AMERICAN SOCIETY FOR MICROBIOLOGY | Journals.ASM.org

ASM Journals: Bringing You the Latest Research in Public Health

I. **Discovery of Bat Coronaviruses through Surveillance and Probe Capture-Based Next-Generation Sequencing - *mSphere*®**
(Bei Li, Hao-Rui Si, Yan Zhu, Xing-Lou Yang, Danielle E. Anderson, Zheng-Li Shi, Lin-Fa Wang, Peng Zhou)

II. **Middle East Respiratory Syndrome Coronavirus Antibodies in Bactrian and Hybrid Camels from Dubai - *mSphere*®**
(Susanna K. P. Lau, Kenneth S. M. Li, Hayes K. H. Luk, Zirong He, Jade L. L. Teng, Kwok-Yung Yuen, Ulrich Wernery, Patrick C. Y. Woo)

III. **Receptor Recognition by Novel Coronavirus from Wuhan: An Analysis Based on Decade-Long Structural Studies of SARS - *Journal of Virology*®**
(Yushun Wan, Jian Shang, Rachel Graham, Ralph S. Baric, Fang Li)

RESEARCH ARTICLE
Clinical Science and Epidemiology

Discovery of Bat Coronaviruses through Surveillance and Probe Capture-Based Next-Generation Sequencing

Bei Li,[a] Hao-Rui Si,[a,b] Yan Zhu,[a] Xing-Lou Yang,[a] Danielle E. Anderson,[c] Zheng-Li Shi,[a] Lin-Fa Wang,[c] Peng Zhou[a]

[a]CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, China
[b]University of Chinese Academy of Sciences, Beijing, China
[c]Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore, Singapore

Bei Li and Hao-Rui Si contributed equally. Author order was determined by the importance of their work.

ABSTRACT Coronaviruses (CoVs) of bat origin have caused two pandemics in this century. Severe acute respiratory syndrome (SARS)-CoV and Middle East respiratory syndrome (MERS)-CoV both originated from bats, and it is highly likely that bat coronaviruses will cause future outbreaks. Active surveillance is both urgent and essential to predict and mitigate the emergence of these viruses in humans. Next-generation sequencing (NGS) is currently the preferred methodology for virus discovery to ensure unbiased sequencing of bat CoVs, considering their high genetic diversity. However, unbiased NGS is an expensive methodology and is prone to missing low-abundance CoV sequences due to the high background level of nonviral sequences present in surveillance field samples. Here, we employ a capture-based NGS approach using baits targeting most of the CoV species. Using this technology, we effectively reduced sequencing costs by increasing the sensitivity of detection. We discovered nine full genomes of bat CoVs in this study and revealed great genetic diversity for eight of them.

IMPORTANCE Active surveillance is both urgent and essential to predict and mitigate the emergence of bat-origin CoV in humans and livestock. However, great genetic diversity increases the chance of homologous recombination among CoVs. Performing targeted PCR, a common practice for many surveillance studies, would not reflect this diversity. NGS, on the other hand, is an expensive methodology and is prone to missing low-abundance CoV sequences. Here, we employ a capture-based NGS approach using baits targeting all CoVs. Our work demonstrates that targeted, cost-effective, large-scale, genome-level surveillance of bat CoVs is now highly feasible.

KEYWORDS bat, coronavirus, genome, enrichment, next-generation sequencing

Coronaviruses (CoVs) have the largest nonsegmented genomes among all RNA viruses, reaching up to 30 kb in length. The large genomes enhance plasticity, thereby allowing modification by mutations and recombination, which in turn leads to greater genetic diversity and high chances of cross-species transmission (1, 2). The major reason for this phenomenon may be the numerous subgenomic RNAs generated during viral replication, which increase the chance of homologous recombination among closely related genes from different lineages of CoVs or other viruses (3, 4). As a result, CoV taxonomy is constantly changing. Currently, there are four genera (*Alpha-*, *Beta-*, *Gamma-*, and *Deltacoronavirus*) consisting of 38 unique species in the CoV subfamily *Orthocoronavirinae*, and the number is still increasing (5). Open reading frame 1b (ORF1b) is the gene used for classification, but viruses in the same species may show great diversity in regions outside ORF1b, confounding the designation (6). Bat CoVs

Citation Li B, Si H-R, Zhu Y, Yang X-L, Anderson DE, Shi Z-L, Wang L-F, Zhou P. 2020. Discovery of bat coronaviruses through surveillance and probe capture-based next-generation sequencing. mSphere 5:e00807-19. https://doi.org/10.1128/mSphere.00807-19.

Editor Matthew B. Frieman, University of Maryland, College Park

Address correspondence to Lin-Fa Wang, linfa.wang@duke-nus.edu.sg, or Peng Zhou, peng.zhou@wh.iov.cn.

Received 6 November 2019
Accepted 19 December 2019
Published 29 January 2020

classed as the same species can differ significantly in terms of receptor usage or virus-host interaction, as observed in bat severe acute respiratory syndrome (SARS)-related CoVs (SARSr-CoVs) (7). This difference would not be reflected by performing targeted PCR on short genomic fragments of ORF1b, currently a common practice for many surveillance studies (8).

Over the past 20 years, two pandemics, SARS and Middle East respiratory syndrome (MERS), have been attributed to CoVs (9, 10). The outbreak in 2018 of swine acute diarrhea syndrome (SADS), another bat CoV, is a timely reminder that CoVs will continue to emerge and cause new outbreaks in the future (11). All three disease agents can be traced back to bats, animals known to harbor other deadly viruses, including Ebola virus, Marburg virus, Nipah virus, and Hendra virus (12). Bat CoVs are highly prevalent around the world and also show great genetic diversity, making up almost 60% of all known *Alpha-* and *Betacoronavirus* species. It is generally believed that some of these bat CoVs have the potential to spill over into humans and other mammalian species, causing another SARS-like pandemic (13). While predicting the potential spillover and emergence of a novel bat coronavirus is difficult, active surveillance is a valuable monitoring mechanism. Surveillance programs have been designed to aid viral discovery in wildlife reservoir hosts to mitigate infection and emergence in the human population. These programs propose to use next-generation sequencing (NGS) and other approaches to ensure unbiased evolutionary analysis of bat CoVs that takes into consideration their high genetic diversity (14). In order to be effective, these types of surveillance programs rely upon processing of samples in a high-throughput manner and require the compilation of whole-genome sequences. Although NGS enables unbiased pathogen discovery, implementation of this methodology for virus surveillance is costly. Additionally, the inherent lack of sensitivity with an unbiased approach increases the burden of data analysis and decreases the chance of detection in field samples with low viral loads.

Strategies to improve the efficiency of NGS have been explored, including subtraction of host genetic material or enrichment of viral nucleic acid through positive selection using a capture-based system, where the latter was proven more cost-effective (15–17). Virus enrichment NGS has been successfully used for various viral families, and the most common protocols rely on predesigned viral probes that share more than 60% homology with the target virus sequence (15–17). In this study, we utilized an enrichment NGS approach with predesigned probes targeting most of the CoV species (18). Our aim was to strategically perform bat CoV surveillance in which high-throughput sample processing for virus discovery would be balanced with cost effectiveness. Ultimately, the aim is to determine the best strategy to mitigate potential virus emergence in the future.

RESULTS

Enrichment NGS aids in the detection and characterization of diverse CoVs. In surveillance studies, detection and characterization are fundamental requirements to fully assess the risk that bat CoVs pose to humans. We aimed to address two main issues encountered during surveillance studies. First, many samples are collected but not all samples harbor viruses. Second, when viruses are detected, the high genetic diversity of CoVs means that full-length genome sequencing is essential to fully characterize viruses. In the context of CoV discovery, it is not time or cost-effective to perform unbiased NGS on all samples. Most data generated from unbiased NGS can be attributed to non-CoV-specific reads. To assess whether samples can be enriched to allow sequencing of only CoV-specific reads, we utilized NGS in conjunction with viral nucleic acid capture specifically targeting most of the known CoVs using a pool of 4,303 unique baits (18). These baits were designed from 90 representative CoV genomes, and *in silico* analysis determined that these baits should target all known CoV species tested here (Table S2 in the supplemental material).

A panel of 5 diverse CoVs (SARSr-CoV, MERS-CoV, porcine epidemic diarrhea virus [PEDV], transmissible gastroenteritis coronavirus [TGEV], and mouse hepatitis virus

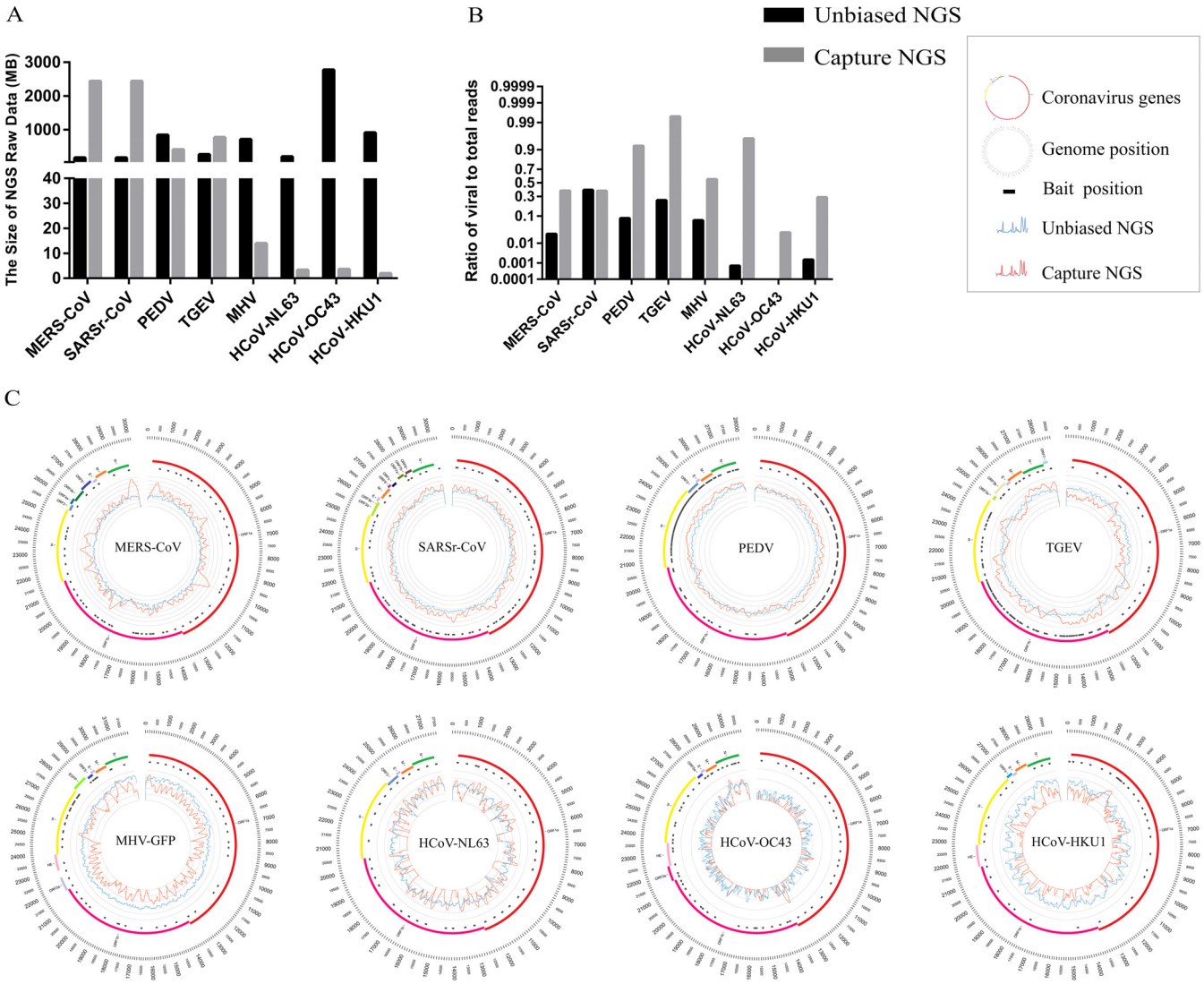


FIG 1 Next-generation sequencing (NGS) using a coronavirus (CoV) enrichment approach. Five cultured viruses (SARSr-CoV, MERS-CoV, PEDV, TGEV, and MHV) and three human clinical samples (HCoV-HKU1, HCoV-OC43, and HCoV-NL63) were used. (A) Amount of NGS data in megabytes (Mb). Data amounts were compared between unbiased and enriched NGS. (B) Ratios of viral to total reads were determined by mapping reads to the respective reference genome using CLC genomics. (C) CoV Circos plots. RNA extracted from cell culture supernatant or human oral swabs was subjected to NGS analysis. Circos plots, from outer to inner circle: CoV genome length (bp), genome annotation, CoV bait positions, read depth from direct NGS (blue lines), and read depth from enrichment NGS (red lines). Scale of read depth is shown as seven thin circular lines and ranges from 0 to 10^6. Sample details can be found in Table S2 in the supplemental material.

[MHV]) were amplified in cell culture, and RNA was extracted from the supernatants. Additionally, to test the robustness of the assay, RNA was extracted from 3 clinical samples (oral swabs from humans infected with human CoV OC43 [HCoV-OC43], HCoV-HKU1, and HCoV-NL63). NGS libraries were constructed and either directly sequenced or subjected to enrichment prior to sequencing. The 17 captured samples were made into two pools (8 or 9 per pool) for sequencing. The total amount of data obtained was variable across samples, but in swabs, unbiased NGS consistently produced more data (Fig. 1A). Within these data sets, the ratios of viral reads to total number of reads increased by almost 100% for captured samples, in contrast to the ratios of less than 1% for most of the unbiased NGS (Fig. 1B). The high ratio of viral to total reads in conjunction with decreased data size reduces the sequencing cost and data analysis burden. This methodology could thus greatly facilitate large-scale surveillance studies.

Once viral reads are detected in a sample, enrichment NGS can be retrospectively complemented with unbiased NGS and/or additional Sanger sequencing to obtain full-length genomes. The full-length genomes were obtained by NGS for the five cultured viruses and with minimal further gap filling for HCoV-HKU1 (240 bp), HCoV-NL63 (566 bp), and HCoV-OC43 (2,465 bp). The efficiency of CoV enrichment NGS was closely related to the number of baits and cycle threshold (C_T) value (Fig. 1C and Table S2). Read depth and genome coverage were compared between unbiased NGS and enrichment NGS. The read depth for SARS-related coronavirus (SARSr-CoV), MERS-CoV, PEDV, and TGEV increased from 10- to 1,000-fold throughout the genome after enrichment. The increase in read depth can be partially attributed to the high viral titers in cultured samples. Sequencing of the full-length genome of green fluorescent protein-labeled MHV (MHV-GFP) was successful, although the read depth was lower than for unbiased NGS (Fig. 1C). An increase in read depth of at least 10-fold was observed in HCoV-NL63 and HCoV-OC43 in regions where baits were present. Sequencing of the partial genome of HCoV-HKU1 was successful with enrichment NGS, but the full-length genome was obtained with unbiased NGS (Fig. 1C). Taken together, these data indicate that enrichment NGS not only decreases the amount of data requiring analysis but can produce full-length genome coverage in both laboratory and clinical samples.

Discovery of bat CoV genomes using capture-based NGS. The NGS pipeline was assessed for CoV discovery in bat samples. Samples from representative bat CoV species were selected based on RNA-dependent RNA polymerase (RdRp) sequence similarity to that of reference genomes. Similar to the human swab samples (Fig. 1), more data were obtained from unbiased NGS, but a higher ratio of viral to total reads was observed after enrichment (Fig. 2A and B). An increase of at least 10-fold in read depth was observed for BtMiCoV-1 (*Miniopterus bat coronavirus 1*), BtMiCoV-HKU8r (*Miniopterus bat coronavirus HKU8* related) (hereinafter, "r" denotes "related"), BtRhCoV-HKU2r (*Rhinolophus bat coronavirus HKU2* related), and BtPiCoV-HKU5r (*Pipistrellus bat coronavirus HKU5* related) in regions where baits were located (Fig. 2C). Although reads were obtained for BtRaCoV-229Er (*Human coronavirus 229E* related; sampled from *Rousettus aegyptiacus* bat) and BtScCoV-512r (*Scotophilus bat coronavirus 512* related) after enrichment, more virus-specific reads were obtained with unbiased NGS. Similarly, the efficiency of unbiased NGS was poor on BtHpCoV-HKU10r (*Bat coronavirus HKU10* related; sampled from *Hipposideros pomona* bat), BtHiCoV-CHB25 (related to *Bat coronavirus HKU10*; sampled from *Hipposideros pomona* bat), and BtTyCoV-HKU4r (*Tylonycteris bat coronavirus HKU4* related). In total, full-length genome coverage was obtained for six of nine genomes without further gap filling. More than 75% genome coverage was obtained for another 3 samples. Although complete genome coverage was obtained mostly from unbiased NGS, targeted enrichment clearly identified the presence of CoVs in bat samples. In a surveillance study, targeted enrichment is a valuable tool to triage samples for further processing.

Diversity of bat CoV genomes. To assess the diversity of nine novel bat CoV genomes, a phylogenetic tree was constructed using the conserved ORF1b protein as a reference (Fig. 3A). The newly identified viruses were most closely related to BtPiCoV-HKU5 (BtPiCoV-HKU5r), BtTyCoV-HKU4 (BtTyCoV-HKU4r), BtMiCoV-1 (BtMiCoV-1r), BtRhCoV-HKU2 (BtRhCoV-HKU2r), BtRaCoV-229E (BtRaCoV-229Er), BtScCoV-512 (BtScCoV-512r), BtMiCoV-HKU8 (BtMiCoV-HKU8r), and BtHpCoV-HKU10 (BtHpCoV-HKU10r and BtHiCoV-CHB25). In addition to the comparison with ORF1b at the protein level, the genomes of the newly identified viruses were compared to their respective reference genomes at the nucleotide level. The nucleotide sequence similarities were 97%, 96%, 96%, 89%, 85%, 91%, and 90% for BtPiCoV-HKU5r, BtTyCoV-HKU4r, BtMiCoV-1r, BtRhCoV-HKU2r, BtRaCoV-229Er, BtScCoV-512r, and BtMiCoV-HKU8r, respectively. The nucleotide sequence similarities of BtHpCoV-HKU10r and BtHiCoV-CHB25 to BtHpCoV-HKU10 were 88% and 73%, respectively (Fig. 3B).

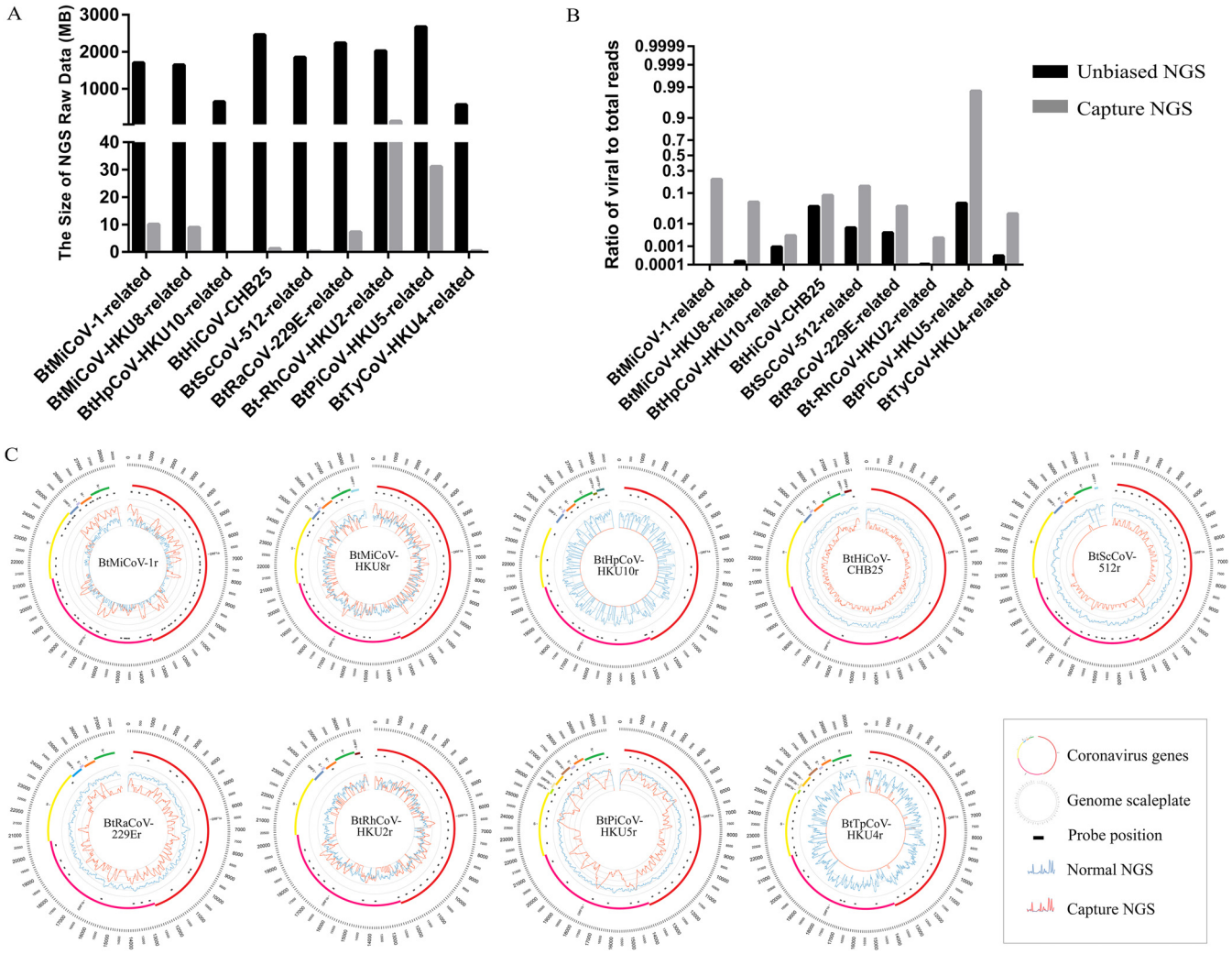


FIG 2 Bat CoV genome discovery using enrichment NGS. Nine bat CoV-positive samples from previous viral surveillances were used. (A) NGS data sizes were compared. (B) Ratios of viral to total reads were determined by mapping reads to the respective reference genome using CLC genomics. (C) CoV Circos plots, from outer to inner circle: CoV genome length (nt), genome annotation, CoV bait positions, read depth from direct NGS (blue lines), and read depth from enrichment NGS (red lines). Scale of read depth is shown as seven thin circular lines and ranges from 0 to 10^6. Sample details can be found in Table S2.

The most divergent region of the genome was the region encoding the N terminus of the spike protein, which is usually responsible for receptor binding. In this region, seven of eight genomes showed less than 90% nucleotide identity, and five were below 40% nucleotide identity, suggesting these viruses may utilize a different receptor than their reference viruses. The most divergent, BtMiCoV-HKU8r, shared less than 10% sequence identity in this region. Another divergent region in the CoV genome is the region encoding the C terminus of the product of the N gene and the 3′ untranslated region (UTR) of the accessory protein gene (Fig. 3B). CoV accessory proteins are responsible for host response modulation and are highly variable among CoVs (1). The diversity observed in the genomes of these newly identified viruses suggests that these CoVs may be quite different in terms of receptor usage or virus-host interaction.

DISCUSSION

Zoonotic viruses have caused most of the emerging viral disease outbreaks in recent years, and global virome surveillance programs were launched to evaluate the feasibility of preemptively mitigating pandemic threats (14). Unbiased approaches like NGS are powerful and effective, but at the same time, these methodologies are not cost-effective for routine or large-scale surveillance. Based on past experience, we

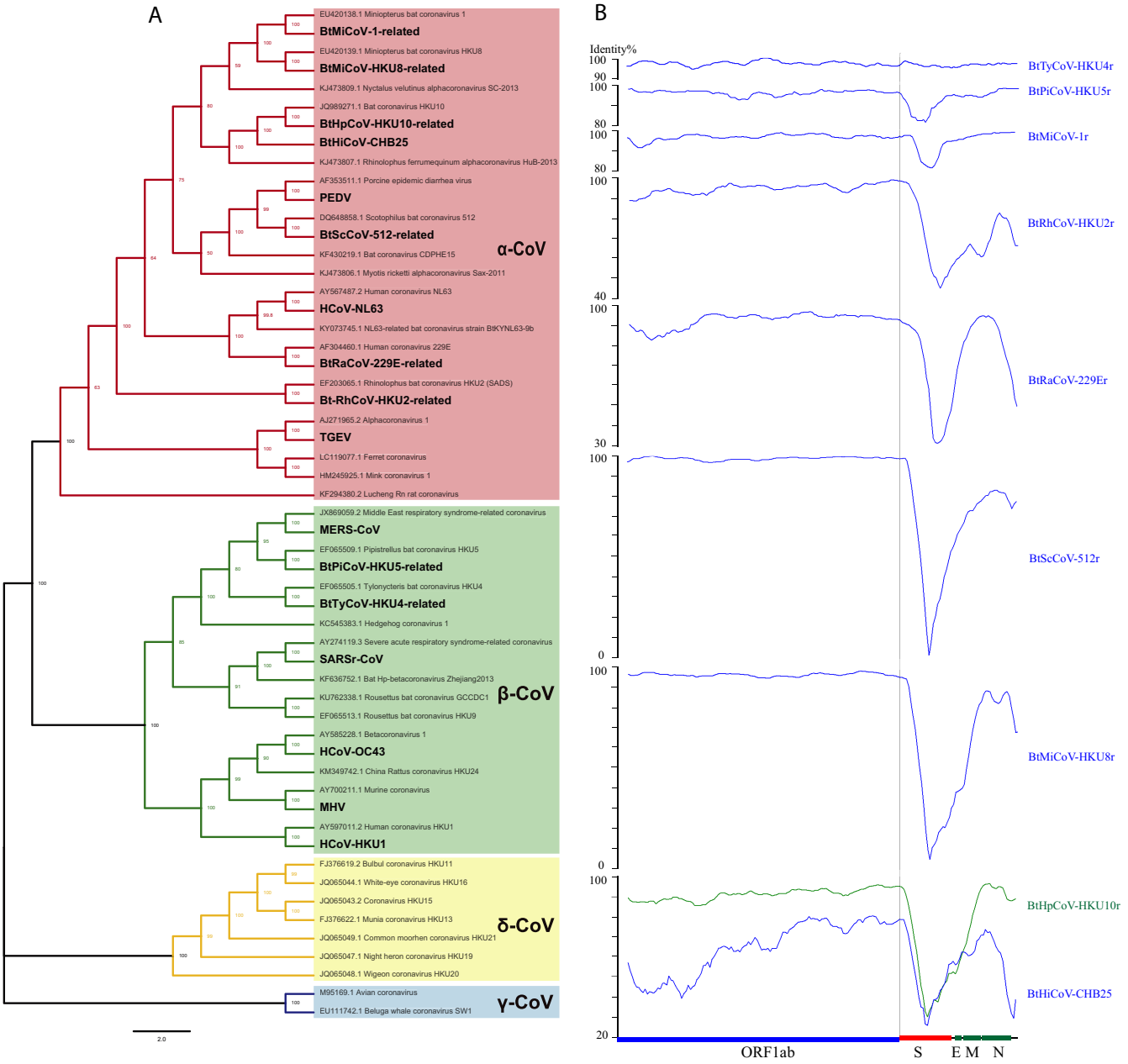


FIG 3 Analysis of bat CoV genomes. (A) Bayesian phylogenetic tree of ORF1b sequences from reference genomes or from CoV genomes analyzed in this study (boldface). NCBI accession numbers of reference genomes are shown. (B) Similarity plot based on the full-length genome sequences of bat CoVs. Bat CoV genomes from this study were compared to their reference genomes. The analysis was performed with the Kimura model, using a window size of 1,500 bp and a step size of 150 bp. The *x* axis illustrates typical genes in a CoV genome, and the genes are drawn to scale. The gray vertical line represents a breakpoint in most of the bat CoV species. Plots were adjusted to be the same length, as some CoVs have longer genomes. S, spike glycoprotein gene; E, small membrane protein gene; M, membrane protein gene; N, nucleocapsid protein gene. The 5′-UTR (before ORF1ab) and 3′-UTR (after N gene) regions are not shown.

expect bat CoVs to cause future outbreaks. The goal of this study was to develop an efficient and cost-effective pipeline to identify and characterize bat CoVs in future surveillance projects. Toward this end, we performed unbiased and targeted NGS on known and unknown CoVs in both laboratory and field samples.

Full-length sequences were obtained for most of the 17 CoVs in this study using unbiased NGS, but the depth of coverage differed between samples. There was an expected correlation between the amount of virus in the sample, as measured by quantitative PCR (qPCR), and the read depth obtained. PCR gap filling could be used to obtain the full-length genomes. We then compared the results of unbiased NGS with

those of enrichment NGS. The CoV enrichment NGS approach with our custom bait panel resulted in good performance in most of the samples tested, but the sequencing cost was dramatically reduced. In our study, the approximate per-sample cost of HiSeq NGS (2 Gb of data) was $100, while the cost of enrichment NGS in a 10-plex sample format was approximately $60, including the hybridization and bait costs (detailed in Table S3 in the supplemental material). The cost (influenced by data size) can be further reduced by multiplexing more samples in one run. Based on the data obtained in this study, we recommend multiplexing 48 samples per sequencing run.

The use of targeted NGS for virus discovery is not new (15–17, 19). Notably, Virocap and VirCapSeq-VERT are two well-established platforms targeting viruses that infect vertebrate hosts (15, 17). Due to the broad nature of these platforms, the libraries only include a relatively small proportion of CoV baits. Furthermore, the included CoV baits are biased toward pandemic viruses, such as SARS-CoV and MERS-CoV, for which more sequence information is available in the NCBI database. The effectiveness of these baits for capturing CoVs was only tested on SARS-CoV and MERS-CoV (15, 17). We specifically designed our library to target major mammalian CoVs.

Like unbiased NGS, the depth of coverage obtained by enrichment NGS was expectedly dependent on the quantity of viral RNA in the sample. Enrichment NGS performed poorly on samples containing low viral titers. Minimal reads were obtained for the genomes of BtTyCoV-HKU4r and BtHpCoV-HKU10r, and both samples had small amounts of viral RNA (C_T value of >30). The technical procedure of the capture is itself a limitation of enrichment NGS (15, 18). Key steps in the hybridization protocol, such as washes, could result in the loss of viral nucleic acid. While this loss is tolerable when the viral titer is high, a low level of viral RNA may give a false-negative result. We observed this situation with BtHpCoV-HKU10r. This virus was only detected by unbiased NGS. Similar observations have been made in previous studies, where full-length genome sequencing of human herpesvirus 1, West Nile virus, and MERS-CoV was achieved only when high viral titers were present (15, 18). We could improve the capture efficiency in two ways in the future: by using newly designed probes that bind better to their targets or by changing the steps that affect binding. For example, 65°C is the preferred temperature for Dynabeads, and thus, any step that affected the temperature would cause loss of yield. We can create a constant work condition for this step in future, or we can use different beads that require a less stringent environment. Above all, although unbiased NGS is a better choice for these samples, large-scale bat surveillance would benefit from the reduced cost of targeted enrichment. We suggest that direct NGS and gap-filling PCR are good complements to enrichment NGS once a positive sample has been identified.

While the CoV enrichment NGS successfully identified nine new CoVs, the CoV-specific enrichment also has limitations. While other enrichment NGS approaches aim to identify a broad range of known viruses across the virome (15–17), our pipeline was designed to identify known and diverse CoVs. The most challenging region to sequence was the spike gene, which has the lowest bait coverage across the genome. The genome references used in bait design do not fully reflect the diversity in this region. This is not unexpected, as this technology was not designed to detect completely novel viruses (15–17). One solution is to constantly update the baits in the CoV library to include sequence variations as they are reported (20).

Once CoVs have been identified in a sample, characterization of the full-length genome is important. Genome recombination has been documented for human CoVs, including OC43, NL63, HKU1, SARS-CoV, and MERS-CoV (2, 4). It has also been suggested that recombination between the bat SARSr-CoV strains WIV16 and Rf1-related generated a new strain, SARSr-Civet CoV SZ3, with a breakpoint at the NSP16/spike and S2 gene region (7). Breakpoints at the NSP16/spike and S2 gene region and nucleoprotein/accessory protein gene region can be found in most of the bat CoV species analyzed, suggesting that recombination is rather common. Recombination in spike or accessory proteins may generate a new virus capable of infecting via a different receptor or lead to different virus-host interactions. Genome diversity has not been assessed for CoV

species like BtRhCoV-HKU2r since they were first discovered (13, 21). We should be alert and vigilant with the knowledge that bat CoVs are likely to cause another disease outbreak, not only because of their prevalence but also because the high frequency of recombination between viruses may lead to the generation of viruses with changes in virulence. BtMiCoV-HKU8r is probably a new recombinant virus that may use a different receptor than the reference virus, considering the low similarity in their spike genes. And yet, we know very little about the functionality of their accessary proteins or the biological significance of this diversity. We previously provided serological evidence that HKU8r-CoV had jumped over from bats to camels and recombined with MERS-CoV, alerting other researchers that the CoV species could be dangerous (22). Therefore, analysis of the short RNA-dependent RNA polymerase region, used in most CoV surveillance studies, is not sufficient and genome-level comparison is needed to monitor the risk of alterations in species tropism and pathogenesis.

In conclusion, we have provided a cost-effective methodology for bat CoV surveillance. The high genetic diversity observed in our newly sequenced samples suggests further work is needed to characterize these bat CoVs prior to or in the early stages of spillover to humans.

MATERIALS AND METHODS

Sample preparation. Control viruses were cultured for RNA extraction. Porcine epidemic diarrhea virus (PEDV), transmissible gastroenteritis coronavirus (TGEV), MERS-CoV, SARSr-CoV, and mouse hepatitis virus (MHV) samples were cultured in Vero, swine testis (ST), Huh7, Vero E6, and DBT cells, respectively. All cells were maintained in Dulbecco modified Eagle medium (DMEM) containing 10% fetal bovine serum (FBS) and incubated at 37°C with 5% CO_2. Once cytopathic effect (CPE) was observed, 140 μl of supernatant was collected for RNA extraction.

To analyze clinical samples, HCoV-NL63, HCoV-OC43, and HCoV-HKU1 were extracted from human oral swabs. RNA for BtMiCoV-1r, BtMiCoV-HKU8r, BtHpCoV-HKU10r, BtHpCoV-CHB25, BtScCoV-512r, BtRaCoV-229Er, BtRhCoV-HKU2r, BtPiCoV-HKU5r, and BtTyCoV-HKU4r, which were collected during previous bat CoV surveillance projects, was extracted from bat rectal swabs (11, 23, 24). To process RNA, 560 μl of buffer AVL (Qiagen) was added to the tube containing 140 μl swab sample or culture supernatant. Samples were vortexed for 15 s and then centrifuged at 12,000 $\times$ *g* for 10 min to obtain a clear supernatant. Viral nucleic acid was extracted using the QIAamp viral RNA minikit (Qiagen) following the manufacturer's instructions.

qPCR. For quantitative PCR (qPCR) analysis, primers based on the CoV *RdRp* gene were used (Table S1 in the supplemental material). RNA was reverse transcribed using PrimeScript RT master mix (TaKaRa). The 10-μl qPCR mixture contained 5 μl 2$\times$ SYBR premix Ex Taq II (TaKaRa), 0.4 μM each primer, and 1 μl cDNA. Amplification was performed as follows: 95°C for 30 s, followed by 40 cycles at 95°C for 5 s and 60°C for 30 s and an additional melt step.

Preparation of Illumina DNA libraries from RNA. Libraries for NGS were constructed from total RNA using the TruSeq stranded mRNA library preparation kit for Illumina (Illumina) according to the manufacturer's instructions. Briefly, 8 μl of total RNA was added to first-strand synthesis buffer and random primers before a 4-min incubation at 94°C to generate RNA fragments larger than 300 nucleotides (nt). Following first- and second-strand cDNA synthesis, double-stranded cDNA was purified using Agencourt AMPure XP beads (Beckman Coulter Genomics) and eluted in 20 μl nuclease-free H_2O. To obtain a library size larger than 300 nt, the library was amplified by PCR using the following conditions: initial denaturation at 98°C for 30 s, 10 cycles of denaturation for 10 s at 98°C, annealing for 30 s at 60°C, and extension for 30 s at 72°C, and then a final extension for 5 min at 72°C. Libraries were purified using Agencourt AMPure XP beads (Beckman Coulter Genomics), eluted in 10 μl nuclease-free H_2O, visualized on a 1.5% agarose gel, and quantified using a Bioanalyzer high-sensitivity DNA assay (Agilent). Once prepared, the libraries were divided in two. Half the library was sequenced directly to obtain the unbiased reads, and half was enriched prior to NGS.

Enrichment of CoV sequences in libraries. Targeted CoV genome enrichment was achieved using our customized biotinylated 120-mer xGen Lockdown baits (Integrated DNA Technologies) (18). Prior to capture of viral sequences, 2 μl of xGen universal blocker-TS mixture (Integrated DNA Technologies), matched according to the library index, was added to 20 μl of library DNA. To block binding of baits to nonviral regions of library fragments, 0.5 μl of 5 μg Cot-1 DNA (Invitrogen) was added. Blocked libraries were ethanol precipitated and resuspended in 2.5 μl of nuclease-free H_2O, 3 μl NimbleGen hybridization solution, and 7.5 μl NimbleGen 2$\times$ hybridization buffer (Roche). Following a 10-min incubation at room temperature, resuspended libraries were denatured at 95°C for 10 min and cooled on ice before the addition of the CoV bait pool. A total of 3 pmol of baits was added and hybridized to the libraries for 4 h at 65°C. To capture virus-specific library fragments, 100 μl of Dynabeads M-270 streptavidin magnetic beads (Life Technologies) was added to the hybridization reaction mixture and the mixture was incubated for a further 45 min at 65°C with shaking at 2,000 rpm on a ThermoMixer C shaker (Eppendorf). Streptavidin beads were washed to remove unbound DNA, using the SeqCap EZ hybridization and wash kit (Roche) according to the manufacturer's instructions. A postcapture PCR amplification with P1 and P2 primers (Illumina) was performed using the following conditions: initial denaturation at 95°C for 2 min,

20 cycles of denaturation for 20 s at 95°C, annealing for 20 s at 65°C, and extension for 15 s at 72°C, and then a final extension step for 3 min at 72°C. The enriched library was purified using Agencourt AMPure XP beads (Beckman Coulter Genomics) and eluted in 10 μl nuclease-free H_2O, visualized on a 1.5% agarose gel, and quantified using a Bioanalyzer high-sensitivity DNA assay (Agilent). All samples were subjected to the same library preparation and enrichment.

Data analysis. Each unbiased NGS library was run on one HiSeq lane. The 17 enriched libraries were made into two pools (8 or 9 per each) and run on HiSeq lanes. NGS reads were assembled into genomes using the Galaxy platform (25). PCR and Sanger sequencing were used to fill the genome gaps. All genomes were interrogated for ORFs using ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/). The search parameters were set to ignore nested ORFs and filter out ORFs of less than 150 bp. The standard genetic code and the "ATG only" rule were selected. Each ORF was identified and annotated through BLASTN and BLASTX using the NCBI database. Read mapping or PCR resequencing was used to verify novel ORFs. Read depth was assessed by mapping reads from direct or enriched NGS to their respective genomes using CLC Genomics Workbench version 12.0 (Qiagen). Bait positions were calculated by aligning baits to each genome by BLASTN. The ratio of viral reads to total reads was calculated for each sample. The ORF1b sequences of 38 ICTV reference genomes and 17 CoV genomes from this study were aligned by ClustalW (version 2.1). The phylogenetic tree was generated using the neighbor-joining method in the maximum-composite-likelihood model in MEGA (version 7.0.18) with nucleotide substitution type and 1,000 bootstrap iterations. The schematic diagrams of CoV genomes, including bait positions and read depths of NGS, were prepared using Circos (version 0.69.8). Graphs displaying the data size and viral read ratios were generated using Prism (GraphPad Prism 7).

Data availability. Viral genome data for new CoVs obtained from this study are available in GenBank under accession numbers MN611517 to MN611525.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TABLE S1, XLSX file, 0.1 MB.

TABLE S2, XLSX file, 0.01 MB.

TABLE S3, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

We thank Zheng Zhenhua at Wuhan Institute of Virology (China) for providing MHV virus.

This study was supported by the China National Science and Technology Major Project on Infectious Diseases (grant number 2018ZX10305409-004-001 to P.Z.), the National Natural Science Foundation of China (Excellent Young Scholars grants number 81822028 and 81661148058 to P.Z.), and the Strategic Priority Research Program of the CAS (grants number XDB29010204 to P.Z. and XDB29010101 to Z.-L.S.). The work conducted at Duke-NUS was supported in part by NRF grant numbers NRF2016NRF-NSFC002-013 and NRF2012NRF-CRP001-056, CD-PHRG grant number CDPHRG/0006/2014, NMRC grant number ZRRF16006, and MINDEF grant number DIRP2015-9016102060 to L.-F.W. and D.E.A.

The authors declare no conflict of interest.

REFERENCES

1. Forni D, Cagliani R, Clerici M, Sironi M. 2017. Molecular evolution of human coronavirus genomes. Trends Microbiol 25:35–48. https://doi.org/10.1016/j.tim.2016.09.001.
2. Woo PC, Lau SK, Huang Y, Yuen KY. 2009. Coronavirus diversity, phylogeny and interspecies jumping. Exp Biol Med (Maywood) 234:1117–1127. https://doi.org/10.3181/0903-MR-94.
3. Lau SK, Woo PC, Yip CC, Tse H, Tsoi HW, Cheng VC, Lee P, Tang BS, Cheung CH, Lee RA, So LY, Lau YL, Chan KH, Yuen KY. 2006. Coronavirus HKU1 and other coronavirus infections in Hong Kong. J Clin Microbiol 44:2063–2071. https://doi.org/10.1128/JCM.02614-05.
4. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, Liu W, Bi Y, Gao GF. 2016. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends Microbiol 24:490–502. https://doi.org/10.1016/j.tim.2016.03.003.
5. ICTV. 2018. Virus taxonomy: 2018 release. ICTV, Washington, DC. https://talk.ictvonline.org/taxonomy/.
6. Woo PC, Huang Y, Lau SK, Yuen KY. 2010. Coronavirus genomics and bioinformatics analysis. Viruses 2:1804–1820. https://doi.org/10.3390/v2081803.
7. Hu B, Zeng LP, Yang XL, Ge XY, Zhang W, Li B, Xie JZ, Shen XR, Zhang YZ, Wang N, Luo DS, Zheng XS, Wang MN, Daszak P, Wang LF, Cui J, Shi ZL. 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog 13:e1006698. https://doi.org/10.1371/journal.ppat.1006698.
8. de Souza Luna LK, Heiser V, Regamey N, Panning M, Drexler JF, Mulangu S, Poon L, Baumgarte S, Haijema BJ, Kaiser L, Drosten C. 2007. Generic detection of coronaviruses and differentiation at the prototype strain level by reverse transcription-PCR and nonfluorescent low-density microarray. J Clin Microbiol 45:1049–1052. https://doi.org/10.1128/JCM.02426-06.
9. Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RAM, Berger A, Burguiere AM, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra JC, Muller S, Rickerts V, Sturmer M, Vieth S, Klenk HD, Osterhaus A, Schmitz H, Doerr HW. 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 348:1967–1976. https://doi.org/10.1056/NEJMoa030747.
10. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus A, Fouchier R. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820. https://doi.org/10.1056/NEJMoa1211721.

11. Zhou P, Fan H, Lan T, Yang XL, Shi WF, Zhang W, Zhu Y, Zhang YW, Xie QM, Mani S, Zheng XS, Li B, Li JM, Guo H, Pei GQ, An XP, Chen JW, Zhou L, Mai KJ, Wu ZX, Li D, Anderson DE, Zhang LB, Li SY, Mi ZQ, He TT, Cong F, Guo PJ, Huang R, Luo Y, Liu XL, Chen J, Huang Y, Sun Q, Zhang XLL, Wang YY, Xing SZ, Chen YS, Sun Y, Li J, Daszak P, Wang LF, Shi ZL, Tong YG, Ma JY. 2018. Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. Nature 556:255–258. https://doi.org/10.1038/s41586-018-0010-9.
12. Wang L-F, Cowled C (ed). 2015. Bats and viruses: a new frontier of emerging infectious diseases. John Wiley & Sons, Inc, Hoboken, NJ.
13. Fan Y, Zhao K, Shi ZL, Zhou P. 2019. Bat coronaviruses in China. Viruses 11:210. https://doi.org/10.3390/v11030210.
14. Carroll D, Daszak P, Wolfe ND, Gao GF, Morel CM, Morzaria S, Pablos-Mendez A, Tomori O, Mazet J. 2018. The Global Virome Project. Science 359:872–874. https://doi.org/10.1126/science.aap7463.
15. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jabado OJ, Lipkin WI. 2015. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. mBio 6:e01491-15. https://doi.org/10.1128/mBio.01491-15.
16. Gaudin M, Desnues C. 2018. Hybrid capture-based next generation sequencing and its application to human infectious diseases. Front Microbiol 9:2924. https://doi.org/10.3389/fmicb.2018.02924.
17. Wylie TN, Wylie KM, Herter BN, Storch GA. 2015. Enhanced virome sequencing using targeted sequence capture. Genome Res 25:1910–1920. https://doi.org/10.1101/gr.191049.115.
18. Lim XF, Lee CB, Pascoe SM, How CB, Chan S, Tan JH, Yang X, Zhou P, Shi Z, Sessions OM, Wang LF, Ng LC, Anderson DE, Yap G. 2019. Detection and characterization of a novel bat-borne coronavirus in Singapore using multiple molecular approaches. J Gen Virol 100:1363–1374. https://doi.org/10.1099/jgv.0.001307.
19. Noyes NR, Weinroth ME, Parker JK, Dean CJ, Lakin SM, Raymond RA, Rovira P, Doster E, Abdo Z, Martin JN, Jones KL, Ruiz J, Boucher CA, Belk KE, Morley PS. 2017. Enrichment allows identification of diverse, rare elements in metagenomic resistome-virulome sequencing. Microbiome 5:142. https://doi.org/10.1186/s40168-017-0361-8.
20. Metsky HC, Siddle KJ, Gladden-Young A, Qu J, Yang DK, Brehio P, Goldfarb A, Piantadosi A, Wohl S, Carter A, Lin AE, Barnes KG, Tully DC, Corleis B, Hennigan S, Barbosa-Lima G, Vieira YR, Paul LM, Tan AL, Garcia KF, Parham LA, Odia I, Eromon P, Folarin OA, Goba A, Viral Hemorrhagic Fever Consortium, Simon-Loriere E, Hensley L, Balmaseda A, Harris E, Kwon DS, Allen TM, Runstadler JA, Smole S, Bozza FA, Souza TML, Isern S, Michael SF, Lorenzana I, Gehrke L, Bosch I, Ebel G, Grant DS, Happi CT, Park DJ, Gnirke A, Sabeti PC, Matranga CB. 2019. Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. Nat Biotechnol 37:160–168. https://doi.org/10.1038/s41587-018-0006-x.
21. Lau SK, Woo PC, Li KS, Huang Y, Wang M, Lam CS, Xu H, Guo R, Chan KH, Zheng BJ, Yuen KY. 2007. Complete genome sequence of bat coronavirus HKU2 from Chinese horseshoe bats revealed a much smaller spike gene with a different evolutionary lineage from the rest of the genome. Virology 367:428–439. https://doi.org/10.1016/j.virol.2007.06.009.
22. Zhang W, Zheng XS, Agwanda B, Ommeh S, Zhao K, Lichoti J, Wang N, Chen J, Li B, Yang XL, Mani S, Ngeiywa KJ, Zhu Y, Hu B, Onyuok SO, Yan B, Anderson DE, Wang LF, Zhou P, Shi ZL. 2019. Serological evidence of MERS-CoV and HKU8-related CoV co-infection in Kenyan camels. Emerg Microbes Infect 8:1528–1534. https://doi.org/10.1080/22221751.2019.1679610.
23. Luo C-M, Wang N, Yang X-L, Liu H-Z, Zhang W, Li B, Hu B, Peng C, Geng Q-B, Zhu G-J, Li F, Shi Z-L. 2018. Discovery of novel bat coronaviruses in South China that use the same receptor as Middle East respiratory syndrome coronavirus. J Virol 92:e00116-18. https://doi.org/10.1128/JVI.00116-18.
24. Ge XY, Wang N, Zhang W, Hu B, Li B, Zhang YZ, Zhou JH, Luo CM, Yang XL, Wu LJ, Wang B, Zhang Y, Li ZX, Shi ZL. 2016. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. Virol Sin 31:31–40. https://doi.org/10.1007/s12250-016-3713-9.
25. Goecks J, Nekrutenko A, Taylor J, Galaxy Team. 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol 11:R86. https://doi.org/10.1186/gb-2010-11-8-r86.

OBSERVATION
Clinical Science and Epidemiology

Middle East Respiratory Syndrome Coronavirus Antibodies in Bactrian and Hybrid Camels from Dubai

Susanna K. P. Lau,[a,b,c,d] Kenneth S. M. Li,[d] Hayes K. H. Luk,[d] Zirong He,[d] Jade L. L. Teng,[a,b,d] Kwok-Yung Yuen,[a,b,c,d] Ulrich Wernery,[e] Patrick C. Y. Woo[a,b,c,d]

[a]State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong
[b]Carol Yu Centre for Infection, The University of Hong Kong, Hong Kong
[c]Collaborative Innovation Centre for Diagnosis and Treatment of Infectious Diseases, The University of Hong Kong, Hong Kong
[d]Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong
[e]Central Veterinary Research Laboratory, Dubai, United Arab Emirates

Susanna K. P. Lau, Kenneth S. M. Li, and Hayes K. H. Luk contributed equally to the manuscript. Author order was determined by order of seniority.

ABSTRACT So far, dromedary camels are the only known animal reservoir for Middle East respiratory syndrome (MERS) coronavirus (MERS-CoV). Previous published serological studies showed that sera of Bactrian camels were all negative for MERS-CoV antibodies. However, a recent study revealed that direct inoculation of Bactrian camels intranasally with MERS-CoV can lead to infection with abundant virus shedding and seroconversion. In this study, we examined the presence of MERS-CoV antibodies in Bactrian and hybrid camels in Dubai, the United Arab Emirates (where dromedaries are also present), and Bactrian camels in Xinjiang, China (where dromedaries are absent). For the 29 serum samples from Bactrian camels in Dubai tested by the MERS-CoV spike (S) protein-based enzyme-linked immunosorbent assay (S-ELISA) and neutralization antibody test, 14 (48%) and 12 (41%), respectively, were positive for MERS-CoV antibodies. All the 12 serum samples that were positive with the neutralization antibody test were also positive for the S-ELISA. For the 11 sera from hybrid camels in Dubai tested with the S-ELISA and neutralization antibody test, 6 (55%) and 9 (82%), respectively, were positive for MERS-CoV antibodies. All the 6 serum samples that were positive for the S-ELISA were also positive with the neutralization antibody test. There was a strong correlation between the antibody levels detected by S-ELISA and neutralizing antibody titers, with a Spearman coefficient of 0.6262 ($P < 0.0001$; 95% confidence interval, 0.5062 to 0.7225). All 92 Bactrian camel serum samples from Xinjiang were negative for MERS-CoV antibodies tested using both S-ELISA and the neutralization antibody test. Bactrian and hybrid camels are potential sources of MERS-CoV infection.

IMPORTANCE Since its first appearance in 2012, Middle East respiratory syndrome (MERS) has affected >25 countries, with >2,400 cases and an extremely high fatality rate of >30%. The total number of mortalities due to MERS is already greater than that due to severe acute respiratory syndrome. MERS coronavirus (MERS-CoV) has been confirmed to be the etiological agent. So far, dromedaries are the only known animal reservoir for MERS-CoV. Previously published serological studies showed that sera of Bactrian camels were all negative for MERS-CoV antibodies. In this study, we observed that 41% of the Bactrian camel sera and 55% of the hybrid camel sera from Dubai (where dromedaries are also present), but none of the sera from Bactrian camels in Xinjiang (where dromedaries are absent), were positive for MERS-CoV antibodies. Based on these results, we conclude that in addition to dromedaries, Bactrian and hybrid camels are also potential sources of MERS-CoV infection.

KEYWORDS Bactrian camel, hybrid camel, MERS coronavirus, antibody

Citation Lau SKP, Li KSM, Luk HKH, He Z, Teng JLL, Yuen K-Y, Wernery U, Woo PCY. 2020. Middle East respiratory syndrome coronavirus antibodies in Bactrian and hybrid camels from Dubai. mSphere 5:e00898-19. https://doi.org/10.1128/mSphere.00898-19.

Editor Matthew B. Frieman, University of Maryland, College Park

Address correspondence to Ulrich Wernery, cvrl@cvrl.ae, or Patrick C. Y. Woo, pcywoo@hku.hk.

Received 3 December 2019
Accepted 26 December 2019
Published 22 January 2020

Since its first appearance in 2012, Middle East respiratory syndrome (MERS) has affected more than 25 countries in 4 continents, with more than 2,400 cases and an extremely high fatality rate of more than 30%. The total number of mortalities due to MERS is already greater than that due to severe acute respiratory syndrome. MERS coronavirus (MERS-CoV), a betacoronavirus from subgenus *Merbecovirus*, has been confirmed to be the etiological agent (1). Human dipeptidyl peptidase 4 was found to be the cellular receptor for MERS-CoV (2). Subsequent detection of MERS-CoV and its antibodies in dromedary, or one-humped, camels (*Camelus dromedarius*) in various countries in the Middle East and North Africa have suggested that these animals are probably the reservoir for MERS-CoV (3–5). Other betacoronaviruses in bats from the subgenus *Merbecovirus* (e.g., *Tylonycteris* bat CoV HKU4, *Pipistrellus* bat CoV HKU5, *Hypsugo* bat CoV HKU25), and hedgehogs were found to be closely related to MERS-CoV (6–9).

In addition to the dromedaries, there are two additional surviving Old World camels, the Bactrian, or two-humped, camels (*Camelus bactrianus*) and the wild Bactrian camels (*Camelus ferus*), both inhabitants of Central Asia. Moreover, a dromedary and a Bactrian camel can mate and result in a hybrid camel offspring. Previous published serological studies showed that sera of Bactrian camels were all negative for MERS-CoV antibodies, suggesting that Bactrian camels may not be a reservoir of MERS-CoV (Fig. 1) (10–14). However, a recent study revealed that direct inoculation of Bactrian camels intranasally with MERS-CoV can lead to infection with abundant virus shedding and seroconversion (15). Therefore, we hypothesize that those Bactrian camels, and even the hybrid camels, that reside in countries where there are dromedaries can be infected with MERS-CoV. To test this hypothesis, we examined the presence of MERS-CoV antibodies in Bactrian and hybrid camels in Dubai, the United Arab Emirates (where dromedaries are also present), and Bactrian camels in Xinjiang, China (where dromedaries are absent), using a MERS-CoV spike (S) protein-based enzyme-linked immunosorbent assay (ELISA) and neutralization antibody test.

A total of 29 and 11 serum samples, respectively, were collected from 29 Bactrian camels and 11 hybrid camels from a private collection in Dubai (April to May 2019) (Table 1), and 92 serum samples were collected from Bactrian camels on a camel farm in Xinjiang (November 2012) (16). Antibodies against the S protein of MERS-CoV were tested using microplates precoated with purified $(His)_6$-tagged recombinant receptor-binding domain of S (RBD-S) of MERS-CoV and detected with 1:8,000 diluted horseradish peroxidase-conjugated goat anti-llama IgG (Life Technologies, Carlsbad, CA, USA) conjugate (S-ELISA) (17). The cutoff of the ELISA was defined as three standard deviations above the mean absorbance value of 10 dromedary serum samples that tested negative with the neutralization antibody test. The neutralization antibody test was performed by incubating serially diluted camel sera with 100 50% tissue culture infective dose ($TCID_{50}$) MERS-CoV for 2 hours before infecting the Vero cells for 1 hour. The cytopathic effect (CPE) was observed for 5 days, and the sera were regarded as positive for neutralizing antibody if no CPE was observed in the infected cells (18). All tests were performed in triplicate.

For the 29 serum samples from Bactrian camels in Dubai tested with the S-ELISA and neutralization antibody test, 14 (48%) and 12 (41%), respectively, were positive for MERS-CoV antibodies (Fig. 2A and Table 1). All the 12 serum samples that were positive with the neutralization antibody test were also positive for the S-ELISA. For the 11 serum samples from hybrid camels in Dubai tested with the S-ELISA and neutralization antibody test, 6 (55%) and 9 (82%), respectively, were positive for MERS-CoV antibodies (Fig. 2A and Table 1). All the 6 serum samples that were positive for the S-ELISA were also positive with the neutralization antibody test. There was a strong correlation between the antibody levels detected by S-ELISA and neutralizing antibody titers, with a Spearman coefficient of 0.6262 ($P < 0.0001$; 95% confidence interval, 0.5062 to 0.7225) (Fig. 2B). All 92 Bactrian camel serum samples from Xinjiang were negative for MERS-CoV antibodies tested with both S-ELISA and the neutralization antibody test (Fig. 2A and Table 1).

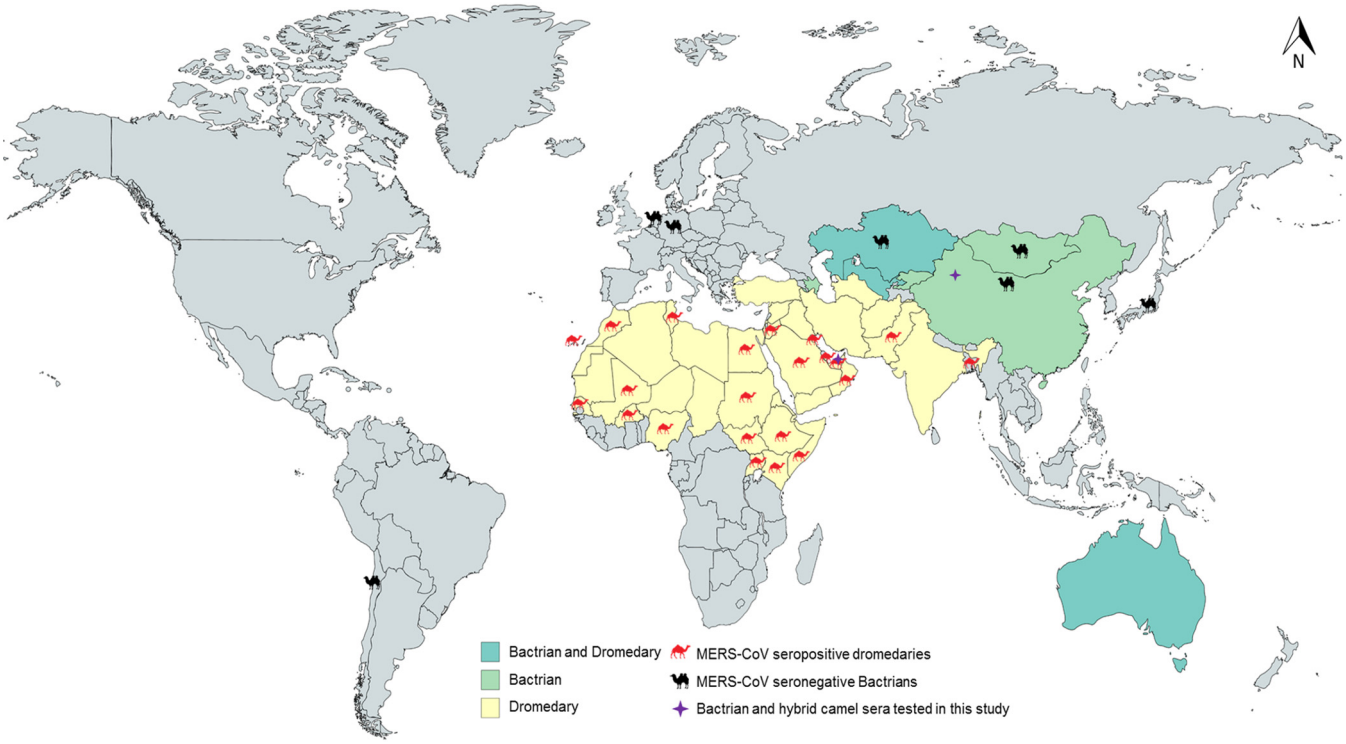


FIG 1 Geographical distribution of dromedaries and Bactrians. Places with MERS-CoV-seropositive dromedaries (red camels) and MERS-CoV-seronegative Bactrians (black camels) from previous studies are labeled.

We showed that MERS-CoV antibodies were present in Bactrian and hybrid camels from Dubai. When the reservoir of MERS-CoV was first discovered in dromedaries (3), it was wondered whether Bactrian camels are also another possible source of the virus. However, in all the studies that looked for MERS-CoV in both wild (China, $n = 190$; Kazakhstan, $n = 95$; Mongolia, $n = 200$) and captive (the Netherlands, $n = 2$; Chile, $n = 2$; Japan, $n = 5$; Germany, $n = 16$) Bactrian camels, it was found that none of the Bactrian camels, as tested with neutralization antibody test, ELISA, immunofluorescence assay, plaque reduction neutralization test, and/or protein microarray, had MERS-CoV antibodies (10–14). However, the sera from all these studies were obtained from Bactrian camels in geographical regions where there were no dromedaries that were positive for MERS-CoV or its antibodies. Interestingly, in a recent study that investigated the susceptibility of Bactrian camels to MERS-CoV, upon intranasal inoculation with 10^7 $TCID_{50}$ of MERS-CoV, the Bactrian camels developed clinical signs of transient upper respiratory tract infections, such as nasal discharge and coughing, with shedding of up to 106.5 to 106.8 plaque forming units/ml of MERS-CoV from the upper respiratory tract and development of neutralizing antibodies against MERS-CoV (15). This suggested that

TABLE 1 MERS-CoV neutralizing antibody titer of Bactrian and hybrid camel sera

Neutralizing antibody titer	No. of samples (%) for:		
	Dubai Bactrian camel ($n = 29$)	Dubai hybrid camel ($n = 11$)	Xinjiang Bactrian camel ($n = 92$)
<10	17 (58.6)	2 (18.2)	92 (100)
10	0 (0)	0 (0)	0 (0)
20	0 (0)	0 (0)	0 (0)
40	0 (0)	0 (0)	0 (0)
80	0 (0)	0 (0)	0 (0)
160	1 (3.4)	1 (9.1)	0 (0)
320	2 (6.9)	2 (18.2)	0 (0)
640	9 (31.1)	6 (54.5)	0 (0)

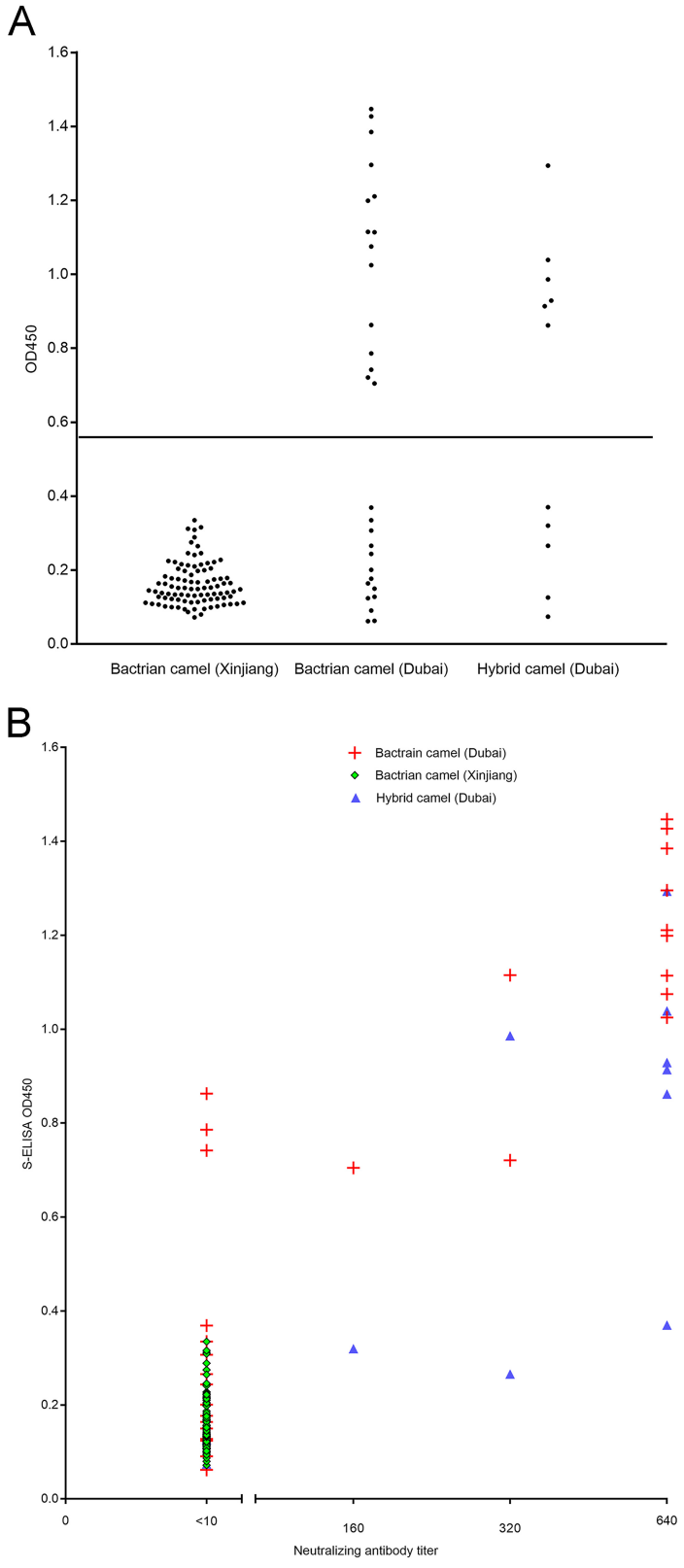


FIG 2 (A) Scatter plot showing MERS-CoV antibody levels detected using S-ELISA in Bactrian and hybrid camel sera from Dubai and Xinjiang. The test results were plotted as optical density at 450 nm (OD_{450}) values. The horizontal line indicates the cutoff value (0.557) for positive diagnosis. (B) Scatter plot showing correlation between antibody levels detected using S-ELISA and neutralizing antibody titers of Bactrian and hybrid camel sera for MERS-CoV.

Bactrian camels can be susceptible to MERS-CoV infection and may actually be another possible reservoir for MERS-CoV. In the present study, we showed that MERS-CoV antibodies were present in 41% of the Bactrian camels and 55% of the hybrid camels from a private collection in Dubai using two independent assays. These camels were kept as hobby animals. They had occasional contacts with dromedaries from camel farms that breed dromedaries for racing. MERS-CoV has been consistently detected in these camel farms in the last few years, and the MERS-CoV seropositive rate increased as the age of the dromedaries increased (17). The Bactrian camels were imported from Kazakhstan more than 10 years ago, whereas the hybrid camels were the offspring of mating between dromedaries and the Bactrian camels. Some of the hybrid camels were sometimes used for camel racing and therefore had more frequent contacts with dromedaries. It is likely that some of the Bactrian and hybrid camels might have acquired the MERS-CoV during their contacts with dromedaries that were shedding the virus, and the virus subsequently infected other Bactrian and hybrid camels in the collection.

To prevent MERS in humans, it might be worthwhile to immunize Bactrian and hybrid camels in addition to the dromedaries. So far, as determined from the results of phylogenomic analyses, several clades of MERS-CoV are circulating in dromedaries. Although it seems that the ultimate origin of MERS-CoV was from bats (6), there are still significant differences between the genome sequences of these betacoronaviruses in bats from the subgenus *Merbecovirus* and MERS-CoV, suggesting that interspecies jumping from bats to camels may not be a very recent event, and hence the dromedaries are probably the reservoir of MERS-CoV where the virus was transmitted to humans. In the last few years, a number of MERS-CoV vaccines have been developed for their potential use in dromedaries (19–23). In this study, our results indicated that Bactrian and hybrid camels are also potential sources of MERS-CoV infection. Therefore, Bactrian and hybrid camels, in addition to the dromedaries, should be immunized in order to reduce the chance of transmitting the virus to humans.

ACKNOWLEDGMENTS

This work is partly supported by the Theme-based Research Scheme (project no. T11-707/15-R), University Grant Committee; the University Development Fund, The University of Hong Kong; and the Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Ministry of Education, China.

REFERENCES

1. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820. https://doi.org/10.1056/NEJMoa1211721.
2. Raj VS, Mou H, Smits SL, Dekkers DH, Müller MA, Dijkman R, Muth D, Demmers JA, Zaki A, Fouchier RA, Thiel V, Drosten C, Rottier PJ, Osterhaus AD, Bosch BJ, Haagmans BL. 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 495:251–254. https://doi.org/10.1038/nature12005.
3. Reusken CB, Haagmans BL, Müller MA, Gutierrez C, Godeke GJ, Meyer B, Muth D, Raj VS, Smits-De Vries L, Corman VM, Drexler JF, Smits SL, El Tahir YE, De Sousa R, van Beek J, Nowotny N, van Maanen K, Hidalgo-Hermoso E, Bosch BJ, Rottier P, Osterhaus A, Gortázar-Schmidt C, Drosten C, Koopmans MP. 2013. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. Lancet Infect Dis 13:859–866. https://doi.org/10.1016/S1473-3099(13)70164-6.
4. Lau SK, Wernery R, Wong EY, Joseph S, Tsang AK, Patteril NA, Elizabeth SK, Chan KH, Muhammed R, Kinne J, Yuen KY, Wernery U, Woo PC. 2016. Polyphyletic origin of MERS coronaviruses and isolation of a novel clade A strain from dromedary camels in the United Arab Emirates. Emerg Microbes Infect 5:e128. https://doi.org/10.1038/emi.2016.129.
5. Chu DK, Hui KP, Perera RA, Miguel E, Niemeyer D, Zhao J, Channappanavar R, Dudas G, Oladipo JO, Traoré A, Fassi-Fihri O, Ali A, Demissié GF, Muth D, Chan MC, Nicholls JM, Meyerholz DK, Kuranga SA, Mamo G, Zhou Z, So RT, Hemida MG, Webby RJ, Roger F, Rambaut A, Poon LL, Perlman S, Drosten C, Chevalier V, Peiris M. 2018. MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity. Proc Natl Acad Sci U S A 115:3144–3149. https://doi.org/10.1073/pnas.1718769115.
6. Lau SK, Li KS, Tsang AK, Lam CS, Ahmed S, Chen H, Chan KH, Woo PC, Yuen KY. 2013. Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of Pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. J Virol 87:8638–8650. https://doi.org/10.1128/JVI.01055-13.
7. Lau SK, Zhang L, Luk HK, Xiong L, Peng X, Li KS, He X, Zhao PS, Fan RY, Wong AC, Ahmed SS, Cai JP, Chan JF, Sun Y, Jin D, Chen H, Lau TC, Kok RK, Li W, Yuen KY, Woo PC. 2018. Receptor usage of a novel bat lineage C betacoronavirus reveals evolution of MERS-related coronavirus spike proteins for human DPP4 binding. J Infect Dis 218:197–207. https://doi.org/10.1093/infdis/jiy018.
8. Corman VM, Kallies R, Philipps H, Göpner G, Müller MA, Eckerle I, Brünink S, Drosten C, Drexler JF. 2014. Characterization of a novel betacoronavirus related to Middle East respiratory syndrome coronavirus in European hedgehogs. J Virol 88:717–724. https://doi.org/10.1128/JVI.01600-13.
9. Lau SK, Luk HK, Wong AC, Fan RY, Lam CS, Li KS, Ahmed SS, Chow FW, Cai JP, Zhu X, Chan JF, Lau TC, Cao K, Li M, Woo PC, Yuen KY. 2019.

Identification of a novel betacoronavirus (Merbecovirus) in Amur hedgehogs from China. Viruses 11:980. https://doi.org/10.3390/v11110980.
10. Chan SM, Damdinjav B, Perera RA, Chu DK, Khishgee B, Enkhbold B, Poon LL, Peiris M. 2015. Absence of MERS-coronavirus in Bactrian camels, Southern Mongolia, November 2014. Emerg Infect Dis 21:1269–1271. https://doi.org/10.3201/eid2107.150178.
11. Liu R, Wen Z, Wang J, Ge J, Chen H, Bu Z. 2015. Absence of Middle East respiratory syndrome coronavirus in Bactrian camels in the West Inner Mongolia Autonomous Region of China: surveillance study results from July 2015. Emerg Microbes Infect 4:e73. https://doi.org/10.1038/emi.2015.73.
12. Meyer B, Müller MA, Corman VM, Reusken CB, Ritz D, Godeke GJ, Lattwein E, Kallies S, Siemens A, van Beek J, Drexler JF, Muth D, Bosch BJ, Wernery U, Koopmans MP, Wernery R, Drosten C. 2014. Antibodies against MERS coronavirus in dromedary camels, United Arab Emirates, 2003 and 2013. Emerg Infect Dis 20:552–559. https://doi.org/10.3201/eid2004.131746.
13. Miguel E, Perera RA, Baubekova A, Chevalier V, Faye B, Akhmetsadykov N, Ng CY, Roger F, Peiris M. 2016. Absence of Middle East respiratory syndrome coronavirus in camelids, Kazakhstan, 2015. Emerg Infect Dis 22:555–557. https://doi.org/10.3201/eid2203.151284.
14. Shirato K, Azumano A, Nakao T, Hagihara D, Ishida M, Tamai K, Yamazaki K, Kawase M, Okamoto Y, Kawakami S, Okada N, Fukushima K, Nakajima K, Matsuyama S. 2015. Middle East respiratory syndrome coronavirus infection not found in camels in Japan. Jpn J Infect Dis 68:256–258. https://doi.org/10.7883/yoken.JJID.2015.094.
15. Adney DR, Letko M, Ragan IK, Scott D, van Doremalen N, Bowen RA, Munster VJ. 2019. Bactrian camels shed large quantities of Middle East respiratory syndrome coronavirus (MERS-CoV) after experimental infection. Emerg Microbes Infect 8:717–723. https://doi.org/10.1080/22221751.2019.1618687.
16. Wang L, Teng JL, Lau SK, Sridhar S, Fu H, Gong W, Li M, Xu Q, He Y, Zhuang H, Woo PC, Wang L. 2019. Transmission of a novel genotype of hepatitis E virus from Bactrian camels to cynomolgus macaques. J Virol 93:e02014-18. https://doi.org/10.1128/JVI.02014-18.
17. Wernery U, Rasoul IH, Wong EY, Joseph M, Chen Y, Jose S, Tsang AK, Patteril NA, Chen H, Elizabeth SK, Yuen KY, Joseph S, Xia N, Wernery R, Lau SK, Woo PC. 2015. A phylogenetically distinct MERS coronavirus detected in a dromedary calf from a closed dairy herd in Dubai with rising seroprevalence along age. Emerg Microbes Infect 4:e74. https://doi.org/10.1038/emi.2015.74.
18. Woo PC, Lau SK, Wernery U, Wong EY, Tsang AK, Johnson B, Yip CC, Lau CC, Sivakumar S, Cai JP, Fan RY, Chan KH, Mareena R, Yuen KY. 2014. Novel betacoronavirus in dromedaries of the Middle East, 2013. Emerg Infect Dis 20:560–572. https://doi.org/10.3201/eid2004.131769.
19. Yong CY, Ong HK, Yeap SK, Ho KL, Tan WS. 2019. Recent advances in the vaccine development against Middle East respiratory syndrome-coronavirus. Front Microbiol 10:1781. https://doi.org/10.3389/fmicb.2019.01781.
20. Haagmans BL, van den Brand JM, Raj VS, Volz A, Wohlsein P, Smits SL, Schipper D, Bestebroer TM, Okba N, Fux R, Bensaid A, Solanes Foz D, Kuiken T, Baumgärtner W, Segalés J, Sutter G, Osterhaus AD. 2016. An orthopoxvirus-based vaccine reduces virus excretion after MERS-CoV infection in dromedary camels. Science 351:77–81. https://doi.org/10.1126/science.aad1283.
21. Adney DR, Wang L, van Doremalen N, Shi W, Zhang Y, Kong WP, Miller MR, Bushmaker T, Scott D, de Wit E, Modjarrad K, Petrovsky N, Graham BS, Bowen RA, Munster VJ. 2019. Efficacy of an adjuvanted Middle East respiratory syndrome coronavirus spike protein vaccine in dromedary camels and alpacas. Viruses 11:212. https://doi.org/10.3390/v11030212.
22. Alharbi NK, Padron-Regalado E, Thompson CP, Kupke A, Wells D, Sloan MA, Grehan K, Temperton N, Lambe T, Warimwe G, Becker S, Hill AV, Gilbert SC. 2017. ChAdOx1 and MVA based vaccine candidates against MERS-CoV elicit neutralising antibodies and cellular immune responses in mice. Vaccine 35:3780–3788. https://doi.org/10.1016/j.vaccine.2017.05.032.
23. Muthumani K, Falzarano D, Reuschel EL, Tingey C, Flingai S, Villarreal DO, Wise M, Patel A, Izmirly A, Aljuaid A, Seliga AM, Soule G, Morrow M, Kraynyak KA, Khan AS, Scott DP, Feldmann F, LaCasse R, Meade-White K, Okumura A, Ugen KE, Sardesai NY, Kim JJ, Kobinger G, Feldmann H, Weiner DB. 2015. A synthetic consensus anti-spike protein DNA vaccine induces protective immunity against Middle East respiratory syndrome coronavirus in nonhuman primates. Sci Transl Med 7:301ra132. https://doi.org/10.1126/scitranslmed.aac7462.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS

Authors: Yushun Wan, Jian Shang, Rachel Graham, Ralph S. Baric, Fang Li

Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS

Yushun Wan [1,*], Jian Shang [1,*], Rachel Graham [2], Ralph S. Baric [2], Fang Li [1,#]

[1] Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, Saint Paul, MN, USA

[2] Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA

* These authors contributed equally to this work.

Correspondence:

Fang Li (lifang@umn.edu)

Keywords: Wuhan coronavirus, SARS coronavirus, angiotensin-converting enzyme 2, animal reservoir, cross-species transmission, human-to-human transmission

Running title: Analyzing receptor usage by Wuhan coronavirus

Abstract

Recently a novel coronavirus (2019-nCoV) has emerged from Wuhan, China, causing symptoms in humans similar to those caused by SARS coronavirus (SARS-CoV). Since SARS-CoV outbreak in 2002, extensive structural analyses have revealed key atomic-level interactions between SARS-CoV spike protein receptor-binding domain (RBD) and its host receptor angiotensin-converting enzyme 2 (ACE2), which regulate both the cross-species and human-to-human transmissions of SARS-CoV. Here we analyzed the potential receptor usage by 2019-nCoV, based on the rich knowledge about SARS-CoV and the newly released sequence of 2019-nCoV. First, the sequence of 2019-nCoV RBD, including its receptor-binding motif (RBM) that directly contacts ACE2, is similar to that of SARS-CoV, strongly suggesting that 2019-nCoV uses ACE2 as its receptor. Second, several critical residues in 2019-nCoV RBM (particularly Gln493) provide favorable interactions with human ACE2, consistent with 2019-nCoV's capacity for human cell infection. Third, several other critical residues in 2019-nCoV RBM (particularly Asn501) are compatible with, but not ideal for, binding human ACE2, suggesting that 2019-nCoV has acquired some capacity for human-to-human transmission. Last, while phylogenetic analysis indicates a bat origin of 2019-nCoV, 2019-nCoV also potentially recognizes ACE2 from a diversity of animal species (except mice and rats), implicating these animal species as possible intermediate hosts or animal models for 2019-nCoV infections. These analyses provide insights into the receptor usage, cell entry, host cell infectivity and animal origin of 2019-nCoV, and may help epidemic surveillance and preventive measures against 2019-nCoV.

Significance

The recent emergence of Wuhan coronavirus (2019-nCoV) puts the world on alert. 2019-nCoV is reminiscent of the SARS-CoV outbreak in 2002-2003. Our decade-long structural studies on the receptor recognition by SARS-CoV have identified key interactions between SARS-CoV spike protein and its host receptor angiotensin-converting enzyme 2 (ACE2), which regulate both the cross-species and human-to-human transmissions of SARS-CoV. One of the goals of SARS-CoV research was to build an atomic-level iterative framework of virus-receptor interactions to facilitate epidemic surveillance, predict species-specific receptor usage, and identify potential animal hosts and animal models of viruses. Based on the sequence of 2019-nCoV spike protein, we apply this predictive framework to provide novel insights into the receptor usage and likely host range of 2019-nCoV. This study provides a robust test of this reiterative framework, providing the basic, translational and public health research communities with predictive insights that may help study and battle this novel 2019-nCoV.

Introduction

A novel coronavirus (2019-nCoV) from Wuhan, China has recently caused over 500 confirmed cases of human infections and at least 17 deaths in China (https://www.cdc.gov/coronavirus/novel-coronavirus-2019.html). There are also numerous confirmed cases of 2019-nCoV infections in other countries including USA. Many of the symptoms caused by 2019-nCoV, such as acute respiratory syndrome, are similar to those caused by SARS coronavirus (SARS-CoV). SARS-CoV emerged in 2002-2003 and transmitted among humans, causing over 8000 confirmed cases of human infections and about 800 deaths (1-4). It briefly re-emerged in 2003-2004, with 4 confirmed cases of mild human infections and no human-to-human transmission (5-7). SARS-CoV has also been isolated from animals and been adapted to lab cell culture (5, 8-11). It is believed that bats and palm civets were the natural and intermediate reservoirs for SARS-CoV, respectively, and that SARS-CV transmitted from palm civets to humans in an animal market in Southern China (12-14). It has been reported that 2019-nCoV also infected humans in an animal market in Wuhan, although the animal source of the outbreak is currently unknown. Moreover, it has been confirmed that 2019-nCoV has the capacity to transmit from human to human.

Coronaviruses are a large family of single-stranded enveloped RNA viruses and can be divided into four major genera (15). Both SARS-CoV and 2019-nCoV belong to the β-genus. An envelope-anchored spike protein mediates coronavirus entry into host cells by first binding to a host receptor and then fusing viral and host membranes (16). A defined receptor-binding domain (RBD) of SARS-CoV spike specifically recognizes its

host receptor angiotensin-converting enzyme 2 (ACE2) (17, 18). Different lines of research have shown that which host is susceptible to SARS-CoV infection is primarily determined by the affinity between the viral RBD and host ACE2 in the initial viral attachment step (19-23). In a span of about 10 years, we determined a series of crystal structures of SARS-CoV RBD complexed with ACE2; the RBDs were from SARS-CoV strains isolated from different hosts species in different years and the ACE2 receptor orthologues were derived from different animal species (18, 24-26). These structures showed that SARS-CoV RBD contains a core structure and a receptor-binding motif (RBM), and that the RBM binds to the outer surface of the claw-like structure of ACE2 (Fig. 1A) (25). Importantly, we identified two virus-binding hotspots on human ACE2 (24, 26). A number of naturally selected RBM mutations occurred near these two virus-binding hotspot and these residues largely determined the host range of SARS-CoV (Fig. 1B, 1C). Furthermore, we discovered specific amino acids at 442, 472, 479, 480 and 487 positions that enhance viral binding to human ACE2, and some other amino acids at these same positions that enhance viral binding to civet ACE2 (Fig. 1C). Importantly, when all human-ACE2-favoring residues were combined into one RBD, this RBD binds to human ACE2 with super affinity and the corresponding spike protein mediates viral entry into human cells with super efficiency (Fig. 1C) (26). An RBD with super affinity for civet ACE2 was also designed and empirically confirmed (Fig. 1C) (26). These gain-of-function data provided strong supporting evidence for the accuracy of our structural predictions. A long-term goal of these earlier studies is to establish a structure-function predictive framework for improved epidemic surveillance. More specifically, we aim to predict the receptor usage and host cell infectivity of future SARS-CoV or SARS-like

viral strains and identify their possible animal origins and animal models, based on the sequences of their spike proteins and the known atomic structures of original SARS-CoV RBD/ACE2 complex. Here, based on the newly released sequence of 2019-nCoV RBD, we reiteratively apply this predictive framework to provide novel insights into the receptor usage and likely host range of 2019-nCoV.

Results

The 2019-nCoV spike phylogeny is firmly rooted among other β-genus lineage b bat SARS-like coronaviruses (Fig. 2), but is ancestral to both human SARS-CoV (epidemic strain isolated in year 2002) and bat SARS-CoV strains that use ACE2 receptor to enter and infect primary host lung cells (11, 17). The overall sequence similarities between 2019-nCoV spike and SARS-CoV spike (isolated from human, civet or bat) are around 76%-78% for the whole protein, around 73%-76% for the RBD, and 50%-53% for the RBM (Fig. 3A, 3B). In comparison, human coronavirus MERS coronavirus (MERS-CoV) and bat MERS-like coronavirus HKU4 share lower sequence similarities in their spikes, RBDs or RBM (Fig. 3C), and yet they recognize the same receptor dipeptidyl peptidase 4 (DPP4) (27, 28). Thus, sequence similarities between 2019-nCoV and SARS-CoV spikes suggest the possibility for them to share the same receptor ACE2. Importantly, compared to SARS-CoV RBM, 2019-nCoV RBM does not contain any deletion or insertion (except for a one-residue insertion on a loop away from the ACE2-binding region) (Fig. 3A), providing additional evidence that 2019-nCoV uses ACE2 as its receptor. Furthermore, among the 14 ACE2-contacting residues in the RBD, 9 are fully conserved and 4 are partially conserved among 2019-nCoV and SARS-CoV

from human, civet and bat (Fig. 3A). A final piece of strong evidence supporting ACE2 as the receptor for 2019-nCoV surrounds the five residues in 2019-nCoV RBM that underwent natural selections in SARS-CoV and played critical roles in the cross-species transmission of SARS-CoV (i.e., residue 442, 472, 479, 480, 487 in SARS-CoV RBD) (Fig. 1B). We discuss these residues in more details below.

First, residue 493 in 2019-nCoV RBD (corresponding to residue 479 in SARS-CoV) is a glutamine (Fig. 1B, 1D). A previously designed SARS-CoV RBD is optimal for binding to human ACE2 (Fig. 1B, 1C) (26). Residue 479 in SARS-CoV RBD is located near virus-binding hotspot Lys31 (i.e., hotspot-31) on human ACE2 (Fig. 1C). Hotspot-31 consists of a salt bridge between Lys31 and Glu35 buried in a hydrophobic environment. In civet SARS-CoV RBD (year 2002), residue 479 is a lysine, which imposes steric and electrostatic interference with hotspot-31. In human SARS-CoV RBD (year 2002), residue 479 becomes an asparagine. The K479N mutation removes the unfavorable interaction at the RBD/human ACE2 interface, enhances viral binding to human ACE2, and played a critical role in the civet-to-human transmission of SARS-CoV (Fig. 1C) (24-26). Importantly, Gln493 in 2019-nCoV RBD is compatible with hotspot-31, suggesting that 2019-nCoV is capable of recognizing human ACE2 and infecting human cells.

Second, residue 501 in 2019-nCoV RBD (corresponding to residue 487 in SARS-CoV) is an asparagine (Fig. 1B, 1D). Based on our previous structural analysis, residue 487 in SARS-CoV is located near virus-binding hotspot Lys353 (i.e., hotspot-353) on human ACE2 (Fig. 1C) (26). Hotspot-353 consists of a salt bridge between Lys353 and Asp38 also buried in a hydrophobic environment. In civet SARS-CoV RBD (year 2002),

155 residue 487 is a serine, which cannot provide favorable support for hotspot-353. In
156 human SARS-CoV isolated in year 2002, residue 487 is a threonine, which strengthens
157 the structural stability of hotspot-353. The S487T mutation adds the favorable interaction
158 at the RBD/human ACE2 interface, enhances viral binding to human ACE2, and played a
159 critical role in the human-to-human transmission of SARS-CoV (24-26). In human
160 SARS-CoV isolated in year 2003, residue 487 is a serine and there was no human-to-
161 human transmission for this SARS-CoV strain. Asn501 in 2019-nCoV RBD provides
162 more support to hotspot-353 than Ser487, but less than Thr487. This analysis suggests
163 that 2019-nCoV recognizes human ACE2 less efficiently than human SARS-CoV (year
164 2002), but more efficiently than human SARS-CoV (year 2003). Hence, at least when
165 considering the ACE2-RBD interactions, 2019-nCoV has gained some capability to
166 transmit from human and human.

167 Third, residues 455, 486 and 494 are leucine, phenylalanine and serine in 2019-
168 nCoV RBD, respectively (corresponding to residues 442, 472 and 480 in SARS-CoV,
169 respectively) (Fig. 1B, 1C, 1D). Based on our previous structural analysis, these three
170 residues in SARS-CoV RBD play significant roles, albeit not as dramatic as residues 479
171 and 487, in ACE2 binding (24-26). More specifically, Tyr442 of human and civet SARS-
172 CoV RBDs provides unfavorable interactions with hotspot-31 on human ACE2 (this
173 residue has been mutated to Phe442 in the optimized RBD); Leu455 of 2019-nCoV RBD
174 provides favorable interactions with hotspot-31, hence enhancing viral binding to human
175 ACE2. Leu472 of human and civet SARS-CoV RBDs provides favorable support for
176 hotspot-31 on human ACE2 through hydrophobic interactions with ACE2 residue Met82
177 and several other hydrophobic residues (this residue has been mutated to Phe472 in the

optimized RBD); Phe486 of 2019-nCoV RBD provides even more support for hotspot-31, hence also enhancing viral binding to human ACE2. Asp480 of human and civet SARS-CoV RBDs provides favorable support for hotspot-353 on human ACE2 through a neighboring tyrosine (this residue remains as an aspartate in the optimized RBD); Ser494 in 2019-nCoV RBD still provides positive support for hotspot-353, but the support is not as favorable as provided by Asp480. Overall, Leu455, Phe486 and Ser494 of 2019-nCoV RBD support that 2019-nCoV recognizes human ACE2 and infects human cells.

Last, having analyzed the interactions between 2019-nCoV RBD and human ACE2, how does 2019-nCoV RBD interact with putative ACE2 receptor orthologues from other animal species? Compared to human ACE2, both hotspot-31 and hotspot-353 on civet ACE2 have changed significantly (Fig. 4A). Specifically, residue 31 of civet ACE2 becomes a threonine, which can no longer form a salt bridge with Glu35; residue 38 of civet ACE2 becomes a glutamate, which forms a strong bifurcated salt bridge with Lys353 and no longer needs strong support from neighboring residues. A previously designed SARS-CoV RBD is optimal for binding to civet ACE2 (Fig. 1B, 4B) (26). In this designed RBD, Tyr442 forms a hydrogen bond with Thr31 of civet ACE2, and Arg479 forms a strong bifurcated salt bridge with Glu35 of civet ACE2. Moreover, in the designed RBD, Pro472 avoids unfavorable interactions with Thr82 of civet ACE2, and Gly480 does not provide unneeded support for hotspot-353. Furthermore, in the designed RBD, Thr487 provides limited but helpful support for hotspot-353. Here we constructed a structural model for the complex of 2019-nCoV RBD and civet ACE2 (Fig. 4C). Based on this model, Phe486 of 2019-nCoV RBD forms moderately unfavorable interaction with the polar side chain of Thr82 of civet ACE2, and Leu455 and Gln493 would lose

favorable interactions with civet ACE2 but they would be still compatible with civet ACE2. Thus, 2019-nCoV likely still uses civet ACE2 as its receptor, although it appears that 2019-nCoV RBD has not evolved adaptively for civet ACE2 binding. Moreover, 2019-nCoV likely does not use mouse or rat ACE2 as its receptor because mouse or rat ACE2 contains a histidine at the 353 position, which does not fit into the virus/receptor interact as well as a lysine does (Fig. 3A). 2019-nCoV RBD likely recognizes ACE2 from pigs, ferrets, cats, orangutans, monkeys and humans with similar efficiency, because these ACE2 molecules are identical or similar in the critical virus-binding residues. The situation involving bat ACE2 is complex because of the diversity of bat species (29). Based on the sequence of ACE2 from *Rhinolophus sinicus* bats (which can be recognized by bat SARS-CoV strain Rs3367), 2019-nCoV RBD likely also recognizes bat ACE2 as its receptor. Overall, 2019-nCoV likely recognizes ACE2 orthologues from a diversity of species, except for mouse and rat ACE2 (which should be poor receptors for 2019-nCoV).

Discussion

Atomic level resolution of complex virus-receptor interactions provides new opportunities for predictive biology. In this instance, we used prior knowledge gleamed from multiple SARS-CoV strains (isolated from different hosts in different years) and ACE2 receptors (from different animal species) to model predictions for novel 2019-nCoV. Our structural analyses confidently predict that 2019-nCoV uses ACE2 as its host receptor, consistent to two other new publications (30, 31). Compared to previously isolated SARS-CoV strains, 2019-nCoV likely uses human ACE2 less efficiently than

human SARS-CoV (year 2002), but more efficiently than human SARS-CoV (year 2003). Because ACE2-binding affinity has been shown to be one of the most important determinants of SARS-CoV infectivity, 2019-nCoV has evolved the capability to infect humans and some capability to transmit among humans. Alarmingly, our data predict that a single N501T mutation (corresponding to the S487T mutation in SARS-CoV) may significantly enhance the binding affinity between 2019-nCoV RBD and human ACE2. Thus, 2019-nCoV evolution in patients should be closely monitored for the emergency of novel mutations at the 501 position (to a lesser extent, also the 494 position).

What is the source of 2019-nCoV and did a key intermediate host play an important role in the current 2019-nCoV outbreak? Similar to SARS-CoV, 2019-nCoV most likely has originated from bats, given its close phylogenetic relationship with other β-genus lineage b bat SARS-CoV (Fig. 2). Moreover, 2019-nCoV likely recognizes ACE2 from a diversity of animal species, including palm civets, as its receptor. In the case of SARS-CoV, some of its critical RBM residues were adapted to human ACE2, while some others were adapted to civet ACE2 (26); this type of partial viral adaptations to two host species promoted virus replication and cross-species transmission between the two host species. In the case of 2019-nCoV, however, there is no strong evidence for adaptive mutations in its critical RBM residues that specifically promote viral binding to civet ACE2. Hence, either palm civets were not intermediate hosts for 2019-nCoV, or they passed 2019-nCoV to humans quickly before 2019-nCoV had any chance to adapt to civet ACE2. Like SARS-CoV, 2019-nCoV will likely replicate inefficiently in mice and rats, ruling them out as intermediate hosts for 2019-nCoV. Moreover, we predict that either 2019-nCoV or laboratory mice and rats would need to be genetically engineered

247 before a robust mouse or rat model for 2019-nCoV would become available. Pigs, ferrets,
248 cats and non-human primates contain largely favorable 2019-nCoV-contacting residues
249 in their ACE2, and hence may serve as animal models or intermediate hosts for 2019-
250 nCoV. It is worth noting that SARS-CoV was isolated in wild palm civets near Wuhan in
251 2005 (9), and its RBD had already been well adapted to civet ACE2 (except for residue
252 487). Thus, bats and other wild animals in and near Wuhan should be screened for both
253 SARS-CoV and 2019-nCoV.

254 These above analyses are based on the modeling of 2019-nCoV RBD/ACE2
255 interactions, heavily grounded in a series of atomic level structures of SARS-CoV
256 isolated from different hosts in different years (18, 24-26). There are certainly other
257 factors that affect the infectivity and pathogenesis of 2019-nCoV and will need to be
258 investigated. Nevertheless, our decade-long structural studies on SARS-CoV have firmly
259 shown that receptor recognition by SARS-CoV is one of the most important determinants
260 of its cross-species transmission and human-to-human transmission, a conclusion that has
261 been confirmed by different lines of research (13, 14). One of the long-term goals of our
262 previous structural studies on SARS-CoV was to build an atomic-level iterative
263 framework of virus-receptor interactions that facilitate epidemic surveillance, predict
264 species-specific receptor usage, and identify potential animal hosts and likely animal
265 models of human diseases. This study provides a robust test of this reiterative framework,
266 providing the basic, translational and public health research communities with predictive
267 insights that may help study and battle this novel 2019-nCoV.
268

Materials and Methods

Structural analysis. Software Coot was used for introducing mutations to structural models (32). Software PyMol was used for preparing structural figures (33).

Phylogenetic analysis. Consensus radial phylograms were generated in Geneious Prime (v.2020.0.3), with the Jukes-Cantor genetic distance model, the Neighbor-Joining build method, and no outgroup, with 100 bootstrap replicates. Phylograms were rendered for publication in Adobe Illustrator CC 2020.

Sequence alignment. Protein sequence alignments were done using Clustal Omega (34).

Acknowledgements

This work was supported by NIH grants R01AI089728 and R01AI110700 (to F.L. and R.S.B.). We would like to thank Shanghai Public Health Clinical Center and School of Public Health, Central Hospital of Wuhan, Huazhong University of Science and Technology, Wuhan Center for Disease Control and Prevention, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control, and University of Sydney Australia for releasing the sequence of 2019-nCoV genome.

References

1. **Lee N, Hui D, Wu A, Chan P, Cameron P, Joynt GM, Ahuja A, Yung MY, Leung CB, To KF, Lui SF, Szeto CC, Chung S, Sung JJY.** 2003. A major outbreak of severe acute respiratory syndrome in Hong Kong. New England Journal of Medicine **348:**1986-1994.

2. **Yu ITS, Li YG, Wong TW, Tam W, Chan AT, Lee JHW, Leung DYC, Ho T.** 2004. Evidence of airborne transmission of the severe acute respiratory syndrome virus. New England Journal of Medicine **350:**1731-1739.

3. **Marra MA, Jones SJM, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YSN, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Coughlin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JE, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, et al.** 2003. The genome sequence of the SARS-associated coronavirus. Science **300:**1399-1404.

4. **Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, Lim W, Nicholls J, Yee WKS, Yan WW, Cheung MT, Cheng VCC, Chan KH, Tsang DNC, Yung RWH, Ng TK, Yuen KY.** 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet **361:**1319-1325.

5. **Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, Luo SW, Li PH, Zhang LJ, Guan YJ, Butt KM, Wong KL, Chan KW, Lim W, Shortridge KF, Yuen KY, Peiris JSM, Poon LLM.** 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. Science **302:**276-278.

6. **Liang GD, Chen QX, Xu JG, Liu YF, Lim W, Peiris JSM, Anderson LJ, Ruan L, Li H, Kan B, Di B, Cheng P, Chan KH, Erdman DD, Gu SY, Yan XG, Liang WL, Zhou DH, Haynes L, Duan SM, Zhang X, Zheng H, Gao Y, Tong SX, Li DX, Fang L, Qin PZ, Xu WB.** 2004. Laboratory diagnosis of four recent sporadic cases of community-acquired SARS, Guangdong Province, China. Emerging Infectious Diseases **10:**1774-1781.

7. **Song HD, Tu CC, Zhang GW, Wang SY, Zheng K, Lei LC, Chen QX, Gao YW, Zhou HQ, Xiang H, Zheng HJ, Chern SWW, Cheng F, Pan CM, Xuan H, Chen SJ, Luo HM, Zhou DH, Liu YF, He JF, Qin PZ, Li LH, Ren YQ, Liang WJ, Yu YD, Anderson L, Wang M, Xu RH, Wu XW, Zheng HY, Chen**

JD, Liang GD, Gao Y, Liao M, Fang L, Jiang LY, Li H, Chen F, Di B, He LJ, Lin JY, Tong SX, Kong XG, Du L, Hao P, Tang H, Bernini A, Yu XJ, Spiga O, Guo ZM, et al. 2005. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proceedings of the National Academy of Sciences of the United States of America **102:**2430-2435.

8. **Sheahan T, Rockx B, Donaldson E, Sims A, Pickles R, Corti D, Baric R.** 2008. Mechanisms of zoonotic severe acute respiratory syndrome coronavirus host range expansion in human airway epithelium. Journal of Virology **82:**2274-2285.

9. **Liu L, Fang Q, Deng F, Wang HZ, Yi CE, Ba L, Yu WJ, Lin RD, Li TS, Hu ZH, Ho DD, Zhang LQ, Chen ZW.** 2007. Natural mutations in the receptor binding domain of spike glycoprotein determine the reactivity of cross-neutralization between palm civet coronavirus and severe acute respiratory syndrome coronavirus. Journal of Virology **81:**4694-4700.

10. **Hu B, Zeng LP, Yang XL, Ge XY, Zhang W, Li B, Xie JZ, Shen XR, Zhang YZ, Wang N, Luo DS, Zheng XS, Wang MN, Daszak P, Wang LF, Cui J, Shi ZL.** 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog **13:**e1006698.

11. **Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, Zhang YJ, Luo CM, Tan B, Wang N, Zhu Y, Crameri G, Zhang SY, Wang LF, Daszak P, Shi ZL.** 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature **503:**535-538.

12. **Cui J, Li F, Shi ZL.** 2019. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol **17:**181-192.

13. **Li F.** 2013. Receptor recognition and cross-species infections of SARS coronavirus. Antiviral Res **100:**246-254.

14. **Li WH, Wong SK, Li F, Kuhn JH, Huang IC, Choe H, Farzan M.** 2006. Animal origins of the severe acute respiratory syndrome coronavirus: Insight from ACE2-S-protein interactions. Journal of Virology **80:**4211-4219.

15. **Perlman S, Netland J.** 2009. Coronaviruses post-SARS: update on replication and pathogenesis. Nature Reviews Microbiology **7:**439-450.

16. **Li F.** 2016. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu Rev Virol **3:**237-261.

17. **Li WH, Moore MJ, Vasilieva N, Sui JH, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M.** 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature **426:**450-454.

18. **Li F.** 2015. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. J Virol **89:**1954-1964.

19. **Li WH, Greenough TC, Moore MJ, Vasilieva N, Somasundaran M, Sullivan JL, Farzan M, Choe H.** 2004. Efficient replication of severe acute respiratory syndrome coronavirus in mouse cells is limited by murine angiotensin-converting enzyme 2. Journal of Virology **78:**11429-11433.

20. **Li WH, Zhang CS, Sui JH, Kuhn JH, Moore MJ, Luo SW, Wong SK, Huang IC, Xu KM, Vasilieva N, Murakami A, He YQ, Marasco WA, Guan Y, Choe HY, Farzan M.** 2005. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. Embo Journal **24:**1634-1643.

21. **McCray PB, Pewe L, Wohlford-Lenane C, Hickey M, Manzel L, Shi L, Netland J, Jia HP, Halabi C, Sigmund CD, Meyerholz DK, Kirby P, Look DC, Perlman S.** 2007. Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. Journal of Virology **81:**813-821.

22. **Moore MJ, Dorfman T, Li WH, Wong SK, Li YH, Kuhn JH, Coderre J, Vasilieva N, Han ZC, Greenough TC, Farzan M, Choe H.** 2004. Retroviruses pseudotyped with the severe acute respiratory syndrome coronavirus spike protein efficiently infect cells expressing angiotensin-converting enzyme 2. Journal of Virology **78:**10628-10635.

23. **Qu XX, Hao P, Song XJ, Jiang SM, Liu YX, Wang PG, Rao X, Song HD, Wang SY, Zuo Y, Zheng AH, Luo M, Wang HL, Deng F, Wang HZ, Hu ZH, Ding MX, Zhao GP, Deng HK.** 2005. Identification of two critical amino acid residues of the severe acute respiratory syndrome coronavirus spike protein for its variation in zoonotic tropism transition via a double substitution strategy. Journal of Biological Chemistry **280:**29588-29595.

24. **Li F.** 2008. Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections. Journal of Virology **82:**6984-6991.

25. **Li F, Li WH, Farzan M, Harrison SC.** 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science **309:**1864-1868.

26. **Wu KL, Peng GQ, Wilken M, Geraghty RJ, Li F.** 2012. Mechanisms of Host Receptor Adaptation by Severe Acute Respiratory Syndrome Coronavirus. Journal of Biological Chemistry **287:**8904-8911.

27. **Yang Y, Du L, Liu C, Wang L, Ma C, Tang J, Baric RS, Jiang S, Li F.** 2014. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc Natl Acad Sci U S A **111:**12516-12521.

28. **Raj VS, Mou HH, Smits SL, Dekkers DHW, Muller MA, Dijkman R, Muth D, Demmers JAA, Zaki A, Fouchier RAM, Thiel V, Drosten C, Rottier PJM, Osterhaus A, Bosch BJ, Haagmans BL.** 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature **495:**251-254.

29. **Hou YX, Peng C, Yu M, Li Y, Han ZG, Li F, Wang LF, Shi ZL.** 2010. Angiotensin-converting enzyme 2 (ACE2) proteins of different bat species confer variable susceptibility to SARS-CoV entry. Archives of Virology **155:**1563-1569.

30. **Michael C Letko VM.** 2020. Functional assessment of cell entry and receptor usage for lineage B β -coronaviruses, including 2019-nCoV. wwwbiorxivorg.

31. **Zheng-Li Shi PZ, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, Hui-Dong Chen, Jing Chen, Yun Luo, Hua Guo, Ren-Di Jiang, Mei-Qin Liu, Ying Chen, Xu-Rui Shen, Xi Wang, Xiao-Shuang Zheng, Kai Zhao, Quan-Jiao Chen, Fei Deng, Lin-Lin Liu, Bing Yan, Fa-Xian Zhan, Yan-Yi Wang, Gengfu Xiao.** 2020. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. wwwbiorxivorg.

32. **Emsley P, Cowtan K.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallographica Section D-Biological Crystallography **60:**2126-2132.

33. **Anonymous.** The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

34. **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG.** 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol **7:**539.

Figure legends:

Figure 1: Structural analysis of human ACE2 recognition by 2019-nCoV and SARS-CoV. (A) Overall structure of human SARS-CoV RBD (year 2002) complexed with human ACE2. PDB ID is 2AJF. ACE2 is in green, the core of RBD (receptor-binding domain) is in cyan, and RBM (receptor-binding motif) is in magenta. (B) Critical residue changes in the RBMs of SARS-CoV and 2019-nCoV. All these five residues in SARS-CoV underwent natural selections and were shown to be critical for ACE2 recognition, cell entry, and host range of SARS-CoV. The residue numbers are shown as in SARS-CoV RBD, with the corresponding residue numbers in 2019-nCoV shown in parentheses. For viral adaption to ACE2, > means "is more adapted" and = means "is similarly adapted". (C) Experimentally determined structure of the interface between a designed SARS-CoV RBD (optimized for human ACE2 recognition) and human ACE2. PDB ID is 3SCI. (D) Modeled structure of the interface between 2019-nCoV RBD and human ACE2. Here mutations were introduced to the RBD region in panel (C) based on sequence differences between SARS-CoV and 2019-nCoV. GenBank accession numbers are: MN908947.1 for 2019-nCoV Spike; NC_004718.3 for human SARS -CoV Spike (year 2002; strain Tor2); AGZ48818.1 for bat SARS-CoV Spike (year 2013; strain Rs3367); AY304486.1 for civet SARS-CoV spike (year 2002; SZ3); AY525636 for human/civet SARS-CoV spike (year 2003; strain GD03). References for the other sequences are: civet SARS-CoV spike (year 2005) (9); human SARS-CoV spike (year 2008) (8).

Figure 2. Spike phylogeny of representative β-genus lineage b coronaviruses. The Spike protein sequences of selected β-genus lineage b coronaviruses were aligned and phylogenetically compared. Sequences were aligned using free end gaps with the Blosum62 cost matrix in Geneious Prime. The tree was constructed using the neighbor-joining method based on the multiple sequence alignment, also in Geneious Prime. Numbers following the underscores in each sequence correspond to the GenBank accession number. The radial phylogram was exported from Geneious and then rendered for publication using EvolView (evolgenius.info) and Adobe Illustrator CC 2020.

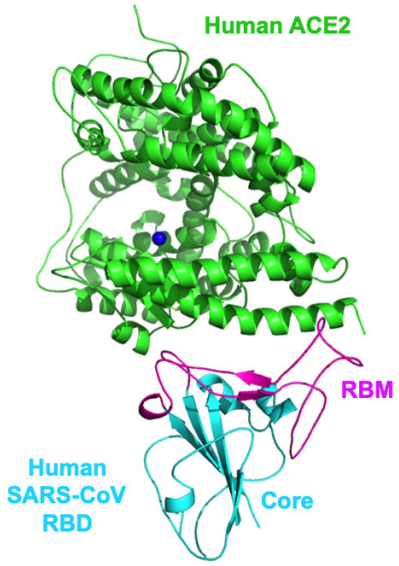
Figure 3: Sequence comparison of 2019-nCoV and SARS-CoV. (A) Sequence alignment of SARS-CoV and 2019-nCoV RBDs. RBM residues are in magenta. The five critical residues in Fig. 1B are in blue. ACE2-contacting residues are shaded. Asterisks indicate positions that have a single, fully conserved residue. Colons indicate positions that have strongly conserved residues. Periods indicate positions that have weakly conserved residues. (B) Sequence similarities of SARS-CoV and 2019-nCoV in the spike protein, RBD and RBM, respectively. (C) Sequence similarities of MERS-CoV and HKU4 virus in the spike protein, RBD and RBM, respectively. GenBank accession numbers are: JX869059.2 for human MERS-CoV Spike; NC_009019.1 for bat HKU4-CoV Spike.

Figure 4: Structural analysis of animal ACE2 recognition by 2019-nCoV and SARS-CoV. (A) Critical changes in virus-contacting residues of ACE2 from different host species. GenBank accession numbers for ACE2 are as follows: NM_001371415.1

Journal of Virology

(human), AAX63775.1 (civet), KC881004.1 (bat), NP_001123985.1 (mouse), AY881244 (rat), NP_001116542.1 (pig), AB208708 (ferret), NM_001039456 (cat), Q5RFN1 (orangutan), and AY996037 (monkey). (B) Experimentally determined structure of the interface between a designed SARS-CoV RBD (optimized for civet ACE2 recognition) and civet ACE2. PDB ID is 3SCK. (C) Modeled structure of the interface between 2019-nCoV RBD and civet ACE2. Here mutations were introduced to the RBD region in panel (B) based on sequence differences between SARS-CoV and 2019-nCoV.

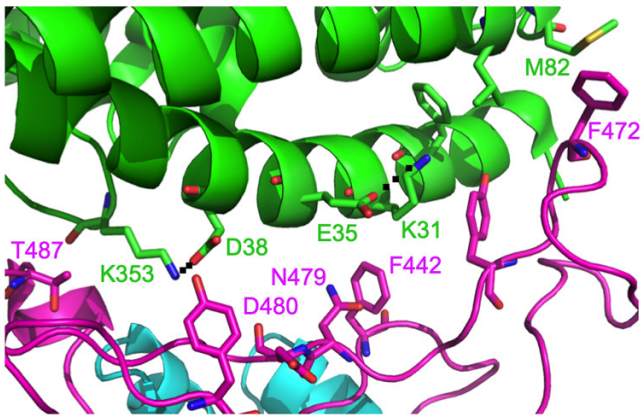
A



B

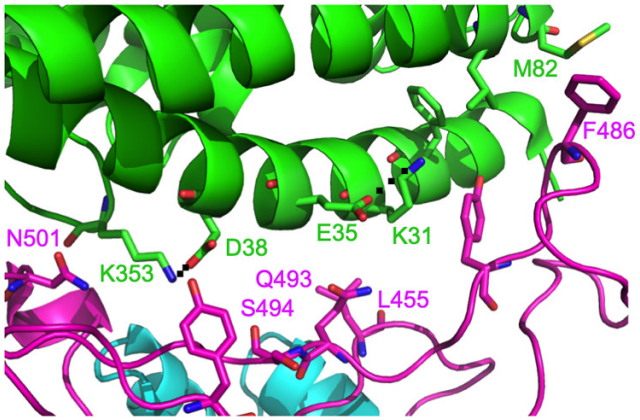
Virus	Year	442	472	479	480	487
SARS - human	2002	Y	L	**N**	D	**T**
SARS - civet	2002	Y	L	**K**	D	**S**
SARS - human/civet	2003	Y	P	**N**	G	**S**
SARS - civet	2005	Y	P	**R**	G	**S**
SARS - human	2008	F	F	**N**	D	**S**
Viral adaption to human ACE2		F > Y	F > L > P	**N = R >>> K**	D > G	**T >>> S**
Optimized - human	In vitro design	F	F	**N**	D	**T**
Viral adaptation to civet ACE2		Y > F	P = L > F	**R > K = N**	G > D	**T > S**
Optimized - civet	In vitro design	Y	P	**R**	G	**T**
SARS - bat	2013	S	F	**N**	D	**N**
2019-nCoV – human	2019	L (455)	F (486)	**Q (493)**	S (494)	**N (501)**

C

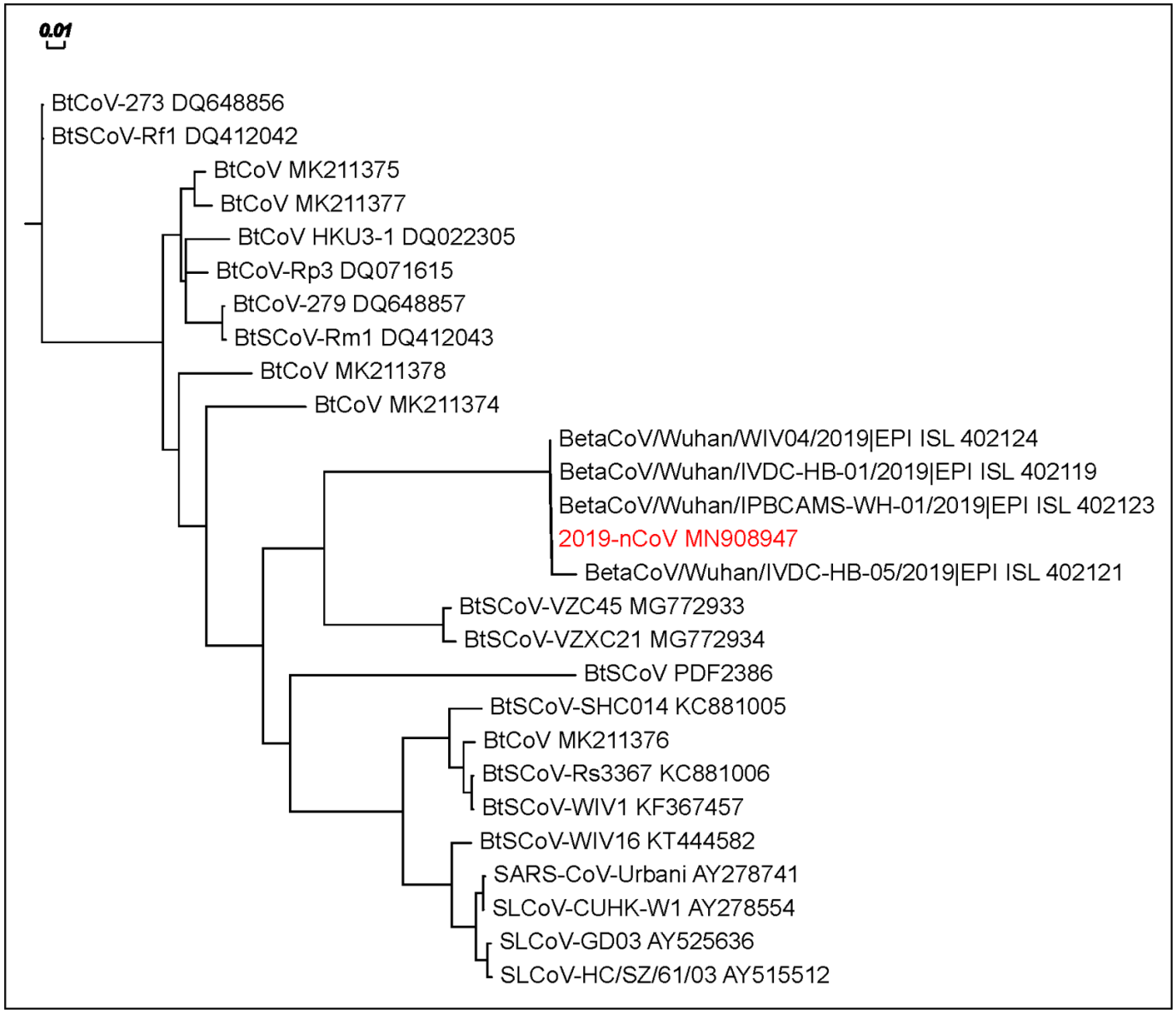


Human SARS-CoV-optimized RBD

D



Human 2019-nCoV RBD (model)



A

```
Human-SARS-2002   306 RVVPS GDVVRFPNIT NLCPFGEVFN ATKFPSVYAW ERKKISNCVA DYSVLYNSTF 360
Civet-SARS-2002   319 RVVPS GDVVRFPNIT NLCPFGEVFN ATKFPSVYAW ERKRISNCVA DYSVLYNSTS 373
Bat-SARS-2013     319 RVAPS KEVVRFPNIT NLCPFGEVFN ATTFPSVYAW ERKRISNCVA DYSVLYNSTS 373
2019-nCoV         319 RVQPT ESIVRFPNIT NLCPFGEVFN ATRFASVYAW NRKRISNCVA DYSVLYNSAS 373
                      ** *:  .:******* ********** ** * ***** :**:****** ********:

Human-SARS-2002  FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG QTGVIADYNY KLPDDFMGCV 420
Civet-SARS-2002  FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG QTGVIADYNY KLPDDFMGCV 433
Bat-SARS-2013    FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG QTGVIADYNY KLPDDFTGCV 433
2019-nCoV        FSTFKCYGVS PTKLNDLCFT NVYADSFVIR GDEVRQIAPG QTGKIADYNY KLPDDFTGCV 433
                 **********  ********: ********:: **:******* *** ****** ****** ***

Human-SARS-2002  LAWNTRNIDA TSTGNYNYKY RYLRHGKLRP FERDISNVPF SPDGKPCTP-P ALNCYWPLND 480
Civet-SARS-2002  LAWNTRNIDA TSTGNYNYKY RYLRHGKLRP FERDISNVPF SPDGKPCTP-P ALNCYWPLKD 493
Bat-SARS-2013    LAWNTRNIDA TQTGNYNYKY RSLRHGKLRP FERDISNVPF SPDGKPCTP-P AFNCYWPLND 493
2019-nCoV        IAWNSNNLDS KVGGNYNYLY RLFRKSNLKP FERDISTEIY QAGSTPCNGVE GFNCYFPLQS 494
                 :***:.*:*: .  ***** * * :*:.:*:* ******.  : . ...**.    .:***:**:.

Human-SARS-2002  YGFYTTTGIG YQPYRVVVLS FELLNAPATV CGPKL 515
Civet-SARS-2002  YGFYTTSGIG YQPYRVVVLS FELLNAPATV CGPKL 528
Bat-SARS-2013    YGFYITNGIG YQPYRVVVLS FELLNAPATV CGPKL 528
2019-nCoV        YGFQPTNGVG YQPYRVVVLS FELLHAPATV CGPKK 529
                 ***  *.*:* ********** ****:***** ****
```

B

Spike / RBD / RBM	SARS-human	SARS-civet	SARS-bat	2019-nCoV
SARS-human	100% / 100% / 100%			
SARS-civet	98.12% / 98.10% / 97.18%	100% / 100% / 100%		
SARS-bat	92.33% / 94.29% / 92.96%	92.75% / 94.76% / 91.55%	100% / 100% / 100%	
2019-nCoV	76.04% / 73.33% / 50.00%	76.78% / 74.29% / 50.00%	77.50% / 75.71% / 52.78%	100% / 100% / 100%

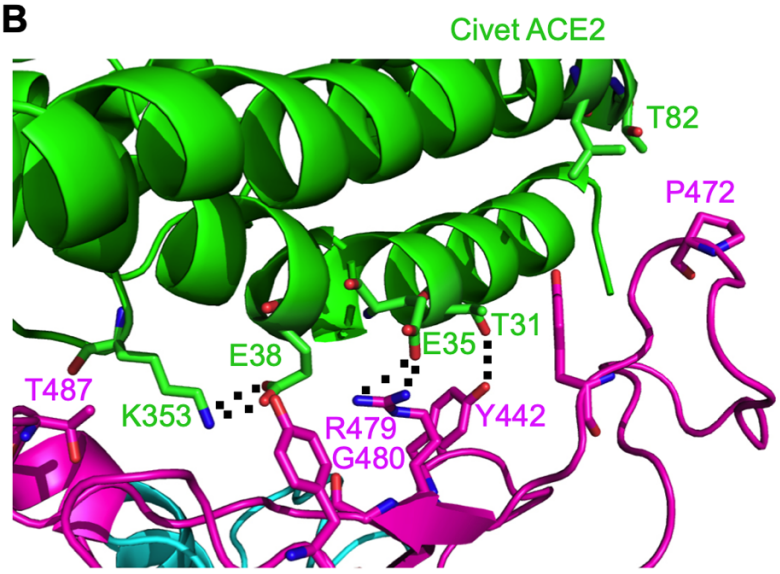
C

Spike /RBD /RBM	MERS-human
HKU4-bat	67.04% /57.69% /40.79%

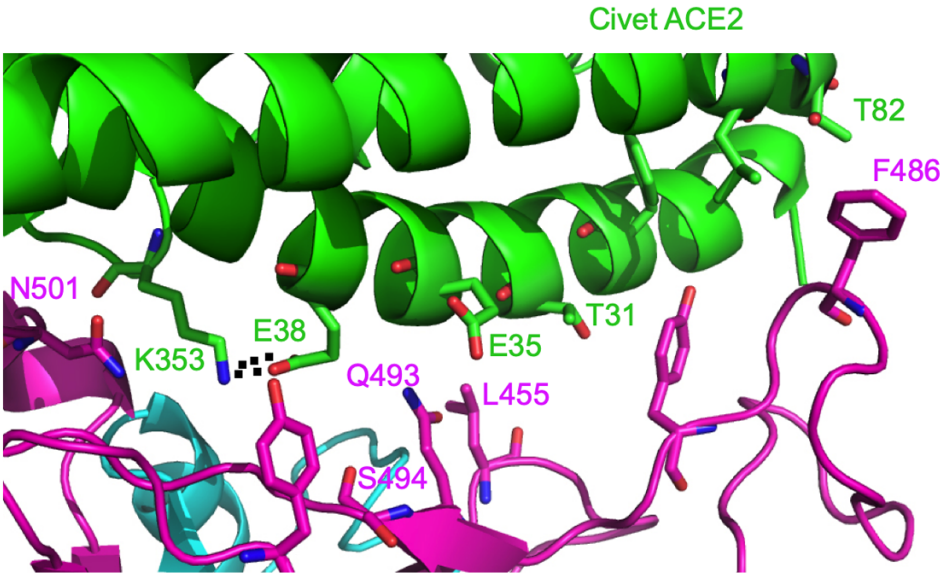
A

ACE2	31	35	38	82	353
Human	K	E	D	M	K
Civet	T	E	E	T	K
Bat	K	K	D	N	K
Mouse	N	E	D	S	H
Rat	K	E	D	N	H
Pig	K	E	D	T	K
Ferret	K	E	E	T	K
Cat	K	E	E	T	K
Orangutan	K	E	D	M	K
Monkey	K	E	D	M	K

B



C



TAKE A CLOSER LOOK AT ASM

YOUR COMMUNITY

Become an ASM member

Join the global community committed to advancing microbial sciences through:

Sharing Science • Building Careers • Networking with Peers • Advocating for Science • Improving Global Health • Publishing Peer Reviewed Research

ASM membership provides you access to:

- Cutting edge science
- Professional development opportunities
- Travel awards
- Career resources
- Directory of your peers
- Curriculum guidelines and lab protocol access
- Discounts on ASM events, books and journals publication fees
- Educational courses and webinars
- CE credits for clinical laboratory scientists (CLS, MT, MLT, MLS)

Learn more and choose the membership that best fits your needs at

asm.org/membership

AMERICAN SOCIETY FOR MICROBIOLOGY

1752 N St. NW
Washington, DC 20036

Questions? Contact service@asmusa.org