

ARTICLE OPEN



Comparative susceptibility of SARS-CoV-2, SARS-CoV, and MERS-CoV across mammals

Meng Li ^{1,6}, Juan Du ^{1,2,6}, Weiqiang Liu ^{1,2}, Zihao Li ^{1,2}, Fei Lv ^{1,2}, Chunyan Hu ^{1,2}, Yichen Dai ¹, Xiaoxiao Zhang ^{1,2}, Zhan Zhang ^{1,2}, Gaoming Liu ¹, Qi Pan ^{1,2}, Yang Yu ³, Xiao Wang ¹, Pingfen Zhu ¹, Xu Tan ⁴, Paul A. Garber ⁵ and Xuming Zhou ^{1✉}

© The Author(s) 2023

Exploring wild reservoirs of pathogenic viruses is critical for their long-term control and for predicting future pandemic scenarios. Here, a comparative *in vitro* infection analysis was first performed on 83 cell cultures derived from 55 mammalian species using pseudotyped viruses bearing S proteins from SARS-CoV-2, SARS-CoV, and MERS-CoV. Cell cultures from Thomas's horseshoe bats, king horseshoe bats, green monkeys, and ferrets were found to be highly susceptible to SARS-CoV-2, SARS-CoV, and MERS-CoV pseudotyped viruses. Moreover, five variants (del69-70, D80Y, S98F, T572I, and Q675H), that beside spike receptor-binding domain can significantly alter the host tropism of SARS-CoV-2. An examination of phylogenetic signals of transduction rates revealed that closely related taxa generally have similar susceptibility to MERS-CoV but not to SARS-CoV and SARS-CoV-2 pseudotyped viruses. Additionally, we discovered that the expression of 95 genes, e.g., *PZDK1* and *APOBEC3*, were commonly associated with the transduction rates of SARS-CoV, MERS-CoV, and SARS-CoV-2 pseudotyped viruses. This study provides basic documentation of the susceptibility, variants, and molecules that underlie the cross-species transmission of these coronaviruses.

The ISME Journal (2023) 17:549–560; <https://doi.org/10.1038/s41396-023-01368-2>

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has posed a considerable threat to public health and the global economy, with more than 500 million cases of human infection and over 6.6 million deaths worldwide as of Jan 2023. SARS-CoV-2 was speculated to have originated in bats and then jump to the human population via an intermediate animal host [1–3]. Although viruses similar to SARS-CoV-2, such as BANAL-20-52 derived from the Malayan horseshoe bat (*Rhinolophus malayanus*), RaTG13 derived from the intermediate horseshoe bat (*Rhinolophus affinis*), and Pangolin-CoV present in Malayan pangolins (*Manis avania*) [4–6] have been identified, at present the exact animal origin of SARS-CoV-2 remains unclear. Two other coronaviruses that caused epidemics prior to COVID-19 are severe acute respiratory syndrome coronavirus (SARS-CoV), which possibly spread from horseshoe bats to humans through infected palm civets, and Middle East respiratory syndrome coronavirus (MERS-CoV), which was possibly a spillover from bats to humans via infected dromedary camels [7–9]. Although the current threat of both SARS-CoV and MERS-CoV is minimal, we must be vigilant about the potential risks of spillback to humans from their natural reservoirs [10, 11].

SARS-CoV-2 is able to infect a broad spectrum of hosts including dogs, mink, ferrets, otters, hamsters, voles, deer, deer mice, bats, small and large felines, and several nonhuman primates [12–14]. However, it remains a challenge to compare the susceptibility of

these species because standardized measurements across species have not been employed. In this regard, several studies have predicted the SARS-CoV-2 infection probability by analyzing angiotensin-converting enzyme 2 (*ACE2*) orthologs. For example, Damas et al. utilized the *ACE2* sequences of 410 vertebrate species and found that certain endangered taxa (e.g., red-shanked doucs, proboscis monkeys, rhesus macaques, and Antarctic minke whales) were at the highest risk for SARS-CoV-2 infection [15]. At the same time, crystal structure resolution, surface plasmon resonance analyses, and molecular dynamic simulations have been used to evaluate the binding affinity of different *ACE2* orthologs to spike proteins [15–17]. While these studies provide valuable information about the likely host range of SARS-CoV-2, their predictions require experimental validation. Moreover, the effects of host factors other than *ACE2* associated with viral invasion have been underappreciated in these predictions [15, 18, 19].

Animal infection experiments provide the best opportunity to understand the susceptibility, pathogenicity, and transmissibility of pathogens across different taxa [20–23]. However, it is impractical to perform *in vivo* inoculation studies in a wide range of species, particularly in wildlife. Alternatively, an *in vitro* infection assay of diverse cell lines has the potential to offer critical insights into the infectivity of SARS-CoV-2 [24, 25]. For instance, Chu et al. assessed the replication kinetics and cytopathic effects of 25 cell lines derived from different species. They demonstrated that

¹Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. ²University of Chinese Academy of Sciences, Beijing 100049, China. ³School of Life Sciences, University of Science and Technology of China, Anhui, China. ⁴Beijing Advanced Center for Structural Biology, Beijing Frontier Innovation Center, School of Pharmaceutical Sciences, Tsinghua-Peking Center for Life Sciences, Tsinghua University, 100084 Beijing, China. ⁵Department of Anthropology, Program in Ecology, Evolution, and Conservation Biology, University of Illinois, Urbana, IL, USA. [✉]These authors contributed equally: Meng Li, Juan Du. [✉]email: zhouxuming@ioz.ac.cn

Received: 15 August 2022 Revised: 10 January 2023 Accepted: 12 January 2023
Published online: 23 January 2023

SARS-CoV-2 can replicate in non-human primate, cat, rabbit, and pig cells [24]. In another study, airway epithelial cells were collected from 12 animal species, and SARS-CoV-2 was found to replicate efficiently in monkey and cat culture models [25]. Aside from the demand for cellular assays to assess the potential host spectrum, it also is urgent to explore how mutations can affect both the infectivity and transmissibility of SARS-CoV-2 to different species. The continuous adaptive evolution of SARS-CoV-2 has resulted in the rapid emergence of novel mutations; however, it remains unknown how many of these mutations have enhanced the susceptibility of animals to SARS-CoV-2.

Here, we first assessed the potential host range of SARS-CoV, MERS-CoV, and SARS-CoV-2 using pseudotyped viruses and cell cultures derived from 55 mammalian species. The ability of site mutations in S proteins to affect the host range of SARS-CoV-2 was determined using site-directed mutagenesis. We then employed a comparative transcriptomics approach to identify gene expressions that were associated with species' susceptibility to these viruses. This study provides new information about the plausible host tropism of SARS-CoV, MERS-CoV, and SARS-CoV-2. Moreover, we uncovered the expression of host factors that are likely to affect the cross-species transmission of SARS-CoV-2 and prevent the future spillback of these coronaviruses from humans to other species.

MATERIAL AND METHODS

Cell culture models

In order to improve the reproducibility of this study, whenever possible, we sampled healthy young adult males of each species to reduce the effects of sex, age, immunity, and related factors. To recapitulate cells with the highest susceptibility, we isolated primary cell cultures used in this study (Table S1) from multiple tissues (i.e., kidney, lung, brain, spleen, and heart) of pets, livestock, and wildlife animals based on protocols previously described [26]. Animals were euthanized using CO₂ or pentobarbital calcium (200 mg/kg). The carcasses were dissected, followed by the collection of kidneys, lungs, hearts, spleens, or brains under aseptic conditions. To avoid drying, all tissues were placed in sterile PBS supplemented with a 2% Penicillin–streptomycin solution (Gibco, USA) until cell extraction. Each tissue was transferred into a 35 mm dish and minced into tiny pieces using dissecting scissors. Tissue fragments were further transferred into a 50 mL conical tube, followed by enzyme digestion using 0.25% EDTA-trypsin (Gibco, USA) at 37 °C for 30 min. During the digestion, the mixture was shaken vigorously every 10 min. Trypsin digestion was stopped by the addition of FBS into the conical tube. The mixture containing tissue fragments was pipetted up and down for 3 min using a 5 mL pipettor to break up the clumps. The resulting solution was spun down at 250 g for 5 min at 4 °C. Pellet cells were then collected, resuspended, counted, and seeded into 100 mm Petri-dishes. All dishes were placed at 37 °C with 5% CO₂. The cells were checked daily for cell growth and contamination. Regular cell passages were performed when the cells reached confluence. All primary cells were isolated and cultured in DMEM-F12 with 10% FBS, a 1% penicillin–streptomycin solution, and 20 mM HEPES. All experiments were conducted in a Biosafety Level 2 (BSL-2) facility and approved by the Animal Ethics Committee of the Institute of Zoology, Chinese Academy of Sciences (permission no: IOZ-IACUC-2021-163).

Cell lines including Huh-7 (*Homo sapiens*, liver), A549 (*Homo sapiens*, lung), SW480 (*Homo sapiens*, colon), HEK-293T (*Homo sapiens*, embryonic kidney), Vero-E6 (*Cercopithecus aethiops*, kidney), Marc-145 (*Cercopithecus aethiops*, kidney), OK (*Monodelphis domestica*, kidney), BHK-21 (*Mesocricetus auratus*, kidney), SP2/O (*Mus musculus*, myeloma cells), NIH3T3 (*Mus musculus*, embryo), MDBK (*Bos taurus*, kidney), PK-15 (*Sus scrofa*, kidney), MDCK (*Canis familiaris*, kidney), F81 (*Felis catus*, kidney), and Mv.1.lu (*Neovison vison*, lung) cells were cultured in DMEM. HT-29 (*Homo sapiens*, colon) cell lines were cultured in RPMI-1640 (Gibco, USA). NEF (Naked-mole rat, embryo) cell lines were cultured in DMEM-F12. Immortalized bat cell lines PaKi (*Pteropus alecto*, kidney), PaLu (*Pteropus alecto*, lung), and PaBr (*Pteropus alecto*, brain) were kindly provided by Dr. Linfa Wang (Duke-NUS Medical School), and cultured in DMEM-F12. All cell lines were incubated at 37 °C in the presence of 5% CO₂ using specified culture mediums that were supplemented with 10% fetal bovine serum (FBS, Gibco, USA), a 1%

penicillin–streptomycin solution, and 20 mM hydroxyethylpiperazine ethane sulfonic acid (HEPES, Gibco, USA).

Plasmids and site-directed mutagenesis

pVSV-eGFP-dG (Addgene #31842), pCAG-VSVN (Addgene #64087), pCAG-VSVP (Addgene #64088), pCAG-VSVL (Addgene #64085), pCAG-VSVG (Addgene #64084), and pCAG-T7pol (Addgene #59926) plasmids were acquired from the Addgene. Human codon-optimized S genes of SARS-CoV (SARS coronavirus Tor2, GenBank accession No. NC_004718.3), SARS-CoV-2 (SARS coronavirus 2 Wuhan-Hu-1, GenBank accession No. NC_045512.2), and MERS-CoV (MERS coronavirus HCoV-EMC, GenBank accession No. NC_019843.3) were synthesized and cloned into a pcDNA3.1 vector for pseudotyped virus generation. The constructed plasmids were named pcDNA3.1-SARS-S, pcDNA3.1-SARS2-S, and pcDNA3.1-MERS-S, respectively. The expression of S proteins in HEK-293T cells was verified using Western Blot. Site-directed mutagenesis was performed using pcDNA3.1-SARS2-S as a template. Forward primers were designed with *Tm* of ~75 °C and the mutations were centered in the middle, while correspondence reverses complementary sequences were selected as reverse primers. A volume of 20 µL PCR mix was prepared to contain 10 µL 2× Phanta Max Master Mix (Vazyme, PR China), 10 pmol of forward and reverse primers, and 10 ng of template plasmid. After 20 cycles of site-directed mutagenesis PCR, the template plasmid was digested using DpnI restriction endonuclease (NEB, USA). Afterward, the PCR product was transformed to *E. coli* DH5a competent cells (Transgen, PR China). Single clones were selected and sequenced. All plasmids were extracted using QIAGEN Plasmid Maxi Kit (Qiagen, USA). To avoid the introduction of contamination in cell culture during the generation of pseudotyped viruses, extracted plasmids were inactivated in a 65 °C water bath for 30 min. The concentrations of plasmids were quantified using Nanodrop2000 (Thermo Scientific, USA).

Pseudotyped virus package

First, we rescued the VSVΔG*-G virus based on methods developed in a previous publication [27]. In detail, HEK-293T cells with 80% confluence were co-transfected with pVSV-eGFP-dG, pCAG-VSVN, pCAG-VSVP, pCAG-VSVL, pCAG-VSVG, and pCAG-T7pol plasmids with a ratio of 10, 3, 5, 1, 3, and 5 µg, respectively, using Lipofectamine 3000 (Invitrogen, Lithuania), following the manufacturer's guidance. The supernatant fluid was harvested 48 h after transfection, and one-half of the fluid was assigned to infect a second plate of HEK-293T cells transfected with pCAG-VSVG plasmid. The HEK-293T cells were examined by fluorescence microscopy for EGFP expression 24 h after infection. Titration of the VSVΔG*-G virus was quantified by the infection of HEK-293T cells, followed by flow cytometry analysis for an EGFP positive signal.

Second, we constructed pseudotyped viruses incorporated with S proteins from SARS-CoV, MERS-CoV, and SARS-CoV-2 using a protocol reported in previous research [28]. HEK-293T cells with 70–90% confluence were transfected with 15 µg of plasmids encoding SARS-CoV, MERS-CoV, and SARS-CoV-2 S proteins using Lipofectamine 3000. Simultaneously, plasmid encoding VSV-G proteins were transfected or mock-transfected into 293 T cells as a positive or negative control. The transfected cells were cultured at 37 °C with 5% CO₂. 24 h later, the cells were infected with VSVΔG*-G viruses at MOI = 3–4 and incubated for 2 h. Afterward, the cell supernatant was discarded, and cells were gently rinsed with warm PBS twice. Then, 10 ml of fresh DMEM supplemented with 10% FBS, a 1% Penicillin–Streptomycin solution, and 20 mM HEPES was added to the dishes, and incubated at 37 °C with 5% CO₂. 24 h later, these cells were checked for an EGFP positive signal, and the supernatant fluid was collected, centrifuged, filtered, and divided into aliquots. All of the pseudotyped viruses were stored at –80 °C. Repeated freezing-thawing cycles were avoided.

Huh-7 cells were used for the quantification of generated pseudotyped viruses. In detail, Huh-7 cells were cultured one day before quantification. Once 80% confluence was reached, the cells were digested, quantified, and seeded into a 48-well plate. After overnight incubation, the cells were infected by a series of 100, 50, 25, 12.5, 6.3, 3.2, and 1.6 µL of pseudotyped viruses with three replicates. The infected cells were placed at 37 °C with 5% CO₂. 24 h post-infection, these cells were examined using a fluorescence microscope, followed by flow cytometry quantification for EGFP-positive cells. Pseudotyped viruses were titrated (transducing units, TU) by counting individual cells infected by the pseudotyped viruses. The final concentration of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS was 4.85 × 10⁶ TU/mL, 5.54 × 10⁶ TU/mL, and 1.14 × 10⁷ TU/mL, respectively.

Infection of cell cultures by VSV pseudotyped viruses

To enhance the reproducibility of this research, P3 and P4 passages of primary cell cultures were used for in vitro infection of pseudotyped viruses. Cell cultures were seeded into 48-well plates (10,000 cells per well), cultured in DMEM-F12 supplemented with 10% FBS, a 1% penicillin–streptomycin solution, and 20 mM HEPES. When the cells reached 70% confluence, culture supernatants were discarded. The cells were rinsed twice using a warm PBS buffer supplemented and a 1% penicillin–streptomycin solution. The cell cultures were infected by pseudotyped viruses, including pseudotyped viruses bearing wild-type and mutated S proteins, VSV-G glycoproteins, and negative control (culture supernatant harvested from mock-transfected cells). Each infection experiment had three independent replicates. The viruses were maintained in the culture medium supplemented with 1% FBS and a 1% penicillin–streptomycin solution for 24 h, followed by microscope observation and flow cytometry analysis for GFP positive cells. For VSVΔG*-SARS and VSVΔG*-SARS2, cell cultures were infected at MOI = 0.15 (referred to Huh-7 cells). For VSVΔG*-MERS, cell cultures were infected at MOI = 0.3 (referred to Huh-7 cells). No correlation was observed between the transduction rates of VSVΔG*-noG and VSVΔG*-SARS2 ($\text{cor} = -0.0637$, $p > 0.05$), between VSVΔG*-noG and VSVΔG*-SARS ($\text{cor} = -0.0324$, $p > 0.05$), and between VSVΔG*-noG and VSVΔG*-MERS ($\text{cor} = -0.0723$, $p > 0.05$).

Because the infection dosage that transduces the same percentage of cells doesn't reflect the same number of viral particles, and the same amount of viral particles of pseudotyped viruses bearing different mutated S proteins does not infect cells with the same efficiencies, it is difficult to adjust pseudotyped viruses into the same infection dosage when cells are infected by pseudotyped viruses bearing mutated SARS-CoV-2 S proteins. And, although we attempted to quantify pseudotyped viruses by RT-qPCR, we found the results contained high systematic errors due to 1) the lack of reasonable reference genes; 2) the fact that the efficiencies of RNA extraction vary among different viruses; 3) the error induced by qPCR technologies; and 4) the yield concentration of pseudotyped viruses was not ideally stable. Therefore, we used a volume of 75 μL pseudotyped viruses to infect each cell culture in each cell infection test.

Flow cytometry analysis

Cells infected or mock-infected with pseudotyped viruses were observed by a fluorescent microscope. After the culture medium was discarded, cells were rinsed twice in warm PBS, followed by 0.25% EDTA-trypsin digestion for 5 min. Then, the cells were collected into 1.5 mL tubes. The cells were washed twice with PBS to remove trypsin, and filtered using a 70 μm cell strainer (BD Falcon, USA) to remove clumps of cells. The final cells were placed on ice in a dark chamber. Flow cytometry analyses were performed (CytoFLEX, Beckman Coulter, USA). Cells infected with VSVΔG*-G viruses or negative control were used to optimize the voltage for FSC, SSC, and FITC detectors, and to quantify the cell populations for positive or negative particles. For each infection assay, at least 5000 total cells were inputted. The percentage of transduction rates was displayed as mean \pm SD based on three independent replicates.

Molecular dynamic simulations and interaction prediction

The interface between the RBD of the S protein of SARS-CoV-2 and ACE2 orthologs from different species was examined by molecular dynamic (MD) simulations using Amber 20 [29–31]. We first downloaded the amino acid sequences from the NCBI (National Center for Biotechnology Information) database or our generated RNA-seq dataset. The RBD-ACE2 crystal structure (PDB ID: 6MOJ) was downloaded from the Protein Data Bank [32], and served as template to prepare 3D models for SARS-CoV-2 S-RBD and ACE2 orthologs on SWISS-MODEL [33]. Next, each system was checked by leap and solvated in a cubic periodic box of TIP3P water extended by 10 Å from the solute. The system was neutralized using a rational number of counter ions of Na^+ or Cl^- , followed by parameterization using an Amber ff14SB force field [34]. Next, 10,000 steps of energy minimization including 5000 steps using the steepest descent method and 5000 steps of conjugate gradient minimization were performed. Each system was then heated to 300 K in the NVT ensemble by 0.2 ns. The minimization, heating, and equilibrium simulations were performed with strong constraints (500 kcal/mol/Å²) on heavy atoms with the sander program in Amber20. The 30 ns MD simulations were performed under constant temperature at 300 K with NPT ensemble and *pmemd.cuda*. The system temperature was maintained using Langevin dynamics. The cutoff distance between the Van der Waals energy and short-range electrostatic energy was 10 Å. The particle mesh Ewald method was used to calculate long-range electrostatic interactions. At least 3000 snapshots were extracted from the equilibrium trajectory for the final average structure of each RBD-ACE2 complex. The

MM/GBSA method was used to calculate the binding free energy (ΔG), and the binding free energy was decomposed into the energy contribution of each residue [35].

We explored whether any spike variant interacted with a specific ACE2 residue based on the transduction rate across species. For each spike variant, the cell lines/species were classified into two groups (high and low infected groups, respectively) using a k-means clustering method. Then, at each position of 42 key residues of ACE2 estimated by MD simulation, the amino acid with the highest number was identified as the major residue for that position, and its frequency was calculated. Lastly, a Fisher's exact test was employed to quantify whether the major residue was evenly distributed between the high and low-infected groups. A residue with a Fisher's exact test p value less than 0.05 indicates that the frequency of the residue significantly differed between high and low infection groups. This implies that it may have interacted with a spike variant.

RNA-seq sequencing, orthologous gene sets, and gene expression profiles

To further improve the reproducibility of this study, we performed RNA-seq analysis for cell cultures that were generated. Total RNAs of the cell cultures that were used for in vitro infection analysis were isolated using TRIzol reagent according to the users' guide (Invitrogen, USA). RNA quality and concentration were assessed using an Agilent 2100 Bioanalyzer (Lexington, USA) prior to library construction. For each RNA sample, a library size of 150 bp was constructed using the library preparation kit NEBNext Ultra RNA Library Prep Kit for Illumina® (NEB, USA). Paired-end sequencing was performed using Novogene Co. Ltd on the "Illumina" NovaSeq 6000 platform. This generated ~2.7 billion reads. We used `IIQC_PRL.pl` in the package "NGSQCtools" to remove low-quality reads of raw data [36]. To generate the species-specific ortholog sets and calculate expression values, we downloaded the reference genomes of species (if available) from the public database. For those without a reference genome, Trinity was used to perform de novo assembly of transcripts [37] based on previously generated liver, kidney, and brain transcriptome to assemble transcripts [38]. The fixed parameter k-mer size was 25, the minimum contig length was 200, and the paired fragment length was 500. We then used "cd-hit-est" to process the assembled transcripts from different tissues of the same species, cluster the sequences with 90% similarity, and leave the longest transcript in each cluster [39, 40]. Finally, Augustus was employed to perform gene prediction on the de-redundant transcripts and to obtain GTF annotation files.

To construct the human reference sequence, we used "gffread" in the "cufflink" package to extract the human CDS sequence, filtered out incomplete ORF transcripts and pseudogene transcripts, and extracted the longest transcript for each gene [41]. BLAST was used to remove highly repetitive and highly similar genes with e -value $< 10^{-6}$ and Identity $> 90\%$ as the filtering threshold [42]. The mapping rates across the dataset ranged from 9.63% to 66.58%. Finally, 18,552 unique protein-coding genes were obtained as reference sequences.

The longest transcript of each gene was extracted and reciprocal "BLAST" was performed with the protein sequences from the human samples [43]. The filtering threshold was set at 10^{-6} for e -values and 30% for identity. Two genes that were best aligned with each other were defined as orthologous genes. Because the complete genome and the de novo genome are quite different when compared, we used the CDS sequence of orthologous genes as the reference genome and generated annotation files in GTF format for RNA-seq data mapping. STAR was used to construct an index based on the sequence size of the orthologous gene set and the read length of different species. We used the default parameters to align the RNA-seq data with the reference genome [44]. We used "featureCounts" in the Subread software package to count reads, and eliminate multiple-matched reads [45].

Finally, we calculated the library size of each sample as a normalization factor. The R software package "edgeR" was used to normalize the library size and gene length (based on humans) by $\log_2(\text{RPKM-TMM} + 1)$ [46, 47]. The sample information, mapping rates, list of orthologs, gene length, and expression levels are available in Tables S12–S17.

Regression analysis and gene set enrichment

In order to identify genes whose expression is associated with the transduction of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS, a generalized linear model was used to test the correlation between gene expression and transduction rates. If one orthologous gene was not identified in a particular species, that gene was assumed to not be expressed in that species (gene expression value = 0). We filtered out

genes that were expressed in less than three samples. The ordinary least squares, Brownian model, and Ornstein-Uhlenbeck model were used to determine the best correlation in the transduction rate of each gene. Because a phylogenetic signal was not detected for the transduction rates of VSVΔG*-SARS2 and VSVΔG*-SARS, phylogenetic relationships were not considered in the regression analyses of transduction rates of the two pseudotyped viruses. However, a two-step method was employed in all regression analyses to correct the p value [47, 48]. In the first step, the species that had the greatest impact on the slope (i.e., potential outliers) was removed by the residuals, and then the regression was performed again. We defined the resulting p value as p_{robust} to remove the influence of the outliers on the regression. The second step was to repeat the regression process for the remaining species and remove one of remaining species each time until all remaining species were removed once. We then scored the largest (least significant) p value in the process as p_{max} to remove the impact of species on the regression. The cutoff of significant genes was set at $p_{robust} < 0.01$ and $p_{max} < 0.05$. Gene enrichment analyses were performed using DAVID 2021 updated (Database for Annotation, Visualization and Integrated Discovery) [49]. The parameter for EASE was 0.3 with Fisher's exact test.

Statistical analysis

R was used for plotting and statistical analysis; the values were expressed as mean \pm SD. When cell cultures were infected by the VSVΔG*-SARS2 bearing the wildtype S protein, the relative transduction rates were classified into four categories (i.e., minimal-, slight-, moderate-, and efficient transduction) based on the following principles. First, we listed the quantile of all transduction rates and calculated the interquartile range (IQR). Then, we calculated the standard deviation (SD) of the transduction rates ranging from minimum to Q1. This was named SD_{Q1} . A cell infection with a transduction rate of less than ten times SD_{Q1} was classified as a negligible infection; a cell infection with a transduction rate between $10 * SD_{Q1}$ to Q3 was classified as a slight infection; a transduction rate between Q3 to $Q3 + 1.5 * IQR$ was classified as a moderate infection; and a transduction rate of more than $Q3 + 1.5 * IQR$ was classified as an efficient infection. This principle also was applied to evaluate the infection results of cell cultures by VSVΔG*-SARS and VSVΔG*-MERS. When cell cultures were infected by VSVΔG*-SARS2-Smut pseudotyped viruses bearing mutated S proteins from SARS-CoV-2, the infection results of each VSVΔG*-SARS2-Smut were normalized to the relative transduction results of A549 cells. The outliers were postulated as mutations that significantly enhanced the tropism of SARS-CoV-2.

RESULTS

In vitro infection assay and transduction rate of VSVΔG*-SARS2

To assess susceptibility to SARS-CoV-2, SARS-CoV, and MERS-CoV, GFP-encoding pseudotyped viruses bearing relevant S proteins, named as VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS, respectively, were generated (Fig. S1). Then, we collected cell cultures derived from 55 mammalian species belonging to 12 orders, including Diprotodontia (one species), Didelphimorphia (one species), Hyracoidea (one species), Scandentia (one species), Primates (three species), Lagomorpha (one species), Rodentia (eight species), Eulipotyphla (one species), Perissodactyla (two species), Artiodactyla (three species), Carnivora (seven species), and Chiroptera (26 species) (Fig. 1, Fig. S1). A total of 83 cell cultures, including 64 primary cell cultures, three immortalized cell cultures, and 16 cell lines were obtained (Table S1). Of these cell cultures, 56 were derived from the kidneys in which ACE2s were abundantly expressed; 10 were derived from the lungs; seven were derived from the heart; and the 10 remaining cell cultures were derived from other tissues, such as the embryo, colon, and spleen (Fig. 1, Table S1). These cell cultures were infected by pseudotyped viruses and subjected to flow cytometry analysis of GFP-positive signals to quantify their transduction rates (Fig. S1). Overall, the cell cultures derived from the kidneys showed relatively higher transduction rates than those from other tissues (Fig. S2A). No significant differences in transduction rates to pseudotyped viruses were observed between primary cell cultures and permanent cell lines (Fig. S2B).

Upon infection with VSVΔG*-SARS2, the 83 cell cultures displayed different sensitivities, with an overall transduction rate of 0–26.7% (Fig. 1, Fig. S3, Table S1). Forty-four cell cultures were minimally transduced by VSVΔG*-SARS2 (transduction rate of <1.2%). A total of 32 cell cultures were slightly (transduction rate of 1.2–3.8%) or moderately (transduction rate of 3.8–8.8%) transduced by VSVΔG*-SARS2, including cell cultures derived from dogs (BcKi, BcLu, and MDCK), cats (F81), ferrets (MpuHe), raccoon dogs (NpKi), minks (Mv.1.Lu), tree shrews (TbKi), and pigs (PK-15) (Fig. 1, Fig. S3). This is consistent with the fact that raccoon dogs, dogs, cats, minks, ferrets, tree shrews, and hamsters are known to be susceptible to SARS-CoV-2 with asymptomatic to moderate clinical signs (Table S2) [22, 50]. We found that VSVΔG*-SARS2 can infect PK-15 cells (4.9%) (Fig. 1), but inconsistent conclusions were drawn by previous experimental infections. Several studies showed that pigs were resistant to SARS-CoV-2 [23, 51, 52], while other studies showed that SARS-CoV-2 could infect pigs and be transmitted to co-housed naive sentinel pigs [53, 54]. Cell cultures from several pets, such as the heart (OdHe), lung (OdLu), and kidney (OdKi) cells of the common degu (*Octodon degus*, a *hystricomorph rodent*) and heart (GkHe) cell cultures of Kellen's dormouse (*Graphiurus kelleni*, a *glirid rodent*) were moderately transduced by VSVΔG*-SARS2 (Fig. 1). Mice were found to be resistant to the SARS-CoV-2 infection; however, cell cultures from mice (NIH3T3 and SP2/0) were slightly transduced by VSVΔG*-SARS2 (Fig. S3, Table S1). Given that SARS-CoV-2 can adapt by serial passaging in the respiratory tract of aged BALB/c mice [55, 56], we surmised that animals who were not susceptible to SARS-CoV-2 could be potential hosts following viral evolution during repeated exposures.

Lastly, seven cell cultures derived from Thomas's horseshoe bats (RtKi), king horseshoe bats, humans, palm civets, Africa green monkeys, and ferrets were found to be highly susceptible to VSVΔG*-SARS2 (transduction rate of >8.8%). Among them, humans, palm civets, and ferrets were identified as susceptible hosts for SARS-CoV and SARS-CoV-2 based on in vivo analyses (Table S2) [57]. In particular, 14.0% of Huh-7 (human cell line), 11.4% of Vero-E6 (green monkey cell line), 11.9% of Marc-145 (green monkey cell line), and 10.0% of MpuKi.2 (ferret kidney cell line) were transduced, which confirmed that these cell lines are highly susceptible to SARS-CoV-2, as found in previous studies [28, 58]. We noted that palm civet kidney cells (PKi) were highly susceptible to VSVΔG*-SARS2, as palm civets have been recognized as one of the replication hosts of SARS-CoV [8, 59] (Fig. 1, Fig. S2). Importantly, ferret (*Mustela putorius furo*) kidney cells (MpuKi.2) also were highly susceptible to SARS-CoV-2. Experimental studies have shown that ferrets are susceptible to SARS-CoV-2 [23, 51, 60–62]. Primary cells cultures derived from Thomas's horseshoe bat and the king horseshoe bat were found to be more sensitive to VSVΔG*-SARS2 than the human Huh-7 cell line (Fig. 1, Table S1), suggesting they may serve as likely hosts for SARS-CoV-2.

Host tropism of VSVΔG*-SARS and VSVΔG*-MERS

The maximum likelihood estimates of Pagel's λ for the transduction rates of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS were 6.74×10^{-5} (phylogenetic signal test: $\log_{\lambda} = -163.95$, $p > 0.05$), 6.33×10^{-5} ($\log_{\lambda} = -192.22$, $p > 0.05$), and 0.74 ($\log_{\lambda} = -174.13$, $\log_0 = 1475.26$, $p = 0.09 \times 10^{-2}$), respectively, indicating that closely-related taxa generally have similar susceptibility to VSVΔG*-MERS but not to VSVΔG*-SARS and VSVΔG*-SARS2. Consistently, VSVΔG*-SARS showed a more similar tropism profile with VSVΔG*-SARS2 ($R^2 = 0.61$, $p < 0.001$, phylogenetic generalized least squares test) than with VSVΔG*-MERS ($R^2 = 0.29$, $p < 0.001$, phylogenetic generalized least squares test). In particular, the 83 cell cultures infected by VSVΔG*-SARS displayed a transduction rate of 0–38.4% (Fig. 1). Overall, 38 cell cultures were minimally (transduction rate of 0–1.1%), 23 slightly (1.1–3.8%), 10 moderately (3.8–8.9%), and 11 highly (8.9–38.4%) transduced by

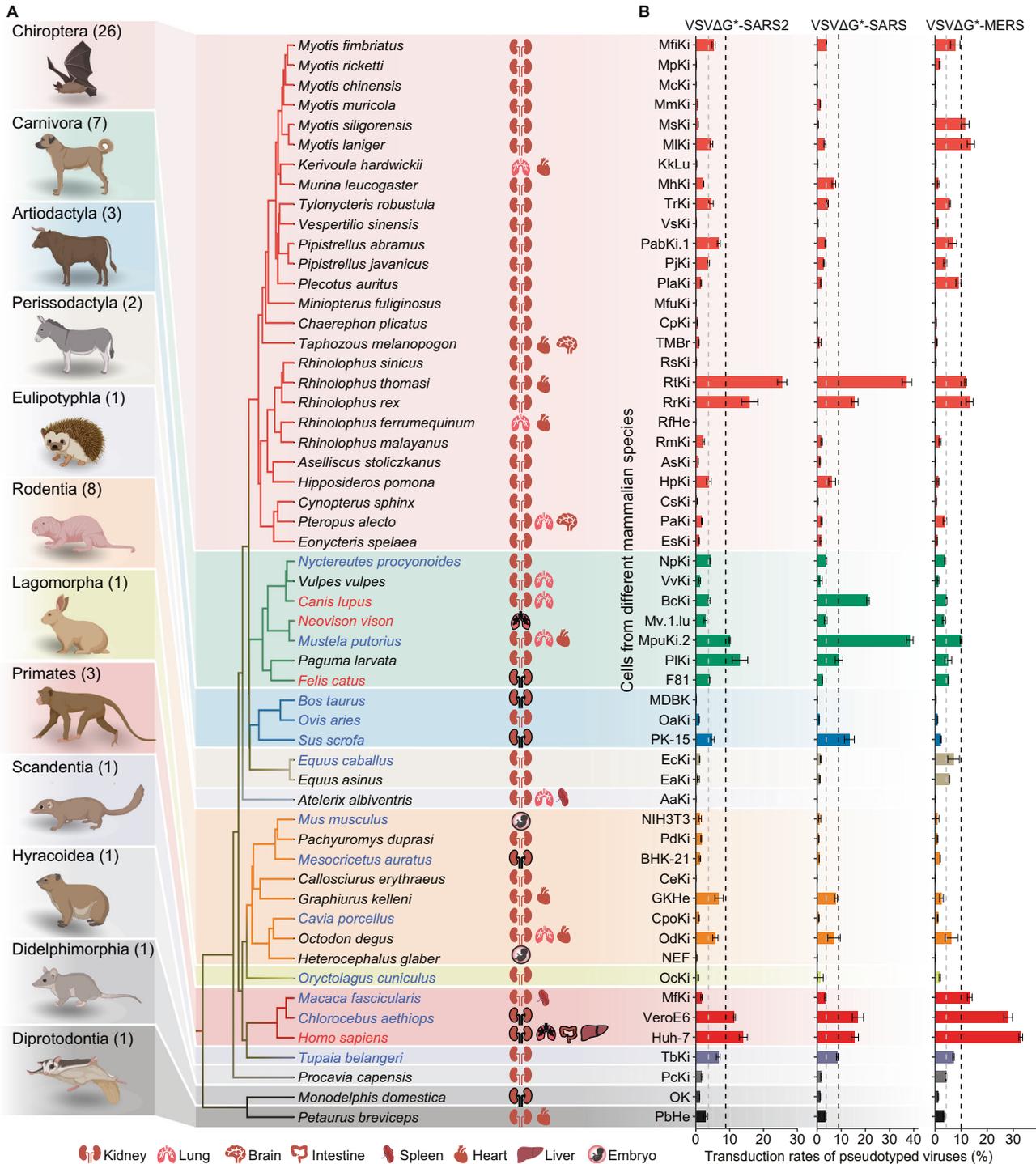


Fig. 1 Transduction rates of SARS-CoV, MERS-CoV, and SARS-CoV-2 pseudotyped viruses to mammalian cell cultures. **A** Phylogenetic tree of 55 mammalian cell cultures was used in this study. A phylogeny was retrieved from TIMETREE (<http://www.timetree.org/>). Taxa indicated in red font are species that can be naturally infected by SARS-CoV-2 based on previous studies, while taxa indicated in blue font are species that were tested by in vivo experimental infection analyses. Moreover, 83 cell cultures were derived from the kidneys, heart, lungs, brain, and other tissues. The cartoon elements with black frames are cell lines derived from these tissues. The tiny cartoons were obtained from BioRender (<https://biorender.com/>). **B** Transduction rates (mean ± SD, $n = 3$) of 83 cell cultures infected by VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS viruses. Some cell cultures were derived from the same host. In that case, cell cultures with the highest transduction rates are displayed in this panel. The gray and black dashed lines represent the standard to define moderate and efficient transduction, respectively.

VSVΔG*-SARS. Cell cultures that showed the highest susceptibility to VSVΔG*-SARS were derived from domesticated ferrets, with 38.4% of ferret kidney MpuKi.2 cells and 31.1% of ferret MpuKi.1 cells transduced. This represents a 2.6-fold and 2.0-fold

increase compared to that of human cells (Huh-7) (Fig. 1, Fig. S4, Table S1). This is noteworthy as ferrets have been used as an animal model for SARS-CoV and experience clinical signs, including sneezing, fever, and diarrhea [63]. The cell culture that

ranked highest after ferret kidney cells was derived from Thomas's horseshoe bat, which had a transduction rate of 37.2%, indicating that this species might also serve as an important host for SARS-CoV.

Palm civets have been recognized as a replication host for SARS-CoV. Our results indicate a consistent and high transduction rate (9.1%) for palm civet PIKi cells. In addition, VSVΔG*-SARS displayed a higher capacity than VSVΔG*-SARS2 in transducing cell cultures from highly susceptible species, such as Rtki (37.2% vs 25.5%), PK-15 (13.4% vs 4.9%), MpuKi.2 (38.4% vs 10.0%), MpuKi.1 (33.3% vs 1.9%), and BcKi (21.1% vs 3.8%) cells (Fig. 1, Fig. S4). This implies that SARS-CoV might be easier to transmit from humans to natural hosts than SARS-CoV-2.

We found that 41 cell cultures were minimally (transduction rate ranging from 0 to 1.3%), 21 slightly (1.3–4.3%), 13 moderately (4.3–10.0%), and 8 efficiently (10.0–32.9%) transduced by VSVΔG*-MERS (Fig. 1, Fig. S3). Strikingly, human cell lines showed the highest transduction rate upon VSVΔG*-MERS infection, which was different than from VSVΔG*-SARS and VSVΔG*-SARS2. Moreover, among the top 10 cell cultures with the highest transduction rates, six (MIKi, RrKi, Rtki, MsKi, PlaKi, and MfKi) were derived from bats (Fig. 1, Fig. S3), suggesting the several bat species serve as potential reservoir hosts for MERS-CoV. In particular, the Japanese pipistrelle bat (*Pipistrellus abramus*) is a species of concern. They live in close proximity to human communities and have the potential to transmit the virus to humans [64]. And, although previous studies have suggested that ferrets are resistant to MERS-CoV infection [65], we found that 10.0% of ferret kidney MpuKi.2 cell cultures were transduced by the MERS-CoV pseudotyped virus (Fig. 1). Moreover, the MERS-CoV receptor *DPP4* [66] is slightly expressed in ferret kidney cells. These results imply that there are factors other than *DPP4* that can aid in MERS-CoV entry and that ferrets are susceptible to SARS-CoV, SARS-CoV-2, and MERS-CoV.

SARS-CoV-2 S protein variants alter the susceptibility of animals

To explore whether SARS-CoV-2 S protein variants alter the susceptibility to SARS-CoV-2, 77,125 spike genes of SARS-CoV-2 were retrieved from GISAID (retrieved on 2020/10/26), and 65 variants with residue frequencies higher than 0.075% were selected (Fig. 2A). In addition, fourteen residue mutations in S proteins from SARS-CoV-2 variants isolated from mouse, cat, dog, mink, and tiger samples were selected, as well as S proteins mutations from bat-CoV RaTG13 (GISAID: EPI_ISL_402131), bat-CoV RmYN02 (GISAID: EPI_ISL_412977), and pangolin-CoV MP789 (GISAID: EPI_ISL_412860). In total, 79 pseudotyped viruses with mutated SARS-CoV-2 S proteins were constructed, followed by the in vitro infection of 44 cell cultures (Fig. 2B).

We found that most of these mutations had little impact on the transduction efficiencies of VSVΔG*-SARS2-S_{mut} pseudotyped viruses to cell cultures (Fig. 2C, Fig. S5). As expected, cell cultures from humans, monkeys, and bats (Huh-7, Vero-E6, Marc-145, and PabKi.1) were more sensitive to pseudotyped viruses compared to other cell cultures (Fig. 2C). Previous studies have shown that substitutions N439K, N501Y, D614G, and H655Y in S proteins enhance the pathogenicity and transmission of SARS-CoV-2 in humans [67–69]. In the present study, we found that these substitutions do not always promote the transduction rate. For example, H655Y increased the transduction of VSVΔG*-SARS2 to tree shrew cells (TbKi) (Fig. S5), while the D614G substitution significantly reduced the capacity of VSVΔG*-SARS2 to transduce in naked mole-rat fibroblast cells and in Japanese pipistrelle bats cells (PabKi.1) (Fig. S6). In addition, the T221 substitution significantly reduced the tropism of SARS-CoV-2 to F81 (cat kidney cell line) and BHK-21 (golden hamster kidney cell line). Moreover, the H519N and T716I substitutions showed a decreased capacity to transduce PK-15 (pig kidney cell line) and PaKi (black flying fox kidney cell line). Finally, substitutions T29I, E281V, and

S939F significantly reduced the tropism of SARS-CoV-2 to MfuKi (eastern bent-wing bat kidney cell) cell cultures (Fig. S5).

Several mutations were found to significantly increase the transduction efficiencies of VSVΔG*-SARS2-S_{mut} to certain cell cultures (Fig. 2C, Fig. S5). In particular, del69–70 in the S protein promoted the capacity of VSVΔG*-SARS2-S_{del69–70} to transduce bat PabKi.1 cells (17.8%) to a level twice that of Huh-7 cells (9.8%) (Fig. S5B). Similarly, VSVΔG*-SARS2-S_{D80Y} caused a 7.2% transduction rate in golden hamster cells, which was higher than that of Huh-7 cells (6.6%) (Fig. S5C). We found that substitution S98F significantly enhanced the transduction rate in raccoon dog kidney cells (62.4%) and Rickett's big-footed bat kidney cells (45.0%) compared to Huh-7 cells (38.9%) (Fig. S5D). Similarly, substitution T572I significantly increased the transduction efficiency of VSVΔG*-SARS2 to pig PK15 cells (11.3%) compared to Huh-7 cells (12.2%) (Fig. S5E). The Q675H mutation also significantly enhanced the ability of the SARS-CoV-2 pseudotyped virus to infect fat-tailed gerbil kidney cells, with a transduction rate of 13.4% compared to 15.7% in Huh-7 cells (Fig. S5F). Notably, these five mutations were located outside the receptor-binding domain (RBD) of the S protein (Fig. 2D). Thus, these results support the idea that different mutations have a unique impact on viral susceptibility in particular mammalian species, and that variants within and outside the RBD region are of equal importance for viral entry and binding.

Molecular dynamic simulations using Amber 20, as well as the crystal complex of RBD-hACE2 (PDB ID: 6M0J), indicate that the binding free energy (ΔG) between RBD and different ACE2 orthologs ranged from -64.7 to -24.8 kcal/mol. ACE2s from green monkeys (*Chlorocebus sabaeus*) and greater horseshoe bats (*Rhinolophus ferrumequinum*) showed the highest and lowest binding free energy with SARS-CoV-2 S-RBD, respectively (Fig. S6A). We further analyzed the free energy contributions based on a per-residue basis (Fig. S6B–D). The results showed that there were 42 amino acid residues that were important to the interactions between RBD and ACE2s (Fig. S6B–E).

In this regard, the cross-species infection spectrum of different spike variants allowed us to explore the possible site-wise "interaction" between ACE2 amino acid residues and spike variants that can affect the transduction rate across species. By testing whether the frequency of the major residues was evenly distributed between high and low infection groups suggested by each spike variant, we identified 98 combinations of spike variants and ACE2 residues that were candidates for such potential "interactions" ("material and methods", Fig. S7, Table S3). The top-ranked ACE2 positions involved in these significant interactions were 49, 31, 354, 82, 75, 35, 353, whereas the top-ranked spike variants included P26L, S254F, H519N, H146Y, A262S, and T478I. The majority of these interactions (86.7%) involved the spike variants of the RBD region (Table S3). These results support the possibility that different spike variants have a unique impact on viral susceptibility by interacting with distinct ACE2 residues and/or other intrinsic factors in particular mammal species.

Genes with expression levels associated with transduction rates across species

Transcriptomic profiles of cell cultures from 42 mammal species were generated, and regression analyses were performed to identify genes whose expressions were significantly correlated with the transduction rates of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS. In total, there were 590 genes whose expression levels were significantly associated with the transduction rates of VSVΔG*-SARS2 (Fig. 3A, Table S4). The most significantly enriched pathway was the Herpes simplex virus 1 infection ($p = 4.30 \times 10^{-23}$, Fisher exact test), in which 93.6% (59 of 63 genes) of associated genes are zinc finger proteins (Fig. 3B, Table S5). The corresponding enriched biological processes that contain zinc finger proteins are DNA-templated transcription

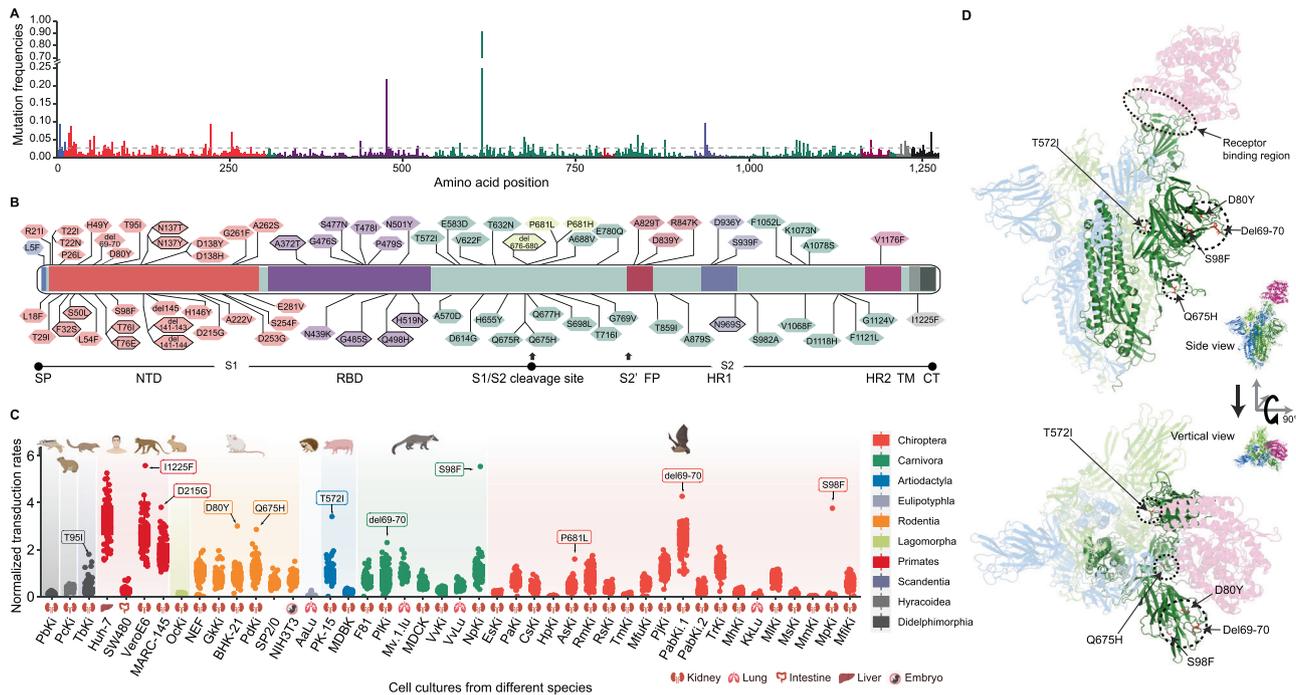
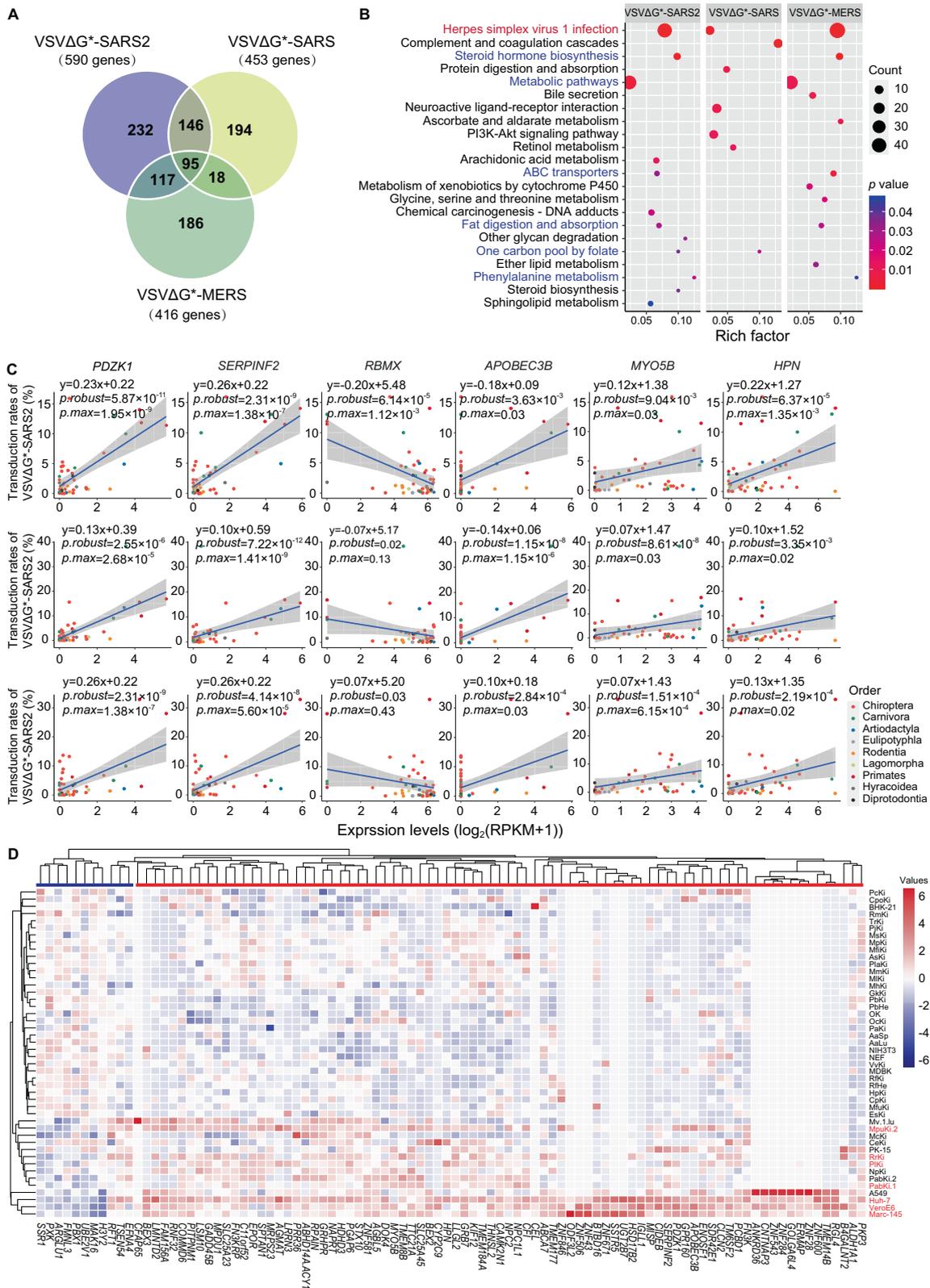


Fig. 2 Characterization of spike variants that impact the host tropism of VSVΔG*-SARS2 across mammals. **A** The substitution frequency for each site was analyzed based on S sequences retrieved from the GISAID database. **B** Distribution of 79 site mutations from the S protein. The site with a black frame means this mutation was selected based on comparative analysis of the viral genome derived from animal hosts. **C** The overall normalized transduction rates of all mutations across different cell cultures. The tiny cartoons shown in the figure represent species from different mammalian orders. The tissue origin of these cell cultures is shown on the x-axis. **D** A 3D structure of the S-ACE2 complex showing that del69-70, D80Y, S98F, T572I, and Q675H are located outside the RBD of the S proteins.

regulation ($p = 2.40 \times 10^{-12}$, Fisher exact test) and RNA polymerase II promoter transcription regulation ($p = 5.10 \times 10^{-10}$, Fisher exact test) (Table S6). Zinc finger proteins appear to function in host-virus interactions and play multiple roles in viral replication by regulating host cell transcription profiles [70]. Many zinc finger proteins are positively correlated with VSVΔG*-SARS2 infection rates. This indicates that the transcription levels of host cells are important for SARS-CoV-2 entry, which usually depends on a mechanism called “cap snatching,” as observed in the influenza virus [71]. The top five genes that were associated with the transduction rates of VSVΔG*-SARS2 in cell cultures were *PDZK1* ($p_{\text{robust}} = 5.87 \times 10^{-11}$), *SERPINF2* ($p_{\text{robust}} = 2.31 \times 10^{-9}$), *SCG5* ($p_{\text{robust}} = 3.51 \times 10^{-9}$), *DEPP1* ($p_{\text{robust}} = 4.57 \times 10^{-9}$), and *ABCC6* ($p_{\text{robust}} = 2.71 \times 10^{-8}$) (Fig. 3C, Fig. S8, Table S4). These genes are worthy of further exploration. For example, *PDZK1* encodes a PDZ (PSD9/DLG/ZO-1) domain-containing scaffolding protein (named NHERF3), which belongs to a group of proteins that mediate cell-cell junctions and is involved in the coordination of a diverse range of regulatory processes. A previous study showed that SARS-CoV and the neurotropic rabies virus are linked to the PDZ-binding functions of their envelope proteins [72]. *PDZK1* can interact with SR-B1 and facilitate hepatitis C virus entry [73]. Moreover, the ACE2 C-terminal PDZ-recognition motif ⁸⁰²QTSF⁸⁰⁵ binds to *NHERF1* and/or *NHERF3* and promotes ACE2-mediated SARS-CoV-2 cell entry [74, 75]. *SERPINF2* encodes one of the serine protease inhibitors. A recent study revealed that *SERPINF2* expression increases in serum levels of patients with COVID-19, along with several other SERPINS (*SERPINA1* and *SERPINA3*) [76]. In addition, several genes, such as *MAK16* ($p_{\text{robust}} = 1.92 \times 10^{-5}$), *H3Y2* ($p_{\text{robust}} = 4.08 \times 10^{-5}$), and *RBMX* ($p_{\text{robust}} = 6.14 \times 10^{-5}$), showed a significant negative correlation with transduction rates (Fig. S8, Table S4). The role of *MAK16* and *H3Y2* in viral infection is unclear, while *RBMX*, which encodes X-ed RNA binding motif proteins, responds to SARS-CoV-2 entry via viral RNA-host protein interactions [77, 78].

We identified a significant association between the expression levels of 453 genes associated with the transduction rate of VSVΔG*-SARS and 416 genes associated with the transduction rate of VSVΔG*-MERS. These genes are primarily enriched in metabolic pathways and Herpes simplex virus 1 infection, consistent with the fact that more than half of all gene expressions were correlated with VSVΔG*-SARS2 transduction (Fig. 3B, Tables S6–S11). Genes whose expressions were associated with VSVΔG*-MERS exhibited a similar profile to VSVΔG*-SARS2 in biological process enrichment, particularly in DNA-templated transcription regulation and RNA polymerase II promoter transcription regulation. However, the top biological process that was enriched by genes whose expressions were associated with VSVΔG*-SARS transduction was proteolysis ($p = 7.7 \times 10^{-4}$, Fisher exact test) (Table S8), highlighting that the proteases of host cells are essential for SARS-CoV pseudotyped virus infectivity [79]. Moreover, a recent study suggests that proteases are central to the infection process of SARS-CoV-2, and drugs targeting proteases can cause a dose-dependent reduction in SARS-CoV-2 titers [80]. This suggests that proteases could be a common target for treating SARS-like diseases. However, the expressions of recognized receptors (*ACE2* and *DPP4*) of SARS-CoV and MERS-CoV did not significantly correlate with transduction rates. Given this result, we examined the potential roles of those related genes. Several gene responses were found (e.g., *HSD17B2*, *IGLL1*) in SARS-CoV-2 infection [81, 82], whereas the role of top-ranked genes associated with VSVΔG*-SARS transduction was more limited.

We compiled a list of 95 genes that were commonly associated with the transduction rates of SARS-CoV, MERS-CoV, and SARS-CoV-2 pseudotyped viruses (Fig. 3D, Table S12). This list not only included genes mentioned above, but also included genes such as *APOBEC3B*, *MYO5B*, and *HPN* (hepsin), which may be involved in SARS-CoV-2 virus–host interactions. For example, *APOBEC3B*, which belongs to a family of proteins in mammals, consists of cellular cytosine



deaminases and serves as a barrier that potentially prevents the cross-species transmission of lentiviruses [83]. In addition, *APOBEC3* can constantly shape the SARS-CoV-2 genome by editing cytosine to uridine [84]. In contrast, *MYO5B* is downregulated in a Spike-RBD-

induced mast cell degranulation model and plays a role in endosomal transport [85]. *MYO5B* also interacts with viral proteins. A previous study suggests that inhibiting MYO5 proteins could be an effective target for COVID-19 treatment [85]. These results

Fig. 3 Comparative transcriptomics analyses identifying genes whose expression levels were correlated with the transduction rates of SARS-CoV, MERS-CoV, and SARS-CoV-2 pseudotyped viruses. **A** A Venn graph showing 590 genes, 453 genes, and 416 genes that were significantly associated with the transduction rates of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS pseudotyped viruses, respectively. A total of 95 genes were commonly associated with the three pseudotyped viruses. **B** The KEGG pathways that were enriched by these genes and had expressions associated with the transduction rates of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS. **C** Plots of several genes whose expressions were associated with the transduction rates of VSVΔG*-SARS2, VSVΔG*-SARS, and VSVΔG*-MERS. **D** Heatmap of 95 genes that were commonly associated with the three pseudotyped viruses. The expression level was represented by $\log_2(\text{RPKM} + 1)$. The cell cultures marked by red fonts were efficiently transduced by VSVΔG*-SARS2. Blue bars under the tree indicate genes that were negatively associated with VSVΔG*-SARS2 transduction, while the red bars indicate genes that were positively associated with VSVΔG*-SARS2 transduction.

broadly implicate these common genes in viral entry and replication, and therefore they remain suitable targets for further evaluation of cross-species transmission risks.

DISCUSSION

Herein, we have demonstrated that SARS-CoV-2, SARS-CoV, and MERS-CoV can infect the cells of dozens of mammal species, indicating they are generalist viruses and not specifically adapted to humans. In addition, cell cultures from different mammals show variable susceptibilities to SARS-CoV-2, SARS-CoV, and MERS-CoV pseudotyped viruses. This implies that SARS-CoV-2 has the capacity to spillover to multiple species and establish natural reservoirs after minor or major adaptive evolutionary changes. Our results highlighted the potential for Thomas's horseshoe bats, king horseshoe bats, green monkeys, and ferrets to serve as reservoir hosts for these coronaviruses. Specially, primary cells cultures derived from Thomas's horseshoe bat and the king horseshoe bat were found to be more sensitive to VSVΔG*-SARS2 than human cell lines. This is important, as horseshoe bats (Rhinolophidae) are considered as a reservoir for many zoonotic viruses and they are assumed to be generally tolerant to infection. However, very rare sarbecoviruses have been reported in both Thomas's horseshoe bats and the king horseshoe bats, but not in other horseshoe bats (e.g., Chinese horseshoe bats) [86]. This not only raises the concern that SARS-CoV-2 might spillover from humans to these two bat species, but also prioritizes the surveillance of virus intolerance in horseshoe bats.

Since ACE2 was found to be the functional receptor for the spike protein of SARS-CoV and SARS-CoV-2, pioneer studies have predicted the susceptibility of various animal species to SARS-CoV-2 using ACE2 sequences and/or their binding affinity to spike proteins [15, 87, 88]. While our experimental assay validated the susceptibility of several species (e.g., rhesus macaque and golden hamster) to SARS-CoV-2, there is a considerable inconsistency between those in silico predictions and our cellular assays. For example, the Chinese tree shrew and ferret were predicted to exhibit low or no risk for SARS-CoV-2 infection [15], possibly due to the fact that their ACE2 sequences diverged from human ACE2. However, our experiments indicated both species are of medium susceptibility. Similarly, bats are generally predicted to be of low infection risk, yet our results suggest that several bat species are likely to be highly infected. Thus, studies that experimentally examine infection susceptibility are extremely valuable for comparison with future in silico predictions. In addition, as indicated in Table S2, 75% of the in vitro susceptibilities presented in our study were supported by in vivo inoculation assays. Differences between in vitro and in vivo assays were observed in the mouse, red fox, and cow. For example, inoculation assay suggested that mice are resistant to SARS-CoV-2 infection, but our study found their cell cultures show moderate transduction rates. This may be due to the fact that cell models may fail to resemble the complexity of a whole organism, or to the particular challenges of translatability of in vitro-generated data into in vivo models [57]. Nonetheless, in vivo testing can benefit from in vitro testing in determining the

susceptibility risk and natural hosts of SARS-CoV-2 and other sarbecoviruses.

We also examined the susceptibilities of several mammalian species to different spike variants and found that each spike variant showed a distinct spectrum of infection. Specially, several spike variants (Del69-70, D80Y, S98F, T572I, and Q675H) that are not in the RBD region were shown to increase the infectivity of SARS-CoV-2. The variant del69-70 was initially identified at the beginning of 2020, and over the past several years it has mutated into three different variants (i.e., Y453F, S493K, and N501Y) [89]. Previous reports have shown that del69-70 increases cleavage in the S protein and infectivity in the SARS-CoV-2 B.1.1.7 variant [90, 91]. This mutation may enhance SARS-CoV-2 spilling over to other animals, such as the Japanese pipistrelle bat, which inhabits urban areas of high human population density. In addition, an interaction between ACE2 residues and spike variants, which potentially affect infection rates, was found in our study. Some of these ACE2 residues are important for the interface between the SARS-CoV RBD and ACE2. For example, Lys31 and Lys353 are virus-binding hotspots, which consist of salt bridges between Lys31 and Glu35 and between Lys353 and Asp38 [92, 93]. However, most of the spike variants in these predicted interactions were located in the N-terminal domain (NTD). Thus, regions inside and outside of RBD may be of equal importance for viral entry and infection. Greater attention should be paid to these spike variants, as they (e.g., P26L, S254F, H146Y, and A262S) constitute the 'NTD supersite', which is recognized by all known NTD-specific neutralizing antibodies [94], and thus could mediate immune escape.

An examination of phylogenetic signals indicates that SARS-CoV exhibits a more similar tropism profile with SARS-CoV-2 than with MERS. This may be due to the fact that SARS-CoV and SARS-CoV-2 use the same receptor. However, the SARS-CoV pseudotyped virus had higher efficiency in transducing several cell cultures from dogs, pigs, and ferrets, suggesting that further investigation may uncover spike variants and other factors between SARS-CoV and SARS-CoV-2 that contribute to this difference (Fig. S4). Regression analyses indicated that the expression of commonly known receptors and factors known to facilitate the entry of SARS-CoV-2, SARS-CoV, and MERS-CoV, did not correlate with cross-species infection (Fig. S8). In particular, the expression levels of *ACE2* ($p_{\text{robust}} = 0.14$), *FURIN* ($p_{\text{robust}} = 0.89$) and *TMPPRSS2* ($p_{\text{robust}} = 0.11$), showed an insignificant correlation with VSVΔG*-SARS2 transduction rates (Figs. S9A-C; S9D, and Table S4, Table S6). In contrast, *CTSL* ($p_{\text{robust}} = 7.4 \times 10^{-3}$) was significantly correlated with VSVΔG*-SARS2 transduction rates.

We found that certain cell lines expressed high levels of *ACE2* (e.g., cells derived from the sugar glider, Nepalese whiskered bat, and greater horseshoe bat), but these were minimally or only slightly transduced by VSVΔG*-SARS2 (Fig. S9A-C and Table S1). However, this may not negate the role of ACE2 in determining SARS-CoV-2 susceptibility in other species because RNA expression cannot fully represent the ACE2 protein on the cell surface. Moreover, subsequent penalized regression model analyses revealed that sequence changes in ACE2s are significantly correlated with transduction rates across phylogenetically diverse

taxa ($F = 15.78$, $p = 8.9 \times 10^{-13}$) (Figs. S10, S11). This suggests that changes in ACE2 sequences rather than expression contribute to differences in susceptibility. However, the total variance of transduction rates explained by ACE2 sequence changes ~28.9%. Therefore, future studies should explore other host intrinsic factors associated with susceptibility, such as genes identified in this study whose expression was significantly correlated with transduction rates.

Finally, our results highlight the fact that cell culture models represent an important method of understanding cross-species transmission of sarbecoviruses by identifying the spectrum of mammalian hosts that are susceptible to SARS-CoV-2, SARS-CoV, and MERS-CoV and their variants.

DATA AVAILABILITY

RNA sequencing data from this study have been deposited in ScienceDB (<https://doi.org/10.57760/sciencedb.j00001.00445>) and in the Genome Sequence Archive in National Genomics Data Center, China National Center for Bioinformatics/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA009056) and NCBI (PRJNA906190).

REFERENCES

- Pekar JE, Magee A, Parker E, Moshiri N, Izhikevich K, Havens JL, et al. The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science*. 2022;377:960–6.
- Garry RF. The evidence remains clear: SARS-CoV-2 emerged via the wildlife trade. *Proc Natl Acad Sci USA*. 2022;119:e2214427119.
- Worobey M, Levy JI, Serrano LM, Crits-Christoph A, Pekar JE, Goldstein SA, et al. The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science*. 2022;377:951–9.
- 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;579:270–3.
- Zhang T, Wu Q, Zhang Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr Biol*. 2020;30:1346–51.
- Temmam S, Vongphayloth K, Baquero E, Munier S, Bonomi M, Regnault B, et al. Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature*. 2022;604:330–6.
- Azhar EI, El-Kafrawy SA, Farraj SA, Hassan AM, Al-Saeed MS, Hashem AM, et al. Evidence for camel-to-human transmission of MERS coronavirus. *N Engl J Med*. 2014;370:2499–505.
- Wang M, Yan M, Xu H, Liang W, Kan B, Zheng B, et al. SARS-CoV infection in a restaurant from palm civet. *Emerg Infect Dis*. 2005;11:1860–5.
- Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*. 2013;503:535–8.
- Sabir JS, Lam TT, Ahmed MM, Li L, Shen Y, Abo-Aba SE, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science*. 2016;351:81–4.
- Menachery VD, Yount BL Jr, Debbink K, Agnihothram S, Gralinski LE, Plante JA, et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med*. 2015;21:1508–13.
- Zhou P, Shi ZL. SARS-CoV-2 spillover events. *Science*. 2021;371:120–2.
- Hale VL, Dennis PM, McBride DS, Nolting JM, Madden C, Huey D, et al. SARS-CoV-2 infection in free-ranging white-tailed deer. *Nature*. 2022;602:481–6.
- McAloose D, Laverack M, Wang L, Killian ML, Caserta LC, Yuan F, et al. From zoonosis to panzootic: natural SARS-CoV-2 infection in tigers and lions at the Bronx Zoo. *mBio*. 2020;11:e02220–20.
- Damas J, Hughes GM, Keough KC, Painter CA, Persky NS, Corbo M, et al. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc Natl Acad Sci USA*. 2020;117:22311–22.
- Wu L, Chen Q, Liu K, Wang J, Han P, Zhang Y, et al. Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. *Cell Discov*. 2020;6:68.
- Fang S, Zheng R, Lei C, Wang J, Zheng R, Li M. Key residues influencing binding affinities of 2019-nCoV with ACE2 in different species. *Brief Bioinform*. 2021;22:963–75.
- Baggen J, Persoons L, Vanstreels E, Jansen S, Van Looveren D, Boeckx B, et al. Genome-wide CRISPR screening identifies TMEM106B as a proviral host factor for SARS-CoV-2. *Nat Genet*. 2021;53:435–44.
- Zhu S, Liu Y, Zhou Z, Zhang Z, Xiao X, Liu Z, et al. Genome-wide CRISPR activation screen identifies candidate receptors for SARS-CoV-2 entry. *Sci China Life Sci*. 2022;65:701–17.
- Schwedhelm P, Kusnick J, Heintz C, Schonfelder G, Bert B. How many animals are used for SARS-CoV-2 research?: An overview on animal experimentation in pre-clinical and basic research. *EMBO Rep*. 2021;22:e53751.
- Sarkar J, Guha R. Infectivity, virulence, pathogenicity, host-pathogen interactions of SARS and SARS-CoV-2 in experimental animals: a systematic review. *Vet Res Commun*. 2020;44:101–10.
- Imai M, Iwatsuki-Horimoto K, Hatta M, Loeber S, Halfmann PJ, Nakajima N, et al. Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proc Natl Acad Sci USA*. 2020;117:16587–95.
- Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science*. 2020;368:1016–20.
- Chu H, Chan JF, Yuen TT, Shuai H, Yuan S, Wang Y, et al. Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. *Lancet Microbe*. 2020;1:e14–e23.
- Gultom M, Licheri M, Laloli L, Wider M, Strassle M, V'Kovski P, et al. Susceptibility of well-differentiated airway epithelial cell cultures from domestic and wild animals to severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis*. 2021;27:1811–20.
- Seluanov A, Vaidya A, Gorbunova V. Establishing primary adult fibroblast cultures from rodents. *J Vis Exp*. 2010:e2033. <https://doi.org/10.3791/2033>.
- Takada A, Robison C, Goto H, Sanchez A, Murti KG, Whitt MA, et al. A system for functional analysis of Ebola virus glycoprotein. *Proc Natl Acad Sci USA*. 1997;94:14764–9.
- Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*. 2020;182:1284–94.
- Case D, Belfon K, Ben-Shalom I, Brozell S, Cerutti D, Cheatham T, et al. AMBER 2020: University of California, San Francisco; 2020.
- Peng C, Zhu ZD, Shi YL, Wang XY, Mu KJ, Yang YQ, et al. Computational insights into the conformational accessibility and binding strength of SARS-CoV-2 spike protein to human angiotensin-converting enzyme 2. *J Phys Chem Lett*. 2020;11:10482–8.
- Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature*. 2020;583:286–9.
- Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020;581:215–20.
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46:W296–W303.
- Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. fF145B: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *J Chem Theory Comput*. 2015;11:3696–713.
- Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov*. 2015;10:449–61.
- Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next-generation sequencing data. *PLoS One*. 2012;7:e30619.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
- Liu W, Zhu P, Li M, Li Z, Yu Y, Liu G, et al. Evolutionary transcriptomics reveals longevity mostly driven by polygenic and indirect selection in mammals. *bioRxiv*. 2023.
- Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22:1658–9.
- Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*. 2012;28:3150–2.
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28:511–5.
- Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res*. 2013;41:W29–33.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
- Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–30.

46. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139–40.
47. Ma S, Upneja A, Galecki A, Tsai YM, Burant CF, Raskind S, et al. Cell culture-based profiling across mammals reveals DNA repair and metabolism as determinants of species longevity. *Elife*. 2016;5:e19130.
48. Ma S, Yim SH, Lee SG, Kim EB, Lee SR, Chang KT, et al. Organization of the mammalian metabolome according to organ function, lineage specialization, and longevity. *Cell Metab*. 2015;22:332–43.
49. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022;28:gakc194.
50. Meekins DA, Gaudreault NN, Richt JA. Natural and experimental SARS-CoV-2 infection in domestic and wild animals. *Viruses* 2021;13:1993.
51. Schlottau K, Rissmann M, Graaf A, Schon J, Sehl J, Wylezich C, et al. SARS-CoV-2 in fruit bats, ferrets, pigs, and chickens: an experimental transmission study. *Lancet Microbe*. 2020;1:e218–e25.
52. Weingartl HM, Copps J, Drebot MA, Marszal P, Smith G, Gren J, et al. Susceptibility of pigs and chickens to SARS coronavirus. *Emerg Infect Dis*. 2004;10:179–84.
53. Pickering BS, Smith G, Pinette MM, Embury-Hyatt C, Moffat E, Marszal P, et al. Susceptibility of domestic swine to experimental infection with severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis*. 2021;27:104–12.
54. Meekins DA, Morozov I, Trujillo JD, Gaudreault NN, Bold D, Arriaga BL, et al. Susceptibility of swine cells and domestic pigs to SARS-CoV-2. *bioRxiv*. 2020.
55. Gu H, Chen Q, Yang G, He L, Fan H, Deng YQ, et al. Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science* 2020;369:1603–7.
56. Dinnon KH 3rd, Leist SR, Schafer A, Edwards CE, Martinez DR, Montgomery SA, et al. A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures. *Nature*. 2020;586:560–6.
57. Rosa RB, Dantas WM, do Nascimento JCF, da Silva MV, de Oliveira RN, Pena LJ *In vitro* and *in vivo* models for studying SARS-CoV-2, the etiological agent responsible for COVID-19 pandemic. *Viruses*. 2021;13:379.
58. Wang L, Fan X, Bonenfant G, Cui D, Hossain J, Jiang N, et al. Susceptibility to SARS-CoV-2 of cell lines and substrates commonly used to diagnose and isolate influenza and other viruses. *Emerg Infect Dis*. 2021;27:1380–92.
59. Lu G, Wang Q, Gao GF. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol*. 2015;23:468–78.
60. Kim YI, Kim SG, Kim SM, Kim EH, Park SJ, Yu KM, et al. Infection and rapid transmission of SARS-CoV-2 in ferrets. *Cell Host Microbe*. 2020;27:704–9.
61. Richard M, Kok A, de Meulder D, Bestebroer TM, Lamers MM, Okba NMA, et al. SARS-CoV-2 is transmitted via contact and via the air between ferrets. *Nat Commun*. 2020;11:3496.
62. Sawatzki K, Hill NJ, Puryear WB, Foss AD, Stone JJ, Runstadler JA. Host barriers to SARS-CoV-2 demonstrated by ferrets in a high-exposure domestic setting. *Proc Natl Acad Sci USA*. 2021;118:e2025601118.
63. Martina BEE, Haagmans BL, Kuiken T, Fouchier RAM, Rimmelzwaan GF, van Amerongen G, et al. SARS virus infection of cats and ferrets. *Nature*. 2003;425:915.
64. Chung CU, Kim SC, Jeon YS, Han SH. Changes in habitat use by female Japanese Pipistrelles (*Pipistrellus abramus*) during different stages of reproduction revealed by radio telemetry. *J Environ Sci Int*. 2017;26:817–26.
65. Raj VS, Smits SL, Provacia LB, van den Brand JM, Wiersma L, Ouwendijk WJ, et al. Adenosine deaminase acts as a natural antagonist for dipeptidyl peptidase 4-mediated entry of the Middle East respiratory syndrome coronavirus. *J Virol*. 2014;88:1834–8.
66. Wang N, Shi X, Jiang L, Zhang S, Wang D, Tong P, et al. Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res*. 2013;23:986–93.
67. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*. 2021;592:116–21.
68. Zhou B, Thao TTN, Hoffmann D, Taddeo A, Ebert N, Labrousseau F, et al. SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature*. 2021;592:122–7.
69. Escalera A, Gonzalez-Reiche AS, Aslam S, Mena I, Laporte M, Pearl RL, et al. Mutations in SARS-CoV-2 variants of concern link to increased spike cleavage and virus transmission. *Cell Host Microbe*. 2022;30:373–87.
70. Wang G, Zheng C. Zinc finger proteins in the host-virus interplay: multifaceted functions based on their nucleic acid-binding property. *FEMS Microbiol Rev*. 2021;45:fuaa059.
71. Walker AP, Fodor E. Interplay between influenza virus and the host RNA polymerase II transcriptional machinery. *Trends Microbiol*. 2019;27:398–407.
72. James CD, Roberts S. Viral interactions with PDZ domain-containing proteins—an oncogenic trait? *Pathogens* 2016;5:8.
73. Eyre NS, Drummer HE, Beard MR. The SR-B1 partner PDZK1 facilitates hepatitis C virus entry. *PLoS Pathog*. 2010;6:e1001130.
74. Zhang Q, Geffer J, Sneddon WB, Mamonova T, Friedman PA. ACE2 interaction with cytoplasmic PDZ protein enhances SARS-CoV-2 invasion. *iScience*. 2021;24:102770.
75. Kliche J, Kuss H, Ali M, Ivarsson Y. Cytoplasmic short linear motifs in ACE2 and integrin beta3 link SARS-CoV-2 host cell receptors to mediators of endocytosis and autophagy. *Sci Signal*. 2021;14:eabf1117.
76. D'Alessandro A, Thomas T, Dzieciatkowska M, Hill RC, Francis RO, Hudson KE, et al. Serum proteomics in COVID-19 patients: altered coagulation and complement status as a function of IL-6 level. *J Proteome Res*. 2020;19:4417–27.
77. Zhang S, Huang W, Ren L, Ju X, Gong M, Rao J, et al. Comparison of viral RNA-host protein interactomes across pathogenic RNA viruses informs rapid antiviral drug discovery for SARS-CoV-2. *Cell Res*. 2022;32:9–23.
78. Flynn RA, Belk JA, Qi Y, Yasumoto Y, Wei J, Alfajaro MM, et al. Discovery and functional interrogation of SARS-CoV-2 RNA-host protein interactions. *Cell*. 2021;184:2394–411.
79. Simmons G, Zmora P, Gierer S, Heurich A, Pohlmann S. Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting edge of antiviral research. *Antivir Res*. 2013;100:605–14.
80. Meyer B, Chiaravalli J, Gellenoncourt S, Brownridge P, Bryne DP, Daly LA, et al. Characterising proteolysis during SARS-CoV-2 infection identifies viral cleavage sites and cellular targets with therapeutic potential. *Nat Commun*. 2021;12:5553.
81. Zolfaghari Emameh R, Nosrati H, Eftekhari M, Falak R, Khoshmirsafa M. Expansion of single cell transcriptomics data of SARS-CoV infection in human bronchial epithelial cells to COVID-19. *Biol Proced Online*. 2020;22:16.
82. Khyber S, Guojun L, Mikhail B, Izzaz A, Muhammad D, Irina T, et al. Is up-regulation gene expression of the certain genes during the viral respiratory tract infection would have any influence in pathogenesis of the SAR-CoV-2 infection? *Acta Med Iran*. 2020;58:246–8.
83. Nakano Y, Aso H, Soper A, Yamada E, Moriwaki M, Juarez-Fernandez G, et al. A conflict of interest: the evolutionary arms race between mammalian APOBEC3 and lentiviral Vif. *Retrovirology* 2017;14:1–12.
84. Wei Y, Silke JR, Aris P, Xia X. Coronavirus genomes carry the signatures of their habitats. *PLoS One*. 2020;15:e0244025.
85. Wu ML, Liu FL, Sun J, Li X, He XY, Zheng HY, et al. SARS-CoV-2-triggered mast cell rapid degranulation induces alveolar epithelial inflammation and lung injury. *Signal Transduct Target Ther*. 2021;6:428.
86. Luk HKH, Li X, Fung J, Lau SKP, Woo PCY. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. *Infect Genet Evol*. 2019;71:21–30.
87. Fischhoff IR, Castellanos AA, Rodrigues J, Varsani A, Han BA. Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2. *Proc Biol Sci*. 2021;288:20211651.
88. Luan J, Lu Y, Jin X, Zhang L. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem Biophys Res Commun*. 2020;526:165–9.
89. Ou J, Lan W, Wu X, Zhao T, Duan B, Yang P, et al. Tracking SARS-CoV-2 Omicron diverse spike gene mutations identifies multiple inter-variant recombination events. *Signal Transduct Target Ther*. 2022;7:138.
90. Meng B, Kemp SA, Papa G, Dahir R, Ferreira I, Marelli S, et al. Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep*. 2021;35:109292.
91. Ren W, Lan J, Ju X, Gong M, Long Q, Zhu Z, et al. Mutation Y453F in the spike protein of SARS-CoV-2 enhances interaction with the mink ACE2 receptor for host adaptation. *PLoS Pathog*. 2021;17:e1010053.
92. Choudhary S, Malik YS, Tomar S. Identification of SARS-CoV-2 cell entry inhibitors by drug repurposing using *in silico* structure-based virtual screening approach. *Front Immunol*. 2020;11:1664.
93. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature*. 2020;581:221–4.
94. McCallum M, De Marco A, Lempp FA, Tortorici MA, Pinto D, Walls AC, et al. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell*. 2021;184:2332–47.e16.

ACKNOWLEDGEMENTS

We thank Pengcheng Wang, Xuanjing Li, Yun Huang, Linfa Wang, and Alice C. Hughes for supporting sample collection and preparation. We also thank Chrissie, Sara, Jenni, and Dax for their support during the writing of this manuscript. This project was funded by grants from the National Natural Science Foundation of China (32070528), the National Key Research and Development Projects of the Ministry of Science and Technology of China (2021YFC2301300), the National Natural Science Foundation of China (82050002), the key program of Chinese Academy of Sciences (KJZD-SW-L11) and the Beijing Natural Sciences Foundation (JQ19022).

AUTHOR CONTRIBUTIONS

XZ and ML conceived the study and designed the project. ML performed experiments, completed analysis and wrote the manuscript; JD, ZL, and FL prepared the cell cultures; WL and CH implement the RNA analysis, data analysis, and generated structural figures; GL, LW., XW, QP, XT, PAG, and XZ discussed the results, and revised this manuscript; All authors contributed to data interpretation.

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/s41396-023-01368-2>.

Correspondence and requests for materials should be addressed to Xuming Zhou.

Reprints and permission information is available at <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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023