Supplementary Materials for

Azacytidine targeting SARS-CoV-2 viral RNA as a potential treatment for COVID-19

This Supplementary file includes:

Materials and Methods

Figures. S1 to S13, and Supplementary Table 1

Materials and Methods

Cells, virus, and reagents

Vero E6 cells were obtained from the American Type Culture Collection (ATCC-1586) and cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco), supplemented with 10% fetal bovine serum (FBS; Gibco), penicillin (100 U/ml), and streptomycin (100 μg/ml), at 37°C in a humidified 5% CO_2 incubator.

SARS-CoV-2 (nCoV-2019BetaCoV/Wuhan/WIV04/2019) was propagated in Vero E6 cells. Viral titer was determined through plaque assay. Prof. Meilin Jin from Huazhong Agricultural University provided mouse-adaptive SARS-CoV-2 (MA-SARS-CoV-2), which was created through continuously passaging wild virus in aged mouse lungs[1]. All infection experiments were performed in a biosafety level-3 (BSL-3) laboratory.

5-Azacytidine (5Aza), Remdesivir (RDV), and Chloroquine (CQ) were obtained from MedChemExpress and dissolved in sterile saline and water, respectively.

Lipo8000 (Beyotime Biotechnology, China) was used to transfect plasmids and siRNA according to the instruction.

Evaluation of cytotoxicity

Vero E6 cells were incubated with different 5Aza concentrations in 96-well plates for 24h. Cytotoxicity was measured using CCK-8 assay. CC_{50} was calculated in Prism (GraphPad).

Quantitative reverse transcription-PCR (qRT-PCR)

Viral RNA was extracted using the TaKaRa MiniBEST Viral RNA/DNA Extraction Kit (Cat No. 9766). RNA was eluted in 40 μl RNase-free water. Viral RNA copy number was determined with TaqMan-based qRT-PCR using the One Step PrimeScript RT-PCR Kit (Takara, RR064A). To generate a standard curve, the receptor binding domain (RBD) of the spike gene was amplified and its copy number determined from serial dilutions. Primers used for qRT-PCR were RBD-F: CAATGGTTTAACAGGCACAGG, RBD-R:

CTCAAGTGTCTGTGGATCACG, and RBD-probe: 5′-FAM-ACAGCATCAGTAGTGTCAGCAATGTCTC-BHQ 3′. The thermocycling schedule was as follows: 95°C for 2 min, followed by 40 cycles of 95°C for 5 s and 60°C for 30 s. The half-maximal inhibitory concentration (IC_{50}) of 5Aza was calculated in GraphPad.

Immunofluorescence microscopy

Cells were fixed overnight with 4% paraformaldehyde, permeabilized with 0.05% Triton X-10 for 10 min, and blocked with 5% bovine serum albumin (BSA) for 1 h. They were then incubated for 2 h at room temperature with a primary antibody against viral N protein (1:500 dilution, Sino Biological), followed by incubation with Alexa 488-labeled goat anti-rabbit IgG (1:500 dilution, Proteintech) for 1 h. Nuclei were stained with DAPI (Beyotime Biotechnology). Images were captured with fluorescence microscopy.

Plaque assay

Vero E6 cells (~1 × 10^5 cells) were incubated in DMEM containing SARS-CoV-2. After 1 h, the supernatant was removed and the cells were incubated in an overlay medium containing DMEM with 2% FBS and 0.9% CMC (Calbiochem). After 96 h, the cells were fixed overnight using 4% paraformaldehyde and stained with crystal violet.

Time-of-drug-addition assay

A time-of-drug-addition assay was performed to determine the SARS-CoV-2 life-cycle stage that 5Aza affects. For “Entry,” Vero E6 cells were pretreated with 24 μM 5Aza for 1 h and infected with 0.01 MOI SARS-CoV-2. Two hours (h) later, the cells were washed twice with DMEM and placed in fresh culture medium. For “Post-entry,” the cells were infected with virus for 2 h, then washed twice and cultured with fresh medium containing 24 μM 5Aza. For “Full-time,” cells were pretreated with 24 μM 5Aza for 1 h and infected with SARS-CoV-2 for 2 h. Subsequently, the cells were washed twice and cultured with fresh medium containing 24 μM 5Aza until the experiment ended. After 24 and 48

hours, supernatants were collected, cells were fixed, and total proteins were extracted for qRT-PCR, fluorescence microscopy, and western blotting, respectively.

RNA immunoprecipitation

Vero E6 cells seed in 10-cm plate were transfected with pCAGGS/HA expressing eGFP, DNMT2, or NSUN2. 24 hours later, cells were infected with 0.2 SARS-CoV-2 for 20 hours in the presence of 16 μM 5Aza or not. Cells were lysed with RIP lysis buffer (gzscbio, Guangzhou, China) with inhibitor cocktail. 1% of the lysates was collected for input analysis, and 3% of lysates were collected for WB analysis. RNAs of remaining lysates were immunopricipated by anti-HA agarose in the presence of RNase inhibitor and 0.2 M EDTA at 4°C overnight. After washing with RIP buffer for 8 times, 20% of the agarose was denatured for IPed protein detection, the remaining was subjected to 10% SDS and proteinase K treatment at 65°C for 45 min. Then the supernatant was collected for RNA extraction, followed by RT-PCR for SARS-CoV-2 RNA detection.

Western blotting

Cells were lysed with RIPA buffer (Thermo, 89900) containing a protease inhibitor cocktail (MCE). Proteins were separated using 10% SDS-PAGE and then transferred onto PVDF membranes (Millipore). Membranes were blocked in Tris Buffered Saline (TBS) buffer with 0.1% Tween-20 (TBST) for 1 h at room temperature, followed by incubation with a primary antibody against viral N protein for 2 h at room temperature. After washing with TBST three times, the membranes were incubated with a secondary peroxidase-conjugated antibody (1:5000 dilution, Proteintech). GAPDH was used as the internal control. Bands were detected using WesternBright ECL (Advansta).

Animal experiments

Female BALB/c mice (8 weeks old) were obtained from Snac Jinda Laboratory Science (Hunan Province, China) and randomly divided into five

groups (Control, 5Aza, Control + SARS-CoV-2, 5Aza + SARS-CoV-2, and Remdesivir + SARS-CoV-2; n = 10 per group). On Day 0, mice were anesthetized using isoflurane. Infection groups were intranasally challenged with 4×10^3 PFU MA-SARS-CoV-2 in 50 μl DMEM, while the un-infection groups were treated with 50 μl DMEM only. One day later, mice were intraperitoneally injected with 2 mg/kg 5Aza (Control and 5Aza + SARS-CoV-2 groups), 25 mg/kg Remdesivir (Remdesivir + SARS-CoV-2 group) or an equivalent volume of sterile saline (Control and Control + SARS-CoV-2 groups), once daily for seven consecutive doses. Body weight was measured daily for 14 d; mouse with more than 30% of body weight loss was immediately euthanized and regarded to be dead, and mouse survival rates were calculated (n = 6). At 4 dpi, four mice from each group were euthanized for lung collecting. Tissues were used for the qRT-PCR quantification of viral RNA copy number and H&E staining for histopathological changes.

All animals were bred in a specific pathogen-free animal facility; experiments were performed in a biosafety level-3 (BSL-3) laboratory and were approved by the Institutional Animal Care and Use Committee of Wuhan Institution of Virology, CAS (No. WIVA04202004).

Bisulfite library preparation, RNA library preparation, and next-generation sequencing

Total RNA (200 ng) was treated with the MGIEasy rRNA Depletion Kit for rRNA removal, bisulfite-converted using the EZ RNA Methylation Kit (Zymo Research), and then fractionated into long and small RNA fractions.

The transcriptome library of bisulfate RNA or non-bisulfate RNA was constructed using the MGIEasy RNA Directional Library Prep Kit (without ligase in the second strand synthesis). Samples were sequenced using the PE100+10 bp method. ERCC was used as negative control and helped to correct false positive sites. The ERCC without any methylation could be used to assay the RNA bisulfite conversation rate.

Nanopore direct RNA sequencing

Total RNA (1 μg) was treated with the MGIEasy rRNA Depletion Kit to remove rRNA. Next, a direct RNA Sequencing kit (SQKRNA002) was used to prepare RNA with 1D sequencing on an Oxford Nanopore equipment. Data were analyzed using an established pipeline.

Quality control

Firstly, a custom Python script was used to demultiplex the barcode list. Secondly, adapter trimming and quality pruning were analyzed with fastp. Bases in the sliding window with mean quality below 20 were cut, and raw reads of those containing 5 or 3 bp adapters were trimmed. mRNA was trimmed to nine bp from the left end.

Mapping and calling

C2T and G2A reference genomes were converted for building indexes using bismark _genome_preparation. Clean sequencing reads were aligned to the two reference genomes with Bismark. Methylation calls were stored in BAM files generated during this step. BAM files were used to ensure that reads were mapped to the same direction and had the same start positions, mapping length, and molecular tag. Bismark_methylation_extractor was used to obtain readable results for downstream analysis.

Quantity and differential expression

Transcript quantification was performed using HTSeq v0.44.0. RNA with zero counts in all samples was removed. Normalization and differential expression were determined using the R package DEseq2 v1.28.1.

Differentially methylated regions

After alignment, we extracted methylation counts from Bismark output; then, DMRfinder was used to cluster CpG sites into test regions for differential methylation.

Sample preparation for nucleotide modification

For DNA and RNA extraction, cells were lysed with TRIzol reagent and digested with proteinase K at 37°C for 30 min. An equal volume of back extraction buffer (4 M guanidine thiocyanate, 50 mM sodium citrate, 1 M Tris, pH 8.0) and 1/10 volume of 3 M sodium acetate (pH 4.0) was mixed with the cell lysate. Samples were then centrifuged at 13,200 rpm for 15 min at 4°C. The upper phase was transferred to a new tube and an equal volume of 100% isopropanol was added. After overnight incubation at -80°C, samples were centrifuged at 13,200 rpm and 4°C for 15 min. The supernatant was removed and pellets were washed twice with 70% ethanol. Samples were eluted with 50 µl nuclease-free buffer and centrifuged at 5,000 rpm for 2 min. The supernatant was then transferred to a new tube. Total RNA was extracted using TRIzol reagent. RNA or DNA/RNA hybrid (500 ng) was digested into nucleotides using 20 µl Nucleoside Digestion Mix (NEB, M0649S). The digestion product was brought to 50 µl with nuclease-free water and filtered using a Microcon-10kDa Centrifugal Filter Unit with Ultracel-10 membrane.

LC-MS/MS

The analysis used an ACQUITY UPLC I-Class system coupled with a QTRAP 6500+ triple quadrupole mass spectrometer (Sciex, Framingham, MA). Nucleotide samples were diluted five times, and 10 µl of each dilution was chromatographed on HSS T3 C18 (100 x 2.1 mm, 1.8 µm, Waters, Milford, MA). Samples were eluted with a linear gradient of A (0.2% formic acid + 5 mM ammonium formate in deionized water) and B (0.2% formic acid + 5 mM ammonium formate in methanol) at a flow rate of 0.3 ml/min and temperature of 45°C. The gradient elution started with 96% of mobile phase A for 2.5 min. Next, mobile phase B was linearly increased up to 45% in 2.5 min and maintained for 0.5 min, then linearly increased up to 100% in 2 min. Finally, the combination was brought back to 96% A in 1 min. Detection was achieved using a QTRAP 6500+ system in positive ion mode and multiple reaction monitoring mode (MRM). Source parameters were set as follows: temperature, 400°C;

curtain gas (CUR), 30; collision gas, low; ion spray voltage, 5000 V; ion source gas (GS1), 45; and drying gas (GS2), 40.

Statistical analysis

Statistical significance was determined by Student's *t* test for two groups. Analysis with a probability value < 0.05 were considered statistically significant.

Reference

[1] Huang K, Zhang Y, Hui X, et al. Q493k and q498h substitutions in spike promote adaptation of sars-cov-2 in mice. EBioMedicine, 67 (2021), p. 103381

Supplementary figures

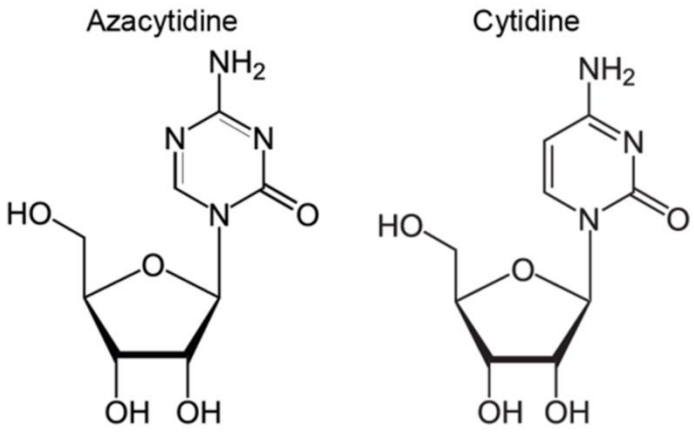


Fig.S1 Molecular structures of cytidine and azacytidine.

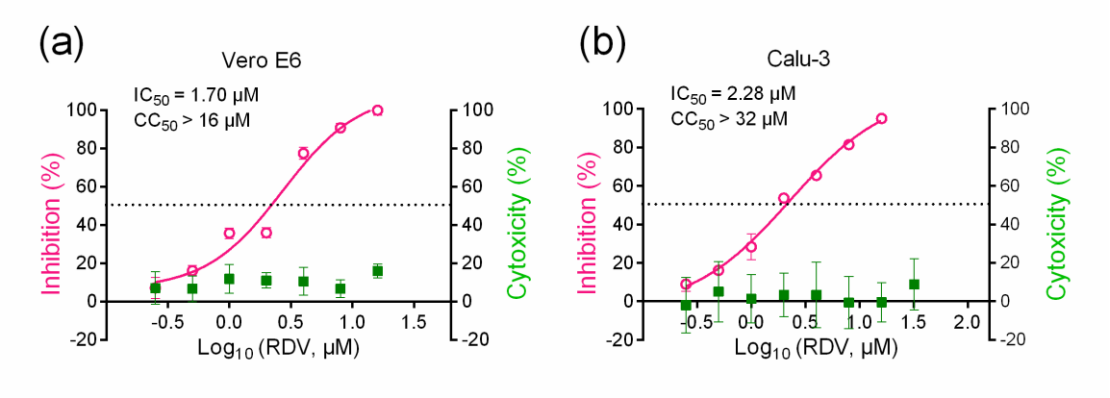
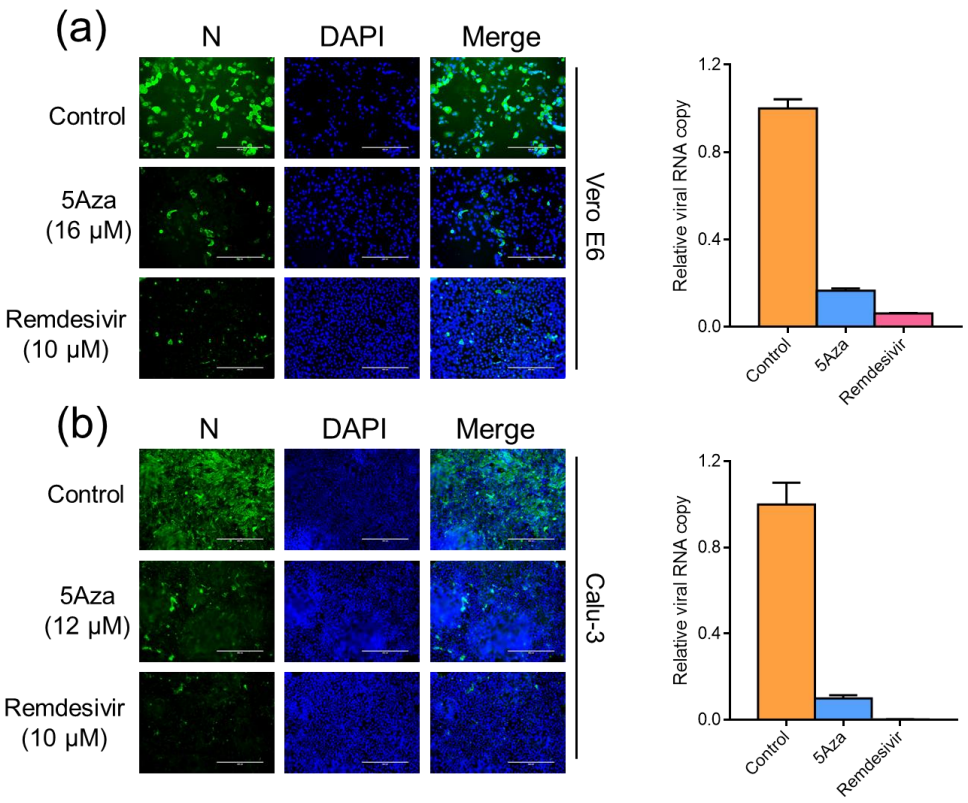


Fig.S2 Remdesivir inhibits SARS-CoV-2 in Vero E6 and Calu-3 cells

Vero E6 (**a**) and Calu-3 (**b**) cells were infected with 0.2 moi SARS-CoV-2, in the presence of various concentration of Remdesivir (RDV). After 24 hpi, supernatant was collected. IC_{50} was calculated by detecting viral RNA copies through qRT-PCR. And cytotoxicity of RDV was determined through CCK-8 assay. Left and right Y-axes represent mean % inhibition of virus and % cytotoxicity of 5Aza, respectively.



206

207 **Fig.S3 Indirect immunofluorescence detecting of SARS-CoV-2.**

208 Vero (**a**) and Calu-3 (**b**) were infected with 0.2 moi SARS-CoV-2 in the presence
209 of indicated drugs. Twenty-four hours later, the supernatant was collected for
210 RNA detection; cells were fixed by 4% paraformaldehyde for
211 immunofluorescence assay of viral N protein. Nuclei were stained with DAPI.
212 Bars, 200 μM. All experiments were performed in triplicate; data are mean ±
213 SD.

214

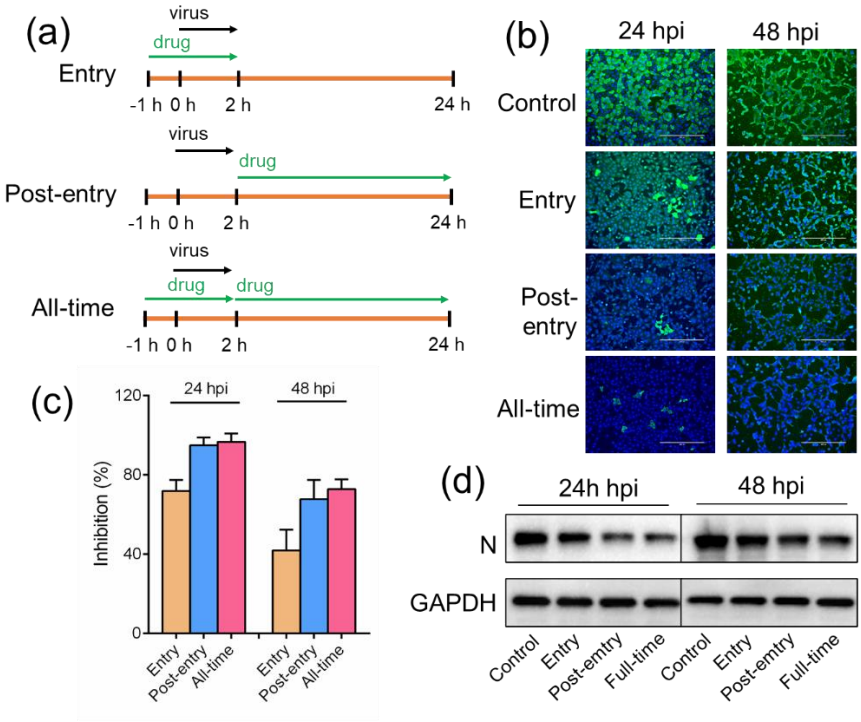


Fig.S4 Time-of-drug-addition experiment of 5Aza.

(**a**) Schematic of time-of-drug-function experiment of 5Aza. Vero E6 cells were treated with 5Aza and then infected with 0.01 moi SARS-CoV-2. (**b**) After 24 and 48 hpi, supernatant was collected for detecting RNA copies using qRT-PCR. (**c**) cells were fixed for immunofluorescence assay of viral N protein, and (**d**) total proteins were extracted for western blotting to detect N protein. Nuclei were stained with DAPI. Bars, 200 μM. Data are mean ± SD.

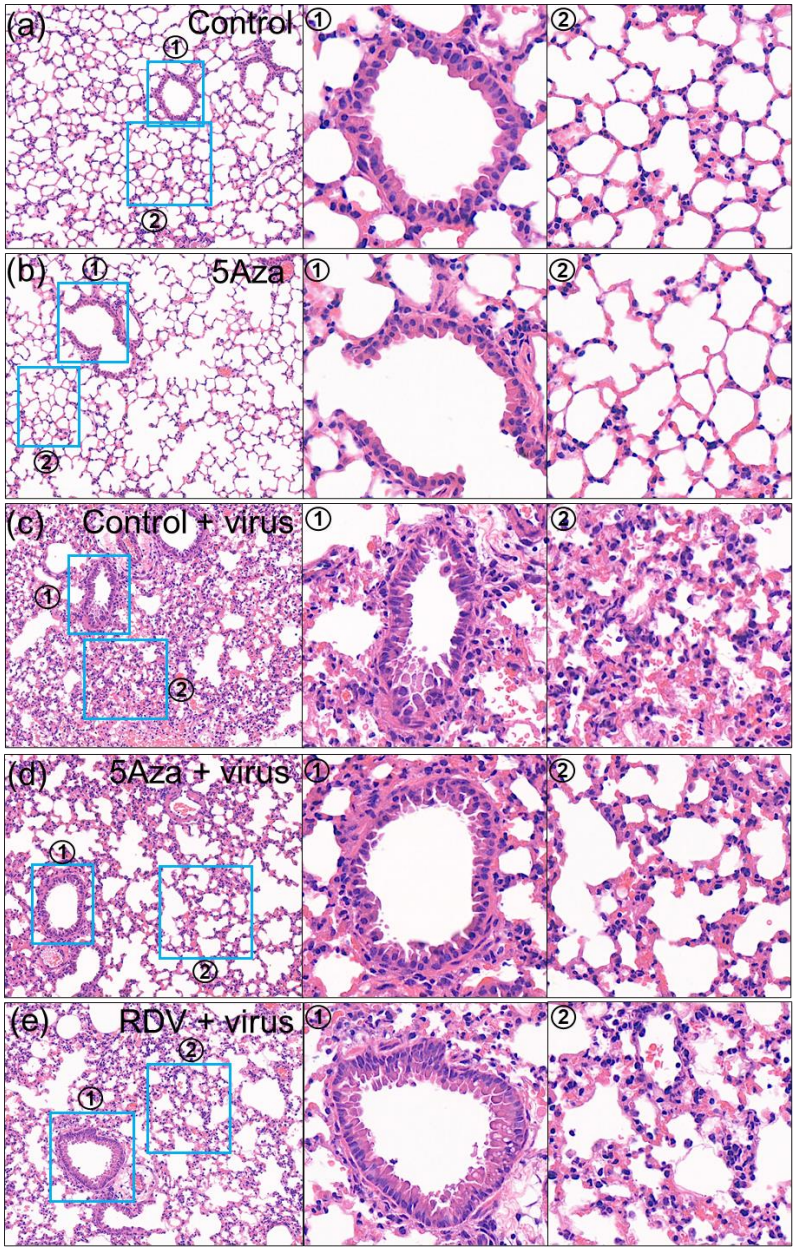


Fig.S5 H&E staining of lung tissues.

At 4 dpi, mice were euthanized, and lungs were collected for histopathological analysis with H&E staining. Numbered blue rectangles are shown in magnified images to the right. (**a**) and (**b**) Lung tissues were normal in the mock infection group. (**c**) Lung tissue section from the control + SARS-CoV-2 group mice displayed (1) bronchiolar epithelium cell death, (2) destruction of alveola with massive infiltration. (**d**) and (**e**) Representative image of lung tissue from 5Aza and Remdesivir treated infection group mice showing alleviated infiltration.

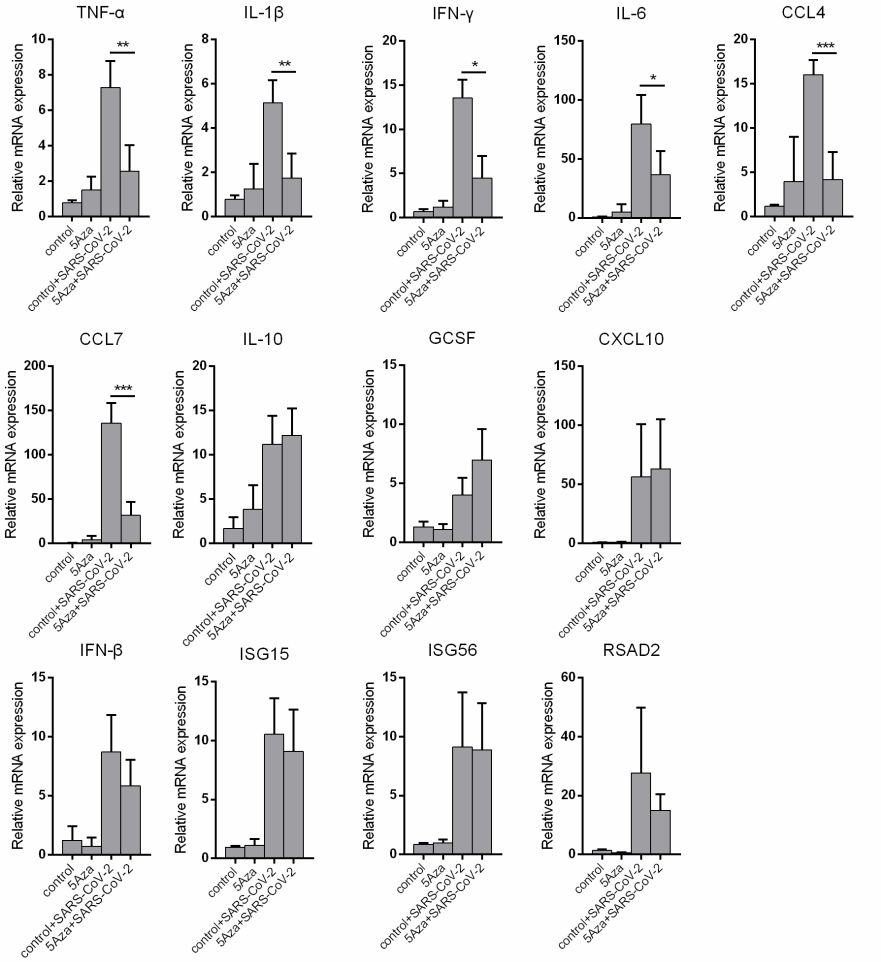


Fig.S6 The cytokines expression in lungs.

Total RNA of lungs was extracted using Trizol. 1 μg RNA was inversely transcribed for detection of indicated genes by qRT-PCR. Gene expression was normalized to that of GAPDH. Data are mean ± SD, and analyzed using Student's *t* test. * $p < 0.05$, ** $p < 0.01$, ** $p < 0.001$.

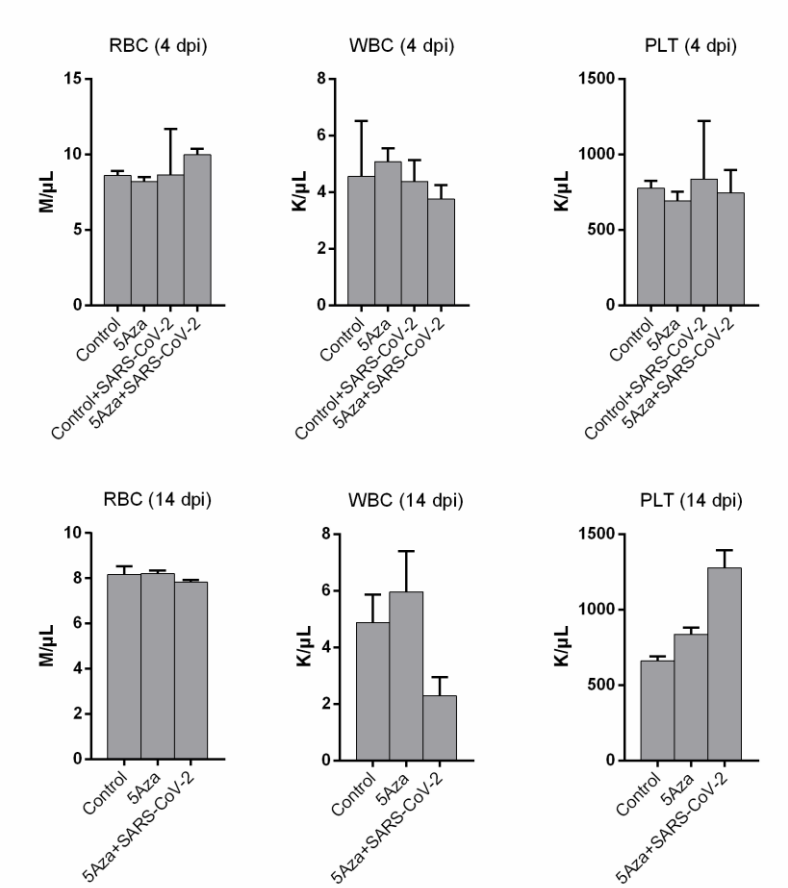


Fig.S7 Detection of RBC, WBC, and PLT in mice blood.

Whole blood from four mice at indicated days was collected with anticoagulation. And Red blood cell (RBC), White blood cell (WBC), and blood platelet (PLT) were determined. At 14 dpi, none mouse survived in Control + SARS-CoV-2 group, and therefore this data is absence.

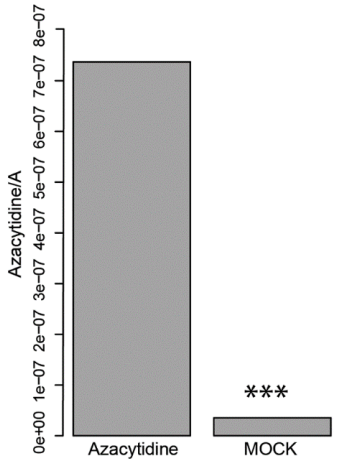


Fig.S8 Profiling and quantification of azacytidine in DNA and RNA using liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Considering azacytidine as an RNA analog with OH-group on ribose 2' carbon (2' C), it could theoretically incorporate into RNAs. “Azacytidine” sample were the cells incubated with 5Aza (10 μM) for 24 h and “Mock” sample were the cells without any treatment. Both DNA and RNA were extracted, followed by single nucleotide digestion (DNA and RNA) and LC-MS/MS. Azacytidine was normalized to adenosine on the y-axis. Data are presented as means of triplicate assays (***$p < 0.001$). Most of the azacytidine incorporated into the RNAs. This was consistent with previous reports, in which radiolabelled azacytidine was measured by scintillation-based quantification or high-resolution mass spectrum termed "AZA-MS".

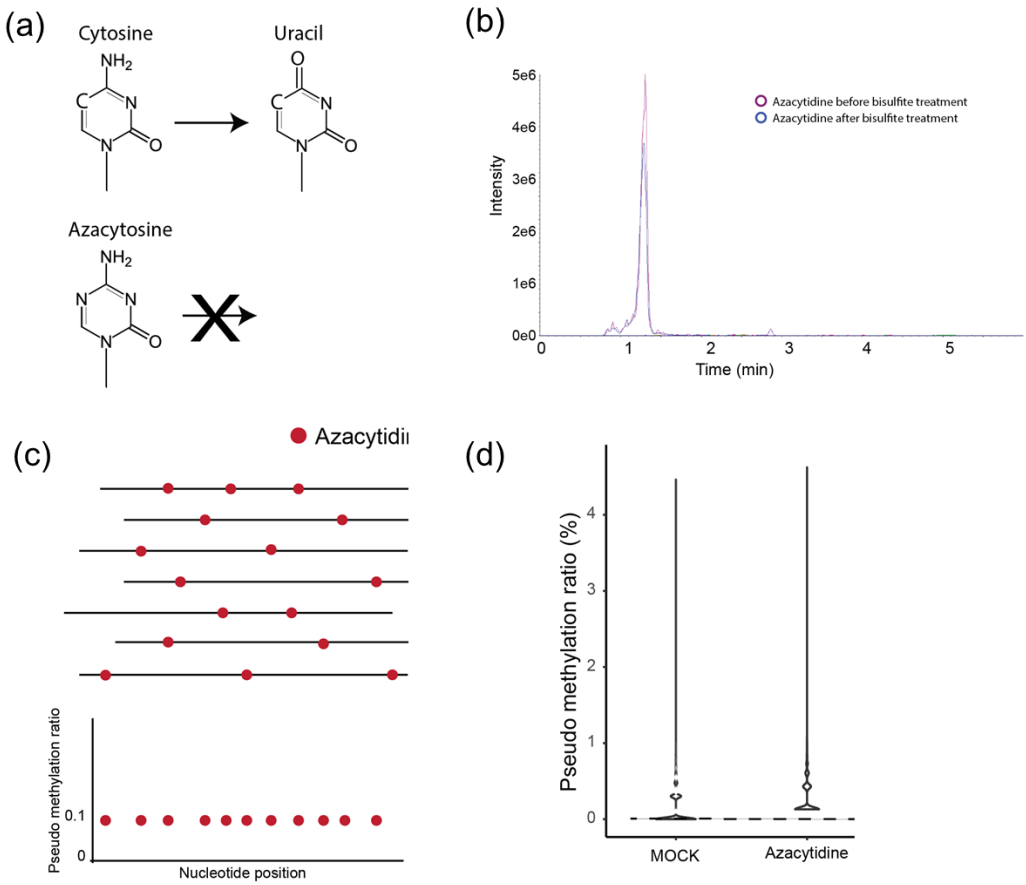


Fig.S9 5Aza-BSseq-identified locations of 5Aza incorporation.

(**a**) We explored whether azacytidine could be converted to uracil under bisulfide treatment and found that was not the case. Bisulfite deamination of cytidine relies on Lewis acid formation and electron transfer from C5. In azacytidine, N5 inhibited this deamination. (**b**) LC-MS chromatograms of 5Aza before and after bisulfite treatment. Supporting the observation in (**a**), LC-MS chromatograms of azacytidine remained unchanged before and after bisulfite treatment. (**c**) Schematic diagram of testing pseudo-methylation sites. This chemical property of azacytidine is similar to the RNA m^5C, which also cannot be converted to uracil in bisulfite sequencing. Unconverted azacytidine may be identified as a pseudo cytidine (pseudo-methylation) site in RNA bisulfite sequencing. Owing to randomized azacytidine incorporation, these azacytidine (pseudo-methylation) sites should be observed as increased ultra-low methylation (methylation ratio <1%). (**d**) The ratio of pseudo-methylation.

Pseudo-methylation distribution analysis revealed that azacytidine treatment increased ultra-low pseudo-methylation (<1%), indicating 5Aza incorporation in RNA.

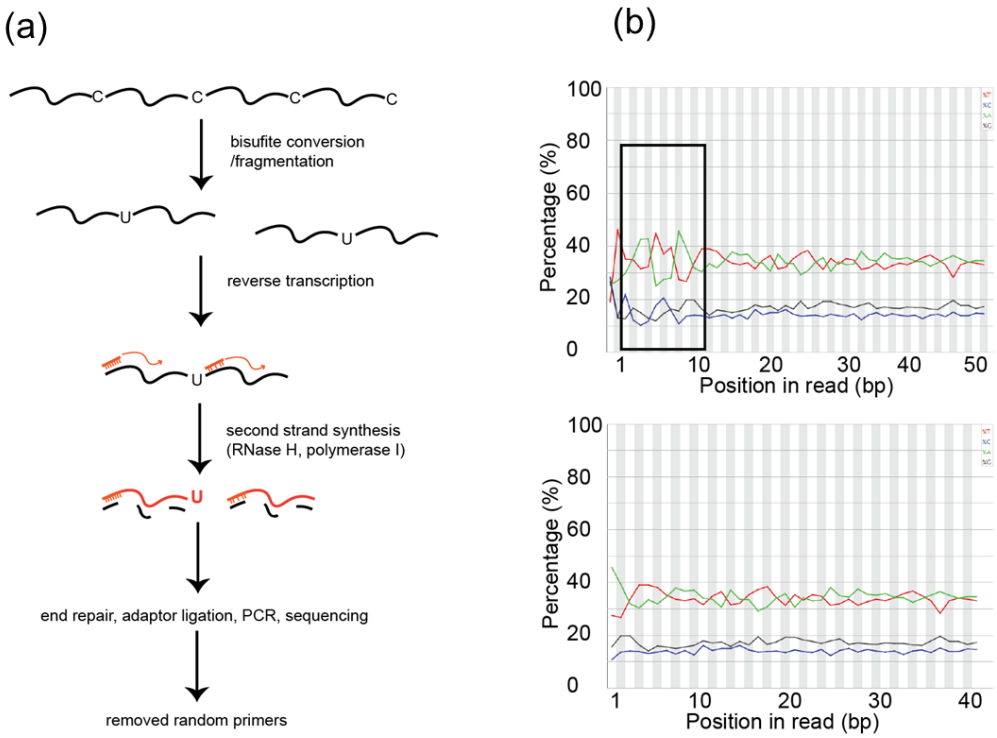


Fig.S10 Experimental flowchart of RNA-BSseq.

Limitations in site-mapping technology hamper the validity of coronavirus RNA methylation. As random priming in RNA bisulfite sequencing introduces mismatch errors, it is difficult to reliably measure m^5C sites. We sought to develop an optimized approach that avoids random primer insertion into sequenced fragments, thus minimizing false positives. (**a**) Without using ligase in second strand synthesis, we could remove random primers in the middle of sequenced fragments. During bioinformatics analysis, we could then easily detect and delete the first six bases affected by random primers initializing reverse transcription. (**b**) Nucleotide density distribution (A, T, C, or G) in sequencing reads. Upper and lower panels show distribution before and after the removal of random primer sequences, respectively. The distribution

fluctuated before and after a smooth line, indicating a few artifact Cs in our modified bisulfite sequencing.

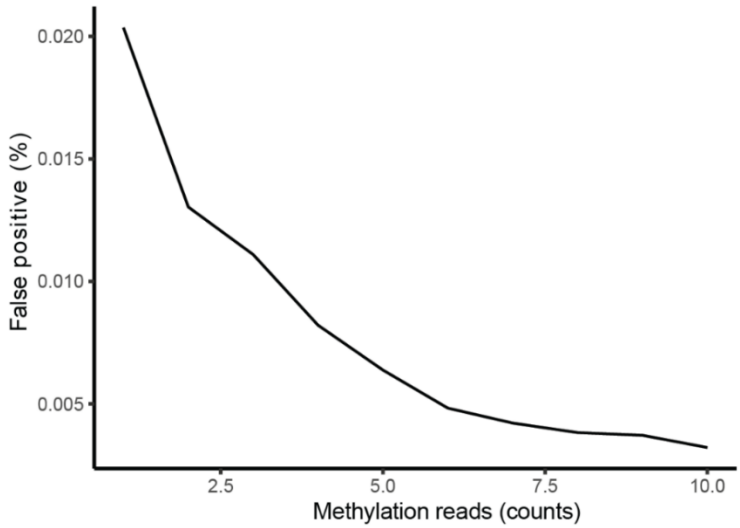


Fig.S11 False-positive rate of ERCC (External RNA Controls Consortium) RNA bisulfite sequencing.

ERCC RNA was synthesized from RNA without m^5C, subjected to bisulfite treatment, and used for library construction. To further enhance the accuracy of RNA-BSseq detection, we used the ERCC to test the false positive rate of our method and to select a suitable cut-off (coverage > 5, methylation ratio > 95%, false positive ~ 0.006).

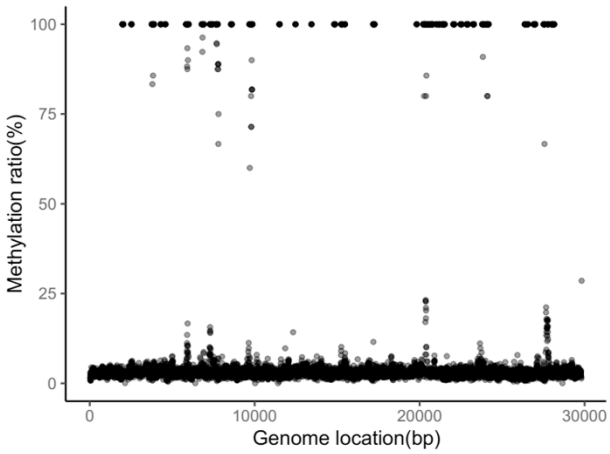


Fig.S12 High confident m^5C methylation sites on a delta variant of SARS-CoV-2

Vero E6 cells were infected with 1 moi delta variant of SARS-CoV-2. 24 hours later, total RNA was extracted. m^5C methylation sites were detected by RNA-BSseq. There were five minimum methylation supporting reads per point. Data are averaged from three technological repeats.

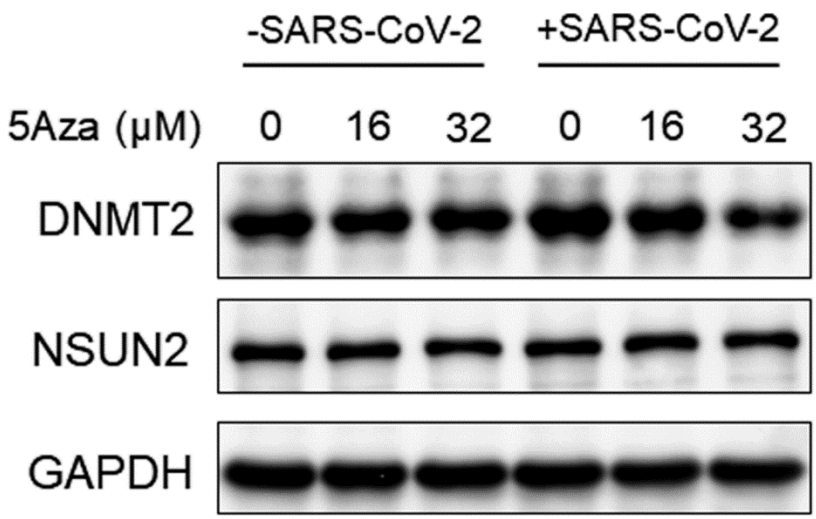


Fig.S13 The effect of 5Aza treatment on the expression of DNMT2 and NSUN2 protein.

Mock or SARS-CoV-2 (moi=0.1) infected Vero E6 cells were treated with indicated concentration of 5Aza. After 24 h, cells were lysed with RIPA buffer, and the lysates were subjected to western blot for detecting indicated proteins.

Supplementary Table 1

Table 1-1 The viral genome RNA mutant analysis comparation to SARS-CoV-2 reference genome after 5Aza treatment

Genome Position	Reference Allele	Alternate Allele	GenoType	REFDepth	ALTDepth	Ref annotation Position	Protein.Position.Amino acids change
10	T	A	T/T	157	29	5'UTR	-
565	T	C	C/C	42	73716	gene-orf1ab	QHD43415.1:p.100S
565	T	G	T/T	42	60	gene-orf1ab	QHD43415.1:p.100S>R
7334	C	T	C/T	30312	47391	gene-orf1ab	QHD43415.1:p.2357H>Y
7334	C	G	C/C	30312	46	gene-orf1ab	QHD43415.1:p.2357H>D
7334	C	A	C/C	30312	6	gene-orf1ab	QHD43415.1:p.2357H>N
17321	C	T	C/T	19976	63384	ene-orf1ab	QHD43415.1:p.5686A>V
17321	C	G	C/C	19976	25	gene-orf1ab	QHD43415.1:p.5686A>G
17825	C	T	T/T	12	79423	gene-orf1ab	QHD43415.1:p.5854T>I
17825	C	A	C/C	12	12	gene-orf1ab	QHD43415.1:p.5854T>N
18844	G	A	G/A	18873	65715	gene-orf1ab	QHD43415.1:p.6194V>I
18844	G	C	G/G	18873	10	gene-orf1ab	QHD43415.1:p.6194V>L
21784	T	A	A/A	308	73557	gene-S	QHD43416.1:p.74N>K
21784	T	G	T/T	308	32	gene-S	QHD43416.1:p.74N>K
21784	T	C	T/T	308	24	gene-S	QHD43416.1:p.74N
23525	C	T	T/T	848	78868	gene-S	QHD43416.1:p.655H>Y
23525	C	A	C/C	848	11	ene-S	QHD43416.1:p.655H>N
23606	C	T	C/T	19736	59756	gene-S	QHD43416.1:p.682R>W
23606	C	G	C/C	19736	24	gene-S	QHD43416.1:p.682R>G
23606	C	A	C/C	19736	6	gene-S	QHD43416.1:p.682R
29573	G	A	A/A	19	73657	gene-ORF10	QHI42199.1:p.6V>I
29573	G	T	G/G	19	6	gene-ORF10	QHI42199.1:p.6V>F

Table 1-2 The viral genome RNA mutant analysis comparation to SARS-CoV-2 reference genome after saline treatment (Control)

Genome Position	Reference Allele	Alternate Allele	GenoType	REFDepth	ALTDepth	Ref annotation Position	Protein.Position.Amino acids change
18	T	G	T/T	1980	166	5'UTR	-
565	T	C	C/C	49	76463	gene-orf1ab	QHD43415.1:p.100S
565	T	G	T/T	49	11	gene-orf1ab	QHD43415.1:p.100S>R
7334	C	T	C/T	35181	48436	gene-orf1ab	QHD43415.1:p.2357H>Y
7334	C	G	C/C	35181	15	gene-orf1ab	QHD43415.1:p.2357H>D
17321	C	T	C/T	22730	62149	gene-orf1ab	QHD43415.1:p.5686A>V
17825	C	T	T/T	10	82113	gene-orf1ab	QHD43415.1:p.5854T>I
17825	C	A	C/C	10	25	gene-orf1ab	QHD43415.1:p.5854T>N
18844	G	A	G/A	21743	66770	gene-orf1ab	QHD43415.1:p.6194V>I
18844	G	C	G/G	21743	18	gene-orf1ab	QHD43415.1:p.6194V>L
18844	G	T	G/G	21743	6	gene-orf1ab	QHD43415.1:p.6194V>L
21784	T	A	A/A	340	80300	gene-S	QHD43416.1:p.74N>K
21784	T	C	T/T	340	24	gene-S	QHD43416.1:p.74N
21784	T	G	T/T	340	24	gene-S	QHD43416.1:p.74N>K
23525	C	T	T/T	995	82974	gene-S	QHD43416.1:p.655H>Y
23525	C	A	C/C	995	10	gene-S	QHD43416.1:p.655H>N
23606	C	T	C/T	22874	61149	gene-S	QHD43416.1:p.682R>W
23606	C	A	C/C	22874	5	gene-S	QHD43416.1:p.682R
29573	G	A	A/A	22	81434	gene-ORF10	QHI42199.1:p.6V>I