Isolation of bat sarbecoviruses of SARS-CoV-2 clade, Japan

Shin Murakami*, Tomoya Kitamura, Hiromichi Matsugo, Haruhiko Kamiki, Ken Oyabu, Wataru Sekine, Akiko Takenaka-Uema, Yuko Sakai-Tagawa, Yoshihiro Kawaoka, Taisuke Horimoto*

Author affiliations:

The University of Tokyo, Tokyo, Japan (S. Murakami, Matsugo, H. Kamiki, W. Sekine, A. Takenaka-Uema, Y. Sakai-Tagawa, Y. Kawaoka, T. Horimoto); National Agriculture and Food Research Organization (NARO), Tokyo, Japan (T. Kitamura); Isumi County-City Nature Preservation Association, Isumi-gun, Chiba, Japan (K. Oyabu)

* Address for correspondence:

Shin Murakami, Laboratory of Veterinary Microbiology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; email: shin-murakami@g.ecc.u-tokyo.ac.jp; phone number: +81-3-5841-5398

19 Taisuke Horimoto, Laboratory of Veterinary Microbiology, Graduate School of

20 Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku,

21 Tokyo 113-8657, Japan; email: taihorimoto@g.ecc.u-tokyo.ac.jp

22 Word count (abstract): 150

23 Word count (text): 2724

24

ABSTRACT

Betacoronaviruses have caused 3 outbreaks in the past 2 decades. SARS-CoV-2, in particular, has caused a serious pandemic. As the betacoronaviruses are considered to originate from bats, surveillance of bat betacoronaviruses is crucial for understanding the mechanism of cross-species transition and potential for future outbreaks. We previously detected and characterized a SARS-CoV-2-related sarbecovirus, Rc-o319, from *Rhinolophus cornutus* in Japan. Here, we detected several bat sarbecoviruses of the SARS-CoV-2 clade from *R. cornutus* in multiple locations in Japan, and successfully isolated them using Vero/TMPRSS2 cells stably expressing *R. cornutus* ACE2 (Vero-RcACE2). The coding sequences of S1 region varied among isolates, whereas other genetic regions were highly conserved. Isolates were efficiently grown in Vero-RcACE2 cells, but did not replicate in Vero/TMPRSS2 cells stably expressing human ACE2, suggesting a narrow host range. Further long-term epidemiological studies of sarbecoviruses in wildlife are expected to facilitate the assessment of the risk of their spillover potential.

Severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2 have caused 3

43 outbreaks in human populations in the past 2 decades. In particular, SARS-CoV-2 has
44 caused a worldwide serious pandemic with devastating damage to human health, many
45 social activities, and the economy. All these betacoronaviruses are considered to be of
46 bat origin. Hence, surveillance of bat betacoronaviruses is crucial for understanding and
47 assessing the spillover potential of betacoronaviruses in humans in the future.
48 SARS-CoV and SARS-CoV-2 belong to the *Betacoronavirus* genus and *Sarbecovirus*
49 subgenus. Bats belonging to the genus *Rhinolophus* are considered natural reservoirs of
50 sarbecoviruses, as most have been detected in this bat species in Asian countries such as
51 China (1-9), Japan (10), South Korea (11, 12), Laos (13), Thailand (14), and Cambodia
52 (15), as well as in European and African countries such as Bulgaria (16), England (17),
53 and Kenya (18). Bat sarbecoviruses closely related to SARS-CoV-2 have also been
54 found in the Yunnan province, China, and Southeast Asian countries (1, 3, 7, 8, 13-15)
55 with some possessing the ability to bind to the human angiotensin-converting enzyme 2
56 (ACE2) as an entry receptor, posing a potential for human infection (3, 8, 13).
57 We previously identified a Japanese bat sarbecovirus, Rc-o319, detected from
58 *Rhinolophus cornutus* bats in the Iwate prefecture, which was shown to
59 phylogenetically belong to the SARS-CoV-2 lineage. Vesicular stomatitis virus
60 (VSV)-based pseudotyped virus possessing the Rc-o319 spike (S) protein was able to

61 infect cells expressing *R. cornutus* ACE2 (RcACE2), but not those expressing human
62 ACE2 (hACE2), suggesting that the Rc-o319 virus uses RcACE2 as its receptor (10).
63 Sarbecoviruses detected in China and other Asian countries were demonstrated to vary
64 genetically. However, the distribution and genetic variation of bat sarbecoviruses in
65 Japan have not yet been determined.

66 Despite the surveillance-based genetic detection of numerous bat
67 sarbecoviruses, infectious viruses have been limitedly isolated to date. Infectious
68 viruses are useful and required for understanding the biological characteristics. Here, we
69 report the detection and isolation of bat sarbecoviruses from several locations in Japan
70 and their genetic and biological characterization.

71

72 **Materials and methods**

73 **Cells and virus**

74 Vero/TMPRSS2 cells (19) were kindly provided by Dr. Makoto Takeda, National
75 Institute of Infectious Diseases, Japan, and maintained in Dulbecco's modified Eagle's
76 medium (DMEM, Nacalai Tesque, Kyoto, Japan) supplemented with 10 % fetal bovine
77 serum (FBS), 100 units/mL of penicillin, and 100 μg/mL of streptomycin. Cells were
78 cultured in 5 % CO2 at 37 °C. SARS-CoV-2 (B.1.1.7, alpha variant,

79 UT-HP127-1Nf/Human/2021/Tokyo) was propagated in Vero/TMPRSS2 cells, and
80 aliquots were stored at -80 °C.
81
82 **Sample collection**
83 We collected 88 fresh fecal samples, 58 from *Rhinolophus cornutus* and 30 from
84 *Rhinolophus ferrumequinum* bats living in caves, abandoned mines, or abandoned
85 tunnels in Niigata, Chiba, and Shizuoka prefectures in Japan (Table 1). When bats were
86 densely packed during daytime roost, plastic sheets were placed under the roost for 1-2
87 h. Fresh feces that dropped onto the sheets were collected. When bats were sporadically
88 placed during daytime roost, we captured bats with permission from the prefectural
89 local governments (No. 962 for Chiba and No. 311 for Shizuoka) and kept each bat in a
90 separate nonwoven fabric bag. Feces excreted by the bat in the bag were collected and
91 bats were released. Fecal samples were transferred into tubes containing
92 phosphate-buffered saline (PBS) supplemented with 200 U/mL penicillin, 200 μg/mL
93 streptomycin, and 0.25 μg/mL amphotericin B, and immediately frozen in dry ice.
94
95 **Reverse transcription-PCR**
96 RNA was extracted from the fecal samples using the RNeasy PowerMicrobiome Kit

(Qiagen KK, Tokyo, Japan) and the partial E gene of sarbecovirus was detected in RNA samples by real-time reverse transcription-PCR (rRT-PCR) using the RNA-direct SYBR Green Realtime PCR Master Mix (Toyobo, Osaka, Japan) and a pair of primers (5′-TCGGAAGAGACAGGTACGTT-3′ and 5′-TCGAAGCGCAGTAAGGATGG-3′) that were designed to target a highly conserved region of the sarbecovirus E gene.

Establishment of ACE2-stably expressing cells

We constructed a plasmid, pCAGGS-blast, by inserting the XhoI-EcoRV fragment of pMXs-IRES-Bsd (20), which contains the encephalomyocarditis virus internal ribosomal entry site and blasticidin-resistant gene, into the XhoI and StuI sites of the pCAGGS-MCS vector. Open reading frame (ORF) sequences of RcACE2 or hACE2 were PCR-amplified from pCAGGS-RcACE2- or pCAGGS-hACE2-expressing plasmids (10) and cloned into EcoRI- and XhoI-digested pCAGGS-blast plasmids using NEBuilder (New England Biolabs, Ipswich, MA, USA). Vero/TMPRSS2 cells were transfected with pCAGGS-blast-RcACE2 or pCAGGS-blast-hACE2 plasmids using the PEI MAX transfection reagent (Polysciences, Warrington, PA, USA). Transfected cells were treated with 10 μg/mL blasticidin S (Kaken Pharmaceutical, Tokyo, Japan) 1 d posttransfection, and blasticidin S-resistant cells were selected and cloned. Highly

susceptible cell clones for the pseudotyped viral infection were selected by screening using GFP-expressing VSV-pseudotyped virus possessing the S protein of Rc-o319 or SARS-CoV-2 (10), generating RcACE2- or hACE2-stably expressing Vero/TMPRSS2 cells (namely Vero-RcACE2 or Vero-hACE2, respectively).

Establishment of ACE2-knockout cells

We generated ACE2-knockout Vero/TMPRSS2 cells (Vero-ACE2KO) by knocking out the corresponding genes using the CRISPR/Cas9 system. The target sequence for the ACE2 gene (5′-TGCTGCTCAGTCCACCATTG-3′) was designed using CRISPR direct (https://crispr.dbcls.jp) and cloned into plentiCRISPR plasmids (21) (Addgene plasmid #52961, a gift from Dr. Feng Zhang) using NEBuilder (NEB). Vero/TMPRSS2 cells were transfected with an ACE2-targeting plasmid using PEI MAX (Polysciences). At 24 h posttransfection, the cell supernatant was replaced with medium containing 10 μg/mL puromycin. Drug-resistant clones were randomly selected and their genomic DNA was sequenced. Cells possessing insertions or deletions (in/dels) in the targeted gene were chosen for further analysis.

Isolation of bat sarbecoviruses

Fecal samples positive for the partial E gene of sarbecovirus were homogenized in TissueLyser II (Qiagen) using 0.1 mm glass beads (Tomy Seiko, Tokyo, Japan) in PBS containing 200 U/mL penicillin, 200 μg/mL streptomycin, and 0.25 μg/mL amphotericin B. The supernatants were collected after centrifugation at 5000 × g for 5 min at 4 °C and diluted 100-fold in cell maintenance medium (DMEM supplemented with 1 % FBS, 200 U/mL penicillin, 200 μg/mL streptomycin, and 0.25 μg/mL amphotericin B). Diluents were inoculated into 6-well plates containing Vero-RcACE2 cells, and plates were incubated for 60 min at 37 °C after removing the inoculum. Wells were then washed once with cell maintenance medium, another 2 mL of cell maintenance medium was added to each well, and incubated at 37 °C. Supernatants from cells that exhibited cytopathic effects (CPE) were collected at 3–4 d postinoculation and were passed through a 0.22 μm filter. The successful isolation of viruses was confirmed by rRT-PCR, and isolates were propagated in Vero -RcACE2 cells, with aliquots being stored at -80 °C.

Next-generation sequencing

A cDNA library was prepared from RNA extracted from bat sarbecoviral isolates using the TruSeq Stranded Total RNA LT Sample Prep Kit Gold (Illumina, San Diego, CA,

151 USA) for Rc-mk2 and Rc-kw8 strains or the MGIEasy RNA Directional Library Prep
152 kit (MGI, Shenzhen, China) for the Rc-os20 strain. Libraries of Rc-mk2 and Rc-kw8
153 strains were sequenced using a Novaseq 6000 sequencer (Illumina), whereas those of
154 Rc-os20 were sequenced using a DNBSEQ-G400RS (MGI) sequencer. Read sequences
155 were mapped to the Rc-o319 genome sequence (GenBank accession No. LC556375)
156 and sarbecoviral sequences were determined using the CLC genomic workbench
157 version 8.0.1 (Qiagen, https://www.qiagen.com) software. The sequences of Rc-os20,
158 Rc-mk2, and Rc-kw8 have been deposited in GenBank (accession Nos. LC663958,
159 LC663959, and LC663793, respectively).
160

161 **Phylogenetic analysis**

162 The nucleotide sequences of sarbecoviruses were aligned using ClustalW version 2.1
163 (Clustal, https://www.clustal.org). Phylogenetic trees were then constructed by
164 performing a maximum-likelihood analysis using MEGA version X (22) in combination
165 with 500 bootstrap replicates.
166

167 **Evaluation of viral growth in cells**

168 Cells were inoculated with viruses at a multiplicity of infection (MOI) of 0.01 and

169 allowed for 1 h for viral adsorption. After removing the inocula, cells were incubated in
170 cell maintenance medium, and the supernatants were collected at 12 h intervals. Viral
171 titers were measured using a plaque assay, in which cells inoculated with diluted viruses
172 were overlaid and incubated with DMEM containing 1 % agarose and 1 % FCS for 2 d,
173 followed by staining with crystal violet before counting plaques.
174
175 **Results**
176 **Detection of bat sarbecoviruses in Japan**
177 We collected fecal samples from bats belonging to the *Rhinolophus cornutus* and
178 *Rhinolophus ferrumequinum* species in Niigata, Chiba, and Shizuoka prefectures
179 (Figure 1). Using real-time RT-PCR, we successfully detected the E gene sequence of
180 sarbecovirus in 1 or 2 *R. cornutus* samples in each prefecture (Table 1). In contrast, all *R.*
181 *ferremuquinum* samples were negative. These data suggested that bat sarbecoviruses are
182 distributed among *R. cornutus* at various locations in Japan.
183
184 **Isolation of bat sarbecoviruses from *R. cornutus***
185 In our previous study, we showed that a VSV-based pseudotyped virus possessing the S
186 protein of the Rc-o319 sarbecovirus from *R. cornutus* only infected RcACE2-expressing

187 cells, but not hACE2- or other *Rhinolophus* ACE2-expressing cells (10), suggesting a
188 narrow host range of Rc-o139. Therefore, to isolate bat sarbecoviruses, we established
189 RcACE2-stably expressing cells (Vero-RcACE2) based on Vero/TMPRSS2 cells (19),
190 which are useful for the isolation of SARS-CoV-2 (23). We inoculated Vero-RcACE2
191 cells with RT-PCR-positive fecal samples from each prefecture and observed any
192 extensive cytopathic effect (CPE) with syncytium formation at 1–2 d postinoculation
193 (Figure 2). Using real-time RT-PCR, we also detected the E gene sequence of
194 sarbecovirus in the supernatants at 3–4 d postinoculation. Following filtration, we
195 inoculated the supernatants into fresh Vero-RcACE2 cells, and observed for CPE,
196 confirming the successful isolation of bat sarbecoviruses. We designated the Niigata,
197 Chiba, and Shizuoka isolates, Rc-os20, Rc-mk2, and Rc-kw8, respectively. Using the
198 same approach, we further isolated the infectious Rc-o319 isolate.
199
200 **Genetic characterization of Japanese isolates**
201 We determined the full genome sequence of all isolates using next-generation
202 sequencing. We detected that the sequence of Rc-o319 was identical to the previously
203 deposited sequence, which was determined using the Sanger method. Next, we mapped
204 the sequence reads of Rc-os20, Rc-mk2, and Rc-kw8 onto the Rc-o319 genome

sequence, and determined their full genomic sequences, which we deposited in GenBank (accession Nos. LC663958, LC663959, and LC663793). We found that sequence homologies were high (ranging 94.8–96.8 %) among all Japanese isolates (Table 2); however, Rc-mk2 and Rc-os20 lacked the entire ORF8 coding region. We also performed similarity plot analysis of entire genome sequence using each isolate as a query, which indicated that the similarities among isolates were high throughout the entire genome sequence, except for coding regions of the N-terminal domain (NTD) and receptor binding domain (RBD) of the S gene, although NTDs of Rc-o319 and Rc-kw8 were conserved (Figure 3A). No clear recombination among the isolates were observed as analyzed by RDP5 software (24). Phylogenetic analysis showed that the Japanese isolates formed a single genetic cluster within the SARS-CoV-2 clade of sarbecoviruses, which might be designated the Japanese clade of bat sarbecoviruses (Figure 3B).

Growth kinetics of Japanese isolates in cell culture

We aligned the receptor binding motif of the S protein of Japanese isolates with that of other sarbecoviruses (Figure 4A). We observed that all isolates had a 9-amino acid deletion in this motif, as previously observed in Rc-o319, and possessed relatively conserved residues with Rc-o319. Therefore, we assumed that the use of hACE2 as an

223 entry receptor is unlikely in these new Japanese isolates. To test this hypothesis, we
224 compared the replication of Japanese isolates and that of a control SARS-CoV-2
225 (B.1.1.7, alpha variant) in Vero-RcACE2, Vero-hACE2, Vero-ACE2KO, and
226 Vero/TMPRSS2 cells. We found that the 4 bat isolates replicated well only in
227 Vero-RcACE2, whereas did not replicate in Vero/TMPRSS2, Vero-hACE2, or
228 Vero-ACE2KO cells, suggesting their RcACE2-dependent infectivity. In contrast, we
229 noticed that SARS-CoV-2 replicated efficiently in Vero/TMPRSS2, Vero-RcACE2, and
230 Vero-hACE2 cells, but not in Vero-ACE2KO cells (Figure 4B), suggesting multiple
231 ACE2-dependent infectivity including that of *R. cornutus*. These data suggested that
232 Japanese bat isolates use only RcACE2 as a receptor, showing narrow host specificity.
233

234 **Discussion**

235 In this study, we successfully isolated 4 bat sarbecoviruses from *R. cornutus*
236 bats in distinct locations in Japan, using Vero-RcACE2 cells. The origin of
237 SARS-CoV-2 is thought to be a sarbecovirus from *Rhinolophus* spp. bats in southern
238 China and the Indochinese peninsula, as bat viruses detected in these regions have been
239 found to have high homology with SARS-CoV-2 and to be able to efficiently bind to
240 human ACE2 (1, 3, 7, 8, 13-15). In contrast, the Japanese bat sarbecovirus isolates were

241 phylogenetically distant from SARS-CoV-2, compared with bat sarbecoviruses from
242 southern China and the Indochinese peninsula, despite belonging to the SARS-CoV-2
243 clade. In addition, the Japanese bat isolates were unable to use hACE2 as a cell entry
244 receptor, thus making it unlikely to be one of the ancestors of SARS-CoV-2. However,
245 phylogenetic analysis indicated that the Japanese viruses were positioned near a
246 branching node between the SARS-CoV and SARS-CoV-2 clades, suggesting that they
247 might be related to the common ancestor of SARS-CoV and SARS-CoV-2. These
248 findings suggested that characterization of bat viruses might provide an understanding
249 of the mechanism and the potential of bat sarbecoviruses to overcome host barriers.

250 Because of the pilot study aspect of this study, the number of samples and the
251 bat species studied were limited. Four species of *Rhinolophus* bats inhabit Japan: *R.*
252 *cornutus*, *R. ferrumequinum*, *R. pumilus*, and *R. perditus*. The former 2 species are
253 widely distributed, exhibiting a major population in Japan, whereas the latter 2 species
254 are mainly confined in Okinawa prefecture, which is located far southwest of the main
255 island of Japan. In this study, the bat sarbecoviruses were isolated from *R. cornutus* but
256 not from *R. ferrumequinum*. Since the virus has been detected from *R. ferrumequinum*
257 in China (5), it is expected that increasing the number of samples will clarify whether *R.*
258 *ferrumequinum*-associated virus existents in Japan. Although there have been no reports

of sarbecovirus detection in *R. pumilus* and *R. perditus*, it is likely that both species might harbor sarbecoviruses with different genetic characteristics from the isolates in our study because of their different niche. Further epidemiological surveys are needed to confirm this hypothesis.

Virus isolates can be used to elucidate their biological characteristics, including pathogenicity and antigenicity. To this date, all detected bat sarbecoviruses are classified into 2 groups based on receptor selectivity: hACE2-binding and non-hACE2-binding types. In particular, hACE2-binding type bat sarbecoviruses were isolated from African green monkey Vero cells, whose ACE2 molecule might be functionally related to hACE2 (9). In contrast, to the best of our knowledge, non-hACE2-binding type bat sarbecoviruses have not been isolated. Comparative analysis between hACE2-binding and non-hACE2-binding viruses would facilitate an understanding of the factors that determine receptor specificity of bat sarbecoviruses.

Both the NTD and RBD of the S protein were highly variable among sarbecoviruses (25-28), probably due to escape from immune pressure, as both regions contain viral neutralizing epitopes (29-31). *Rhinolophus* spp. bats are relatively short-distance migrants (32, 33); hence, the migration of individuals between bat groups is infrequent. Most genome sequences, except for the RBD- and NTD-coding regions of

277 the S gene, were highly conserved among Japanese strains, suggesting that Japanese
278 sarbecoviruses diverged relatively recently from the undefined ancestral virus, and
279 consecutively rapidly accumulated mutations in the NTD and RBD regions due to
280 strong selection pressure. Several studies have been reported that in addition to
281 sarbecoviruses *Rhinolophus* bats harbor alphacoronaviruses (8, 34), suggesting that a
282 virus could undergo major changes by recombination with other coronaviruses in the
283 NTD/RBD regions, leading to the emergence of a novel coronavirus with zoonotic
284 potential.

285 Various sarbecoviruses lacking ORF8 protein have been identified in bats (6,
286 35) and humans (36, 37). Although the function of ORF8 protein in SARS-CoV-2 is not
287 fully understood, it might act as an inhibitor of IFN-I signaling (38) and a
288 downregulator of MHC-I (39). Hence, an ORF8-deleted virus shows lower
289 pathogenicity in humans (40). Among our bat isolates, 2 strains possessed the ORF8
290 gene, but their homology to that of SARS-CoV-2 was very low (approximately 27.5 %
291 in amino acid sequence), suggesting functional differences. Likewise, no appreciable
292 difference was detected in growth in cell culture between viruses possessing or lacking
293 ORF8 protein (e.g., Rc-o319 versus Rc-os20). Further studies are required to clarify the
294 biological characteristics of ORF8 protein in bat sarbecoviruses.

In conclusion, we isolated bat sarbecoviruses from *R. cornutus* in several locations in Japan that were phylogenetically positioned in the same cluster of the SARS-CoV-2 clade. These isolates did not replicate in hACE2-expressing cells, suggesting low potential for human infection. However, sarbecoviruses might mutate and infect humans via intermediate hosts in wildlife or livestock. Therefore, epidemiological studies of sarbecoviruses in wildlife, including bats, need to be conducted on a long-term basis for risk assessment of their zoonotic potential.

Dr. Murakami is an associate professor at the Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo, Japan. His research interests include epidemiologic and molecular biological studies of animal viruses, including coronaviruses and influenza viruses.

Acknowledgements

This work was supported by the Japan Agency for Medical Research and Development (AMED) under the grant number JP21fk0108615. We thank Ms. Satomi Kato for the technical assistance.

313 **Disclosure statement**

314 The author(s) declare that they have no potential conflict of interest.

315

316 **References**

317 1. Li LL, Wang JL, Ma XH, Sun XM, Li JS, Yang XF, et al. A novel SARS-CoV-2
318 related coronavirus with complex recombination isolated from bats in Yunnan province,
319 China. Emerg Microbes Infect. 2021 Dec;10(1):1683-90.

320 2. Guo H, Hu B, Si HR, Zhu Y, Zhang W, Li B, et al. Identification of a novel
321 lineage bat SARS-related coronaviruses that use bat ACE2 receptor. Emerg Microbes
322 Infect. 2021 Dec;10(1):1507-14.

323 3. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia
324 outbreak associated with a new coronavirus of probable bat origin. Nature. 2020
325 Mar;579(7798):270-3.

326 4. Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, et al. Severe acute
327 respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad
328 Sci U S A. 2005 Sep;102(39):14040-5.

329 5. Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, et al. Bats are natural
330 reservoirs of SARS-like coronaviruses. Science. 2005 Oct;310(5748):676-9.

331 6. Hu B, Zeng LP, Yang XL, Ge XY, Zhang W, Li B, et al. Discovery of a rich
332 gene pool of bat SARS-related coronaviruses provides new insights into the origin of
333 SARS coronavirus. PLoS Pathog. 2017 Nov;13(11):e1006698.
334 7. Zhou H, Chen X, Hu T, Li J, Song H, Liu Y, et al. A novel bat coronavirus
335 closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of
336 the spikepProtein. Current Biology. 2020 Jun;30(11):2196-203.e3.
337 8. Zhou H, Ji J, Chen X, Bi Y, Li J, Wang Q, et al. Identification of novel bat
338 coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related
339 viruses. Cell. Aug;184(17):4380-91.e14.
340 9. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and
341 characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature.
342 2013 Nov;503(7477):535-8.
343 10. Murakami S, Kitamura T, Suzuki J, Sato R, Aoi T, Fujii M, et al. Detection and
344 Characterization of Bat Sarbecovirus Phylogenetically Related to SARS-CoV-2, Japan.
345 Emerg Infect Dis. 2020 Dec;26(12):3025-9.
346 11. Kim Y, Son K, Kim Y-S, Lee S-Y, Jheong W, Oem J-K. Complete genome
347 analysis of a SARS-like bat coronavirus identified in the Republic of Korea. Virus
348 Genes. 2019 Aug;55(4):545-9.

12. Lo VT, Yoon SW, Noh JY, Kim Y, Choi YG, Jeong DG, et al. Long-term surveillance of bat coronaviruses in Korea: Diversity and distribution pattern. Transbound Emerg Dis. 2020 Nov;67(6):2839-48.

13. 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 Apr;604(7905):330-6.

14. Wacharapluesadee S, Tan CW, Maneeorn P, Duengkae P, Zhu F, Joyjinda Y, et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. Nat Commun. 2021 Feb;12(1):972.

15. Delaune D, Hul V, Karlsson EA, Hassanin A, Ou TP, Baidaliuk A, et al. A novel SARS-CoV-2 related coronavirus in bats from Cambodia. Nat Commun. 2021 Dec;12(1):6563.

16. Drexler JF, Gloza-Rausch F, Glende J, Corman VM, Muth D, Goettsche M, et al. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. J Virol. 2010 Nov;84(21):11336-49.

17. Crook JM, Murphy I, Carter DP, Pullan ST, Carroll M, Vipond R, et al. Metagenomic identification of a new sarbecovirus from horseshoe bats in Europe. Sci

Rep. 2021Dec;11(1):14723.

18. Tao Y, Tong S. Complete Genome Sequence of a Severe Acute Respiratory Syndrome-Related Coronavirus from Kenyan Bats. Microbiol Resour Announc. 2019 Jul;8(28):e00548-19.

20. Shirogane Y, Takeda M, Iwasaki M, Ishiguro N, Takeuchi H, Nakatsu Y, et al. Efficient multiplication of human metapneumovirus in Vero cells expressing the transmembrane serine protease TMPRSS2. J Virol. 2008 Sep;82(17):8942-6.

21. Kitamura T, Koshino Y, Shibata F, Oki T, Nakajima H, Nosaka T, et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. Exp Hematol. 2003 Nov;31(11):1007-14.

22. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014 Aug;11(8):783-4.

23. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018 06;35(6):1547-9.

24. Matsuyama S, Nao N, Shirato K, Kawase M, Saito S, Takayama I, et al. Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. Proc Natl Acad Sci U S A. 2020 Mar;117(13):7001-3.

385 25. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, et al.
386 RDP5: a computer program for analyzing recombination in, and removing signals of
387 recombination from, nucleotide sequence datasets. Virus Evol. 2021 Jan;7(1):veaa087.
388 26. Wec AZ, Wrapp D, Herbert AS, Maurer DP, Haslwanter D, Sakharkar M, et al.
389 Broad neutralization of SARS-related viruses by human monoclonal antibodies. Science.
390 2020 Aug;369(6504):731-6.
391 27. Pinto D, Park Y-J, Beltramello M, Walls AC, Tortorici MA, Bianchi S, et al.
392 Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody.
393 Nature. 2020 Jul;583(7815):290-5.
394 28. Rappazzo CG, Tse LV, Kaku CI, Wrapp D, Sakharkar M, Huang D, et al. Broad
395 and potent activity against SARS-like viruses by an engineered human monoclonal
396 antibody. Science. 2021 Feb;371(6531):823-9.
397 29. Onodera T, Kita S, Adachi Y, Moriyama S, Sato A, Nomura T, et al. A
398 SARS-CoV-2 antibody broadly neutralizes SARS-related coronaviruses and variants by
399 coordinated recognition of a virus-vulnerable site. Immunity. 2021
400 Oct;54(10):2385-98.e10.
401 30. Suryadevara N, Shrihari S, Gilchuk P, Vanblargan LA, Binshtein E, Zost SJ, et
402 al. Neutralizing and protective human monoclonal antibodies recognizing the

403 N-terminal domain of the SARS-CoV-2 spike protein. Cell. 2021
404 Apr;184(9):2316-31.e15.
405 31. Mccallum M, De Marco A, Lempp FA, Tortorici MA, Pinto D, Walls AC, et al.
406 N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2.
407 Cell. 2021 Apr;184(9):2332-47.e16.
408 32. Liu L, Wang P, Nair MS, Yu J, Rapp M, Wang Q, et al. Potent neutralizing
409 antibodies against multiple epitopes on SARS-CoV-2 spike. Nature. 2020
410 Aug;584(7821):450-6.
411 33. Jones G, Duverge PL, Ransome RD. Conservation biology of an endangered
412 species: Field studies of greater horseshoe bats. Sym Zool S. 1995(67):309-24.
413 34. Bontadina F, Schofield H, Naef-Daenzer B. Radio-tracking reveals that lesser
414 horseshoe bats (*Rhinolophus hipposideros*) forage in woodland. J Zool. 2002
415 Nov;258:281-90.
416 35. Lau SKP, Wong ACP, Zhang L, Luk HKH, Kwok JSL, Ahmed SS, et al. Novel
417 Bat Alphacoronaviruses in Southern China Support Chinese Horseshoe Bats as an
418 Important Reservoir for Potential Novel Coronaviruses. Viruses. 2019 May;11(5):423.
419 36. Wu Z, Yang L, Ren X, Zhang J, Yang F, Zhang S, et al. ORF8-Related Genetic
420 Evidence for Chinese Horseshoe Bats as the Source of Human Severe Acute

421 Respiratory Syndrome Coronavirus. J Infect Dis. 2016 Feb;213(4):579-83.

422 37. Gong Y-N, Tsao K-C, Hsiao M-J, Huang C-G, Huang P-N, Huang P-W, et al.
423 SARS-CoV-2 genomic surveillance in Taiwan revealed novel ORF8-deletion mutant
424 and clade possibly associated with infections in Middle East. Emerg Microbes Infect.
425 2020 Jan;9(1):1457-66.

426 38. Su YCF, Anderson DE, Young BE, Linster M, Zhu F, Jayakumar J, et al.
427 Discovery and Genomic Characterization of a 382-Nucleotide Deletion in ORF7b and
428 ORF8 during the Early Evolution of SARS-CoV-2. mBio. 2020 Jul 21;11(4):e01610-20.

429 39. Li JY, Liao CH, Wang Q, Tan YJ, Luo R, Qiu Y, et al. The ORF6, ORF8 and
430 nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. Virus
431 Res. 2020 Sep;286:198074.

432 40. Zhang Y, Chen Y, Li Y, Huang F, Luo B, Yuan Y, et al. The ORF8 protein of
433 SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. Proc Natl
434 Acad Sci U S A. 2021 Jun;118(23):e2024202118.

435 41. Young BE, Fong S-W, Chan Y-H, Mak T-M, Ang LW, Anderson DE, et al.
436 Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and
437 the inflammatory response: an observational cohort study. The Lancet. 2020
438 Aug;396(10251):603-11.

Figure legends

Figure 1. Sampling locations in Japan. Prefectures, where bats were captured, are indicated in red. Prefectures, where bat sarbecoviruses were detected in our previous study, are indicated in grey.

Figure 2. Images of bat sarbecovirus-inoculated cells. Vero-RcACE2 cells were inoculated with fecal samples from *Rhinolophus cornutus* from several prefectures of Japan. After 1–2 d of inoculation, cytopathic effects with extensive syncytium formation (arrow heads) were observed. Scale bar, 200 μm.

Figure 3. Genetic analysis of bat sarbecovirus isolates in Japan. (A) Similarity plot analysis of isolates was performed using the full-length genome sequence of Rc-o319, Rc-os20, Rc-kw8, or Rc-mk2 as query. SARS-CoV-2 virus was used as a reference. (B) Phylogenetic tree of bat sarbecoviruses was generated using the full genome nucleotide sequences with the maximum-likelihood analysis combined with 500 bootstrap

457 replicates. Red text indicates the isolates in this study. Bootstrap values are shown
458 above and to the left of the major nodes. Scale bars indicate nucleotide substitutions per
459 site.
460
461 **Figure 4**. Receptor binding motif (RBM) and growth kinetics of bat sarbecovirus
462 isolates. (A) Alignment of the RBM sequence of S proteins. Amino acid positions of
463 RBMs contacting human ACE2 identified in SARS-CoV-2 are shaded in yellow, while
464 residues identical to SARS-CoV or SARS-CoV-2 are shown in red. Strains that are
465 capable to bind to human ACE2 are shaded in orange, whereas those that are incapable
466 to bind to human ACE2 are given in blue. (B) Growth kinetics of isolates. Bat isolates
467 Rc-o319, Rc-os20, Rc-mk2, and Rc-kw8 or SARS-CoV-2 (B.1.1.7) were inoculated into
468 Vero/TMPRSS2 (WT), Vero-RcACE2 (RcACE2), Vero-hACE2 (hACE2), or
469 Vero-ACE2KO (ACE2KO) cells at an MOI of 0.01. The culture supernatants were
470 collected at the indicated time points, and viral titers were determined using a plaque
471 assay. Data were reported as the mean titer with standard deviations from 3 independent
472 experiments.
473

Table 1 Detection of sarbecoviruses in Japan using RT-PCR

Location	Bat species	Number of samples	Number of positive RT-PCR samples*
Niigata	*R. cornutus*	26	2
	R. ferrumequinum	1	0
Chiba	*R. cornutus*	11	1
	R. ferrumequinum	16	0
Shizuoka	*R. cornutus*	21	2
	R. ferrumequinum	13	0

*RT-PCR was performed using sarbecovirus consensus primers targeting the E gene.

Table 2 Full genome nucleotide identity among Japanese isolates

	Rc-o319	Rc-os20	Rc-kw8	Rc-mk2
Rc-o319	-	95.6 %	96.8 %	94.8 %
Rc-os20	-	-	95.4 %	95.4 %
Rc-kw8	-	-	-	95.1 %
SARS-CoV-2	81.5 %	80.7 %	81.4 %	80.7 %

Fig. 1

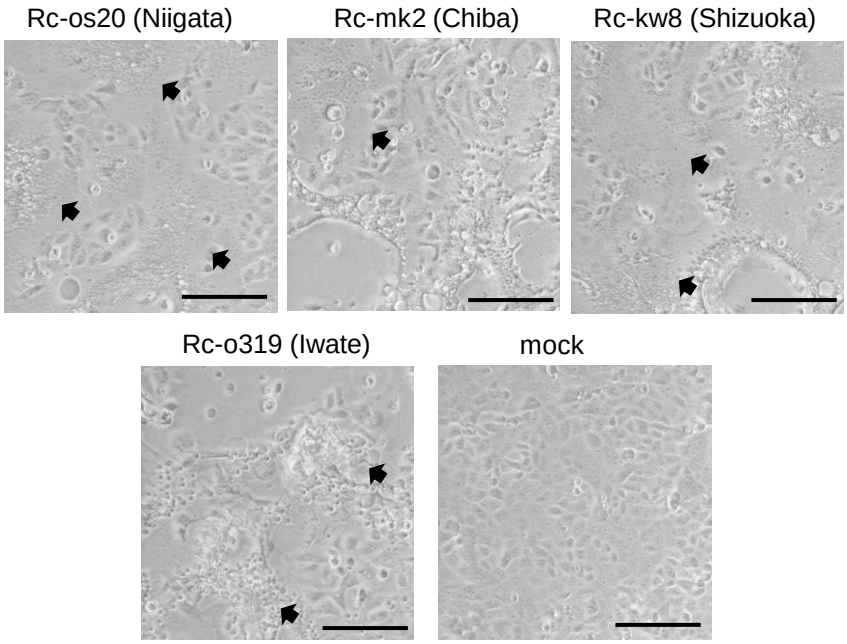
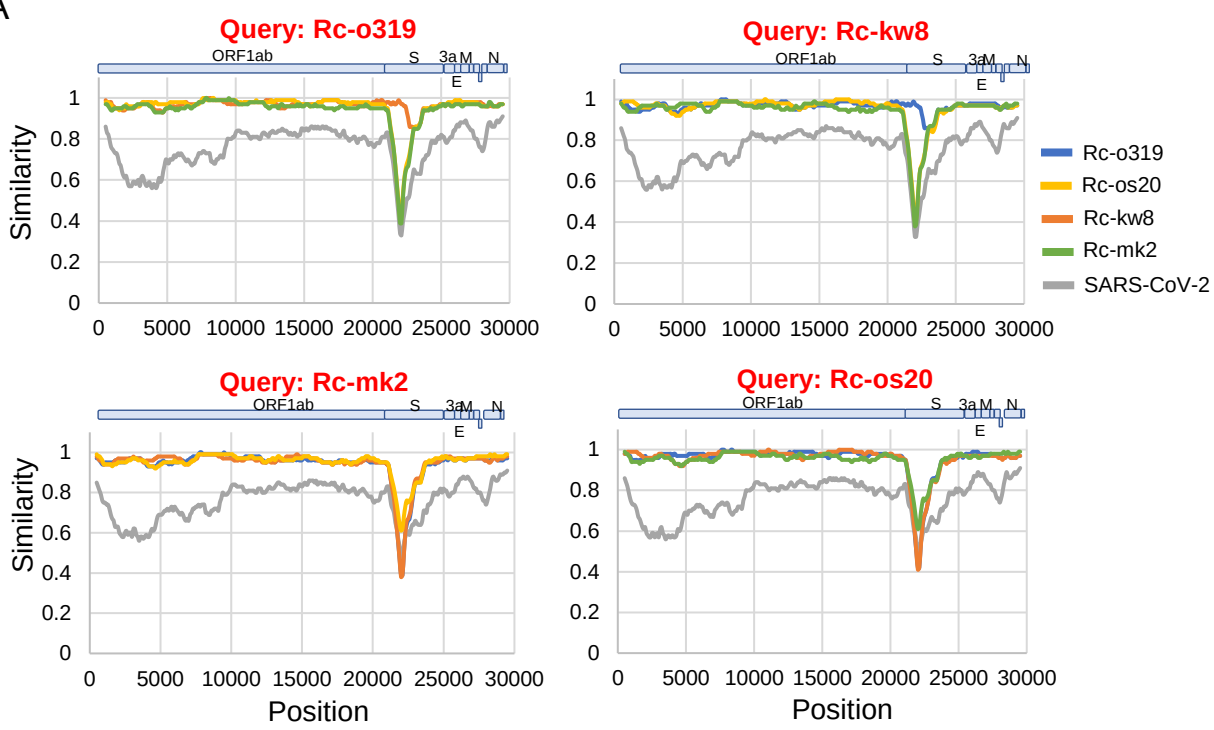

Rc-os20 (Niigata)
Rc-mk2 (Chiba)
Rc-kw8 (Shizuoka)
Rc-o319 (Iwate)
mock

Fig. 2

A


Query: Rc-o319
Query: Rc-kw8
Query: Rc-mk2
Query: Rc-os20
ORF1ab
S
3a
M
N
E
Similarity
Position
0
0.2
0.4
0.6
0.8
1
0
5000
10000
15000
20000
25000
30000
Rc-o319
Rc-os20
Rc-kw8
Rc-mk2
SARS-CoV-2

B

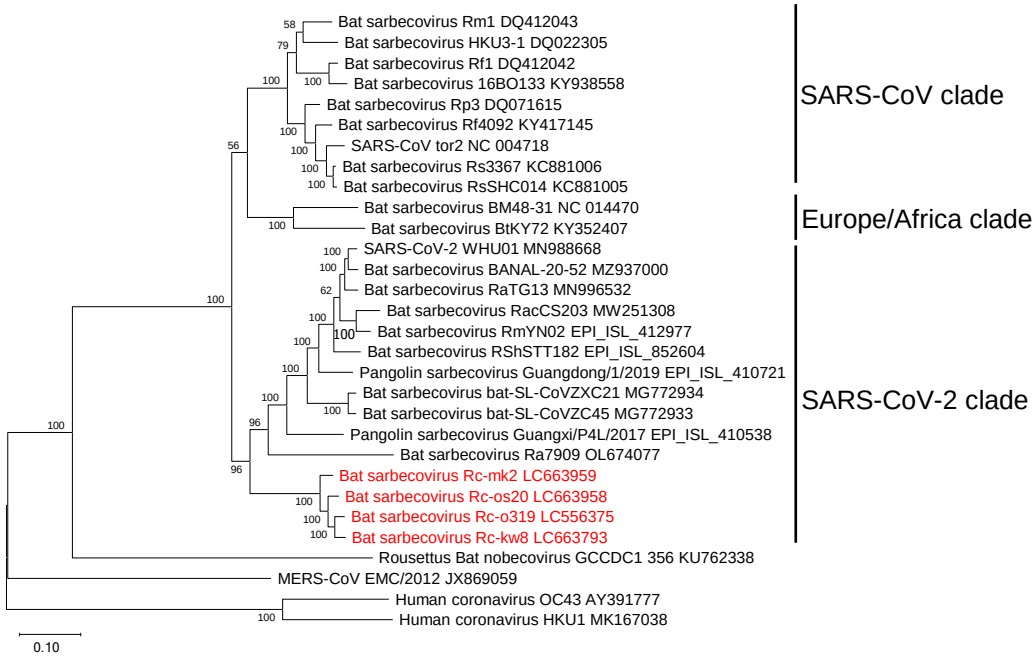

Bat sarbecovirus Rm1 DQ412043
Bat sarbecovirus HKU3-1 DQ022305
Bat sarbecovirus Rf1 DQ412042
Bat sarbecovirus 16BO133 KY938558
Bat sarbecovirus Rp3 DQ071615
Bat sarbecovirus Rf4092 KY417145
SARS-CoV tor2 NC 004718
Bat sarbecovirus Rs3367 KC881006
Bat sarbecovirus RsSHC014 KC881005
Bat sarbecovirus BM48-31 NC 014470
Bat sarbecovirus BtKY72 KY352407
SARS-CoV-2 WHU01 MN988668
Bat sarbecovirus BANAL-20-52 MZ937000
Bat sarbecovirus RaTG13 MN996532
Bat sarbecovirus RacCS203 MW251308
Bat sarbecovirus RmYN02 EPI_ISL_412977
Bat sarbecovirus RShSTT182 EPI_ISL_852604
Pangolin sarbecovirus Guangdong/1/2019 EPI_ISL_410721
Bat sarbecovirus bat-SL-CoVZXC21 MG772934
Bat sarbecovirus bat-SL-CoVZC45 MG772933
Pangolin sarbecovirus Guangxi/P4L/2017 EPI_ISL_410538
Bat sarbecovirus Ra7909 OL674077
Bat sarbecovirus Rc-mk2 LC663959
Bat sarbecovirus Rc-os20 LC663958
Bat sarbecovirus Rc-o319 LC556375
Bat sarbecovirus Rc-kw8 LC663793
Rousettus Bat nobecovirus GCCDC1 356 KU762338
MERS-CoV EMC/2012 JX869059
Human coronavirus OC43 AY391777
Human coronavirus HKU1 MK167038
0.10
SARS-CoV clade
Europe/Africa clade
SARS-CoV-2 clade

Fig. 3

A

```
Bat Rc-o319          413:SRNQDASTSGNFNYYYRIWRSEKLRPFERDIAHYDYQVGTQFKS---------SLKNYGFYSSAGDSHQPY
Bat Rc-os20          430:SKDKDATSSGNYNYLYRMWRPAKLNPFERDIAHYNYLVGVQNKS---------TLYRYGFFSTSGTGKLPY
Bat Rc-kw8           411:NSKAQDAIGINYNYRYRIWRSSKLNPFERDIGVYDYTVGGQQRP---------TIKSYTFTSTVGVDYQPY
Bat Rc-mk2           426:TKTQDSSQSGNFNYYYRSWRPSKLNPFERDIAHYTHSVGGVQKS---------TLTSYGFYSTAGVGYQPY
SARS-CoV-2           439:SNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
Bat BANAL-20-52      438:SNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFHPTNGVGYQPY
Bat RaTG13           439:SKHIDAKEGGNFNYLYRLFRKANLKPFERDISTEIYQAGSKPCNGQTGLNCYYPLYRYGFYPTDGVGHQPY
SARS-CoV             426:TRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTP-PALNCYWPLNDYGFYTTTGIGYQPY
Bat Rs3367           427:TRNIDATQTGNYNYKYRSLRHGKLRPFERDISNVPFSPDGKPCTP-PAFNCYWPLNDYGFYITNGIGYQPY
Bat CoVZXC21         434:TAKQDTG-----HYFYRSHRSTKLKPFERDLSSDE-------------NGVRTLSTYDFNPNVPLEYQAT
Bat Rp3              430:TAKQDQG-----QYYYRSHRKTKLKPFERDLSSDE-------------NGVRTLSTYDFYPSVPVAYQAT
Bat Rm1              430:TAQQDQG-----QYYYRSYRKEKLKPFERDLSSDE-------------NGVYTLSTYDFYPSIPVEYQAT
Bat HKU3-1           430:TAKHDTG-----NYYYRSHRKTKLKPFERDLSSDDG------------NGVYTLSTYDFNPNVPVAYQAT
Bat Rf1              430:TAKQDVG-----SYFYRSHRSSKLKPFERDLSSEE-------------NGVRTLSTYDFNQNVPLEYQAT
Bat Rf4092           423:TAKQDVG-----SYFYRSHRSSKLKPFERDLSSDE-------------NGVRTLSTYDFNPNVPLDYQAT
Bat Ra7909           425:SKAKDE----NGQYFYRLFRKSKLLPFQRDVSNVTYGSGKNDGCNPSEADCYWPLLKYGFTGSVSQDYQPY
Bat BM48-31          431:TNSLDSSN----EFFYRRFRHGKIKPYGRDLSNVLFNPSGGTCSAEG-LNCYKPLASYGFTQSSGIGFQPY
```

Human ACE2 binding +

Human ACE2 binding -

B

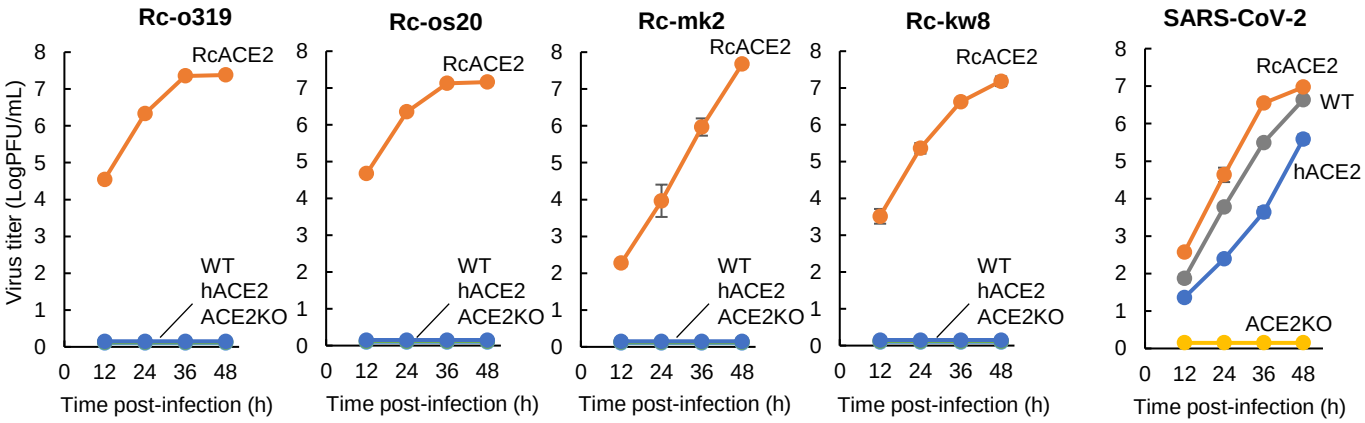


Fig. 4