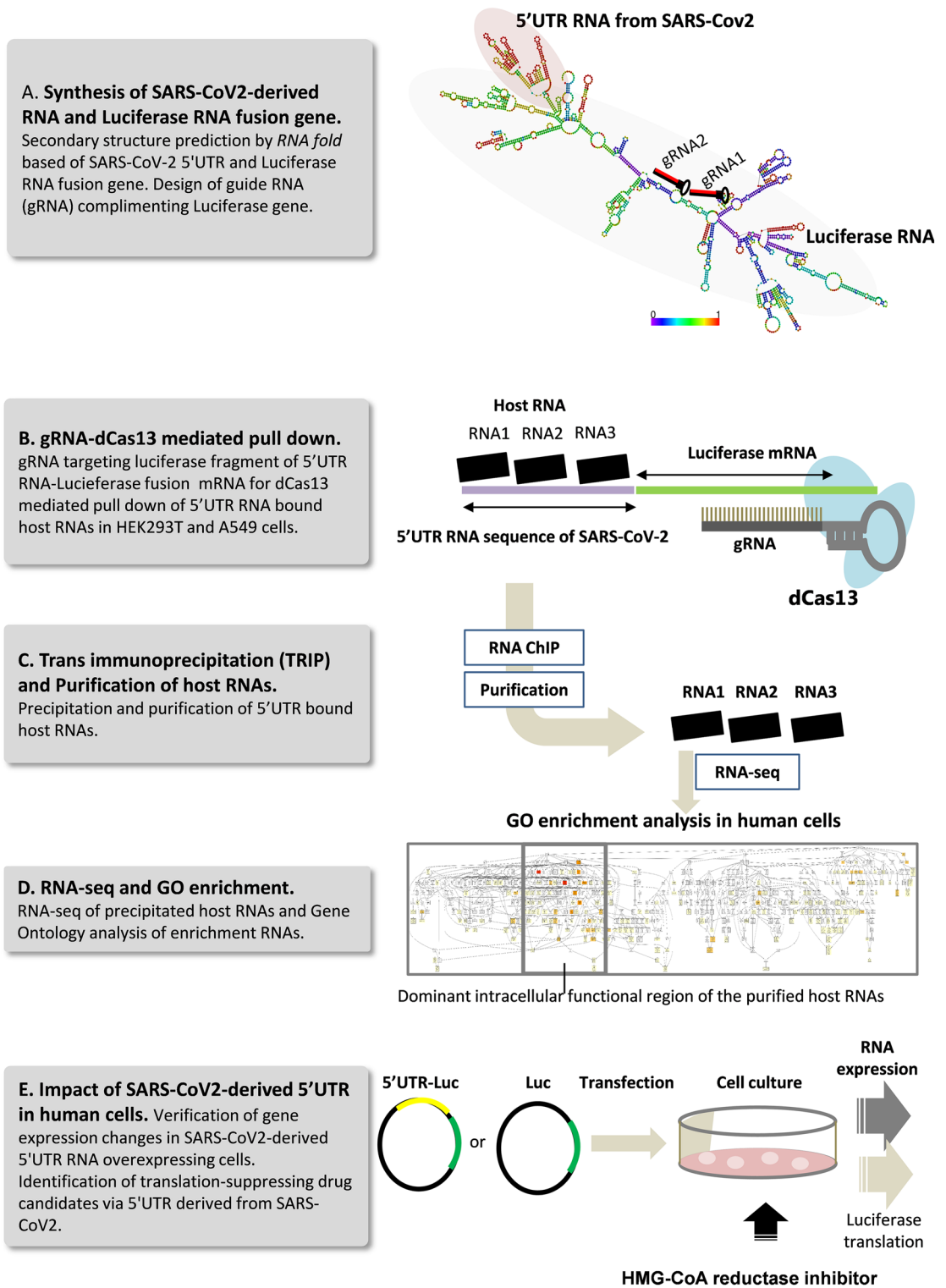
## Results

**Selection of high prevalent 5'UTR region of SARS-CoV-2.** In order to understand the molecular influence of SARS-CoV-2 in human cells, we introduced the 5'UTR leader sequence of SARS-CoV-2 into the host cell (A549 and HEK293T cells). This region is important to facilitate the understanding of (1) the array of host RNAs that would interact with 5'UTR leader in regulating translation of the viral genome and (2) the host cell types that are more susceptible for SARS-CoV-2 infection and replication. However, considering the fact that hypermutation in 5'UTR regions is typical to viral genomes it was essential to use only those 5'UTRs in the present experiment which have large geographic expansions i.e. prevalent across the diverse ethnicities and conserved since the first outbreak.

Numerous reports have been made available on sequence variation and phylogenetics of SARS-CoV-2, and have been constantly updated on <https://nextstrain.org/sars-cov-2/>. The data snapshot containing a quality filtered subset of 3,434 available data points from the site is provided in Supplementary Fig. 1. Among the 264nt sequences in the 5'UTR region of SARS-CoV-2, the two derived from the 210 G>T and 241C>T point mutations were proven to be the most prevalent in almost all the infected cases worldwide. Interestingly, the 241 T mutant was also detected at the first outbreak in Wuhan China in 2019, while other 5'UTR mutations, although been reported, did not lead to clade formations. It also deviates from the binding segment of the virus-derived Nsp1 protein, presumably it benefits viral survival and replication in host cells. Given this literature, we hence selected 5'UTR sequence for our study and henceforth pursued to identify the cluster of host RNAs that interact with this 5'UTR that aids translation and regulation of SARS-CoV-2 genome in human cells. The complete workflow of this study is demonstrated in Fig. 1.

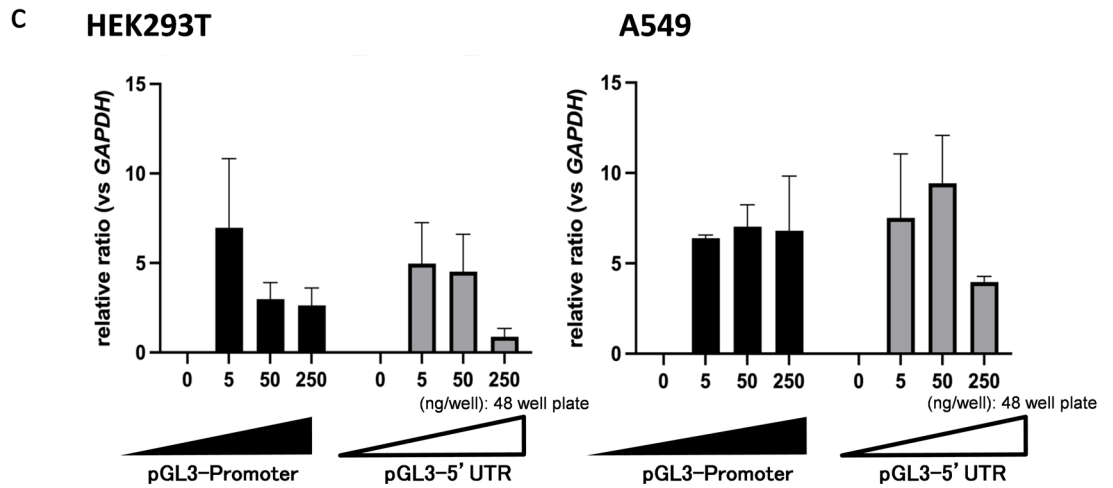
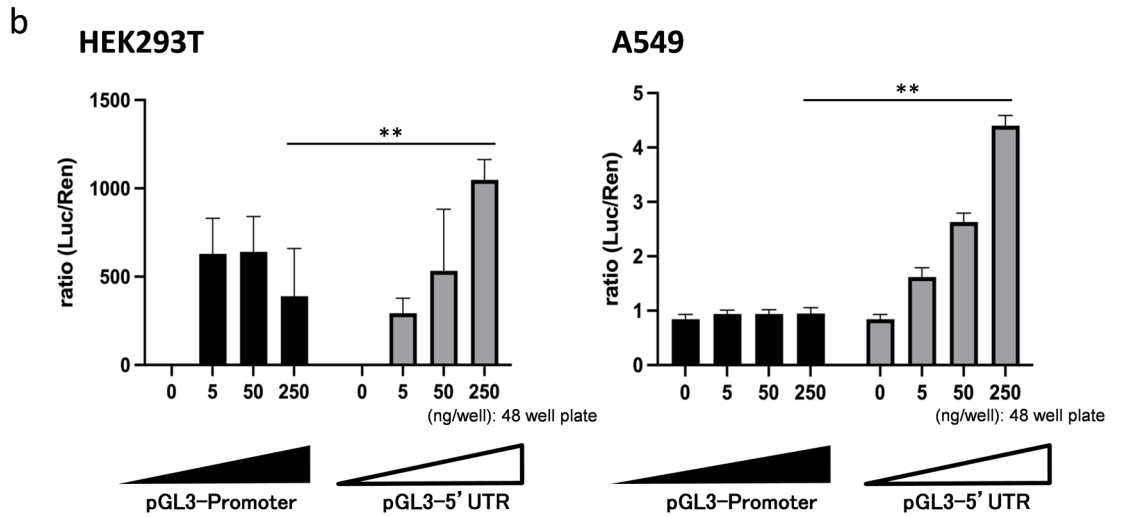
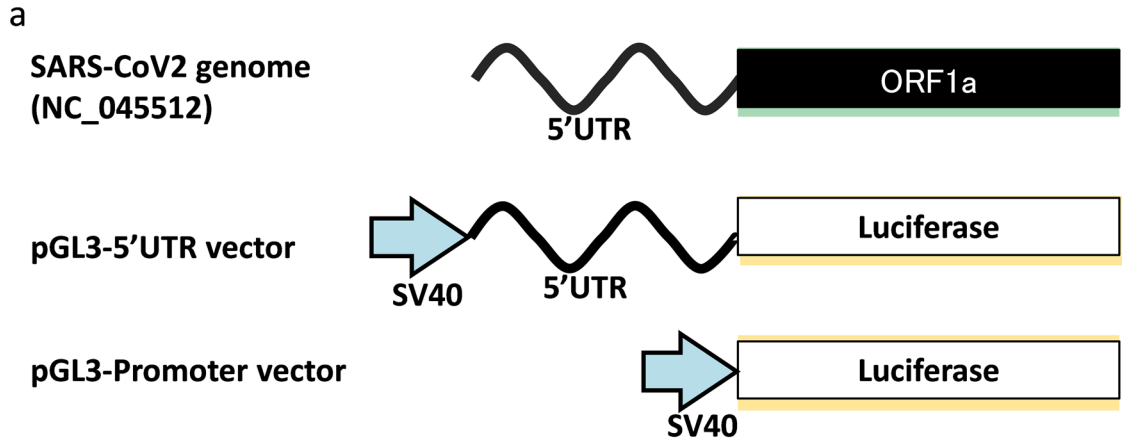
**Translation activation by 5'UTR RNA derived from SARS-CoV-2.** The sequence of the 5'UTR region of SARS-CoV-2 (isolated from the Wuhan-Hu-1 genome (NC\_045512)) was oligo-synthesized. The 5'UTR sequence of SARS-CoV-2 (Wuhan-Hu-1, complete genome, Sequence ID: NC\_045512.2) consists of the following 264 nucleotides, including the Kozak sequence: "AUUAAAGUUUAUACCUUCCCAGGUAA CAAACCAACCAACUUUCGAUCUCUUGUAGAUCUGUUCUCUAAACGAACUUUAAAAUCUGUGUG GCUGUCACUCGGCUGCAUGCUUAGUGCACUCACGCAGUAUAAUUAUAAACUAAUUAACUGUCGU UGACAGGACACGAGUAAUCGUCUACUUCUGCAGGUCGUCUACGGUUUCGUCCGUGUUGCAG CCGAUCACAGCAUCUAGGUUUUCGUCCGGUGUGACCCGAAAGGUAAGCCACC". The synthesized sequence was cloned in Luciferase reporter pGL3-Promoter Vector (Promega: E1761) specifically between SV40 promoter sequence and luciferase gene sequence to express the 5'UTR-fused-luciferase RNA (Fig. 2a). This bonafide vector was transfected into human cell lines HEK293T and A549 with a concentration gradient and the expression was measured as a function of luminescence from the translation of luciferase protein. In the present study, HEK293T cells were used as reference cells for being easy to transfect with the pGL3 vector and suitable for overexpressing 5'UTR RNA of SARS-CoV-2. Interestingly, luciferase RNA fused to SARS-CoV-2-derived 5'UTR showed stronger luciferase activity in HEK293T cells and A549 cells in contrast to cells expressing luciferase RNA alone (Fig. 2b). Intriguingly, there is no discernible difference between the luciferase transcripts produced by the SARS-CoV-2 5'UTR and the pGL3 luciferase vectors (in both A549 and HEK-293 T cells), suggesting that the expression of the luciferase transcript is not dependent on the presence of the 5'UTR in the pGL3 vector. This finding suggests that SARS-CoV-2 5'UTR is solely responsible to enhance translational efficiency by serving as a motif for binding of host translational factors to code viral proteins in the human cells (Fig. 2c). Further, neither the introduction of the pGL3-promoter vector nor the pGL3-5'UTR vector affected cell viability (Supplemental Fig. 2).

**gRNA-dCas13 assisted Trans RNA Immunoprecipitation (TRIP) of host regulatory RNA bound to 5'UTR of SARS-CoV-2.** Following our investigation which showed the significant roles of 5'UTR in enhancing translational efficiency, our subsequent pursuit was to determine the cluster of host cellular RNAs that regulate SARS-CoV-2 genome by interacting to 5'UTR sequence. To achieve this, we first predicted the structure of 5'UTR fused luciferase RNA and confirmed that 5'UTR RNA formed a secondary structure independent of luciferase RNA. Next, in order to pull down host RNAs bound to 5' UTR (by TRIP), we targeted luciferase fraction of 5'UTR-luciferase fused RNA by gRNA-dCas13. We designed two candidate single-strand regions for designing guide RNAs (gRNA) in the luciferase RNA. The gRNAs targeting these regions were named Luc1 and Luc2, respectively (Fig. 3a). The dCas13 gene sequence was inserted into the AAVS1 gene region of HEK293T cells and A549 cells by homologous recombination, followed by the selection of vector inserted cells with 10ug/ml puromycin for 24 h incubation. Next, lentivirus vectors expressing RFP and Luc1-gRNA or Luc2-gRNA were inserted in each cell, and the cells expressing RFP were collected by cell sorting. In the following experiment, we introduced pGL3-5'UTR vector or pGL3-promoter vector into cell lines expressing dCas13 and Luc1-gRNA



**Figure 1.** Workflow of the study.

or Luc2-gRNA, and performed RNA Immunoprecipitation (RIP) using AM-Tag antibody (pAb) (Active Motif: 61678). With qPCR we confirmed that the immunoprecipitated RNA libraries contained transcribed SARS-CoV-2-derived 5'UTR RNA and/or luciferase RNA. SARS-CoV-2-derived 5'UTR RNA and luciferase RNAs were not detected in cells those expressing non-targeting gRNAs (gRNAs not target-specific for luciferase RNA) (Fig. 3b). Hence with gRNA-dCas13 assisted TRIP method we extracted 5' UTR bound host RNA clusters that are crucial in the translation of SARS-CoV-2 genome.



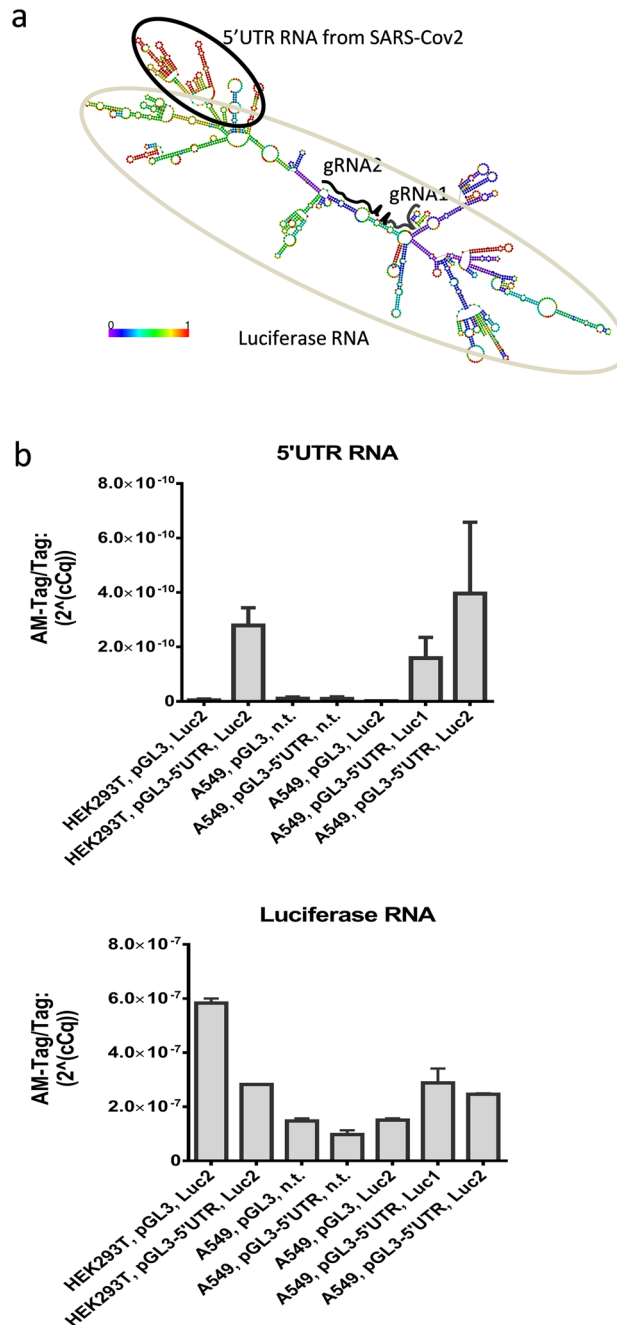
◀**Figure 2.** pGL3-5'UTR construct and Luciferase reporter assay. (a) Vector construction with SARS-CoV-2 5'UTR insert in pGL3 luciferase vector. The genomic architecture of SARS-CoV-2 starts with 5' UTR followed by ORF1 and structural genes. To elucidate the translational mechanism in host cells, 5' UTR segment of the SARS-CoV-2 was inserted in the pGL3 vector between SV40 promoter and luciferase gene. Luciferase gene is a reporter for the transfection efficiency and expression of 5' UTR. Further, luciferase RNA forms the target for crRNA-dCAS13 facilitating TRIP for 5'UTR bound host translational RNAs. (b) Luciferase protein expression was measured as a luminescence activity of luciferase protein. Upon introducing the pGL3-5'UTR vector, relative luciferase activity increased more than ten times to the equivalent amount of pGL3-promoter. Note that the SARS-CoV-2-derived 5'UTR synergistically activated translation in an intracellular factor-dependent manner. The relative luciferase activity was calculated by dividing the activity of Renilla reniformis luciferase as internal control. The mean and standard error of the three independent experiments is shown. The Mann-Whitney U test was used for statistical significance ( $p < 0.05$ ). (c) The expression of luciferase RNA in both SARS-CoV-2 5'UTR and pGL3 luciferase vectors brought about by the SV40 promoter was measured by qPCR. Note that The luciferase transcript expressed from SARS-CoV-2 5'UTR and pGL3 luciferase vectors (in both A549 and HEK-293 T cells) does not quantitatively differ significantly, implying that the expression of luciferase transcript is independent of the inclusion of 5'UTR in the pGL3vector. The expression results were expressed as fold relative to the housekeeping internal control GAPDH gene expression. The degradation of the RNA progressed when 250 ng of plasmid was introduced.

**5'UTR bound host regulatory RNA predominantly contains genes for lipid metabolism.** Host regulatory RNAs bound to 5'UTR extracted from TRIP were further subjected to ultra-low input RNA sequencing. In the cluster analysis of differentially enriched genes, we found that the RNA group that interact with SARS-CoV-2 5'UTR RNA in A549 cells belong to a different cluster for the RNA group that binds to Luciferase RNA (Fig. 4a). Similar enrichment differences between RNA group that binds to the 5'UTR RNA of SARS-CoV-2 and the control group that binds only to the RNA of luciferase could be confirmed in HEK293T cells as well (We confirmed the duplication of genes that are differentially enriched in each group compared to the control group or genes that are differentially enriched in multiple groups). We found that more than 1000 RNA genes that bind to the 5'UTR of SARS-CoV-2 are common to A549 cells and HEK293T cells (Fig. 4b). Further, we performed Gene Ontology enrichment analysis of RNA genes that bind predominantly ( $p$ -value  $< 0.01$ ) to 5'UTR of SARS-CoV-2 extracted from HEK293T cells and A549. The top 5 GO terms identified in the analysis were (a) lipid metabolic process, (b) cellular lipid metabolic process, (c) small molecule metabolic process, (d) organic acid metabolic process and (e) oxoacid metabolic process (Fig. 4c and Supplement Table 1).

Next, we performed peak calling followed by motif analysis in order to find the sequence motifs in the 5'UTR which binds to host proteins (Supplement Fig. 3 and Supplementary Table 2). From the concentrated RNA motif, respectively 12 and 14 types of RNA-binding proteins were extracted from HEK293T and A549 cells. The proteins commonly observed in both cells were RBM38, BRUNOL4, BLUNOL5, and TARDBP. Although indirectly, these proteins have been shown to be involved in lipid metabolism and fat biosynthesis<sup>20–24</sup>. However to establish the specific role of these proteins in connection to 5'UTR of SARS-Cov2 RNA in regulating lipid metabolism still remains elusive and requires additional targeted metabolic investigations, nevertheless, this observation suggests that SARS-CoV-2 5'UTR may alter the lipid metabolic process in the host. Therefore from the given analysis it is suggestive that SARS-CoV-2 virus uses 5'UTR to interact with various protein and regulatory RNAs to sneak into the human intracellular environment. It may be thought of as "the takeover of RNA metabolism" mechanism in the human cell by the SARS-CoV-2 virus.

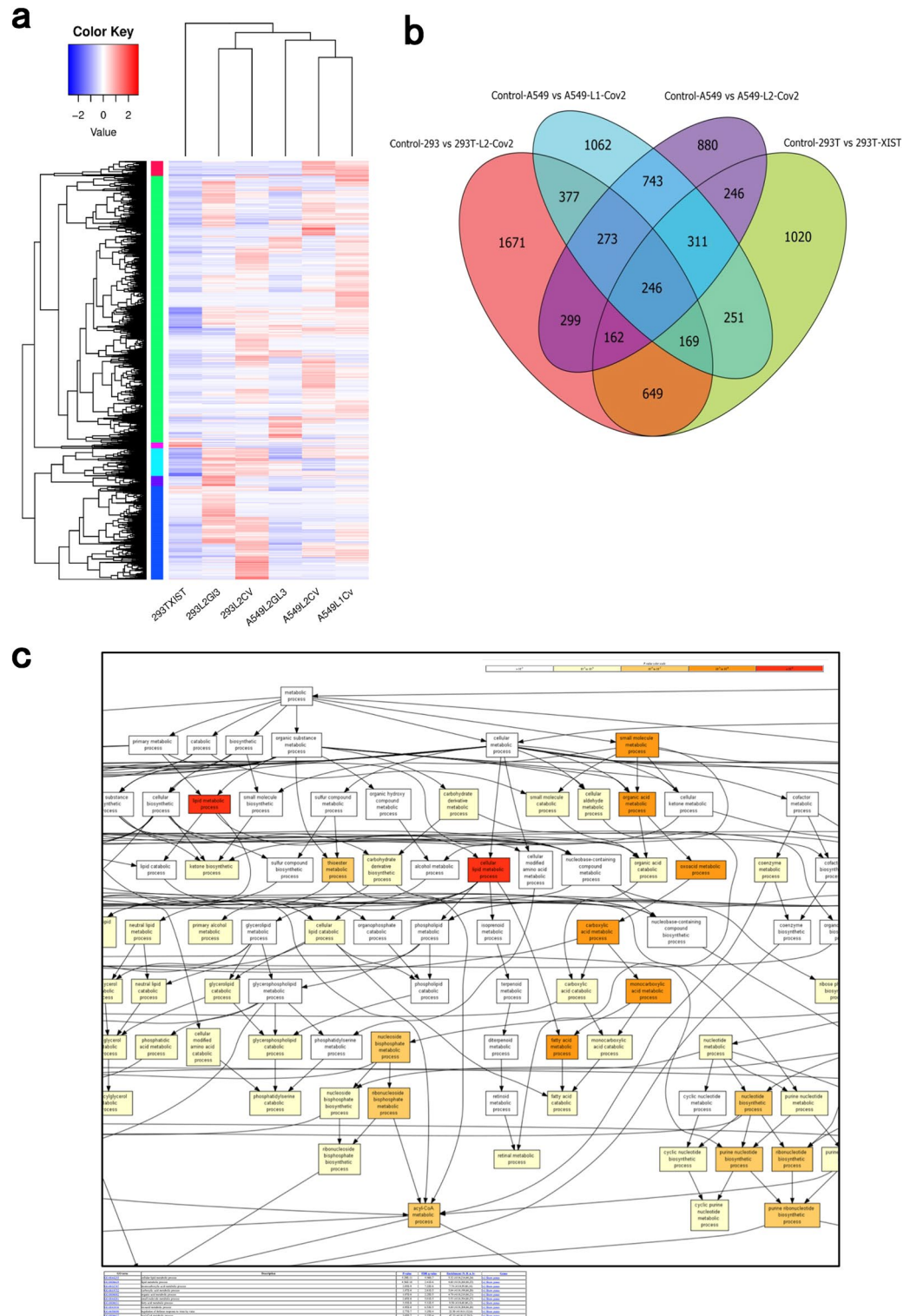
*5'UTR-mediated translation influences Mevalonate pathway.* Following our previous investigation, which revealed that RNAs interacting with the 5'UTR are enriched in lipid metabolism, we investigated whether factors involved in lipid metabolism are altered when SARS-CoV-2-derived 5'UTR RNAs are abundantly expressed in human cells. Therefore, we introduced pGL3-5'UTR or pGL3-Promoter vectors into HEK293T cells and A549 cells, and quantified the RNAs. Surprisingly, significant changes in the mRNA expression level were observed for ACAA2, HMGCS, and FADS1/2 genes which are enriched in the Mevalonate and Polyunsaturated fatty acid biosynthesis pathway, and remarkably, HMGCS expression was highly upregulated at HEK293T cell (Fig. 5a). The Mevalonate pathway was predominant in HEK293T cells, whereas the Polyunsaturated fatty acid biosynthesis pathway was predominant in A549 cells (Fig. 5a). Since we had a set of enriched genes that bind to 5'UTR of SARS-CoV-2 in human cells, we further proceeded for drug repositioning by target identification. We used TargetMine program (<https://targetmine.mizuguchilab.org>) which aided the identification of suitable targets. As a result, four interacting compounds (Atorvastatin, Simvastatin, Rosuvastatin and Pravastatin) that can be expected to act ( $p$ -value  $< 0.05$ ) on the enriched RNA were identified. Amongst all, statins ( $p$ -value  $< 0.01$ ) have been identified as optimal and ready-to-use compounds to treat antiviral growth in COVID-19 patients (Fig. 5b). To investigate, whether suppression of translational activity by 5'UTR of SARS-CoV-2 can be introduced, we added over-the-counter drug HMG-CoA reductase inhibitors after overexpressing the pGL3-5'UTR or pGL3-Promoter vector into HEK293T cells and A549 cells. It was interesting to observe that 5'UTR SARS-CoV-2-mediated translational activity in both HEK293T cells and A549 cells were significantly suppressed upon treatment with 100 nM Atorvastatin (Fig. 5c). In addition in HEK293T cells, a translation inhibitory effect was observed upon treatment with 100nM Rosuvastatin (Fig. 5c). Since both Atorvastatin and Rosuvastatin are hydrogen-bonded to Ser565 of HMG-CoA, it is thought that they have an auxiliary effect that other drugs do not have, which may be one of the reasons for their efficacies. Furthermore, to identify factors that affect the translational activity of luciferase proteins via the 5'UTR, we knocked down ACAA2 and HMGCS2 using MISSION\*





**Figure 3.** RNA-immunoprecipitation of 5-UTR bound host RNA. **(a)** Prediction of luciferase mRNA-5'UTR fusion. Two regions in the luciferase RNA were selected for target by 28 nt gRNAs—gRNA1 and gRNA2. **(b)** gRNA-dCas13 mediated RNA immunoprecipitation of host RNA bound to 5'UTR sequences. RT-qPCR was performed to determine the accumulation of immunoprecipitated 5'UTR sequence (Top) and luciferase RNA (bottom) via gRNA-dCas13 mediated targeting of luciferase RNA. The host RNA bound to 5'UTR was immunoprecipitated and the enrichment factor was calculated as  $2^{-\Delta\Delta C_t}$  [RIP/background]. The mean and standard error of the three independent experiments is shown. Statistical analysis was performed using a two-sided paired t-test. \*Indicates a significant difference from the control sample at a value of  $p < 0.05$ . HEK293T cells (first and second white bars from the left) and A549 cells (third to seventh grey bars from the left).

esiRNA (Sigma). Interestingly, knockdown of ACAA2 repressed the translational activity of luciferase proteins, whereas this was less pronounced for knockdown of HMGCS2 (Fig. 5d, Supplemental Fig. 4). Interestingly in close parallels to our study emphasizing relationship between the mevalonate pathway and viral infection, previous reports have shown that suppressing HMGCS2 or ACAA2 expression may suppress papillomavirus or poxvirus infection<sup>25,26</sup>.



**Figure 4.** Differential expression by RNA seq analysis. **(a)** A set of four samples was analyzed for RNA-seq analysis. Each lane represents Luciferase RNA-immunoprecipitation sample having no 5'UTR sequence (293L2GL3 and A549L2GL3) as controls, and Luciferase-RNA fused with 5'UTR (293L2CV, A549L2CV, and A549L1CV) as test samples. (L1 or L2 are gRNAs targeting Luciferase RNA). Heatmap analysis of differentially enriched genes shows enrichment in 293L2CV, A549L2CV, and A549L1CV immunoprecipitated samples. **(b)** Common and differentially expressed genes in test and control are shown in the Venn diagram. Set of unique genes in the test sample (293L2CV, A549L2CV, and A549L1CV) were significantly enriched ( $p < 0.05$ ) compared to the control (293L2GL3 or A549L2GL3). 293T-XIST is an RNA-immunoprecipitation sample which served as a negative control (HEK293T cells expressing lncRNA-XIST immunoprecipitated by gRNA-dCas13 pull down). The genes enriched in Control-293 vs. 293T-L2-CoV2, Control-A549 vs. A549-L1-CoV2, and Control-A549-L2-CoV2 are human intracellular factors that 5'UTR of SARS-CoV-2. **(c)** Genes enriched in Control-293 vs 293T-L2-CoV2 dataset were ranked according to their differential enrichment and the resulting enriched GO terms were visualized using a DAG graphical representation (increase in enrichment degree is shown as a function of white to red gradient).

**Figure 5.** Overexpression of SARS-CoV-2-5'UTR of alters genes expression involved in the mevalonate pathway in HEK293T and A549 cells. **(a)** 1 [micro]g of pGL3-promoter and pGL3-5'UTR vectors were transfected into cells and cultured in 6 well plates (cell density:  $0.3 \times 10^6$  cells/ml). Total RNA was extracted from the transfected cells and the expression of lipid metabolism genes was analyzed by qPCR. As shown, upon overexpression of 5'UTR, the expression of ACAA2 and HMGCS2 is significantly altered ( $p < 0.05$ ). Three independent trials were performed. Results are shown as mean  $\pm$  SD, and Student's t-test was used as a test of significance. The expression results were expressed as fold relative to house-keeping internal control GAPDH gene expression. The y-axis is fold change which is the ratio of cells transfected with pGL3-promoter only to cells transfected with pGL3-5'UTR. **(b)** Pathways showing overexpression of 5'UTR of SARS-CoV-2 triggers Acetyl-CoA pathway, which can be repressed by HMG-CoA reductase inhibitors like statins. **(c)** 1 [micro]g of pGL3-promoter and pGL3-5'UTR vectors was introduced in cells cultured in 24 well plates. HMGCoA reductase inhibitors were added to the culture solution at a concentration range of 1 nM, 10 nM, 100 nM and 1000 nM. Following incubation for 48 h, the luciferase activity was measured. Y axis is relative luciferase activity which was calculated using the ratio of luciferase to Renilla reniformis luciferase as the internal control. **(d)** The luciferase activity was measured by transfecting 60 ng and 120 ng of esiRNA of HMGCS2 and ACAA2 into HEK293T and A549 cells, respectively. To assess the effect on the 5'UTR, the pGL3-promoter and luciferase activity from each adjacent well into which the pGL3-5'UTR vector was introduced were subtracted. The internal control used was the sum of each value. The results are shown as mean  $\pm$  SD and tested for significance by Student's t-test. The internal control was the sum of the respective values. Y axis is relative luciferase activity which was calculated using the ratio of luciferase to Renilla reniformis luciferase as the internal control. Results are shown as mean  $\pm$  SD, and Student's t-test examined significance. Statistical analysis data are in Supplement Information Table 3.

## Discussion

Since the outbreak of the COVID-19 pandemic in December 2019, its underlying molecular pathogenesis and clinical complications following the SARS-CoV-2 infection have been unceasingly investigated. Previous studies have shown that the COVID-19 patients suffer from excessive immune response termed "cytokine storm" with infiltration of cytokines, mononuclear cells, and monocytes<sup>27,28</sup>. However, in patients with bariatric diseases, the infiltration is extreme and prolonged that leads to acute tissue damage resulting in respiratory distress syndrome (ARDS)<sup>7-9</sup>. Although our in vitro studies were an artificial validation of the fusion of the 5'UTR with luciferase RNA, where in we found significant alteration in lipid metabolism in cells treated with 5'UTR of SARS-CoV-2. It remains obscure why lipid metabolism is altered upon SARS-CoV-2 infection making obese people more vulnerable to disease, however recent studies suggest preferential loss of lipids and amino acids associated enhanced infiltration of cytokines to be responsible for disease severity in COVID-19<sup>29,30</sup>. It is anticipated that further understanding of altered lipid metabolism in the virus-infected cells may help uncover the severity pattern in patients.

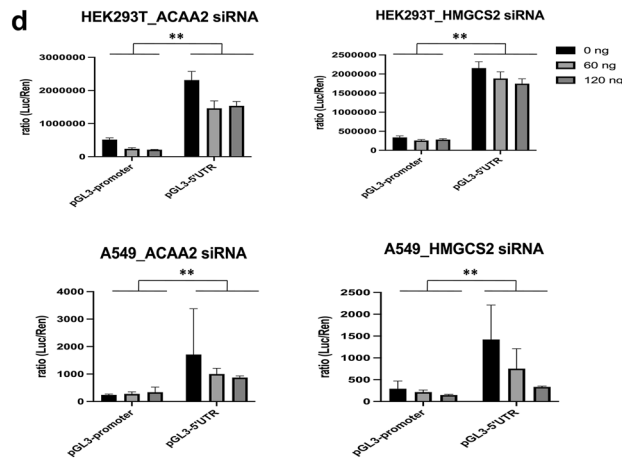
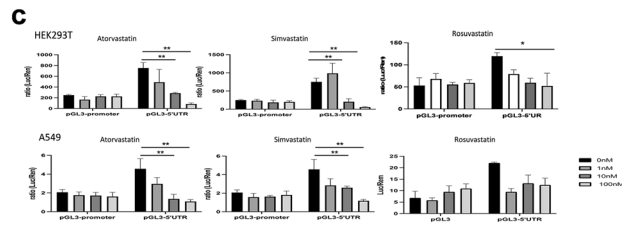
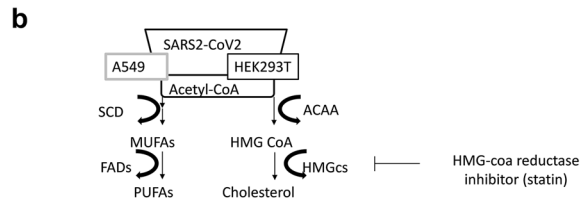
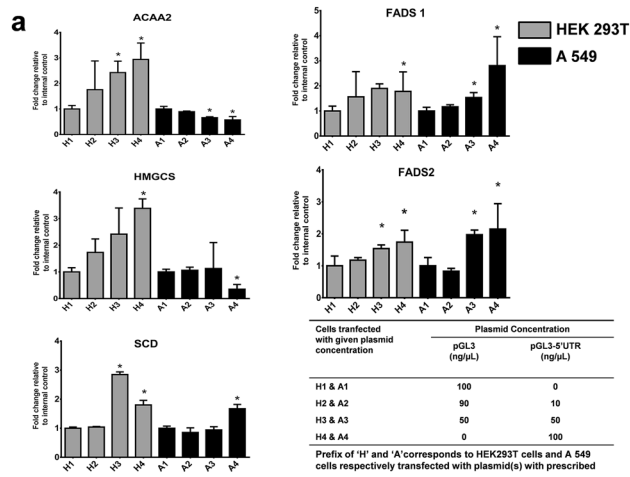
Recent structural and biochemical insights have shown 10- to 20-fold higher affinity of SARS-CoV-2 spike proteins to human ACE2 receptors than SARS-CoV<sup>31</sup> which possibly explains the higher transmissibility of COVID-19<sup>32</sup>. Therefore, an antibody-based therapeutic strategy that blocks the interaction site of SARS-CoV-2 and ACE2 receptor is anticipated to be effective<sup>33</sup>. Further, two phase-I trials (NCT04392219 and NCT04746183) have demonstrated that oral molnupiravir, a Novel Oral Anti-SARS-CoV-2 agent, is safe and well-tolerated at therapeutic doses. However, the latest analysis of a late-stage clinical trial of molnupiravir has shown to reduce the risk of hospitalisation or death from serious illness by 30%<sup>34</sup>. Despite all curative interventions, providing preliminary supportive care to the patients forms an effective disease management approach during the growing pandemic. Drug repositioning forms the rescue approach until a safe and reliable cure becomes available.

Elucidating the function of the 5'UTR region stands important as it forms the translational initiation region for host cellular factors to translate SARS-CoV-2 RNA. Although limited to translational activity from luciferase RNA fused to the 5'UTR of SARS-CoV-2, we observed significant changes in expression for genes related to mevalonate pathways and lipid metabolism. We herein show that HMG-CoA reductase inhibitors like Atorvastatin suppress the translational activation function of the 5'UTR region. Our cell-based investigation interestingly finds a good correlation with the previous clinical studies which showed that patients under the administration of HMG-CoA reductase inhibitors showed less aggravated symptoms for COVID-19<sup>35-37</sup>. We anticipate that short-term administration of HMG-CoA reductase inhibitors may likely suppress the exacerbation of SARS-CoV-2-infected individuals. When the expression of ACAA2 was suppressed by siRNA instead of statin, it had a strong inhibitory effect on the translational activity of luciferase protein via 5'UTR in both cells. The proliferative capacity of HEK293T and A549 cells was also good in a statin or siRNA-treated environment with 5'UTR-mediated inhibition of translation of the luciferase protein (Supplementary Fig. 1). In terms of drug repositioning the use of statins in SARS-CoV-2 infected patients may be reasonable, however, the indication of drugs that inhibit ACAA2 expression may be more effective in repressing translation through the 5'UTR of SARS-CoV-2.

## Conclusion

In conclusion from the present study we herein elucidate the function of the SARS-CoV-2 5'UTR region in human cells which helps in understanding of the host transcriptome changes which may in turn help to predict the severity of the infection. Further elucidation of regulation of HMGCS2 or ACAA2 expressions including the results obtained from the present study may provide multiple antiviral drug development opportunities in the near future. We anticipate that the altered expression of genes as a result of infection may serve as a biomarker







Carlsbad, California, USA). RNA immunoprecipitation was performed as described previously<sup>25</sup> except that the AM-tag fused dCas13-RNA complexes were eluted from the Protein A Dynabeads (Thermo Fisher Scientific, 10001D, Waltham, Massachusetts, USA) and treated with 20 µg of proteinase K at 45 °C for 1 h, then de-crosslinked at 65 °C for 1 h. The RNA was isolated using TRI reagent (Ambion, AM9738, Austin, Texas, USA) according to the manufacturer's instructions. After treatment with Turbo DNase (Ambion, AM2238), the RNAs were subjected to RT-qPCR using primers to detect RNA sequence of Luciferase and 5'UTR of SARS-CoV-2. The RNAs were further subjected to Ultra Low Input RNA-seq (GeneWiz, <https://www.genewiz.com>). The DNase treatment of the extracts was performed at 25 °C for 30 min using 20 U of Turbo DNase (Ambion, AM2238).

**Quantitative polymerase chain reaction.** HEK293T cells or A549 cells were transfected with pGL3-promoter and pGL3-5'UTR plasmids. The amount of pGL3-promoter was reduced according to the amount of pGL3-5'UTR in order to equalize the amount of plasmid introduced. Poly(A)'RNA and total RNA were extracted from HEK293T and A549 using the RNeasy minikit (Qiagen, 74104, Hilden, Germany). cDNA was prepared by reverse transcription of 500 ng of total RNA using the PrimeScrip RT Master Mix (Takara, RR036A, Kusatsu, Japan). The resulting cDNAs were amplified using the QuantiTect SYBR Green PCR kit (Qiagen, 204143). The RT-PCR reaction was performed on StepOne Real-Time PCR Systems (Applied Biosystems, Foster City, California, USA). Quantitation, and calling of real-time amplification values were performed on StepOne Software Version 2.2 (<https://www.thermofisher.com/>). Expression data was normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) expression to the corresponding sample. Primer sequences are included in the supplemental information (Table 4).

**Luciferase reporter assay.** pGL3-promoter vector and pGL3-5'UTR vector were used as a reporter gene for investigating the 5'UTR dependent translation activity. Plasmids were transiently transfected into HEK293T and A549 cells. The cells were harvested for 48 h after transfection, followed by which luciferase activity was analyzed with Dual-Luciferase reporter assay system (Promega, E1910) using MiniLumat LB9506 (Berthold Technologies). Luciferase activity and reporter activities were normalized in reference to cotransfected Rous sarcoma virus (RSV)-Renillareniformis luciferase expression plasmid as described previously<sup>39</sup>. The SD values were obtained from three trials of each test.

**Drug-induced translation suppression.** For translation suppression effect induced by HMG-CoA reductase inhibitors, cells were maintained in the presence or absence of Atorvastatin Calcium, Simvastatin, Rosuvastatin Calcium Salt, Pravastatin Sodium Salt (FUJIFILM Wako Chemicals, 104451679902-63-9, 187-03361, 162-19821, Osaka, Japan) diluted in DMSO (1 nM–1000 nM) for 48 h.

**RNA seq and data analysis.** Total RNA was sequenced on illuminaNovaSeq 6000 sequencing platform on a 2 × 150 bp flow cell. Reads generated from sequencing were saved in Illumina specific (Binary Base Call (BCL) file format. All the BCL files were converted to fastq files with Bcl2fastq v2.17.1.14 (<https://sapac.support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html>) compatible for downstream sequence analysis programs. Prior to data analysis, the reads were filtered by eliminating low-quality reads (Phred score < 30) by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), trimming reads those less than 75 bp by Trimmomatic v 0.36 (<http://www.usadellab.org/cms/index.php?page=trimmomatic>)<sup>40</sup>. Cutting adapter sequences by Cutadapt version 1.9.1 (<https://cutadapt.readthedocs.io/en/v1.9.1/>)<sup>41</sup>. The filtered reads were subsequently aligned to the Human reference GRCh38.p12 using Hisat2 v2.0.1<sup>42</sup>, followed by assigning the reads onto genomic features—exons, introns and intergenic regions. The mapped reads were visualized by Integrative Genomics Viewer Version 2.5.2 (<https://software.broadinstitute.org/software/igv/>)<sup>43</sup>. For assembly and predict alternative splicing StringTie v1.3.3b (<https://ccb.jhu.edu/software/stringtie/>)<sup>44</sup> was used and ASprofile V1.0.4 (<https://ccb.jhu.edu/software/ASprofile/>)<sup>45</sup> was employed for classification and quantification of the spliced reads. For de novo transcript assembly StringTie v1.3.3b was employed using the alignment bam files, the result of which was then subjected to comparison with existing annotation reference (gtf file) using Cuffcompare V2.2.1 (<https://www.genepattern.org/modules/docs/Cufflinks.cuffcompare/7>)<sup>46</sup>. Gene expression calculation was performed as a function of FPKM with a threshold of 0.1–1 based on read counts from HT-seq v 0.6.1 (<https://bioweb.pasteur.fr/packages/pack@HTSeq@0.6.1>)<sup>47</sup>. Differential gene analysis was performed using the Bioconductor package DESeq2 V1.6.3 and edgeR V3.4.6 (<https://bioconductor.org/packages/release/bioc/html/edgeR.html>)<sup>48</sup>. The results from EdgeR analysis were further analyzed to determine genes with significant differential expression according to the criteria of fold change greater than 2 and qvalue (fdr, padj) less than 0.05. The GO functional enrichment analysis was performed by Goseq 1.42.0 (<https://bioconductor.org/packages/release/bioc/html/goseq.html>)<sup>49</sup>. TopGO (<https://bioconductor.org/packages/release/bioc/html/topGO.html>)<sup>50</sup> was applied for generating Directed Acyclic Graph (DAG) for results of enrichment analysis of the differentially expressed genes. DEXSeq (V1.18.4) (<https://www.rdocumentation.org/packages/DEXSeq/versions/1.18.4>) was used to assess differential expression of exons and alternative exon usage in alternative splicing<sup>51</sup>. Enriched gene list is included in the supplemental information.

Received: 26 April 2021; Accepted: 8 June 2023

Published online: 27 June 2023

## References

- Iype, E. & Gulati, S. Understanding the asymmetric spread and case fatality rate (CFR) for COVID-19 among countries. medRxiv 2020.04.21.20073791; <https://doi.org/10.1101/2020.04.21.20073791>.
- Chen, N. *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *Lancet (London, England)* **395**(10223), 507–513. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7) (2020).
- Buonsenso, D. *et al.* Novel coronavirus disease-19 pneumoniae: A case report and potential applications during COVID-19 outbreak. *Eur. Rev. Med. Pharmacol. Sci.* **24**, 2776–2780 (2020).
- Zhang, J. J. *et al.* Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy* **75**(7), 1730–1741. <https://doi.org/10.1111/all.14238> (2020).
- Wang, C., Horby, P. W., Hayden, F. G. & Gao, G. F. A novel coronavirus outbreak of global health concern. *The Lancet* **395**(10223), 470–473 (2020).
- Honce, R. *et al.* Obesity-related microenvironment promotes emergence of virulent influenza virus strains. *MBio* **11**(2), e03341–e3419. <https://doi.org/10.1128/mBio.03341-19> (2020).
- Huttunen, R. & Syrjänen, J. Obesity and the risk and outcome of infection. *Int. J. Obes.* **37**(3), 333–340. <https://doi.org/10.1038/hjo.2012.62> (2013).
- Jones, C. Y. *et al.* Overweight and human immunodeficiency virus (HIV) progression in women: Associations HIV disease progression and changes in body mass index in women in the HIV epidemiology research study cohort. *Clin. Infect. Dis.* **37**(Suppl 2), S69–S80. <https://doi.org/10.1086/375889> (2003).
- Falagas, M. E., Athanasoulia, A. P., Peppas, G. & Karageorgopoulos, D. E. Effect of body mass index on the outcome of infections: A systematic review. *Obes. Rev.* **10**(3), 280–289 (2009).
- Muscogiuri, G., Pugliese, G., Barrea, L., Savastano, S., & Colao, A. (2020). Commentary: obesity: The “Achilles heel” for COVID-19?. *Metab. Clin. Exp.* 108.
- Popkin, B. M. *et al.* Individuals with obesity and COVID-19: A global perspective on the epidemiology and biological relationships. *Obes. Rev.* **21**(11), e13128 (2020).
- Dietz, W. & Santos-Burgoa, C. Obesity and its implications for COVID-19 mortality. *Obesity* **28**(6), 1005–1005 (2020).
- Hodcroft, E. B., Zuber, M., Nadeau, S., Crawford, K. H., Bloom, J. D., Velesler, D., *et al.* (2020). Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. MedRxiv.
- Jaafar, Z. A. & Kieft, J. S. Viral RNA structure-based strategies to manipulate translation. *Nat. Rev. Microbiol.* **17**(2), 110–123 (2019).
- Wu, A. *et al.* Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **27**(3), 325–328. <https://doi.org/10.1016/j.chom.2020.02.001> (2020).
- Miao, Z., Tidu, A., Eriani, G., & Martin, F. (2020). Secondary structure of the SARS-CoV-2 5'-UTR. *RNA Biol.*, 1–10.
- Babendure, J. R., Babendure, J. L., Ding, J. H. & Tsien, R. Y. Control of mammalian translation by mRNA structure near caps. *RNA* **12**(5), 851–861 (2006).
- Sample, P. J. *et al.* Human 5' UTR design and variant effect prediction from a massively parallel translation assay. *Nat. Biotechnol.* **37**(7), 803–809. <https://doi.org/10.1038/s41587-019-0164-5> (2019).
- Vankadari, N., Jeyasankar, N. N. & Lopes, W. J. Structure of the SARS-CoV-2 Nsp1/5'-untranslated region complex and implications for potential therapeutic targets, a vaccine, and virulence. *J. Phys. Chem. Lett.* **11**(22), 9659–9668. <https://doi.org/10.1021/acs.jpcclett.0c02818> (2020).
- Jiang, Y. *et al.* The Rbm38-p63 feedback loop is critical for tumor suppression and longevity. *Oncogene* **37**(21), 2863–2872 (2018).
- Lee, H., Lee, J. J., Park, N. Y., Dubey, S. K., Kim, T., Ruan, K., *et al.* (2021). Multi-omic analysis of selectively vulnerable motor neuron subtypes implicates altered lipid metabolism in ALS. *Nat. Neurosci.*, 1–13.
- Zhao, L. *et al.* TDP-43 facilitates milk lipid secretion by post-transcriptional regulation of Btn1a1 and Xdh. *Nat. Commun.* **11**(1), 1–15 (2020).
- Anant, S. *et al.* Novel role for RNA-binding protein CUGBP2 in mammalian RNA editing: CUGBP2 modulates C to U editing of apolipoprotein B mRNA by interacting with apobec-1 and ACF, the apobec-1 complementation factor. *J. Biol. Chem.* **276**(50), 47338–47351 (2001).
- Chen, Z., Eggerman, T. L. & Patterson, A. P. ApoB mRNA editing is mediated by a coordinated modulation of multiple apoB mRNA editing enzyme components. *Am. J. Physiol. Gastrointest. Liver Physiol.* **292**(1), G53–G65. <https://doi.org/10.1152/ajpgi.00118.2006> (2007).
- Filone, C. M. *et al.* The master regulator of the cellular stress response (HSF1) is critical for orthopoxvirus infection. *PLoS Pathog* **10**(2), e1003904 (2014).
- Lipovsky, A. *et al.* Genome-wide siRNA screen identifies the retromer as a cellular entry factor for human papillomavirus. *Proc. Natl. Acad. Sci.* **110**(18), 7452–7457 (2013).
- Diao, B., Feng, Z., Wang, C., Wang, H., Liu, L., Wang, C., *et al.* (2020). Human kidney is a target for novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. MedRxiv.
- Mehta, P. *et al.* COVID-19: Consider cytokine storm syndromes and immunosuppression. *The Lancet* **395**(10229), 1033–1034 (2020).
- Alqahtani, S. A. & Schattenberg, J. M. Liver injury in COVID-19: The current evidence. *United Eur. Gastroenterol. J.* **8**(5), 509–519 (2020).
- Cecere, G., Hoersch, S., Jensen, M. B., Dixit, S. & Grishok, A. The ZFP-1 (AF10)/DOT-1 complex opposes H2B ubiquitination to reduce Pol II transcription. *Mol. Cell* **50**(6), 894–907 (2013).
- Wan, Y., Graham, R., Baric, R., & Li, F. (2020). An analysis based on decade-long structural studies of SARS 3, JVI Accepted Manuscript Posted Online 29 January 2020. *J. Virol.*
- Bourgonje, A. R. *et al.* Angiotensin-converting enzyme 2 (ACE2), SARS-CoV-2 and the pathophysiology of coronavirus disease 2019 (COVID-19). *J. Pathol.* **251**(3), 228–248 (2020).
- Gue, Y. X., Kanji, R., Markides, V. & Gorog, D. A. Angiotensin Converting Enzyme 2 may mediate disease severity in COVID-19. *Am. J. Cardiol.* **130**, 161–162 (2020).
- Couzin-Frankel, J. Antiviral pills could change pandemic's course. *Science* **374**(6569), 799–800. <https://doi.org/10.1126/science.acx9605> (2021).
- Tan, W. Y., Young, B. E., Lye, D. C., Chew, D. E. & Dalan, R. Statin use is associated with lower disease severity in COVID-19 infection. *Sci. Rep.* **10**(1), 1–7 (2020).
- Song, S. L. *et al.* Statin use is associated with decreased risk of invasive mechanical ventilation in COVID-19 patients: A preliminary study. *Pathogens* **9**(9), 759 (2020).
- De Spiegeleer, A. *et al.* The effects of ARBs, ACEis, and statins on clinical outcomes of COVID-19 infection among nursing home residents. *J. Am. Med. Dir. Assoc.* **21**(7), 909–914.e2. <https://doi.org/10.1016/j.jamda.2020.06.018> (2020).
- Dalvai, M. *et al.* A scalable genome-editing-based approach for mapping multiprotein complexes in human cells. *Cell Rep.* **13**(3), 621–633 (2015).
- Tsuda, M., Takahashi, S., Takahashi, Y. & Asahara H. Transcriptional co-activators CREB-binding protein and p300 regulate chondrocyte-specific gene expression via association with Sox9. *J Biol Chem.* **278**(29), 27224–27229 (2003).

40. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**(15), 2114–2120 (2014)
41. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**(1), 10–12 (2011).
42. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**(8), 907–915 (2019).
43. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**(1), 24–26 (2011).
44. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**(1), 1–13 (2019).
45. Florea, L., Song, L. & Salzberg, S. L. Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000Research* **2**, 188 (2013).
46. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**(5), 511–515 (2010).
47. Anders, S., Pyl, P. T. & Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**(2), 166–169 (2015).
48. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1), 139–140 (2010).
49. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **11**(2), 1–12 (2010).
50. Alexa, A., & Rahnenfuhrer, J. (2020). TopGo: Enrichment analysis for gene ontology; R Package Version 2.42. 0. 2020.
51. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**(10), 2008–2017. <https://doi.org/10.1101/gr.133744.111> (2012).

### Author contributions

T.I.: Principal investigator of the project, conducted in vitro experiments, conceptualized the idea, written original draft, performed formal analysis and generated figures. Y.S.: Co-investigated the project, conducted in vitro experiments and validated the results. S.B., M.H.: Conducted NGS data analysis and performed statistical analysis. S.Y., T.S., K.U.: Performed additional in vitro experiments including cell culture and transfection assays. Y.K., S.L.: Performed drug induced translation assays in cells. N.K.T., S.T.: Validated and curated the data and additional programmer of the project. S.S.: Chief project administrator.

### 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/s41598-023-36680-6>.

**Correspondence** and requests for materials should be addressed to T.I.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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) 2023