

1

2 Discovery of a novel coronavirus associated with the recent pneumonia outbreak in
3 humans and its potential bat origin

4

5 Peng Zhou^{1*}, Xing-Lou Yang^{1*}, Xian-Guang Wang^{2*}, Ben Hu¹, Lei Zhang¹, Wei
6 Zhang¹, Hao-Rui Si^{1,3}, Yan Zhu¹, Bei Li¹, Chao-Lin Huang², Hui-Dong Chen², Jing
7 Chen^{1,3}, Yun Luo^{1,3}, Hua Guo^{1,3}, Ren-Di Jiang^{1,3}, Mei-Qin Liu^{1,3}, Ying Chen^{1,3}, Xu-
8 Rui Shen^{1,3}, Xi Wang^{1,3}, Xiao-Shuang Zheng^{1,3}, Kai Zhao^{1,3}, Quan-Jiao Chen¹, Fei
9 Deng¹, Lin-Lin Liu⁴, Bing Yan¹, Fa-Xian Zhan⁴, Yan-Yi Wang¹, Geng-Fu Xiao¹,
10 Zheng-Li Shi^{1†}

11

12 **Affiliations:**

13 ¹CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for
14 Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, People's Republic
15 of China

16 ²Wuhan Jinyintan hospital, Wuhan, China

17 ³University of Chinese Academy of Sciences, Beijing, People's Republic of China

18 ⁴Hubei Provincial Center for Disease Control and Prevention, Wuhan, People's
19 Republic of China

20 *These authors contributed equally.

21 †To whom correspondence should be addressed: zlshi@wh.iov.cn

22

23 **Since the SARS outbreak 18 years ago, a large number of severe acute**
24 **respiratory syndrome related coronaviruses (SARSr-CoV) have been discovered**
25 **in their natural reservoir host, bats¹⁻⁴. Previous studies indicated that some of**
26 **those bat SARSr-CoVs have the potential to infect humans⁵⁻⁷. Here we report the**
27 **identification and characterization of a novel coronavirus (nCoV-2019) which**
28 **caused an epidemic of acute respiratory syndrome in humans, in Wuhan, China.**
29 **The epidemic, started from December 12th, 2019, has caused 198 laboratory**
30 **confirmed infections with three fatal cases by January 20th, 2020. Full-length**
31 **genome sequences were obtained from five patients at the early stage of the**
32 **outbreak. They are almost identical to each other and share 79.5% sequence**
33 **identify to SARS-CoV. Furthermore, it was found that nCoV-2019 is 96%**
34 **identical at the whole genome level to a bat coronavirus. The pairwise protein**
35 **sequence analysis of seven conserved non-structural proteins show that this virus**
36 **belongs to the species of SARSr-CoV. The nCoV-2019 virus was then isolated**
37 **from the bronchoalveolar lavage fluid of a critically ill patient, which can be**
38 **neutralized by sera from several patients. Importantly, we have confirmed that**
39 **this novel CoV uses the same cell entry receptor, ACE2, as SARS-CoV.**

40

41 Coronavirus has caused two large-scale pandemic in the last two decades, SARS and
42 MERS (Middle East respiratory syndrome)^{8,9}. It was generally believed that SARSr-
43 CoV, mainly found in bats, might cause future disease outbreak^{10,11}. Here we report
44 on a series of unidentified pneumonia disease outbreaks in Wuhan, Hubei province,
45 central China (Extended Data Figure 1). Started from a local fresh seafood market, the
46 epidemic has resulted in 198 laboratory confirmed cases with three death according to
47 authorities so far¹². Typical clinical symptoms of these patients are fever, dry cough,

48 dyspnea, headache, and pneumonia. Disease onset may result in progressive
49 respiratory failure due to alveolar damage and even death. The disease was
50 determined as viral induced pneumonia by clinicians according to clinical symptoms
51 and other criteria including body temperature rising, lymphocytes and white blood
52 cells decreasing (sometimes normal for the later), new pulmonary infiltrates on chest
53 radiography, and no obvious improvement upon three days antibiotics treatment. It
54 appears most of the early cases had contact history with the original seafood market,
55 and no large scale of human-to-human transmission was observed so far.
56
57 Samples from seven patients with severe pneumonia (six are seafood market peddlers
58 or delivers), who were enrolled in intensive unit cares at the beginning of the outbreak,
59 were sent to WIV laboratory for pathogen diagnosis (Extended Data Table 1). As a
60 CoV lab, we first used pan-CoV PCR primers to test these samples¹³, considering the
61 outbreak happened in winter and in a market, same environment as SARS. We found
62 five PCR positive. A sample (WIV04) collected from bronchoalveolar lavage fluid
63 (BALF) was analysed by metagenomics analysis using next-generation sequencing
64 (NGS) to identify potential etiological agents. Of the 1582 total reads obtained after
65 human genome filtering, 1378 (87.1%) matched sequences of SARSr-CoV (Fig. 1a).
66 By *de novo* assembly and targeted PCR, we obtained a 29,891-bp CoV genome that
67 shared 79.5% sequence identity to SARS-CoV BJ01 (GenBank accession number
68 AY278488.2). This sequence has been submitted to GISAID (accession no.
69 EPI_ISL_402124). Following the name by WHO, we tentatively call it novel
70 coronavirus 2019 (nCoV-2019). Four more full-length genome sequences of nCoV-
71 2019 (WIV02, WIV05, WIV06, and WIV07) (GISAID accession nos.

72 EPI_ISL_402127-402130) that were above 99.9% identical to each other were
73 subsequently obtained from other four patients (Extended Data Table 2).

74

75 The virus genome consists of six major open reading frames (ORFs) common to
76 coronaviruses and a number of other accessory genes (Fig. 1b). Further analysis
77 indicates that some of the nCoV-2019 genes shared less than 80% nt sequence
78 identity to SARS-CoV. However, the seven conserved replicase domains in ORF1ab
79 that were used for CoV species classification, are 94.6% aa sequence identical
80 between nCoV-2019 and SARS-CoV, implying the two belong to same species
81 (Extended Data Table 3).

82

83 We then found a short RdRp region from a bat coronavirus termed BatCoV RaTG13
84 which we previously detected in *Rhinolophus affinis* from Yunnan Province showed
85 high sequence identity to nCoV-2019. We did full-length sequencing to this RNA
86 sample. Simplot analysis showed that nCoV-2019 was highly similar throughout the
87 genome to RaTG13 (Fig. 1c), with 96.2% overall genome sequence identity. The
88 phylogenetic analysis also showed that RaTG13 is the closest relative of the nCoV-
89 2019 and form a distinct lineage from other SARSr-CoVs (Fig. 1d). The receptor
90 binding protein spike (S) gene was highly divergent to other CoVs (Extended Data
91 Figure 2), with less than 75% nt sequence identity to all previously described SARSr-
92 CoVs except a 93.1% nt identity to RaTG13 (Extended Data Table 3). The S genes of
93 nCoV-2019 and RaTG13 S gene are longer than other SARSr-CoVs. The major
94 differences in nCoV-2019 are the three short insertions in the N-terminal domain, and
95 four out of five key residues changes in the receptor-binding motif, in comparison

96 with SARS-CoV (Extended Data Figure 3). The close phylogenetic relationship to
97 RaTG13 provides evidence for a bat origin of nCoV-2019.
98
99 We rapidly developed a qPCR detection based on the receptor-binding domain of
100 spike gene, the most variable region among genome (Fig. 1c). Our data show the
101 primers could differentiate nCoV-2019 with all other human coronaviruses including
102 bat SARSr-CoV WIV1, which is 95% identity to SARS-CoV (Extended Data Figure
103 4a and 4b). From the seven patients, we found nCoV-2019 positive in six BALF and
104 five oral swab samples during the first sampling by qPCR and conventional PCR
105 (Extended Data Figure 4c). However, we can no longer find viral positive in oral
106 swabs, anal swabs, and blood from these patients during the second sampling (Fig.
107 2a). Based on these findings, we conclude that the disease should be transmitted
108 through airway, yet we can't rule out other possibilities if the investigation extended
109 to include more patients.
110
111 For serological detection of nCoV-2019, we used previously developed bat SARSr-
112 CoV Rp3 nucleocapsid protein (NP) as antigen in IgG and IgM ELISA test, which
113 showed no cross-reactivity against other human coronaviruses except SARSr-CoV⁷.
114 As a research lab, we were only able to get five serum samples from the seven viral
115 infected patients. We monitored viral antibody levels in one patient (ICU-06) at seven,
116 eight, nine, and eighteen days after disease onset (Extended Data Table 2). A clear
117 trend of IgG and IgM titre (decreased at the last day) increase was observed (Fig. 2b).
118 For a second investigation, we tested viral antibody for five of the seven viral positive
119 patients around twenty days after disease onset (Extended Data Table 1 and 2). All

120 patient samples, but not samples from healthy people, showed strong viral IgG
121 positive (Fig. 2b). We also found three IgM positive, indicating acute infection.
122
123 We then successfully isolated the virus (named nCoV-2019
124 BetaCoV/Wuhan/WIV04/2019), in Vero and Huh7 cells using BALF sample from
125 ICU-06 patient. Clear cytopathogenic effects were observed in cells after three days
126 incubation (Extended Data Figure 5a and 5b). The identity of the strain WIV04 was
127 verified in Vero E6 cells by immunofluorescence microscopy using cross-reactive
128 viral NP antibody (Extended Data Figure 5c and 5d), and by metagenomic sequencing,
129 from which most of the reads mapped to nCoV-2019 (Extended Data Figure 5e and
130 5f). Viral partials in ultrathin sections of infected cells displayed typical coronavirus
131 morphology under electron microscopy (Fig. 3). To further confirm the neutralization
132 activity of the viral IgG positive samples, we conducted serum-neutralization assays
133 in Vero E6 cells using the five IgG positive patient sera. We demonstrate that all
134 samples were able to neutralize 120 TCID₅₀ nCoV-2019 at a dilution of 1:40-1:80.
135 We also show that this virus could be cross-neutralized by horse anti-SARS-CoV
136 serum at dilutions 1:80, further confirming the relationship of the two viruses
137 (Extended Data Table 4).
138
139 Angiotensin converting enzyme II (ACE2) was known as cell receptor for SARS-
140 CoV¹⁴. To determine whether nCoV-2019 also use ACE2 as a cellular entry receptor,
141 we conducted virus infectivity studies using HeLa cells expressing or not expressing
142 ACE2 proteins from humans, Chinese horseshoe bats, civet, pig, and mouse. We
143 show that nCoV-2019 is able to use all but mouse ACE2 as an entry receptor in the
144 ACE2-expressing cells, but not cells without ACE2, indicating which is likely the cell

145 receptor of nCoV-2019 (Fig. 4). We also proved that nCoV-2019 does not use other
146 coronavirus receptors, aminopeptidase N and dipeptidyl peptidase 4 (Extended Data
147 Figure 6).

148

149 The study provides the first detailed report on nCoV-2019, the likely etiology agent
150 responsible for ongoing acute respiratory syndrome epidemic in Wuhan, central China.

151 Viral specific nucleotide positive and viral protein seroconversion observed in all
152 patients tested provides evidence of an association between the disease and the

153 presence of this virus. However, there are still many urgent questions to be answered.

154 We need more clinical data and samples to confirm if this virus is indeed the etiology
155 agent for this epidemic. In addition, we still don't know if this virus continue evolving

156 and become more transmissible between human-to-human. Moreover, we don't know
157 the transmission routine of this virus among hosts yet. We showed viral positive in

158 oral swabs, implying nCoV-2019 may be transmitted through airway. However, this

159 needs to be confirmed by extending detection range. Finally, based on our results, it

160 should be expected and worth to test if ACE2 targeting or SARS-CoV targeting drugs

161 can be used for nCoV-2019 patients. At this stage, we know very little about the virus,

162 including basic biology, animal source or any specific treatment. The almost identical

163 sequences of this virus in different patients imply a probably recent introduction in

164 humans, thus future surveillance on viral mutation and transmission ability and

165 further global research attention are urgently needed.

166

167 **ACKNOWLEDGEMENTS:** We thank the Pei Zhang and An-na Du from WIV core

168 facility and technical support for their help with producing TEM micrographs. This

169 work was jointly supported by the Strategic Priority Research Program of the Chinese

170 Academy of Sciences (XDB29010101 to ZLS and XDB29010104 to PZ), China
171 Natural Science Foundation for excellent scholars (81822028 to PZ, 31770175 to ZLS
172 and 31800142 to BH), Mega-Project for Infectious Disease from Minister of Science
173 and Technology of the People's Republic of China (2020ZX09201001 to DYZ and
174 2018ZX10305409-004-001 to PZ), Youth innovation promotion association of CAS
175 (2019328 to XLY).

176

177 **AUTHOR CONTRIBUTIONS:** Z.L.S., P.Z., Y.Y.W., and G.F.X. conceived the
178 study. G.S.W., C.L.H., H.D.C., F.D., Q.J.C., F.X.Z., and LLL., collected patient
179 samples. X.L.Y., B.Y., W.Z., B.L., J.C., X.S.Z., Y.L., H.G., R.D.J., M.Q.L., Y. Chen,
180 X.W., X.R.S., and K.Z. performed qPCR, serology, and virus culturing. L.Z., Y.Z.,
181 H.R.S., and B.H. performed genome sequencing and annotations. The authors declare
182 no competing financial interests. Correspondence and requests for materials should be
183 addressed to ZLS (zlshi@wh.iov.cn).

184

- 185 1 Li, W. D. *et al.* Bats are natural reservoirs of SARS-like coronaviruses.
186 *Science* **310**, 676-679, (2005).
187 2 Ge, X. Y. *et al.* Isolation and characterization of a bat SARS-like
188 coronavirus that uses the ACE2 receptor. *Nature* **503**, 535-+, (2013).
189 3 Yang, L. *et al.* Novel SARS-like Betacoronaviruses in Bats, China, 2011.
190 *Emerg Infect Dis* **19**, 989-991, (2013).
191 4 Hu, B. *et al.* Discovery of a rich gene pool of bat SARS-related
192 coronaviruses provides new insights into the origin of SARS coronavirus.
193 *PLoS pathogens* **13**, e1006698, (2017).
194 5 Menachery, V. D. *et al.* A SARS-like cluster of circulating bat coronaviruses
195 shows potential for human emergence. *Nat Med* **21**, 1508-1513, (2015).
196 6 Menachery, V. D. *et al.* SARS-like WIV1-CoV poised for human emergence.
197 *Proc Natl Acad Sci U S A* **113**, 3048-3053, (2016).
198 7 Wang, N. *et al.* Serological Evidence of Bat SARS-Related Coronavirus
199 Infection in Humans, China. *Virol Sin* **33**, 104-107, (2018).
200 8 Drosten, C. *et al.* Identification of a novel coronavirus in patients with
201 severe acute respiratory syndrome. *New Engl J Med* **348**, 1967-1976,
202 (2003).

- 203 9 Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E. &
204 Fouchier, R. A. M. Isolation of a Novel Coronavirus from a Man with
205 Pneumonia in Saudi Arabia. *New Engl J Med* **367**, 1814-1820, (2012).
206 10 Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses.
207 *Nat Rev Microbiol* **17**, 181-192, (2019).
208 11 Fan, Y., Zhao, K., Shi, Z. L. & Zhou, P. Bat Coronaviruses in China. *Viruses* **11**,
209 (2019).
210 12 Wuhan Municipal Health Commission. *Press statement related to novel*
211 *coronavirus infection (in Chinese)*, 2020).
212 13 Poon, L. L. *et al.* Identification of a novel coronavirus in bats. *Journal of*
213 *virology* **79**, 2001-2009, (2005).
214 14 Li, W. *et al.* Angiotensin-Converting Enzyme 2 Is a Functional Receptor for
215 the Sars Coronavirus. *Nature* **426**, 450-454, (2003).
216

217 **Supplementary Information is available in the online version of the paper.**

218 **Main Figure Legend**

219

220

221

222

223

224

225

226

227

228

229

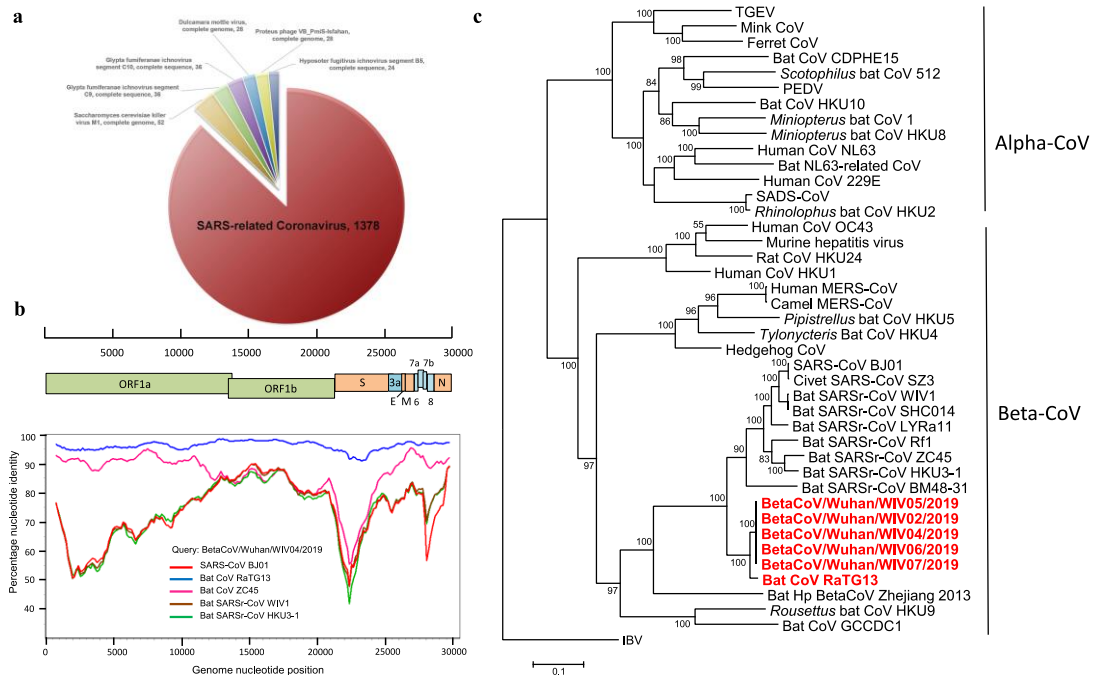
230

231

232

233

234 **Fig. 1 | Genome characterization of nCoV-2019. a,** pie chart showing
 235 metagenomics analysis of next-generation sequencing of bronchoalveolar lavage fluid
 236 from patient ICU06. **b,** Genomic organization of nCoV-2019 WIV04. **c,** Similarity
 237 plot based on the full-length genome sequence of nCoV-2019 WIV04. Full-length
 238 genome sequences of SARS-CoV BJ01, bat SARSr-CoV WIV1, bat coronavirus
 239 RaTG13 and ZC45 were used as reference sequences. **d,** Phylogenetic tree based on
 240 nucleotide sequences of complete ORF1b of coronaviruses. Software used and
 241 settings can be found in material and method section.
 242



243

244

245

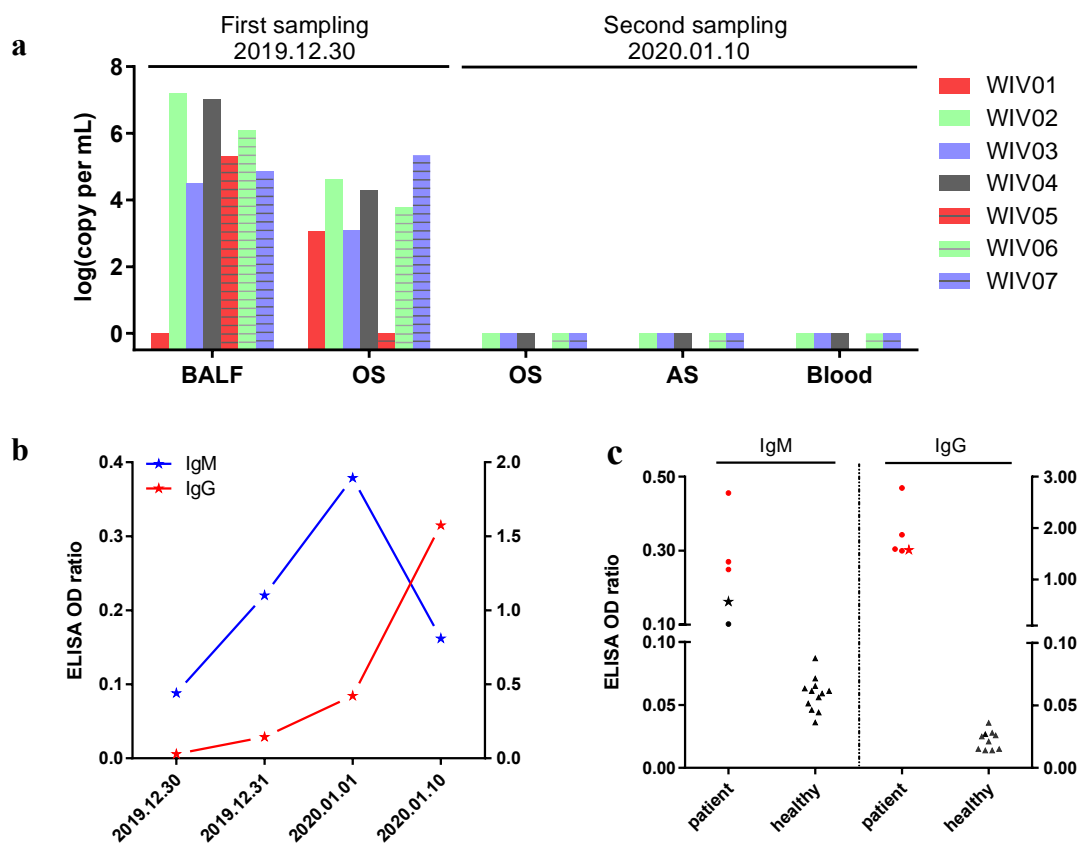
246

247

248

249

250 **Fig. 2 | Molecular and serological investigation of patient samples. a**, molecular
251 detection of nCoV-2019 in seven patients during two times of sampling. Patient
252 information can be found in Extended Data Table 1 and 2. Details on detection
253 method can be found in material and methods. BALF, bronchoalveolar lavage fluid;
254 OS, oral swab; AS, anal swab. **b**, dynamics of nCoV-2019 antibodies in one patient
255 who showed sign of disease on 2019.12.23 (ICU-06). **c**, serological test of nCoV-
256 2019 antibodies in five patients (more information can be found in Extended Data
257 Table 2). Star indicates data collected from patient ICU-06 on 2020.01.10. For b and c,
258 cut-off was set up as 0.2 for IgM test and 0.3 for IgG test, according to healthy
259 controls.
260

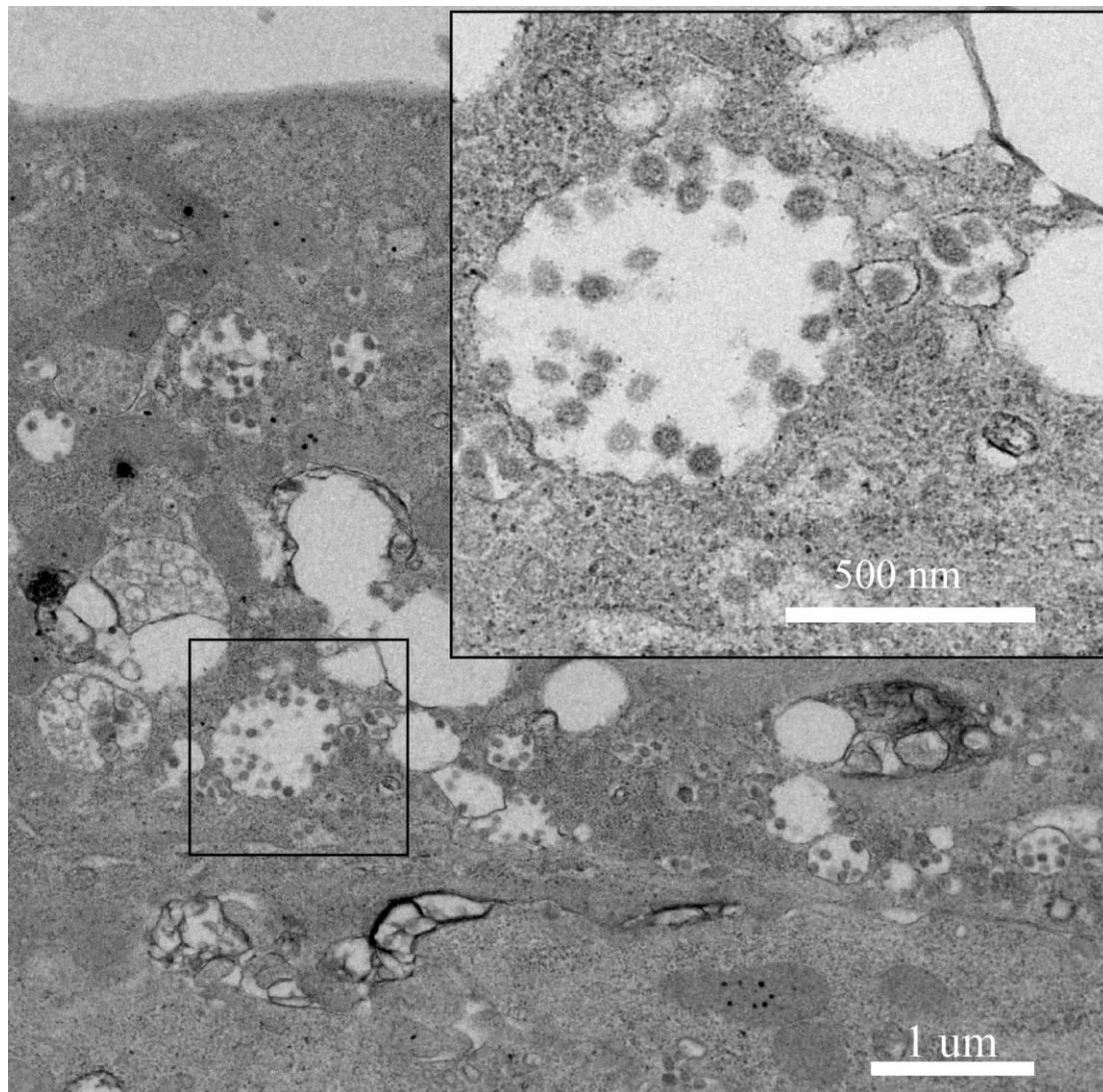


261

262

263 **Fig. 3 | Virions. a**, viral particles in the ultrathin sections under electron microscope

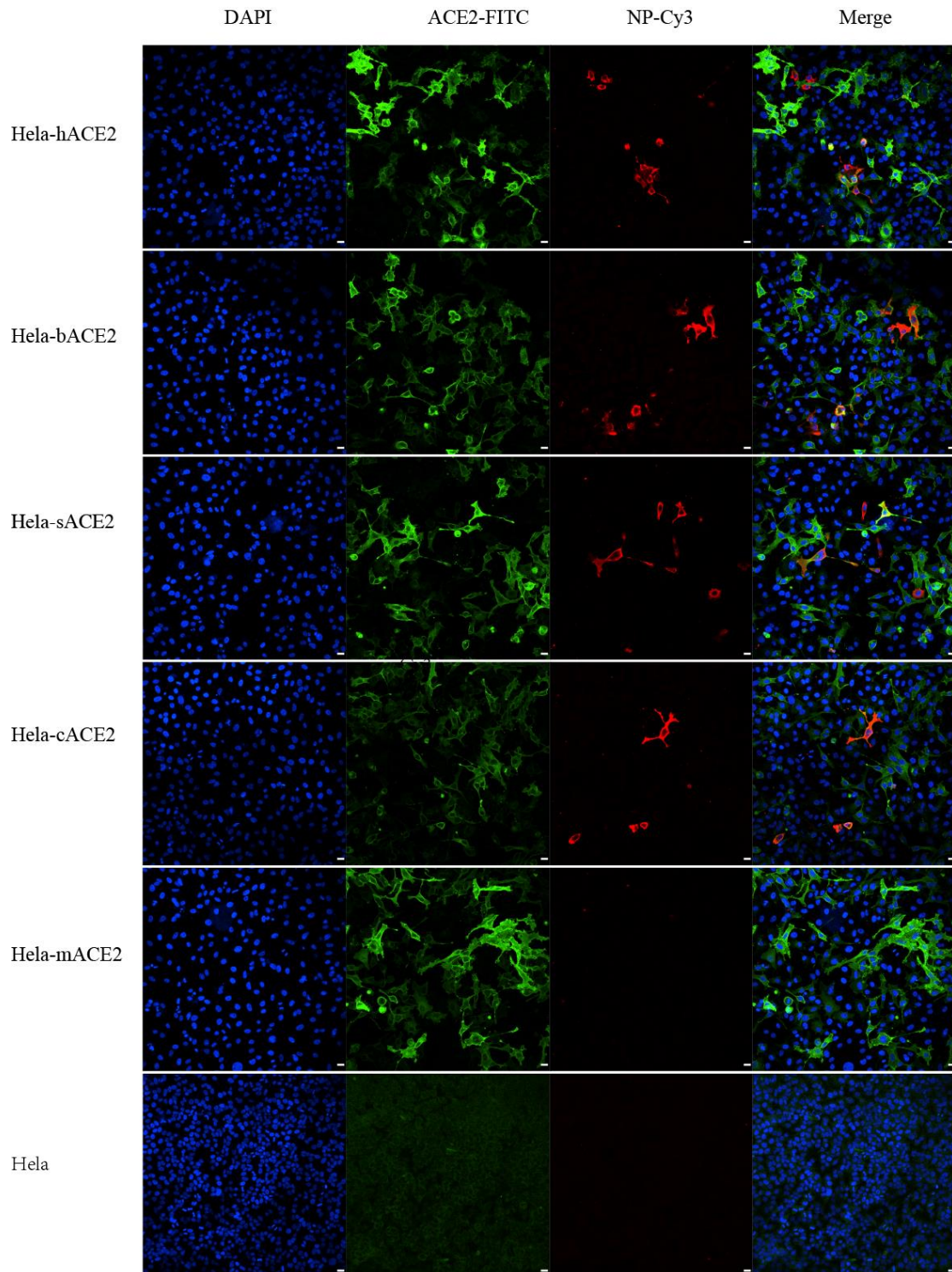
264 at 200 kV, sample from viral infected Vero E6 cells



265

266

267 **Fig. 4 | Analysis of nCoV-2019 receptor usage.** Determination of virus infectivity in
268 HeLa cells with or without the expression of ACE2. h, human; b, *Rhinolophus sinicus*
269 bat; c, civet; s, swine (pig); m, mouse. ACE2 protein (green), viral protein (red) and
270 nuclei (blue) was shown. Scale bar=10 um.



271

272

273 **METHODS**

274 **Sample collection.** Human samples, including oral swabs, anal swabs, blood, and
275 BALF samples were collected by Jinyintan hospital (Wuhan) with the consent from
276 all patients. Patients were sampled without gender or age preference unless where
277 indicated. For swabs, 1.5 ml DMEM+2% FBS medium was added each tube.
278 Supernatant was collected after 2500 rpm, 60 s vortex and 15-30 min standing.
279 Supernatant from swabs or BALF (no pretreatment) was added to either lysis buffer
280 for RNA extraction or to viral transport medium (VTM) for virus isolation. VTM
281 composed of Hank's balanced salt solution at pH7.4 containing BSA (1%),
282 amphotericin (15 µg/ml), penicillin G (100 units/ml), and streptomycin (50 µg/ml).
283 Serum was separated by centrifugation at 3,000 g for 15 min within 24 h of collection,
284 followed by 56 °C 30 min inactivation, and then stored at 4 °C until use.

285

286 **Virus isolation, cell infection, electron microscope and neutralization assay.** The
287 following cells were used for virus isolation in this study: Vero, Vero E6, and Huh7
288 that were cultured in DMEM +10% FBS. A list of cells were used for susceptibility
289 test (Extended Data Fig. 6). All cell lines were tested free of mycoplasma
290 contamination, applied to species identification and authenticated by microscopic
291 morphologic evaluation. None of cell lines was on the list of commonly misidentified
292 cell lines (by ICLAC).

293

294 Cultured cell monolayers were maintained in their respective medium. PCR-positive
295 BALF sample from ICU-06 patient was spin at 8,000 g for 15 min, filtered and
296 diluted 1:2 with DMEM supplied with 16 µg/ml trypsin before adding to cells. After
297 incubation at 37 °C for 1 h, the inoculum was removed and replaced with fresh culture

298 medium containing antibiotics (below) and 16 µg/ml trypsin. The cells were incubated
299 at 37 °C and observed daily for cytopathic effect (CPE). The culture supernatant was
300 examined for presence of virus by qRT-PCR developed in this study, and cells were
301 examined by immunofluorescent using SARSr-CoV Rp3 NP antibody made in house
302 (1:100). Penicillin (100 units/ml) and streptomycin (15 µg/ml) were included in all
303 tissue culture media.

304

305 The Vero E6 cells were infected with new virus at MOI of 0.5 and harvested 48 hpi.
306 Cells were fixed with 2.5% (wt/vol) glutaraldehyde and 1% osmium tetroxide, and
307 then dehydrated through a graded series of ethanol concentrations (from 30 to 100%),
308 and embedded with epoxy resin. Ultrathin sections (80 nm) of embedded cells were
309 prepared, deposited onto Formvar-coated copper grids (200 mesh), stained with
310 uranyl acetate and lead citrate, then observed under 200 kV Tecnai G2 electron
311 microscope.

312

313 The virus neutralization test was carried out in a 48-well plate. The patient serum
314 samples were heat-inactivated by incubation at 56 °C for 30 min before use. The
315 serum samples (5 µL) were diluted to 1:10, 1:20, 1:40 or 1:80, and then an equal
316 volume of virus stock was added and incubated at 37 °C for 60 min in a 5% CO₂
317 incubator. Diluted horse anti SARS-CoV serum or serum samples from healthy
318 people were used as control. After incubation, 100 µL mixtures were inoculated onto
319 monolayer Vero E6 cells in a 48-well plate for 1 hour. Each serum were repeated
320 triplicate. After removing the supernatant, the plate was washed twice with DMEM
321 medium. Cells were incubated with DMEM supplemented with 2% FBS for 24 hours.
322 Then the cells were fixed with 4% formaldehyde. And the virus were detected using

323 SL-CoV Rp3 NP antibody followed by Cy3-conjugated mouse anti-rabbit IgG. Nuclei
324 were stained with DAPI. Infected cell number was counted by high-content
325 cytometers.

326

327 **RNA extraction and PCR.** Whenever commercial kits were used, manufacturer's
328 instructions were followed without modification. RNA was extracted from 200 µl of
329 samples with the High Pure Viral RNA Kit (Roche). RNA was eluted in 50 µl of
330 elution buffer and used as the template for RT-PCR.

331

332 For qPCR analysis, primers based on nCoV-2019 S gene was designed: RBD-qF1: 5'-
333 CAATGGTTTAAACAGGCACAGG-3'; RBD-qR1: 5'-

334 CTCAAGTGTCTGTGGATCACG-3'. RNA extracted from above used in qPCR by
335 HiScript® II One Step qRT-PCR SYBR® Green Kit (Vazyme Biotech Co.,Ltd).

336 Conventional PCR test was also performed using the following primer pairs: ND-

337 CoVs-951F TGTKAGRTTYCCTAAYATTAC; ND-CoVs-1805R

338 ACATCYTGATANARAACAGC¹³. The 20 µl qPCR reaction mix contained 10 µl 2×

339 One Step SYBR Green Mix, 1 µl One Step SYBR Green Enzyme Mix, 0.4 µl 50 ×

340 ROX Reference Dye 1, 0.4 µl of each primer (10 uM) and 2 µl template RNA.

341 Amplification was performed as follows: 50 °C for 3 min, 95 °C for 30 s followed by

342 40 cycles consisting of 95 °C for 10 s, 60 °C for 30 s, and a default melting curve step

343 in an ABI 7700 machine.

344

345 **Serological test.** In-house anti-SARSr-CoV IgG and IgM ELISA kits were developed

346 using SARSr-CoV Rp3 NP as antigen, which shared above 90% amino acid identity

347 to all SARSr-CoVs². For IgG test, MaxiSorp Nunc-immuno 96 well ELISA plates

348 were coated (100 ng/well) overnight with recombinant NP. Human sera were used at
349 1:20 dilution for 1 h at 37 °C. An anti-Human IgG-HRP conjugated monoclonal
350 antibody (Kyab Biotech Co., Ltd, Wuhan, China) was used at a dilution of 1:40000.
351 The OD value (450–630) was calculated. For IgM test, MaxiSorp Nunc-immuno 96
352 wellELISA plates were coated (500 ng/well) overnight with anti-human IgM (μ
353 chain). Human sera were used at 1:100 dilution for 40 min at 37 °C, followed by anti-
354 Rp3 NP-HRP conjugated (Kyab Biotech Co., Ltd, Wuhan, China) at a dilution of
355 1:4000. The OD value (450–630) was calculated.

356

357 **Examination of ACE2 receptor for nCoV-2019 infection.** HeLa cells transiently
358 expressing ACE2 were prepared by a lipofectamine 3000 system (Thermo Fisher
359 Scientific) in 96-well plate, with mock-transfected cells as controls. nCoV-2019
360 grown from Vero E6 cells was used for infection at multiplicity of infection 0.05.
361 Same for testing of APN and DPP4. The inoculum was removed after 1 h absorption
362 and washed twice with PBS and supplemented with medium. At 24 hpi, cells were
363 washed with PBS and fixed with 4% formaldehyde in PBS (pH 7.4) for 20 min at
364 room temperature. ACE2 expression was detected using mouse anti-S tag monoclonal
365 antibody followed by FITC-labelled goat anti-mouse IgG H&L (Abcam, ab96879).
366 Viral replication was detected using rabbit antibody against the Rp3 NP protein (made
367 in house, 1:100) followed by cyanin 3-conjugated goat anti-rabbit IgG (1:50, Abcam,
368 ab6939). Nucleus was stained with DAPI (Beyotime). Staining patterns were
369 examined using the FV1200 confocal microscopy (Olympus).

370

371 **High throughput sequencing, pathogen screening and genome assembly.** Samples
372 from patient BALF or from virus culture supernatant were used for RNA extraction

373 and next-generation sequencing using Illumina MiSeq 3000 sequencer. Metagenomic
374 analysis was carried out mainly base on the bioinformatics platform MGmapper
375 (PE_2.24 and SE_2.24). The raw NGS reads were firstly processed by Cutadapt
376 (v1.18) with minimum read length of 30bp. BWA (v0.7.12-r1039) was utilized to
377 align reads to local database with a filter hits parameter at 0.8 FMM value and
378 minimum alignment score at 30. Parameters for post-processing of assigned reads was
379 set with minimum size normalized abundance at 0.01, minimum read count at 20 and
380 other default parameters. A local nucleic acid database for human and mammals was
381 employed to filter reads of host genomes before mapping reads to virus database. The
382 results of metagenomic analysis were displayed through pie charts using WPS Office
383 2010. NGS reads were assembled into genomes using Geneious (v11.0.3) and
384 MEGAHIT (v1.2.9). PCR and Sanger sequencing was performed to fill gaps in the
385 genome. 5'-RACE was performed to determine the 5'-end of the genomes using
386 SMARTer RACE 5'/3' Kit (Takara). Genomes were annotated using Clone Manager
387 Professional Suite 8 (Sci-Ed Software).

388

389 **Phylogenetic analysis.** Routine sequence management and analysis was carried out
390 using DNASTar. Sequence alignment and editing were conducted using ClustalW and
391 GeneDoc. Maximum Likelihood phylogenetic trees based on nucleotide sequences of
392 full-length ORF1b and S genes were constructed using the Jukes-Cantor model with
393 bootstrap values determined by 1000 replicates in the MEGA6 software package.

394

395 **Data Availability statement.** Sequence data that support the findings of this study
396 have been deposited in GISAID with the accession no. EPI_ISL_402124 and
397 EPI_ISL_402127-402130.