

# Exogenous coronavirus interacts with endogenous retrotransposon in human cells

Ying Yin

Huazhong University of Science and Technology

Xiao-zhao Liu

Huazhong University of Science and Technology

Ximiao He (✉ [XimiaoHe@hust.edu.cn](mailto:XimiaoHe@hust.edu.cn))

Huazhong University of Science and Technology

Li-quan Zhou (✉ [zhouliquan@hust.edu.cn](mailto:zhouliquan@hust.edu.cn))

Huazhong University of Science and Technology

---

## Research Article

**Keywords:** Coronavirus, Retrotransposon, SARS-CoV-2, TET, LINE

**DOI:** <https://doi.org/10.21203/rs.3.rs-40063/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** There is an increased global outbreak of diseases caused by coronaviruses affecting respiratory tracts of birds and mammals. Recent particularly dangerous coronaviruses are MERS-CoV, SARS-CoV and SARS-CoV-2, causing respiratory illness and even failure of several organs. However, profound impact of coronavirus on host cells remains elusive.

**Results:** Here, we go deep into transcriptome of MERS-CoV, SARS-CoV and SARS-CoV-2 infected human lung-derived cells, and observed that infection of these coronaviruses all induced increase of retrotransposon expression through upregulation of TET genes. Similar upregulation of retrotransposon was also observed in SARS-CoV-2 infected human intestinal organoids. Retrotransposon upregulation will lead to increased genome instability and more frequent readthrough from retrotransposon to dysregulate gene expression. People with higher basal level of retrotransposon like cancer patients and aged people will have increased risk of symptomatic infection. Additionally, we show evidence supporting long-term epigenetic inheritance of retrotransposon upregulation. We also observed significant amount of chimeric transcripts of retrotransposon and SARS-CoV-2 RNA for potential human genome invasion of viral fragments, with the front and the rear part of SARS-CoV-2 genome being easier to form chimeric RNA, and this may apply for other coronaviruses. Here we suggest that primers and probes for nucleic acid detection should be designed in the middle of virus genome to identify live virus with higher probability.

**Conclusions:** In summary, we propose that infection of coronaviruses especially SARS-CoV-2 induce retrotransposon activation, formation of chimeric coronavirus-retrotransposon RNA, and elicits more severe symptoms in patients with underlying diseases. More attention may need to be paid to potential harm contributed by retrotransposon dysregulation in treatment of coronavirus-infected patients.

## Background

Emerging coronaviruses often spread rapidly from person to person and there seems to be an increased global outbreak of related diseases. MERS-CoV and SARS-CoV are two identified rare coronavirus strains which cause not only severe lung infection but also serious complications (1-4). More recently, coronavirus disease named COVID-19 caused by a novel coronavirus SARS-CoV-2 is expanding globally and rapidly, resulting in emerging health issues (5-8). Although cell receptors and the routes of infection of these coronaviruses have been identified (8-11) complicated impact on human cells is far from clear.

Transposable Elements (TEs) are mobile DNA elements in virtually all eukaryotes and comprise more than 40% of human genome (12). They can self-replicate and insert into various locations inside genome. Dysregulation of TE may lead to various illnesses like inflammatory diseases (13). The only active member in TE is retrotransposon which can “copy and paste” themselves through RNA intermediate. Expression of most of retrotransposon members is suppressed in somatic cells and they are only active in brains, germ cells, early embryos and pathological conditions (14). About 5% of newborn babies show

a new retrotransposon integration event (15). Abnormally upregulation of retrotransposons cause insertions, deletions and inversions in genome (16, 17), resulting in compromised genetic stability and even cell death (18, 19). Accumulated evidence in recent years also proved their importance in orchestration of gene expression (20), regulation of chromatin structure (21) and modulation of developmental program (22, 23).

Long interspersed nuclear elements (LINEs) are common autonomous retrotransposons and comprise about 17% of human genome (24). Some LINE-1 elements can be transcribed and translated in cells. After reverse transcription of LINE-1 RNA, they can be integrated back into genome (25). Naturally, LINEs expression is repressed in most cell types. Its RNA is mainly heritable during early embryogenesis because of its enrichment and high retrotransposition activity in early embryos (26). Transgenic mouse model carrying mouse/human LINE-1 retrotransposition reporter demonstrated that this activity creates somatic mosaicism during development (27). Besides LINEs, short interspersed nuclear elements (SINEs) and long terminal repeats (LTRs) are also enriched retrotransposons in human genome, and mobilization of SINEs relies on LINE-1-encoded proteins (12).

In our study, we analyzed publicly available transcriptome data of human cells infected with coronavirus MERS-CoV, SARS-CoV and SARS-CoV-2, and observed enhanced expression of TEs including several retrotransposons, as well as inflammation, immunity and apoptosis related genes. We further noticed potential fusion of SARS-CoV-2 RNA with retrotransposon transcripts especially LINEs and SINEs. Therefore, further examinations on genome and transcriptome of cells from patients and studying models will be valuable to evaluate potential crosstalk between coronavirus and retrotransposons.

## Results And Discussion

### **Coronavirus infection disturbs diverse biological processes in human cells and can stimulate ACE2 expression through IRF1 and STAT1**

Coronaviral infection led to not only respiratory failure but also multiple organ dysfunction syndromes, indicating that there are common pathways for coronavirus to impact human cells (28). Transcriptome analysis may provide valuable information on how human cells react with coronavirus entry.

To examine whether coronavirus infection disturbs expression of specific gene sets in human cells, we analyzed public available RNA-seq data of human lung-derived cells with infection of MERS-CoV, SARS-CoV, and SARS-CoV-2. Through comparison of transcriptomes before and after infection, we identified thousands of dysregulated genes (adjusted p-value < 0.05) for each group (Fig. 1A). Among those dysregulated genes, we found that 26 genes were commonly upregulated after infection of the three coronaviruses (Fig. 1B), but very few genes were identified to be commonly downregulated (Fig. 1C). GO analysis of the 26 commonly upregulated genes demonstrated enrichment on inflammation, immunity and apoptosis related pathways (Fig. 1B). Through relative viral sequence content in transcriptome, we found that the three coronaviruses can infect various human lung-derived cells (Fig. 1D), however, low

dose of coronavirus or using NHBE cells for infection were not successful to support coronavirus replication (Fig. S1).

ACE2 is the cell receptor of SARS-CoV-2 (8, 9). Differently from robust expression of ACE2 in Calu-3 cells, ACE2 expression was undetectable in A549 cells, but after SARS-CoV-2 infection, low level of ACE2 was observed (Fig. 1E). This indicates that transcription factors responding to coronavirus infection induced ACE2 expression. Recent report showed that ACE2 can be stimulated by interferon, and proposed IRF1 and STAT1-binding sites near ACE2 transcription start site (Fig. S2) (29). Here, we noticed that expression of both IRF1 and STAT1 were increased after SARS-CoV-2 infection, and ACE2 expression was significantly reduced when IRF1 was depleted in virus-infected human cells or STAT1 was depleted in interferon-treated human cells (Fig. 1E). These results confirmed that IRF1 and STAT1 are essential upstream activators of ACE2 upon virus infection. So we propose that SARS-CoV-2 might enter human cells with low efficiency by bulk-phase endocytosis in A549 cells, inducing IRF1 and STAT1 expression which further enhances ACE2 expression to facilitate receptor-mediated viral entry. Therefore, IRF1 and STAT1 seem to be two promising drug targets to limit coronavirus entry through ACE2.

### **Coronavirus infection enhanced retrotransposon expression in human lung-derived cells**

Next, we ask whether TE expression is impacted by coronavirus infection. We first examined transcriptome of human lung adenocarcinoma cell line Calu-3 after 24-hr infection of MERS-CoV (30). We observed that TE expression was generally activated after coronavirus infection (Fig. 2A). Further examination documented that subfamilies of LINEs, SINEs, LTRs were differentially upregulated by coronavirus (Fig. 2B). LINE-1 is the mostly well-studied autonomous retrotransposon. Most LINE-1 elements are inactivated in somatic cells, but some escape variously evolved silencing mechanisms. Hence, we ask whether evolutionarily old and young retrotransposons were impacted by coronavirus infection differently. We compared the ratio of fold change of specific LINE-1 element expression ordered by predicted evolutionary ages, and found that older and younger LINE-1 elements were similarly influenced (Fig. 2C) (31). One of the major mechanisms for LINE-1 silencing is DNA methylation, and we examined expression of genes encoding DNA methyltransferases (DNMTs) and Ten-eleven translocation (TET) enzymes mediating active DNA demethylation. We observed that Tet genes were generally upregulated after coronavirus infection (Fig. 2D), and upregulated DNA demethylation activity may lead to demethylation of retrotransposon promoters. This result supports that increased retrotransposon expression was caused by genome-wide DNA demethylation. We obtained similar results in MERS-CoV/SARS-CoV infected MRC5 cells which are noncancerous human lung fibroblast cells (Fig. 2A-D).

Recent COVID-19 outbreak is caused by the novel coronavirus SARS-CoV-2. Here, we explored transcriptomes of SARS-CoV-2 infected A549 and Calu-3 cells. Similar to MERS-CoV and SARS-CoV infection, we found general increase of multiple transposable elements (Fig. 2A-B), no biased impact of older and younger LINE-1 elements by SARS-CoV-2 infection (Fig. 2C). SARS-CoV-2 infection also causes upregulation of TET gene expression (Fig. 2D). Similarly, SARS-CoV-2 was identified to have the

capability of infecting human intestinal organoids (Fig. 2E) and increased retrotransposon expression can also be observed post infection in a time-dependent manner (Fig. 2F).

Therefore, upregulation of retrotransposon seems to be a common event induced by coronavirus infection, possibly through enhancing global DNA demethylation activity. Despite of similar upregulation of retrotransposon families triggered by the three coronaviruses, individual retrotransposons are differently dysregulated, and this may cause various phenotypes in human cells. Note that above results were from 24-hr infection of coronaviruses, and impact of long-term infection should be more severe. Moreover, retrotransposon is able to encode proteins and can form retrovirus-like particles (26), so electron microscopy examination of coronavirus-infected samples may need to discriminate coronavirus from retrovirus-like particles because of upregulation of retrotransposons.

### **Upregulation of retrotransposon may be long-term memorized epigenetically**

We then ask whether retrotransposon upregulation can be long-term inherited through several generations of cell divisions. We found the mouse model of transgenerational epigenetic inheritance of acquired traits may provide molecular insights into this question.

tRNA-derived small RNAs (tsRNAs) in sperm were reported to transmit abnormal epigenetic information into preimplantation embryo, and epigenetic abnormality was further inherited to adult tissue, causing metabolic disorders (32). Two kinds of tsRNAs were previously identified to regulate retrotransposon LTR (33), so we ask whether abnormal retrotransposon activity is inheritable during this process. We analyzed the transcriptome of cleavage mouse embryo and adult islet originated from zygote with injection of tsRNA of sperm from normal or high-fat diet (HFD) male mice. We found that LINE, SINE and LTR retrotransposons were all upregulated in 8-cell embryo when HFD tsRNA was injected (Fig. 3A). Notably, LTR retrotransposon also showed upregulation in adult islet (Fig. 3B). Further analysis on LTR families supported that upregulation of ERV1 expression was inherited from early embryo (Fig. 3C) to adult islet (Fig. 3D), probably through DNA methylation inheritance at ERV1 locus. Therefore, above result indicates that enhancement of retrotransposon expression, ERV1 in this case, can be long-term inherited through several generations of cell cycles, even from cleavage-stage early embryos to adult tissues, with change of DNA methylation as the potential molecular mechanism (Fig. 3E).

### **SARS-CoV-2 RNA forms chimeric transcripts with retrotransposon RNA especially LINE for potential insertion into host genome**

Coronaviruses are RNA viruses and are not supposed to integrate into host genome by themselves. However, it was reported that several RNA viruses have capacity to recombine with retrotransposons to invade host genome (34, 35). Regarding contribution of SARS-CoV-2 RNA to total transcriptome in infected Calu-3 cells to be as high as 15.32% (Fig. 1D), we explored in the transcriptome the potential chimeric transcripts of SARS-CoV-2 and cellular RNA, and obtained subtranscriptome with chimeric reads.

We found that 0.23% of SARS-CoV-2 RNA formed chimeric transcripts with non-TE genes and 0.14% with TE (Fig. 4A). Surprisingly, TE-virus chimeric reads contribute 37.36% to total mapped chimeric reads, while TE reads are only 2.83% in total mapped reads (Fig. 4B), indicating that TE is much more efficient to form chimeric transcripts with SARS-CoV-2 RNA than non-TE genes. We randomly extracted reads from subtranscriptome of chimeric transcripts of SARS-CoV-2 and cellular RNA, and confirmed identity of the chimeric reads (Fig. 4C).

We further analyzed distribution of TE subfamilies in total transcriptome and subtranscriptome with chimeric reads, and found that reads of retrotransposon LINE, SINE and LTR were all enriched in the subtranscriptome of chimeric reads (Fig. 4D). Unexpectedly, only LINE RNA was overrepresented in subtranscriptome with chimeric reads than in total transcriptome, and further analysis showed that virus-LINE-1 was overrepresented in virus-LINE reads (Fig. 4E). This demonstrates high efficiency of LINE family especially LINE-1 in forming chimeric transcript with SARS-CoV-2 RNA. LINE-1 is autonomous retrotransposon with retrotransposition activity, and RNA-RNA ligation mediated by endogenous RNA ligase RtcB was previously reported for LINE-1 to carry other types of RNA for host genomic invasion (36), so similar mechanisms may apply for SARS-CoV-2 transcripts. Further analysis of human genome from SARS-CoV-2 infected human cells or biopsies will be particularly important to identify existence of integration of coronavirus RNA into human genome.

Moreover, to identify which region of SARS-CoV-2 RNA prone to form chimeric transcripts with cellular RNA, we aligned total transcriptome and subtranscriptome to SARS-CoV-2 genome, and viewed on IGV to find that the front and the rear parts, especially the rear part of coronavirus RNA were biased in forming chimeric transcripts (Fig. 4F). That means the front and the rear parts of SARS-CoV-2 fragments are easier to be inserted into human genome for prolonged expression, indicating that people even positive for Nucleic Acid Test may just have infection history, and not really carry live coronavirus but only silent viral fragments. Taken together, we suggest that primers and probes for SARS-CoV-2 testing are designed in middle of the SARS-CoV-2 genome.

### **The model of coronavirus-retrotransposon interaction**

Based on above analysis, we propose that coronavirus infection may increase retrotransposon expression through modulating TET activity to reduce global DNA methylation. Increased retrotransposon RNA may further form chimeric transcripts with coronavirus RNA, and integrate viral genomic fragments into human genome. Moreover, enforced retrotransposon expression may be harmful and probably long-term inherited (Fig. 5A).

TE is widely expressed in human tissues (Fig. 5B), with highest enrichment in early human embryos (Fig. 5C). The cells used in this study are mainly derived from human lung and also robustly express TE (Fig. 5D). Moreover, TE subfamilies are variable in different cell types (Fig. 5E-G), suggesting extensive but specific phenotype upon global retrotransposon upregulation.

The first concern regarding global retrotransposon upregulation is genome instability. Retrotransposition activity is high in early embryo (26) and brain (37) during normal development, so potential integration of coronavirus sequence into human genome is suggested to be scrutinized for these cells. It was also reported that retrotransposon upregulation is positively correlated with tumor progression (38), causing genomic deletion, translocation and duplication (39). What's more, increased expression of retrotransposon LINE-1 contributes to age-associated inflammation in several tissues (40). Additionally, vapers and smokers demonstrated higher retrotransposon expression and hypomethylation at associated loci (41). Also, people with neurological disorders may have higher retrotransposon expression and retrotransposition activity (42). These reports not only show that upregulation of retrotransposon expression may cause several diseases, but also indicate that persons with higher basal level of retrotransposons are supposed to be more susceptible to coronavirus infection and have increased risk of symptomatic infection. In support of this, recent analysis of SARS-CoV-2 patients showed that cancer patients (43) and aged people (44) get more severe symptoms after infection. Therefore, inhibition of reverse transcriptase activity in human cells may be necessary during pharmaceutical treatment of coronavirus-infected patients, especially those with higher basal level of retrotransposons.

The second concern regarding global retrotransposon upregulation is disturbance of retrotransposon adjacent gene expression. Accumulated evidence shows that retrotransposons are not just genomic fossils, but have molecular functions. For example, physically adjacent retrotransposon activates gene promoter of TMEM156 by readthrough mechanism (Fig. 5H, Fig. S3). Also, transcripts of LINEs, SINEs and low-complexity repeats physically interacted with specific genomic areas to play distinct roles (45).

The third concern regarding global retrotransposon upregulation is whether coronavirus RNA can enter nucleus and associate with specific genomic regions through sequence homology, similar like the behavior of retrotransposon RNA (21, 45). Blast analysis in NCBI using SARS-CoV-2 genome showed no similar sequence in human genome. We further used CENSOR program (46) to analyze the SARS-CoV-2 genome and all predicted candidate repetitive elements are less than 200bp. Therefore, no evidence supports that SARS-CoV-2 RNA has the ability to recognize human genome by homologous sequence even these transcripts enter nucleus by chance.

## Conclusions

Taken together, we demonstrate that coronavirus infection increases retrotransposon expression in human cells, possibly through global DNA hypomethylation, and increased retrotransposon RNA may further form chimeric transcripts with coronavirus RNA for integration of viral genomic fragments into human genome. These enhanced retrotransposon transcripts may be long-term inherited to harm host organs. Therefore, we propose that retrotransposon upregulation induced by coronavirus infection has significant contributions to coronavirus caused symptoms, and suggest careful transcriptome examination and genetic tests in future investigations on coronavirus-infected patients.

## Methods

## Cell types used for transcriptome study of coronavirus infection

Cell types below are used in this study. Calu-3, human lung cancer cell; MRC5, human fetal lung strain; A549, human adenocarcinomic alveolar basal epithelial cell; NHBE, primary human bronchial epithelial cell. Each group above has three replicates. For human intestinal organoids, each group has two replicates.

## RNA-seq data processing

Raw reads were processed with cutadapt v1.16 to perform quality trimming with default parameters except for: quality-cutoff =20, pair-filter=both. To include as many non-uniquely mapped reads as possible, trimmed reads were firstly aligned to human/mouse genome (hg19/mm10) by STAR (v2.5.1b) with default settings including parameters '-winAnchorMultimapmax 2000 -outFilterMultimapNmax 1000'. RSEM was used to calculate FPKM value of genes. The annotation and fasta sequences for consensus transposable element sequences were downloaded from Repbase (version 20.01) (47). TE transcript with default parameters was used to get counts for transposable elements. For RNA-seq alignment of coronavirus genomes, MERS-CoV (NC\_019843), SARS-CoV (NC\_004718) and SARS-CoV-2 (NC\_045512) genomes were downloaded from NCBI, and trimmed reads were aligned to coronavirus genome by STAR (v2.5.1b) using default parameters. To identify potential chimeric transcripts of coronavirus and cellular transcripts from single-end RNA-seq data, 30nt fastq reads from each end were extracted from raw fastq reads and both were aligned to human and SARS-CoV-2 genomes respectively. Non-viral end of the chimeric reads were mapped to consensus transposable element sequences using STAR with parameters '-winAnchorMultimapmax 2000 -outFilterMultimapNmax 1000' to get counts of transposons. Integrative Genomics Viewer (IGV) was used for snapshot of transcriptome. R package Deseq2 was used to get differential expressed genes. Metascape was used to visualize functional profiles of genes and gene clusters (48). Graphs were created by R or Excel. Images were organized by Adobe Illustrator.

## Accession number

RNA sequencing data of coronavirus-infected human lung-derived cells are from GSE122876 (transcriptome of MERS-CoV-infected Calu-3 cells; single-read; MOI 2, 24 hours) (30), GSE56192 (transcriptome of MERS-CoV and SARS-CoV infected MRC5 cells; paired-end; MOI 2, 24 hours), GSE147507 (transcriptome of SARS-CoV-2 infected A549 cells, Calu-3 cells, and NHBE cells; MOI 2, 24 hours) (49). RNA sequencing data of SARS-CoV-2-infected human intestinal organoids are from GSE149312 (MOI 1, 24 and 60 hours, grown in differentiation medium) (50). SARS-CoV-2 infected Calu-3 cells were used to identify chimeric transcripts of coronavirus and cellular RNA. RNA sequencing data of IRF1 knockout and control human hepatocytes infected with hepatitis A virus are from GSE114916. RNA sequencing data of STAT1 knockout and control human HepG2 cells treated by IFN are from GSE98372 (51). RNA sequencing data of human tissues and cell types are from GSE83115 (52). RNA sequencing data of human early embryos and embryonic stem cells are from GSE36552 (53). RNA sequencing data

of 8-cell mouse embryos and adult mouse islet developed from zygotes with injection of sperm tsRNAs from high-fat-diet males are from GSE75544 (32).

## List Of Abbreviations

TE, transposable element

LINE, long interspersed nuclear element

SINE, short interspersed nuclear element

LTR, long terminal repeat

## Declarations

### Acknowledgements

We thank Dr. Bing Li from Shanghai Jiao Tong University for assistance of data analysis on retrotransposon expression.

### Authors' contributions

Li-quan Zhou and Ximiao He conceived and designed the project. Ying Yin analyzed the data and wrote the manuscript. Xiao-zhao Liu performed analysis on chimeric transcripts. Li-quan Zhou and Ximiao He revised the manuscript. All authors have read and approved the final manuscript.

### Funding

This work was supported by National Key R&D Program of China [2018YFC1004502, 2018YFC1004001], National Natural Science Foundation of China [NSFC 31771661].

### Availability of data and materials

RNA sequencing data of MERS-CoV-infected Calu-3 cells (30) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA506733 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA506733/>).

RNA sequencing data of MERS-CoV-infected and SARS-CoV-infected MRC5 cells are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA233943 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA233943/>). RNA sequencing data of SARS-CoV-2 infected A549 cells, Calu-3 cells, and NHBE cells (49) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA615032 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA615032/>). RNA sequencing data of SARS-CoV-2-infected human intestinal organoids (50) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA628628 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA628628/>). RNA sequencing data of IRF1 knockout and control human hepatocytes infected with hepatitis A virus are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA473130

(<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA473130/>). RNA sequencing data of STAT1 knockout and control human HepG2 cells treated by IFN (51) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA384926 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA384926/>). RNA sequencing data of human tissues and cell types (52) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA324812 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA324812/>). RNA sequencing data of human early embryos and embryonic stem cells (53) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA153427 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA153427/>). RNA sequencing data of 8-cell mouse embryos and adult mouse islet developed from zygotes with injection of sperm tsRNAs from high-fat-diet males (32) are obtained from NCBI Sequence Read Archive with BioProject ID: PRJNA304514 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA304514/>).

### **Competing interests**

The authors declare that they have no competing interests.

### **Ethics approval and consent to participate**

Not applicable

### **Consent for publication**

Not applicable

## **References**

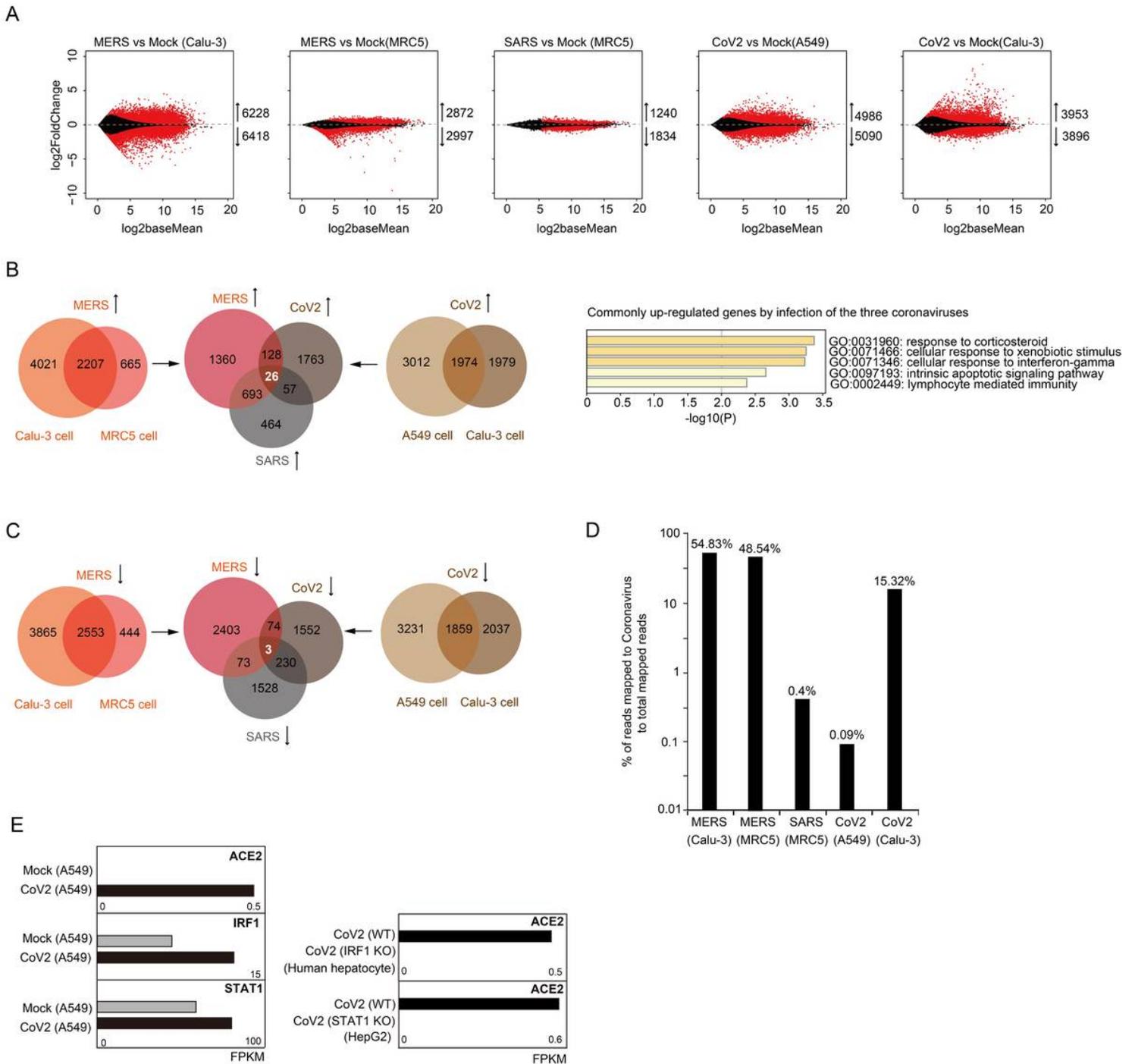
1. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med.* 2003;348(20):1953-66.
2. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science.* 2003;300(5624):1394-9.
3. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med.* 2012;367(19):1814-20.
4. Arabi YM, Balkhy HH, Hayden FG, Bouchama A, Luke T, Baillie JK, et al. Middle East Respiratory Syndrome. *N Engl J Med.* 2017;376(6):584-94.
5. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* 2020;395(10223):497-506.
6. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med.* 2020.
7. Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family

- cluster. *Lancet*. 2020;395(10223):514-23.
8. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270-3.
  9. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;367(6483):1260-3.
  10. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*. 2003;426(6965):450-4.
  11. Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, Dijkman R, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*. 2013;495(7440):251-4.
  12. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet*. 2003;35(1):41-8.
  13. Saleh A, Macia A, Muotri AR. Transposable Elements, Inflammation, and Neurological Disease. *Front Neurol*. 2019;10:894.
  14. Munoz-Lopez M, Vilar-Astasio R, Tristan-Ramos P, Lopez-Ruiz C, Garcia-Perez JL. Study of Transposable Elements and Their Genomic Impact. *Methods Mol Biol*. 2016;1400:1-19.
  15. Cordaux R, Hedges DJ, Herke SW, Batzer MA. Estimating the retrotransposition rate of human Alu elements. *Gene*. 2006;373:134-7.
  16. Gilbert N, Lutz-Prigge S, Moran JV. Genomic deletions created upon LINE-1 retrotransposition. *Cell*. 2002;110(3):315-25.
  17. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, et al. Human I1 retrotransposition is associated with genetic instability in vivo. *Cell*. 2002;110(3):327-38.
  18. Newkirk SJ, Lee S, Grandi FC, Gaysinskaya V, Rosser JM, Vanden Berg N, et al. Intact piRNA pathway prevents L1 mobilization in male meiosis. *Proc Natl Acad Sci U S A*. 2017;114(28):E5635-E44.
  19. Malki S, van der Heijden GW, O'Donnell KA, Martin SL, Bortvin A. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev Cell*. 2014;29(5):521-33.
  20. Izsvak Z, Wang J, Singh M, Mager DL, Hurst LD. Pluripotency and the endogenous retrovirus HERVH: Conflict or serendipity? *Bioessays*. 2016;38(1):109-17.
  21. Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, et al. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol*. 2013;20(3):332-8.
  22. Lu JY, Shao W, Chang L, Yin Y, Li T, Zhang H, et al. Genomic Repeats Categorize Genes with Distinct Functions for Orchestrated Regulation. *Cell Rep*. 2020;30(10):3296-311 e5.
  23. Percharde M, Lin CJ, Yin Y, Guan J, Peixoto GA, Bulut-Karslioglu A, et al. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell*. 2018;174(2):391-405 e19.
  24. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet*. 2009;10(10):691-703.

25. Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH, Jr. L1 integration in a transgenic mouse model. *Genome Res.* 2006;16(2):240-50.
26. Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature.* 2015;522(7555):221-5.
27. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, et al. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev.* 2009;23(11):1303-12.
28. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet.* 2020;395(10223):470-3.
29. Ziegler CGK, Allon SJ, Nyquist SK, Mbanjo IM, Miao VN, Tzouanas CN, et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell.* 2020.
30. Yuan S, Chu H, Chan JF, Ye ZW, Wen L, Yan B, et al. SREBP-dependent lipidomic reprogramming as a broad-spectrum antiviral target. *Nat Commun.* 2019;10(1):120.
31. Khan H, Smit A, Boissinot S. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res.* 2006;16(1):78-87.
32. Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, et al. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science.* 2016;351(6271):397-400.
33. Schorn AJ, Gutbrod MJ, LeBlanc C, Martienssen R. LTR-Retrotransposon Control by tRNA-Derived Small RNAs. *Cell.* 2017;170(1):61-71 e11.
34. Geuking MB, Weber J, Dewannieux M, Gorelik E, Heidmann T, Hengartner H, et al. Recombination of retrotransposon and exogenous RNA virus results in nonretroviral cDNA integration. *Science.* 2009;323(5912):393-6.
35. Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, Oshida T, et al. Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature.* 2010;463(7277):84-7.
36. Moldovan JB, Wang Y, Shuman S, Mills RE, Moran JV. RNA ligation precedes the retrotransposition of U6/LINE-1 chimeric RNA. *Proc Natl Acad Sci U S A.* 2019;116(41):20612-22.
37. Zhao B, Wu Q, Ye AY, Guo J, Zheng X, Yang X, et al. Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett patients and healthy individuals. *PLoS Genet.* 2019;15(4):e1008043.
38. Jung H, Choi JK, Lee EA. Immune signatures correlate with L1 retrotransposition in gastrointestinal cancers. *Genome Res.* 2018;28(8):1136-46.
39. Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, Zamora J, Supek F, Demeulemeester J, et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nat Genet.* 2020;52(3):306-19.
40. De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature.* 2019;566(7742):73-8.

41. Caliri AW, Caceres A, Tommasi S, Besaratinia A. Hypomethylation of LINE-1 repeat elements and global loss of DNA hydroxymethylation in vapers and smokers. *Epigenetics*. 2020;1-14.
42. Terry DM, Devine SE. Aberrantly High Levels of Somatic LINE-1 Expression and Retrotransposition in Human Neurological Disorders. *Front Genet*. 2019;10:1244.
43. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol*. 2020;21(3):335-7.
44. Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med*. 2020;26(4):506-10.
45. Ding Y, He L, Zhang Q, Huang Z, Che X, Hou J, et al. Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: implications for pathogenesis and virus transmission pathways. *J Pathol*. 2004;203(2):622-30.
46. Jurka J. Repeats in genomic DNA: mining and meaning. *Curr Opin Struct Biol*. 1998;8(3):333-7.
47. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA*. 2015;6:11.
48. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
49. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Moller R, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. 2020;181(5):1036-45 e9.
50. Lamers MM, Beumer J, van der Vaart J, Knoops K, Puschhof J, Breugem TI, et al. SARS-CoV-2 productively infects human gut enterocytes. *Science*. 2020.
51. Chen K, Liu J, Liu S, Xia M, Zhang X, Han D, et al. Methyltransferase SETD2-Mediated Methylation of STAT1 Is Critical for Interferon Antiviral Activity. *Cell*. 2017;170(3):492-506 e14.
52. Zhu J, Chen G, Zhu S, Li S, Wen Z, Bin L, et al. Identification of Tissue-Specific Protein-Coding and Noncoding Transcripts across 14 Human Tissues Using RNA-seq. *Sci Rep*. 2016;6:28400.
53. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol*. 2013;20(9):1131-9.

## Figures

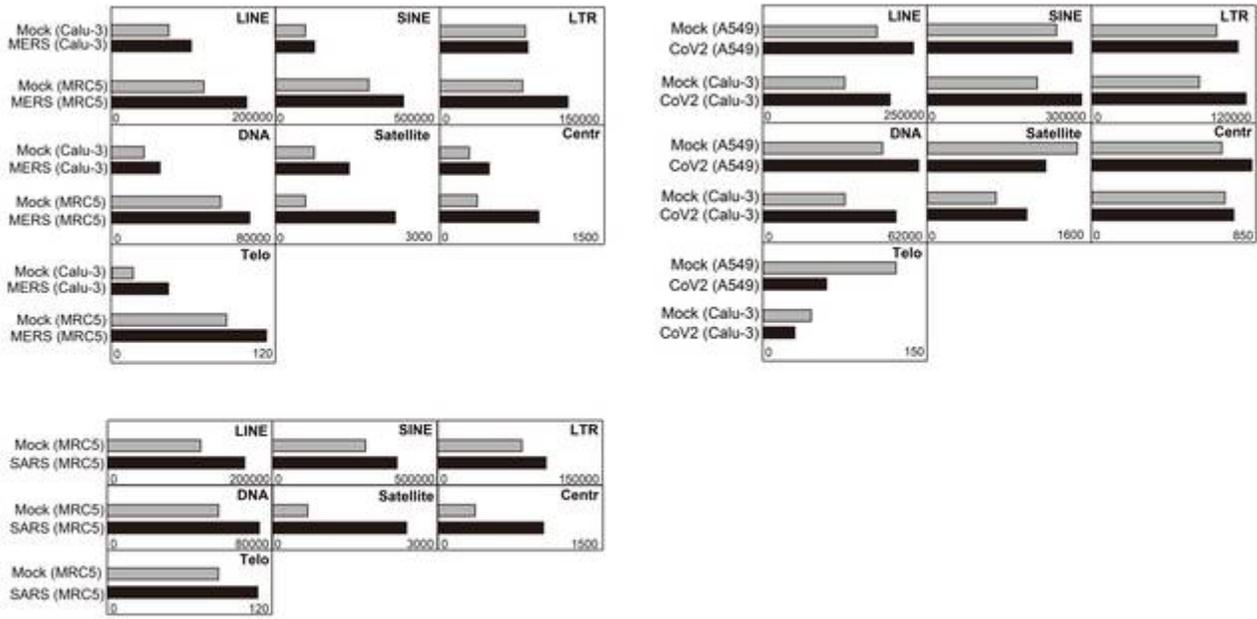


**Figure 1**

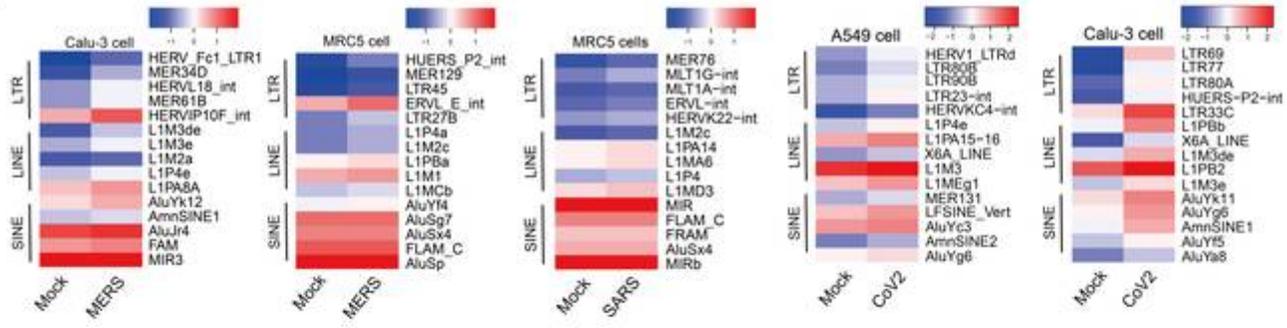
Analysis of transcriptome alteration induced by infection of various coronaviruses. A: MA plot (log ratio RNA abundance versus log abundance) of RNA-seq data comparing control and coronavirus-infected cells. Differentially expressed genes with adjusted  $P < 0.05$  are highlighted in red. Numbers of up/down-regulated genes are indicated. Calu-3, MRC5 and A549 are all human cells with lung origin. MERS, MERS-CoV; SARS, SARS-CoV; CoV2, SARS-CoV-2. B: Venn diagrams document 26 commonly upregulated genes by different coronavirus infection (left panel). Gene ontology analysis of the 26 genes for enriched biological processes (right panel). C: Venn diagrams document only 3 commonly downregulated genes by SARS-CoV-2 infection in two cell types. D: Bar graph indicates percentage of reads mapped to

coronavirus genome to total mapped reads in human cells infected with coronavirus. E: Bar graphs demonstrate that SARS-CoV-2 infection caused change of ACE2 expression from below detection to low level in A549 cells. SARS-CoV-2 infection also caused upregulation of IRF1 and STAT1. IRF1 knockout in human hepatocytes infected with hepatitis A virus significantly decreased ACE2 expression. STAT1 knockout in IFN-treated human HepG2 cells significantly decreased ACE2 expression.

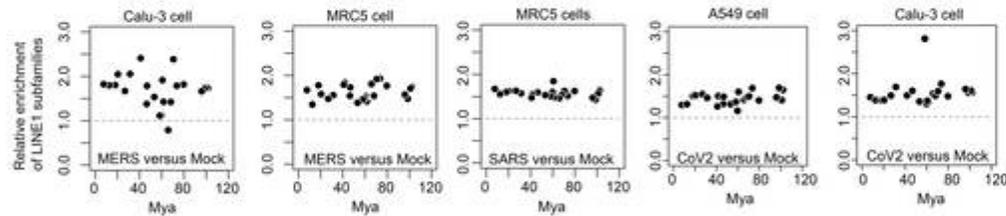
A



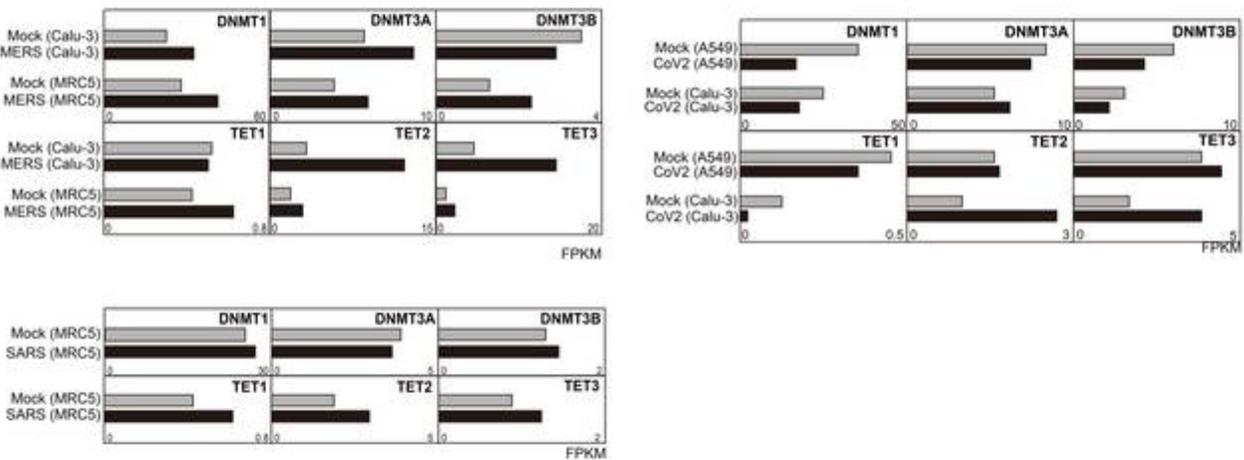
B



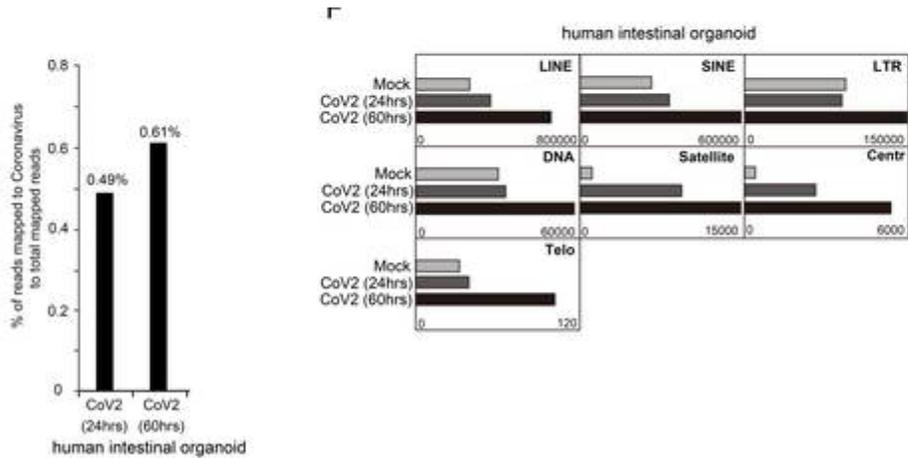
C



D

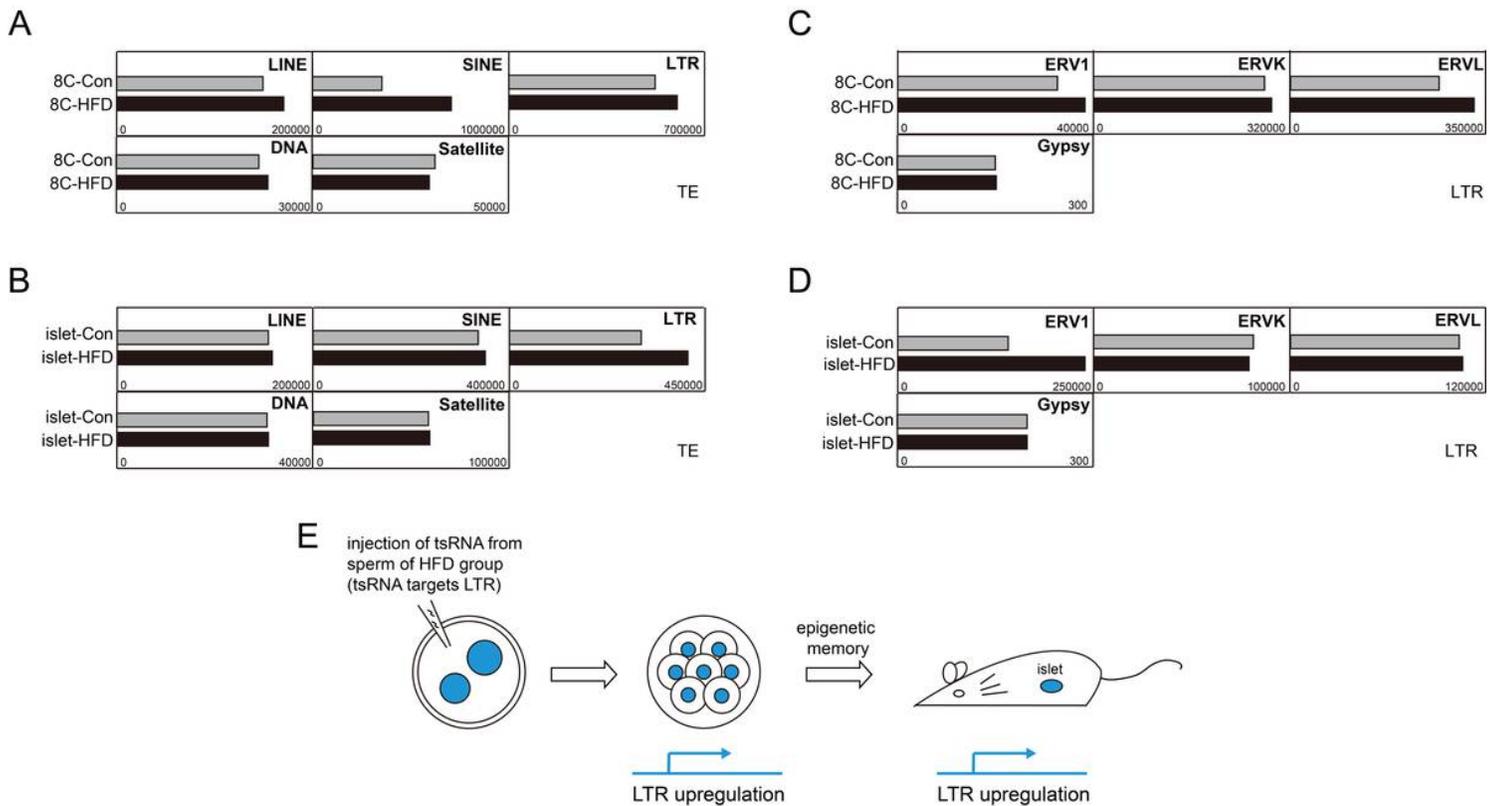


E



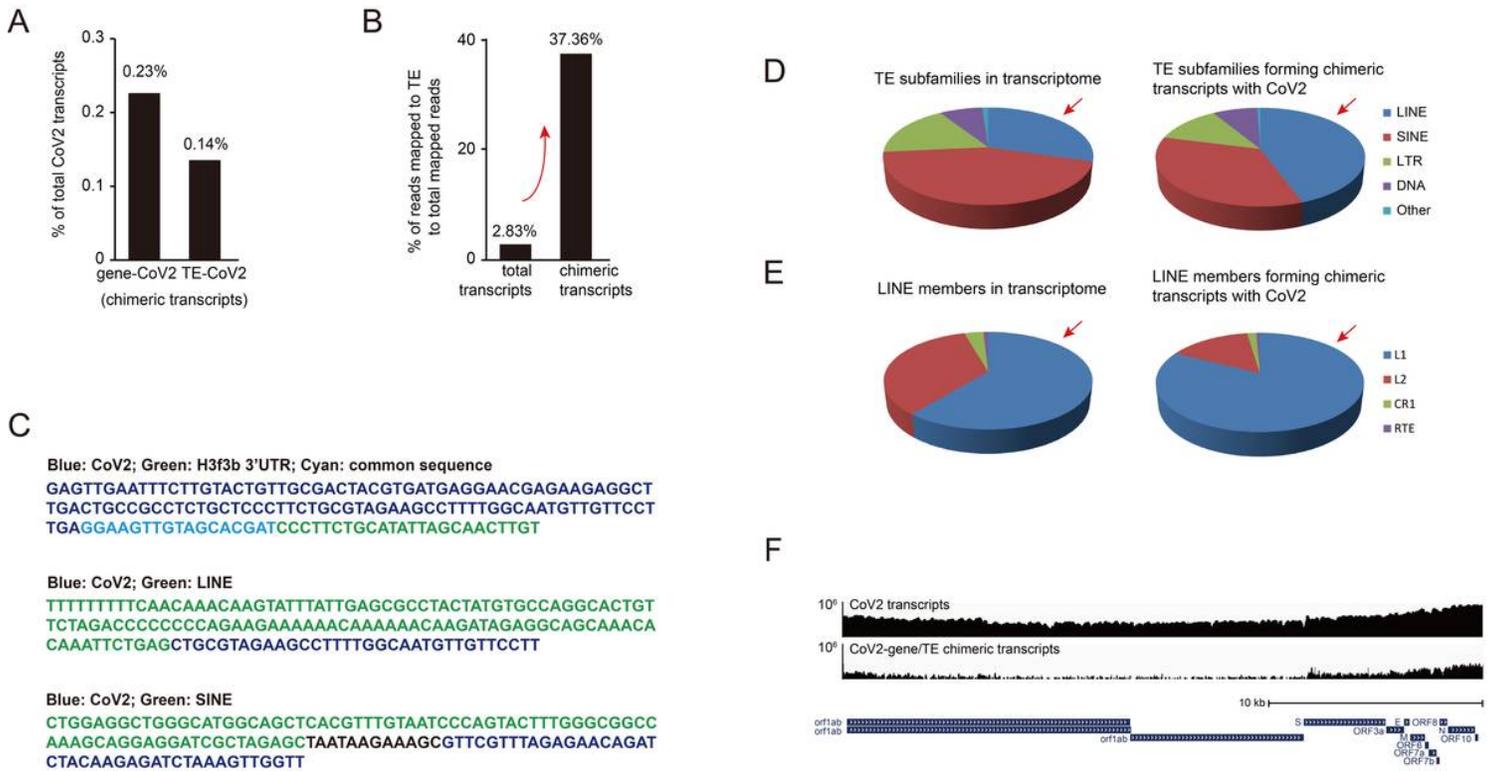
**Figure 2**

Coronavirus infection in human cells enhanced retrotransposon expression. A: Bar graphs show that after 24-hr MERS-CoV/SARS-CoV/SARS-CoV-2 infection in Calu-3/MRC5/A549 cells, expression of TE was generally increased. B: Heatmap indicates upregulation of several LTR, LINE and SINE elements induced by MERS-CoV/SARS-CoV/SARS-CoV-2 infection. C: Ratio of LINE1 upregulation was not determined by evolutionary age of LINE1 elements. Evolutionary age was calculated based on a substitution rate of 0.17%/million years. Mya, million years ago. D: Bar graphs depict expression of genes encoding enzymes controlling DNA methylation status before and after MERS-CoV/SARS-CoV/SARS-CoV-2 infection in Calu-3/MRC5/A549 cells. E: Bar graphs show that 24hrs or 60hrs post SARS-CoV-2 infection in human intestinal organoid, expression of TE was generally increased. F: Bar graph indicates percentage of reads mapped to SARS-CoV-2 genome to total mapped reads in transcriptome of human intestinal organoids infected with SARS-CoV-2.



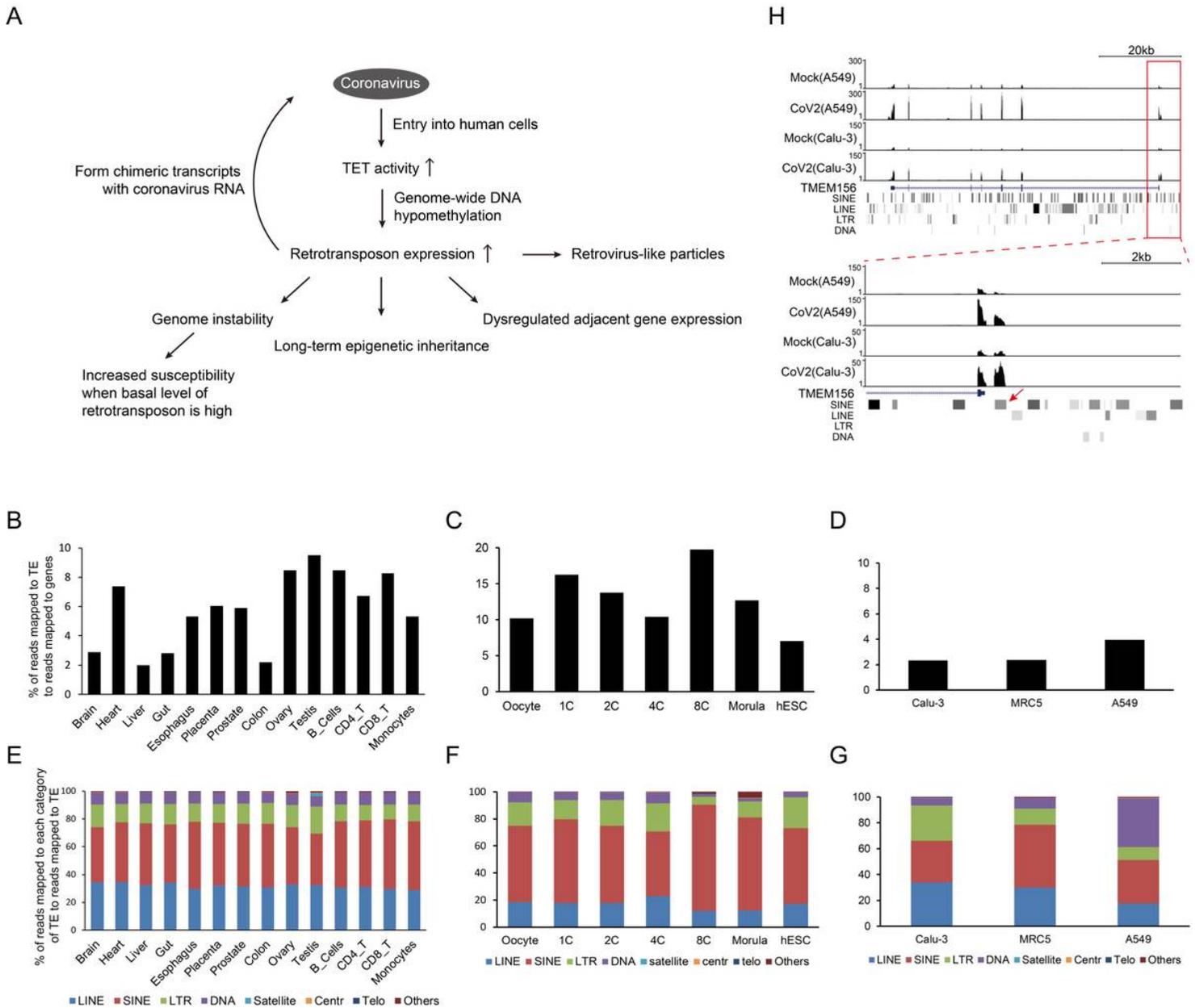
**Figure 3**

Potential long-term memory of TE upregulation. A: Expression of TE subfamilies in 8-cell (8C) embryos developed from normal mouse zygote injected with tsRNA derived from sperm of high-fat-diet (HFD) male. B: Expression of TE subfamilies in adult islet developed from normal mouse zygote injected with tsRNA derived from sperm of HFD male. C: Expression of LTR elements in 8C embryos developed from normal mouse zygote injected with tsRNA derived from sperm of HFD male. D: Expression of LTR elements in adult islet developed from normal mouse zygote injected with tsRNA derived from sperm of HFD male. E: Scheme of the process from obtaining tsRNA injected zygote to examining TE expression in 8C embryos and adult islet.



**Figure 4**

Retrotransposon-coronavirus chimeric transcripts were observed in SARS-CoV-2 infected human cells. A: Bar graph shows relative enrichment of chimeric transcripts of coronavirus and cellular transcripts to total coronavirus transcripts. B: Examination of ratio of mapped TE reads to non-TE gene reads in total transcriptome and in subtranscriptome of chimeric reads (between viral and cellular transcripts). C: Example of chimeric reads with junctions of coronavirus-gene, coronavirus-LINE and coronavirus-SINE. D: Pie charts demonstrate distribution of TE subfamilies in total transcripts (left panel) and coronavirus-retrotransposon chimeric transcripts (right panel) in SARS-CoV-2 infected Calu-3 cells. Red arrow indicates overrepresentation of LINE reads. E: Pie charts demonstrate distribution of LINE members in total transcripts (left panel) and coronavirus-retrotransposon chimeric transcripts (right panel) in SARS-CoV-2 infected Calu-3 cells. Red arrow indicates overrepresentation of LINE-1 (L1) reads. F: IGV snapshot of SARS-CoV-2 transcripts (upper) and chimeric transcripts (between viral and cellular transcripts, lower) identified in infected Calu-3 cells. SARS-CoV-2 genome was used for alignment. Logarithmic scale is displayed. The reference panel was obtained from UCSC genome browser.



**Figure 5**

Model of how coronavirus may impact retrotransposons to harm human cells. **A:** Generally, entry of coronavirus into human cells enhances TET activity for genome-wide DNA hypomethylation to facilitate retrotransposon enhancement. This may lead to increased fusion between viral and retrotransposon transcripts, reduced genome stability and increased susceptibility of aged people and cancer patients, dysregulated TE-adjacent gene expression, and this influence may be inherited for a long term. Meanwhile, increased retrovirus-like particles of retrotransposons may be induced. **B-G:** Endogenous retrotransposon expression and distribution of subfamilies are variable in human tissues and cells. Bar graphs indicate percentages of reads mapped to TE to reads mapped to genes in human tissues and immunocytes (**B**), human oocytes and early embryos (**C**), and normal Calu-3, MRC5 and A549 cells (**D**). Bar charts (**E-G**) demonstrate distribution of individual subfamilies of TE in tissues or cell types shown above. **H:** An example of retrotransposon-initiated gene expression by readthrough mechanism.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalFile.docx](#)