AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

Discovery of Novel Bat Coronaviruses in South China That Use the Same Receptor as Middle East Respiratory Syndrome Coronavirus

Chu-Ming Luo,[a,b,c] Ning Wang,[a,b] Xing-Lou Yang,[a] Hai-Zhou Liu,[a] Wei Zhang,[a] Bei Li,[a] Ben Hu,[a] Cheng Peng,[a] Qi-Bin Geng,[c] Guang-Jian Zhu,[d] Fang Li,[c] Zheng-Li Shi[a]

[a]CAS Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, Hubei, China

[b]University of Chinese Academy of Sciences, Beijing, China

[c]Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, Minnesota, USA

[d]EcoHealth Alliance, New York, New York, USA

ABSTRACT Middle East respiratory syndrome coronavirus (MERS-CoV) has represented a human health threat since 2012. Although several MERS-related CoVs that belong to the same species as MERS-CoV have been identified from bats, they do not use the MERS-CoV receptor, dipeptidyl peptidase 4 (DPP4). Here, we screened 1,059 bat samples from at least 30 bat species collected in different regions in south China and identified 89 strains of lineage C betacoronaviruses, including *Tylonycteris pachypus coronavirus HKU4*, *Pipistrellus pipistrellus coronavirus HKU5*, and MERS-related CoVs. We sequenced the full-length genomes of two positive samples collected from the great evening bat, *Ia io*, from Guangdong Province. The two genomes were highly similar and exhibited genomic structures identical to those of other lineage C betacoronaviruses. While they exhibited genome-wide nucleotide identities of only 75.3 to 81.2% with other MERS-related CoVs, their gene-coding regions were highly similar to their counterparts, except in the case of the spike proteins. Further protein-protein interaction assays demonstrated that the spike proteins of these MERS-related CoVs bind to the receptor DPP4. Recombination analysis suggested that the newly discovered MERS-related CoVs have acquired their spike genes from a DPP4-recognizing bat coronavirus HKU4. Our study provides further evidence that bats represent the evolutionary origins of MERS-CoV.

IMPORTANCE Previous studies suggested that MERS-CoV originated in bats. However, its evolutionary path from bats to humans remains unclear. In this study, we discovered 89 novel lineage C betacoronaviruses in eight bat species. We provide evidence of a MERS-related CoV derived from the great evening bat that uses the same host receptor as human MERS-CoV. This virus also provides evidence for a natural recombination event between the bat MERS-related CoV and another bat coronavirus, HKU4. Our study expands the host ranges of MERS-related CoV and represents an important step toward establishing bats as the natural reservoir of MERS-CoV. These findings may lead to improved epidemiological surveillance of MERS-CoV and the prevention and control of the spread of MERS-CoV to humans.

KEYWORDS MERS-related coronavirus, bat, dipeptidyl peptidase 4, virus discovery

Coronaviruses (CoVs) infect a wide range of mammalian and avian hosts, causing respiratory, enteric, hepatic, or neurological diseases of varying severity. These viruses have the largest genomes among all RNA viruses, leading to an increased number of replication errors compared to the host genome (1). Different CoVs can also recombine their genomes upon infecting the same host cell, contributing substantially

Received 22 January 2018 **Accepted** 3 April 2018

Accepted manuscript posted online 18 April 2018

Citation Luo C, Wang N, Yang X, Liu H, Zhang W, Li B, Hu B, Peng C, Geng Q, Zhu G, Li F, Shi Z. 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.

Editor Tom Gallagher, Loyola University Medical Center

Address correspondence to Fang Li, lifang@umn.edu, or Zheng-Li Shi, zlshi@wh.iov.cn.

to their rapid evolution (2). Hence, CoVs have the ability to adapt to different hosts with relative ease and often cross species barriers for host range expansion.

The envelope-anchored spike (S) protein is responsible for binding to a receptor on the host cell surface and guiding CoV entry into host cells (1). The spike protein contains the receptor-binding subunit S1 and the membrane fusion subunit S2 (3). A distinct receptor-binding domain (RBD) in the S1 subunit specifically recognizes a host receptor, and this RBD-receptor interaction is a critical determinant of CoV host ranges. Among all CoV genes, the gene encoding the spike protein is the most diverse due to heavy immune pressure from and adaptation to different hosts. A virus would likely need to recognize receptor analogs in order to be transmitted between different host species. Therefore, to examine whether the CoVs isolated from different host species have the same evolutionary origin, the sequences of their genomes and selected genes, RNA-dependent RNA polymerase (*RdRp*), *S*, and nucleocapsid (*N*), as well as their receptor usages, are typically compared and evaluated.

Since its discovery in Middle Eastern countries in 2012, Middle East respiratory syndrome (MERS)-CoV has infected over 2,000 people with a fatality rate of ~36% (http://www.who.int/emergencies/mers-cov/en/). MERS-CoV belongs to lineage C in the *Betacoronavirus* genus of the *Coronavinae* subfamily and *Coronaviridae* family (4). It is phylogenetically related to two other distinct species belonging to the same lineage, *Tylonycteris pachypus coronavirus HKU4* (HKU4) and *Pipistrellus pipistrellus coronavirus HKU5* (HKU5) (5). In addition, highly diverse MERS-related CoVs have been identified from a wide range of vespertilionid bats inhabiting Africa, Asia, Europe, and North America (6–15). Among them, MERS-related CoVs from *Neoromicia capensis* bats in South Africa share a genome-wide nucleotide identity of 85% with MERS-CoV and are considered the closest relatives of MERS-CoV discovered so far (6, 7). However, the spike proteins of these viruses do not use human or bat dipeptidyl peptidase 4 (hDPP4 or bDPP4, respectively), which is the MERS-CoV receptor (9, 16, 17). Two other lineage C betacoronaviruses (BetaCoVs), HKU4 and HKU5, from vespertilionid bats in China, have lower genetic similarities to MERS-CoV genome wide (82% identical) and in the replicase genes (86% identical) (5). However, the spike protein from HKU4, but not HKU5, recognizes human and bat DPP4 as its receptor (17). Hence, further study is needed to understand the evolutionary relationships between MERS-CoV and bat CoVs, which will be essential for identifying the natural reservoir of MERS-CoV, conducting epidemiological surveillance, and preventing and controlling the spread of MERS-CoV to humans.

Here, we investigated the distribution of lineage C BetaCoVs in three provinces of south China using a territory-wide molecular epidemiology study. We identified 89 strains of lineage C BetaCoVs from vespertilionid bats, two of which are genetically similar to MERS-CoV and dependent on DPP4 for cell entry. Our study helps bridge the gap in our knowledge regarding the natural reservoir of MERS-CoV.

RESULTS

Prevalence and genetic diversity of lineage C BetaCoVs. Between the years 2012 and 2015, we surveyed 1,059 bats belonging to 13 genera collected from Guangdong, Guangxi, and Sichuan provinces in China. We performed reverse transcription-PCR (RT-PCR) on these specimens and found 89 strains of lineage C BetaCoVs from among 88 positive specimens (Table 1 and Fig. 1A). These lineage C BetaCoVs included (i) 23 (25%) strains closely related to HKU4 from 115 samples from bamboo bats (*Tylonycteris pachypus* and *Tylonycteris* spp.); (ii) 15 (10%) strains closely related to HKU5 from 149 samples from pipistrelle bats (*Pipistrellus abramus*, *Pipistrellus pipistrellus*, *Pipistrellus minus*, and *Pipistrellus* spp.); and (iii) 50 (14%) strains closely related to MERS-CoV from 364 samples, including pipistrelle bats (*P. abramus* and *P. pipistrellus*), great evening bats (*Ia io*), and particolored bats (*Vespertilio superans*) (Table 1 and Fig. 1B). Lineage C BetaCoVs were not detected from other tested bat species (Table 1). We selected four of these strains that exhibited >90% sequence identity in the 440-bp *RdRp* region to that of MERS-CoV. After sequencing the complete *RdRp* and *S* genes of these viral

TABLE 1 Prevalence of lineage C BetaCoVs in bats

Animal type and scientific name	Common name	No. (%) of positive anal swabs for lineage C BetaCoV	Sampling location[a]
Molossidae			
Chaerephon plicatus	Wrinkle-lipped bat	0/6	11
Tadarida spp.		0/5	6
Pteropodidae			
Cynopterus sphinx	Indian short-nosed fruit bat	0/39	3, 8
Rhinolophidae			
Hipposideros armiger	Himalayan leaf-nosed bat	0/107	6, 9, 10, 11
Hipposideros larvatus	Intermediate leaf-nosed bat	0/28	6, 9, 11
Hipposideros pomona	Pomona leaf-nosed bat	0/18	10
Hipposideros pratti	Pratt's leaf-nosed bat	0/55	9, 11
Rhinolophus luctus	Woolly horseshoe bat	0/3	9
Rhinolophus macrotis	Big-eared horseshoe bat	0/3	9
Rhinolophus pearsonii	Pearson's horseshoe bat	0/25	5, 9, 11
Rhinolophus pusillus	Least horseshoe bat	0/15	9, 11
Rhinolophus rex	Rex horseshoe bat	0/8	9
Rhinolophus sinicus	Chinese horseshoe bat	0/73	5, 6, 9, 10, 11
Vespertilionidae			
Eptesicus spp.		0/1	8
Ia io	Great evening bat	4/53 (8)	6, 9, 11
Miniopterus schreibersii	Common bent-wing bat	0/1	6
Myotis adversus	Large-footed bat	0/9	7, 10
Myotis altarium	Szechwan myotis	0/2	9, 11
Myotis chinensis	Large myotis	0/21	6, 9, 10
Myotis daubentonii	Daubenton's bat	0/57	4, 8, 9, 11
Myotis longipes	Kashmir cave myotis	0/17	9, 11
Myotis ricketti	Rickett's big-footed Myotis	0/47	6, 9
Myotis spp.		0/18	6, 9
Nyctalus plancyi	Chinese noctule	0/1	6
Pipistrellus abramus	Japanese pipistrelle	15/75 (20)	4, 6, 7, 8, 9, 10, 11
Pipistrellus minus	Lesser pipistrelle	1/7 (15)	6, 8, 11
Pipistrellus pipistrellus	Common pipistrelle	6/42 (14)	4, 6, 9, 10, 11
Pipistrellus spp.		4/21 (24)	6, 9
Scotomanes ornatus	Harlequin bat	0/3	6
Scotophilus kuhlii	Lesser Asiatic yellow house bat	0/3	6, 10
Tylonycteris pachypus	Lesser bamboo bat	11/70 (16)	2
Tylonycteris spp.		13/45 (11)	6, 8, 10
Vespertilio murinus	Particolored bat	0/3	11
Vespertilio superans	Asian particolored bat	34/159 (21)	1
	Unclassified bat	0/19	6, 9, 10, 11

[a]Abbreviation of sampling locations: 1, Zigong (Sichuan Province); 2, Chongzuo (Guangxi Zhuang Autonomous Region); 3, Beihai (Guangxi Zhuang Autonomous Region); 4, Xuwen (Guangdong Province); 5, Lechang (Guangdong Province); 6, Yingde (Guangdong Province); 7, Huadu (Guangdong Province); 8, Baiyun (Guangdong Province); 9, Ruyuan (Gunagdong Province); 10, Shenzhen (Guangdong Province); 11, Yangshan (Guangdong Province).

strains, we selected the two that were the most closely related to MERS-CoV for sequencing the entire full-length genome.

Characterization of full-length genomes of novel bat MERS-related CoVs. Both of the sequenced strains were derived from *I. io* samples and were named BtCoV/Ii/GD/2013-845 and BtCoV/Ii/GD/2014-422. Both genomes are 30,113 nucleotides long, with G+C contents of 40.80% and 40.98%, respectively. They exhibit genomic structures and transcription regulatory sequences (TRS) that are identical to those of other lineage C BetaCoVs (Fig. 1C and Table 2). The only exceptions are *ORF4b* and *ORF8b*, which do not have TRS preceding them; hence, these two open reading frames (ORFs) may be translated from bicistronic mRNAs (Table 2; subgenome analysis data not shown). The replicase *ORF1ab* is ~21,300 nucleotides long and encodes a number of putative proteins, including Nsp3 (papain-like protease; PLpro), Nsp5 (chymotrypsin-like protease; 3CLpro), Nsp12 (RdRp), Nsp13 (helicase), and other proteins of unknown function (Table 3). A separate comparison of the amino acid sequences of seven conserved *ORF1ab* domains is presented in Table 4, as suggested by the International

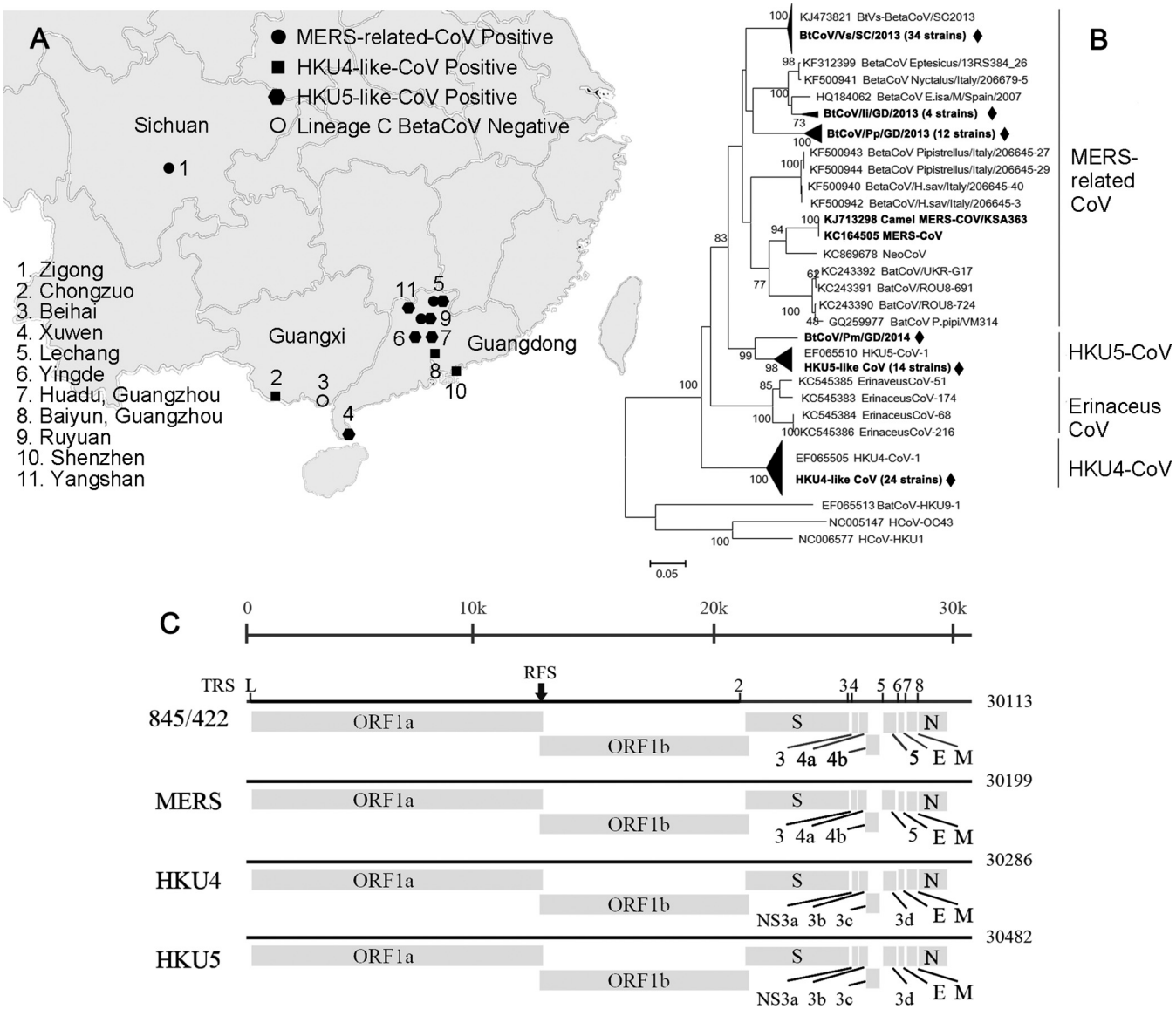


FIG 1 Geographic locations of collection sites, phylogeny of lineage C betacoronaviruses, and genomic organization of novel bat MERS-related CoVs. (A) Map of sampling locations and lineage C betacoronaviruses detected. Names of these counties, districts, and cities are noted. (B) Phylogenetic analysis of the 228-bp *RdRp* genes of the newly detected lineage C betacoronaviruses. The tree was constructed using the maximum likelihood method, with the Kimura 2-parameter model with bootstrap values calculated from 100 trees, and was midpoint rooted. Scale bar indicates the estimated number of substitutions per 10 nucleotides. Viruses characterized in this study are shown in boldface and marked with black diamonds. MERS-CoVs are shown in boldface. (C) Genomic organization of BtCoV/li/GD/2013-845 (845) and BtCoV/li/GD/2014-422 (422). Genomes are represented by black lines, and ORFs are indicated as gray blocks. Transcription regulatory core sequence (TRS) locations are marked with labeled dots. The nucleotide position of the ribosomal frameshift site (RFS) is marked with a black arrow.

Committee on Taxonomy of Viruses (ICTV) for formal CoV species delineation. The sequence identities of the seven concatenated domains of *ORF1ab* between these two novel MERS-related CoVs and MERS-CoV are around 90%, which is the threshold for defining a single viral species as proposed by the ICTV.

We next examined the complete genomic sequences of BtCoV/li/GD/2013-845 and BtCoV/li/GD/2014-422 (Table 4; see also Fig. 3A). The two genomes share only 88.1% nucleotide identity, suggesting considerable polymorphism between these two strains. The K_a/K_s ratios (where K_s is the number of synonymous substitutions per synonymous site and K_a is the number of nonsynonymous substitutions per nonsynonymous site) are generally low in the two MERS-related CoVs for all genes except *ORF8b*, suggesting that the remaining genes have evolved under purifying selection (data not shown).

TABLE 2 Coding of potential and putative transcription regulatory sequence of the novel MERS-related CoV genome sequences

Strain, protein/ORF	Nucleotide position	Length (aa)	Frame	TRS location	TRS sequence distance bases to AUG[a]
BtCoV/li/GD/2013-845 (30,113 nt, 40.98% G+C content)					
ORF1ab	237–21517	7,093	+1, +3	61	**AACGAA**(170)AUG
S	21459–25529	1,356	+1	21407	**AACGAA**(46)AUG
ORF3	25539–25841	100	+1	25528	**AACGAA**(5)AUG
ORF4a	25850–26134	94	+3	25840	**AACGAA**(4)AUG
ORF4b	26055–26783	242	+1		
ORF5	26790–27473	227	+1	26782	**AACGAA**(2)AUG
E	27551–27799	82	+3	27543	**AACGAA**(2)AUG
M	27814–28470	218	+2	27798	**AACGAA**(10)AUG
N	28526–29824	432	+2	28503	**AACGAA**(17)AUG
ORF8b	28725–29156	143	+1		
BtCoV/li/GD/2014-422 (30,113 nt, 40.80% G+C content)					
ORF1ab	238–21560	7,107	+1, +3	62	**AACGAA**(170)AUG
S	21502–25551	1,349	+1	21450	**AACGAA**(46)AUG
ORF3	25559–25879	106	+2	25550	**AACGAA**(3)AUG
ORF4a	25857–26141	94	+2	25847	**AACGAA**(4)AUG
ORF4b	26081–26779	232	+2		
ORF5	26786–27469	227	+2	26778	**AACGAA**(2)AUG
E	27547–27795	82	+1	27539	**AACGAA**(2)AUG
M	27810–28466	218	+3	27794	**AACGAA**(10)AUG
N	28521–29822	433	+3	28497	**AACGAA**(18)AUG
ORF8b	28567–29148	194	+1		

[a]The TRS are shown in boldface.

Particularly divergent are the 5′-terminal regions of the *S* gene, as well as of *ORF4a*, *ORF4b*, and *ORF5*. As the 5′-terminal region of the *S* gene encodes the RBD, it is interesting that two CoVs isolated from the same host species in the same sampling location contain divergent RBDs. Overall, the genomic sequences of both BtCoV/li/GD/

TABLE 3 Prediction of the putative polyprotein pp1a/pp1ab cleavage sites of novel MERS-related CoVs based on comparison with MERS-CoV strain EMC/2012

	BtCoV/li/GD/2013-845		BtCoV/li/GD/2014-422		
Nsp[a]	First-last amino acid residues	Protein size (aa)	First-last amino acid residues	Protein size (aa)	Putative functional domain(s)[b]
1	Met^{1}-Gly^{196}	196	Met^{1}-Gly^{196}	196	IFN antagonist
2	Asp^{197}-Gly^{856}	660	Asp^{197}-Gly^{856}	660	
3	Ala^{857}-Gly^{2754}	1,898	Ala^{857}-Gly^{2768}	1,912	ADRP, PL2pro
4	Ala^{2755}-Gln^{3261}	507	Ala^{2769}-Gln^{3275}	507	TM-2
5	Ser^{3262}-Gln^{3567}	306	Ser^{3276}-Gln^{3581}	306	3CLpro
6	Ser^{3568}-Gln^{3859}	292	Ser^{3582}-Gln^{3873}	292	TM-3
7	Ser^{3860}-Gln^{3942}	83	Ser^{3874}-Gln^{3956}	83	
8	Ala^{3943}-Gln^{4141}	199	Ala^{3957}-Gln^{4155}	199	Primase
9	Asn^{4142}-Gln^{4251}	110	Asn^{4156}-Gln^{4265}	110	RNA-binding protein
10	Ala^{4252}-Gln^{4391}	140	Ala^{4266}-Gln^{4405}	140	ZBD, cofactor of OMT
11	Ser^{4392}-Ile^{4395}	14	Ser^{4406}-Ile^{4419}	14	Short peptide
12	Ser^{4392}-Gln^{5325}	934	Ser^{4406}-Gln^{5339}	934	RdRp
13	Ala^{5326}-Gln^{5923}	598	Ala^{5340}-Gln^{5937}	598	Hel, NTPase
14	Ser^{5924}-Gln^{6447}	524	Ser^{5938}-Gln^{6461}	524	ExoN, NMT
15	Gly^{6448}-Gln^{6970}	343	Gly^{6462}-Gln^{6804}	343	NendoU
16	Ala^{6969}-Arg^{7093}	303	Ala^{6805}-Cys^{7107}	303	OMT

[a]Nsp, nonstructural protein.
[b]ADRP, ADP-ribose 1″-phosphatase; PL2pro, papain-like protease 2; TM, transmembrane domain; 3CLpro, coronavirus Nsp5 protease; ZBD, zinc-binding domain; Hel, helicase; NTPase, nucleoside triphosphatase; ExoN, exoribonuclease; NMT, N7 methyltransferase; NendoU, endoribonyuclease; OMT, 2′O-methyltransferase.

TABLE 4 Comparison of novel MERS-related CoV genomic features and amino acid identities to other lineage C BetaCoVs with complete genome

Feature	% Nucleotide or amino acid identity[f]					
	MERS-CoV	Bat MERS-related CoV[a]	HKU4[b]	HKU5[c]	HeCoV-1[d]	Within 845 and 422[e]
Genome	75.3	74.7–81.2	70.7–70.9	70.6–71.0	70.0–70.4	88.1
ADRP	70.6	69.5–85.1	56.7–57.8	55.9–59.0	66.0–66.5	91.0
3CLpro	89.2–89.9	89.5–95.1	78.8–79.7	83.3–84.0	81.7–82.0	97.7
RdRp	94.4–94.9	95.0–97.0	89.0–89.3	92.2–92.8	88.8–89.0	98.6
Hel, NTPase	94.8–95.2	95.0–98.3	93.1–93.8	94.5–95.3	92.0–92.5	99.5
ExoN	94.5–95.0	94.5–96.9	84.7–85.9	91.2–92.0	89.7–90.3	98.7
NendoU	87.8–88.0	88.0–90.1	75.1–77.4	80.5–81.2	85.4–86.3	97.1
OMT	88.4–88.7	88.4–95.7	84.6	86.1–86.4	85.8–87.7	96.7
S	64.7–64.9	43.3–68.3	65.0–67.8	58.5–60.6	45.9–46.1	76.2
S-RBD[g]	61.9–63.6	34.6–57.6	71.3–72.9	56.7–60.3	41.1	83.8
E	80.5–82.9	82.9–85.4	79.7–72.0	69.5	74.4–75.6	90.2
M	86.2–88.5	88.1–92.7	78.4–80.3	80.7–82.6	80.3–81.2	95.4
N	79.0–79.1	78.8–90.5	73.1–73.8	71.1–71.5	72.2–73.3	85.6
ORF3	41.6–45.0	44.6–57.0	31.9–40.0	37.3–45.8	20.6–23.0	50.0
ORF4a	51.1–52.1	54.3–71.3	44.7	40.9–47.3	43.2	80.9
ORF4b	38.7–43.3	41.3–66.5	30.3–31.6	26.5–34.8	39.6–46.3	58.8
ORF5	66.1–67.0	65.2–78.9	43.1–45.8	55.7–56.6	52.3–53.3	89.0

[a]Sequences accession numbers KJ473821, KX574227, and KC869678.
[b]Sequences accession numbers NC_009019, EF065506, EF065507, EF065508, and KJ473822.
[c]Sequences accession numbers NC_009020, EF065510, EF065511, EF065512, KJ473820.
[d]Sequences accession numbers KC545383 and KC545386.
[e]845 and 422 are abbreviations for BtCoV/li/GD/2013-845 and BtCoV/li/GD/2014-422, respectively.
[f]Nucleotide identities are shown for genome. Amino acid identities are shown for other proteins. Method for nucleotide sequences analysis involved the bootstrap method with 100 replications with Kimura 2-parameter model; for amino acid sequences, the bootstrap method with 100 replications, JTT model, was used.
[g]Counterparts to MERS-CoV S-RBD (accession number AFS88936, aa 367 to 588) through alignment.

2013-845 and BtCoV/li/GD/2014-422 are more similar to that of MERS-related CoV than to those of other lineage C BetaCoVs.

We further investigated the individual genes and predicted protein sequences of the two novel MERS-related CoVs (Table 4). The membrane (M) protein is the most conserved, with an identity of 95.4% between the two novel strains and MERS-CoV. ORF3 is the least conserved, with 50.0% identity between the two novel strains and MERS-CoV. Phylogenetic analysis of polyprotein 1 and the E, M, and N proteins suggests that the two novel strains are more closely related to MERS-related CoV than to other lineage C BetaCoVs (Fig. 2). The predicted RdRp sequences of the two newly detected strains are closely related to MERS-CoV and other MERS-related CoVs, with 94.4 to 97.0% amino acid identities. In contrast, their spike proteins are highly divergent from those of MERS-CoV and other MERS-related CoVs, with 43.3 to 64.9% amino acid identities (Table 4). Phylogenetic analysis suggests that the RdRp genes form an independent cluster closely related to MERS-CoV. In comparison, the S genes form an independent cluster closely related to HKU4 (Fig. 2).

Recombination analysis of novel bat MERS-related CoV *S* gene. We further focused on the spike protein, which is the most divergent among all CoV proteins and responsible for receptor binding and virus entry. The S proteins of BtCoV/li/GD/2013-845 and BtCoV/li/GD/2014-422 are 1,349 and 1,356 residues long, respectively. Their ectodomains include residues 1 to 1293 and 1 to 1298, respectively, while their transmembrane anchors include residues 1294 to 1316 and 1299 to 1321, respectively, and their RBDs include residues 371 to 592 and 372 to 593, respectively. Based on our analysis, at least two recombination events have occurred among the *S* genes of these two viral strains and other lineage C BetaCoVs (Fig. 3A). The recombination event involving the RBD (BtCoV/li/GD/2013-845 breakpoints at nucleotide [nt] 22694/23467; BtCoV/li/GD/2014-422 breakpoints at nt 22732/23306) includes a bootscan value of $<10^{-38}$, indicating significant evidence for this event (Fig. 3B). Unlike the full-length S

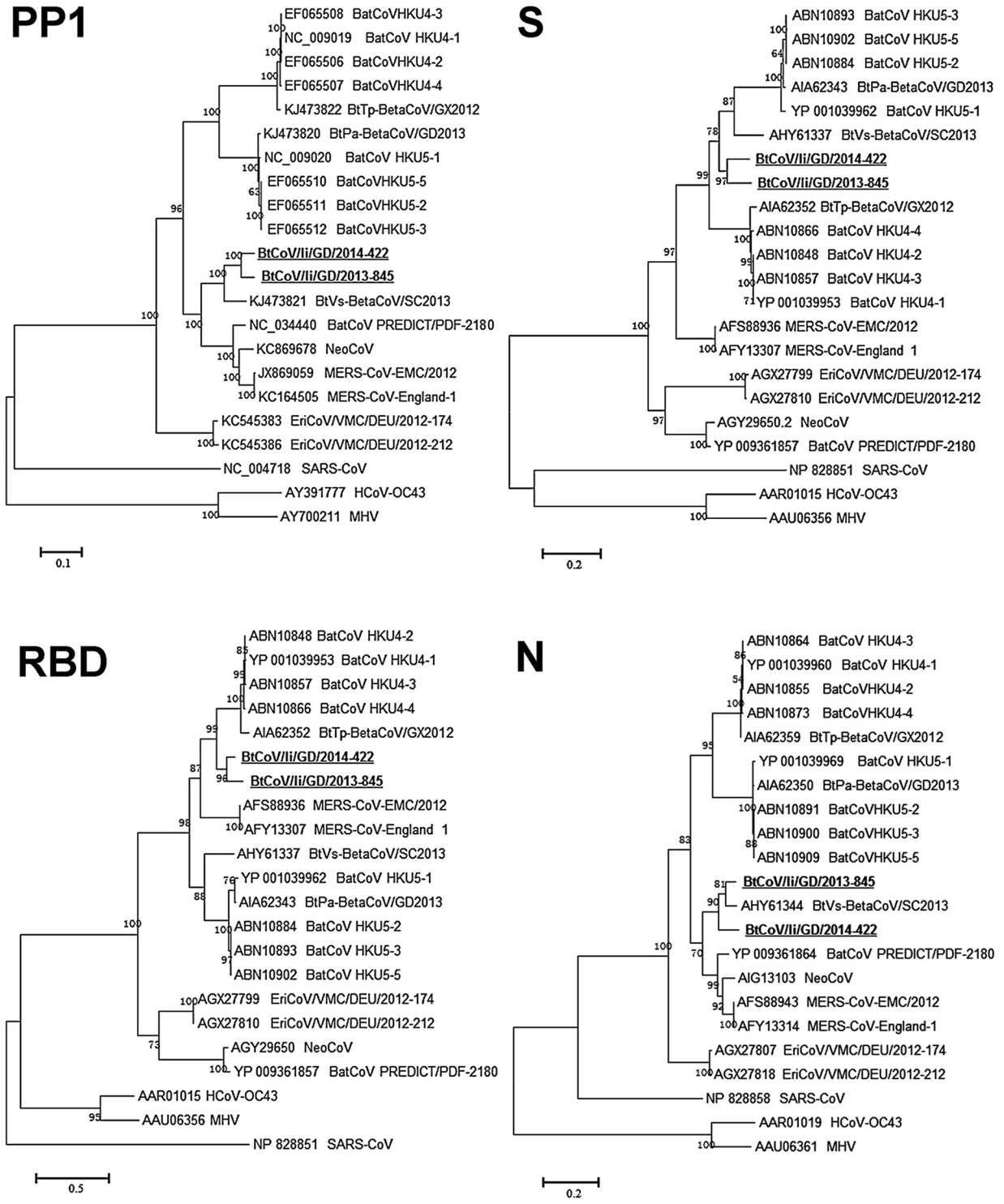


FIG 2 Phylogenetic trees derived from the amino acid sequences of lineage C betacoronaviruses for PP1, S, RBD, and N. Trees were constructed using the maximum-likelihood method, with the JTT model with bootstrap values calculated from 100 trees, and were midpoint rooted. For PP1, S/N, and RBD, scale bars indicate the estimated number of substitutions per 10, 5, or 2 amino acids, respectively. Viruses characterized in this study are shown in boldface.

protein, their spike RBDs are more similar to the HKU4 RBD than to the MERS-CoV or HKU5 RBDs. Therefore, recombination may have occurred between the ancestor of the two novel MERS-related CoVs and HKU4. Moreover, the bootscan results also showed that the S1 subunits of the two viruses were derived from the ancestor of HKU4, while

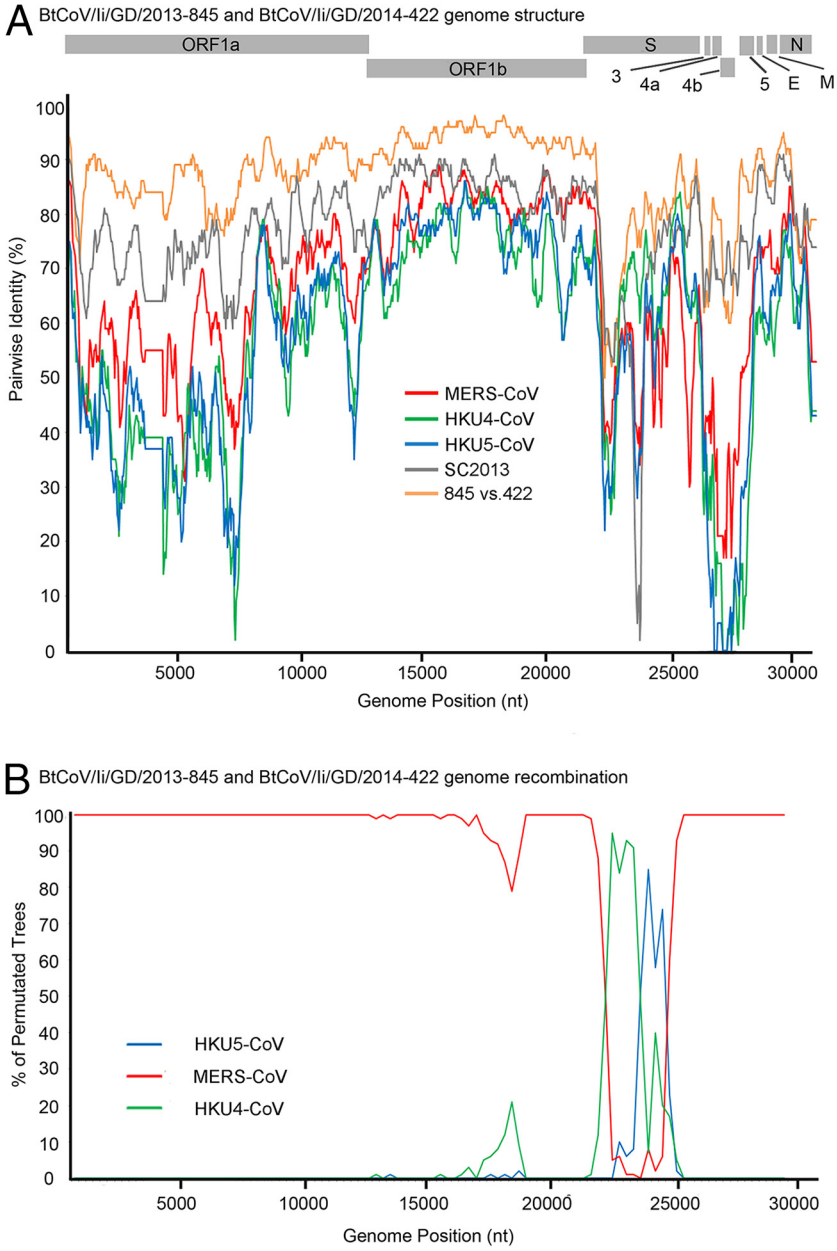


FIG 3 Genomic sequence identities between novel MERS-related CoVs and other lineage C betacoronaviruses (A) and evidence for recombination in BtCoV/li/GD/2013-845 and BtCoV/li/GD/2014-422 (B). Similarity (window of 400 nt, step size of 40 nt) and recombination (window of 1,500 nt, step size of 300 nt) plots were generated using Simplot (v3.5.7) with default settings (25).

their S2 subunits were derived from the ancestor of HKU5 (Fig. 3). The complicated recombination history between these lineage C BetaCoVs suggests frequent gene transfers, especially of the *S* gene, among different CoVs, which may be responsible for the cross-species transmission of these CoVs.

Binding assay of BtCoV/li/GD/2014-422 spike protein to DPP4. Thirteen residues in the MERS-CoV S-RBD were shown to directly interact with human DPP4 as part of the RBD-human DPP4 complex. The S-RBDs of the novel MERS-related CoVs are identical to that of MERS-CoV at six residues, making them more similar to MERS-CoV than any other known lineage C BetaCoV (Fig. 4A).

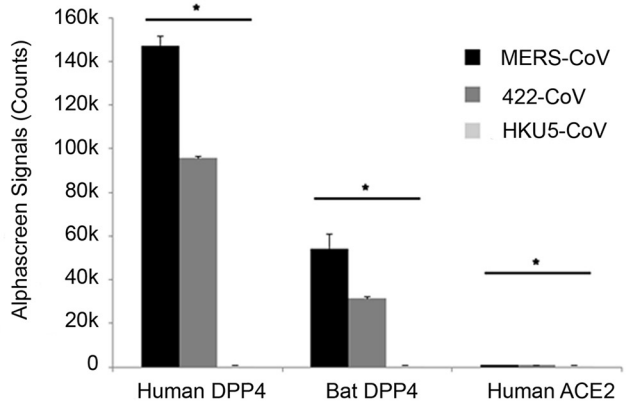
To understand the receptor usage of the newly detected novel MERS-related CoVs, we selected the BtCoV/li/GD/2014-422 spike protein for detailed functional studies, as

A

```
                467
MERS             FNYKQSFSNP TCLILATVPH NLTT---ITK PLKYSYINKC SRLLSDD-RT  512
422              YNYKQSFANP TCRIFATAPA NLT----ITK PSSYSFISKC SRLTGDNSHI  516
845              FNYKQSFANP TCRIFATAPA NLT----ISK PSSYSYISKC SRLTGDNQHI  517
HKU4             YNYKQSFANP TCRVMASVLA NVT----ITK PHAYGYISKC SRLTGANQDV  517
SC2013           FNYKQDFSNP TCRILATVPA NLSASGLLPK PSNYVWLSEC YQNSFTG---  488
Neo              FNYNQDYSNP SCRIHSKVNS SIG----ISY AGAYSYITNC NYGATNK---  512
PDF-2180         FNYNQDYSNP SCRIHSKVNS SVG----ISY SGLYSYITNC NYGGFNK---  513
HKU5             FNYKQDFSNP TCRVLATVPQ NLTT---ITK PSNYAYLTEC YKTSAYG---  518
                 :**:*.::** :* : :..   .:     :.  .  * ::.:*          .

                513
MERS             EVPQLVNANQ YSPCVSIVPS TVWEDGDYYR KQLSPLEGGG WLVASGSTVA  562
422              ETPIVINPGE YSICKNFAPN GFSQDGDYFT RQLSQLEGGG ILVGVGSVTP  566
845              ETPITINPGE YSICRGFAPN GLSEDGQVFT RQLSDYEGGG TLVGVGNTVP  567
HKU4             ETPLYINPGE YSICRDFSPG GFSEDGQVFK RTLTQFEGGG LLIGVGTRVP  567
SC2013           KNFQYVKAGQ YTPCLGLAAN GFEKSYQTHR DPV------S KLAVTGVVTP  532
Neo              DDVVKPGGRA SQQCITGALN S-PTTGQLWA YNF-----GG VPYRVSRLTY  556
PDF-2180         DDVVKPGGRA SQPCVTGALN S-PTNGQVWS FNF-----GG VPYRTSRLTY  557
HKU5             KNYLYNAPGA YTPCLSLASR GFSTKYQSHS D--------G ELTTTGYIYP  561
                       *             :             .      .
```

B



C

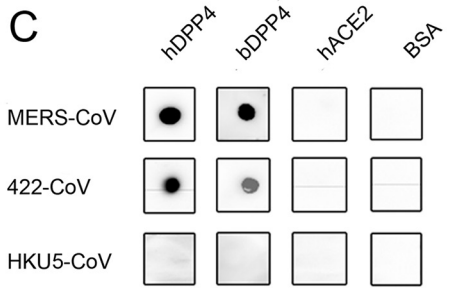


FIG 4 BtCoV/li/GD/2014-422 RBD analysis and DPP4-binding assay. (A) Sequence alignment of the partial S1 domains (including all positions with direct interactions with human DPP4) of selected lineage C betacoronaviruses. Asterisks indicate positions with fully conserved residues. Colons indicate positions with strongly conserved residues. Periods indicate positions with weakly conserved residues. Positions that have direct interactions with human DPP4 according to data from MERS-CoV are in gray bars. Residues identical to corresponding MERS-CoV residues are in red. (B) AlphaScreen assay showing the direct binding interactions between the coronavirus spike RBD and hDPP4 or bDPP4. Binding affinity was characterized as AlphaScreen counts. Error bars indicate standard errors of the means (SEM) (*, $P < 0.05$ by two-tailed t test; $n = 3$). (C) Dot blot hybridization assay showing the direct binding interactions between the coronavirus spike RBD and hDPP4 or bDPP4. His_8-tagged hDPP4, bDPP4, or hACE2 was dotted and then incubated with each of the Fc-tagged coronavirus RBDs, followed by anti-IgG4 monoclonal antibody detection.

its spike RBD is slightly more identical to MERS-CoV than BtCoV/li/GD/2013-845 (64.9% versus 64.7%). Using four alternative approaches, we investigated the binding interactions between the BtCoV/li/GD/2014-422 spike protein and DPP4. First, an AlphaScreen protein-protein binding assay showed that the BtCoV/li/GD/2014-422 RBD binds both

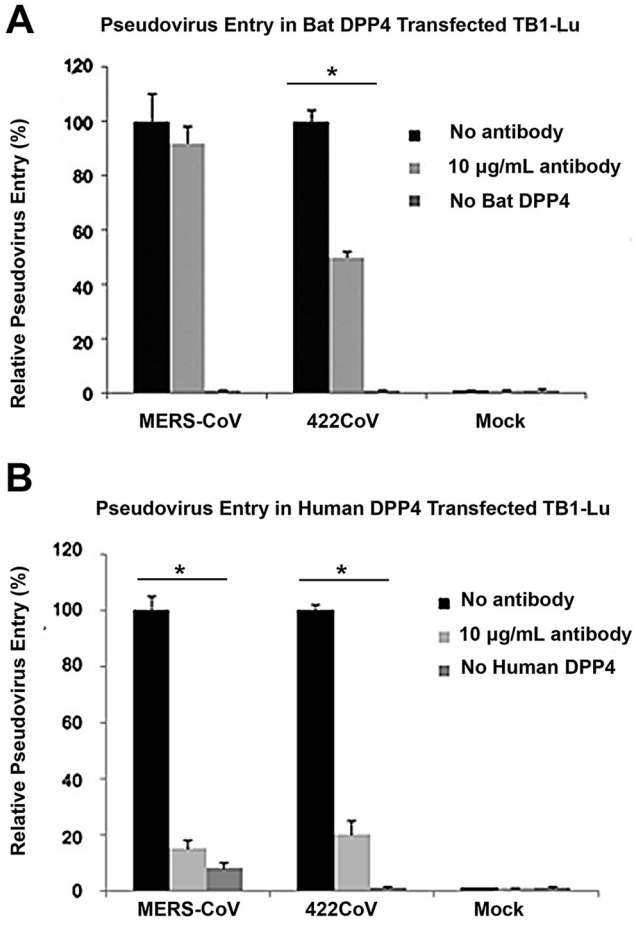


FIG 5 BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry and inhibition assay. Anti-hDPP4 antibodies competitively block interactions between BtCoV/li/GD/2014-422 spike and hDPP4 but not bat DPP4. Error bars indicate SEM (*, $P < 0.05$ by two-tailed t test; $n = 4$). (A) Anti-hDPP4 antibodies did not neutralize BtCoV/li/GD/2014-422 spike-mediated and MERS-CoV spike-mediated pseudovirus entry into bDPP4-expressing cells. (B) Anti-hDPP4 antibodies strongly neutralized BtCoV/li/GD/2014-422 spike-mediated and MERS-CoV spike-mediated pseudovirus entry into hDPP4-expressing Tb1-Lu cells.

human and bat DPP4 (Fig. 4B). Both MERS-CoV and HKU4 RBDs, used as controls, bound to human and bat DPP4, but the HKU5 RBD did not. Second, dot blot hybridization assays confirmed the above-described results (Fig. 4C). Third, we established a BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry assay. To this end, retroviruses pseudotyped with BtCoV/li/GD/2014-422 spike were incubated with bat cells exogenously expressing or not expressing DPP4. The results showed that bat cells exogenously expressing human or bat DPP4 support entry of BtCoV/li/GD/2014-422 pseudoviruses, while bat cells not expressing DPP4 do not (Fig. 5A and B). Lastly, the entry of BtCoV/li/GD/2014-422 pseudoviruses into hDPP4-expressing bat cells was inhibited by anti-DPP4 antibodies (Fig. 5B). Taken together, these results reveal that the BtCoV/li/GD/2014-422 spike protein recognizes the human and bat DPP4 receptors via the RBD.

Cross-neutralization assay of anti-MERS-spike antibodies to novel bat MERS-related CoV. We used BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry assay to test the cross-neutralization reactivity of MERS-CoV antibodies. Neither antiserum nor monoclonal antibody could neutralize the entry of BtCoV/li/GD/2014-422 pseudoviruses (Fig. 6), suggesting no cross-neutralization reaction between BtCoV/li/GD/2014-422 and MERS-CoV.

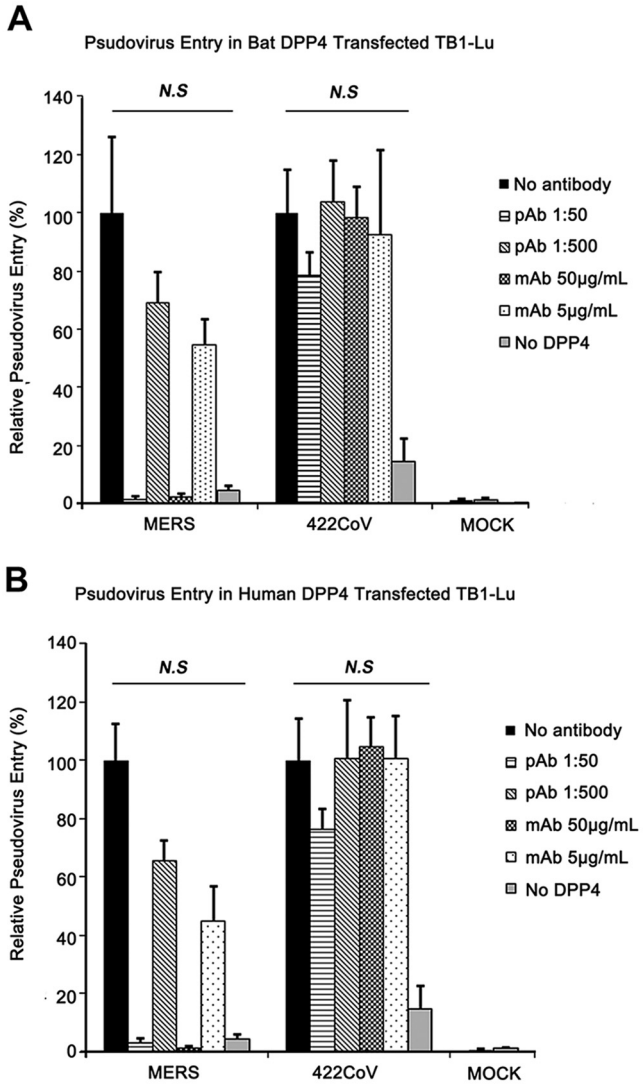


FIG 6 Cross-neutralization assay of MERS-CoV antibodies to BtCoV/li/GD/2014-422 spike-mediated pseudovirus. Anti-MERS-spike antibodies could not inhibit the entry of BtCoV/li/GD/2014-422 spiked pseudovirus. Error bars indicate SEM (*N.S* indicates a *P* value of >0.05 by two-tailed *t* test; $n = 4$). (A) Anti-MERS-spike antibodies neutralize MERS-CoV spike-mediated pseudovirus entry into bDPP4-expressing cells but not BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry. (B) Anti-MERS-spike antibodies neutralize MERS-CoV spike-mediated pseudovirus entry into hDPP4-expressing cells but not BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry.

DISCUSSION

In this study, we identified and characterized two novel MERS-related CoVs from great evening bats in south China. Lineage C BetaCoVs are widely distributed across old-world continents and North America (6–15). Along with the newly identified BtCoV/li/GD/2013-845 and BtCoV/li/GD/2014-422 strains, there are three different clusters of lineage C BetaCoVs that are closely related to MERS-CoV, including HKU4 and HKU5 (38). Our study shows that MERS-related CoVs are highly associated with vespertilionid bats, a widely distributed and highly diverse family within the *Microchioptera* suborder (39). These viruses are not present in other bat families at the same or nearby sampling locations. However, MERS-related CoVs have the widest tropism among lineage C BetaCoV in bats (9). Complete genomes were acquired from *Hypsugo* (40), *Ia*, *Neoromicia* (7), *Pipistrellus* (9), and *Vespertilio* (8) genera. In comparison, CoVs in the HKU4 and HKU5 clusters are highly prevalent among *Tylonycteris* and *Pipistrellus* bats, respectively (5, 26). Moreover, we detected different clusters of CoVs at the same or

nearby sampling locations within the same bat species. For example, viruses in the MERS-CoV and HKU5 clusters were both detected in *Pipistrellus* bats at the same or nearby sampling locations in south China. These complex combinations of bat and lineage C BetaCoV species provide opportunities for these viruses to undergo interspecies transmissions, coinfections, and genetic recombination.

Furthermore, our study provides evidence for genetic recombination among bat CoVs of the lineage C BetaCoVs. More specifically, evidence for at least two recombination events was observed in the *S* genes of the two novel MERS-related CoVs, including a recombination event between viruses in the MERS-related CoV and HKU4 clusters that allowed the two viruses to acquire a DPP4-binding S1 subunit from HKU4 (17) and another recombination between viruses in the MERS-related CoV and HKU5 clusters that allowed the two viruses to acquire an S2 subunit from HKU5. These recombination events likely occurred in bat hosts because of the above-mentioned complex combinations of bat and lineage C BetaCoV species.

To date, bat viruses in the MERS-related CoV cluster that do not use DPP4 as their receptor have all been identified in Africa (9), whereas one that uses DPP4 as its receptor has been detected in Asia (17, 26, 40). The current study provides new evidence that bat MERS-related CoV identified from *I. io* uses DPP4 as its host receptor, demonstrated by using a combination of experimental approaches, including AlphaScreen, dot blot hybridization, pseudovirus entry, and antibody neutralization. It should be noted that receptor analyses in previous reports and the current study were all based on protein-protein interaction and pseudovirus techniques due to the lack of virus isolation. To fully understand entry and replication efficiency and pathogenesis of these MERS-related CoVs *in vivo* and *in vitro*, study of virus isolation or construction of infectious clones will be essential in the future.

Despite having receptor usage similar to that of MERS-CoV, anti-MERS-spike polyclonal and monoclonal antibodies could not block the bat MERS-related CoV spike-mediated pseudovirus entry. These results indicate that the preventive vaccine against MERS-CoV is not applicable for these bat viruses. Considering the diversity of MERS-related CoVs in natural reservoirs and potential spillover, the strategy and alternative vaccine for prevention and control of diseases potentially caused by these bat viruses should be prepared in advance.

In summary, the current study indicates that there are more and highly diverse MERS-related CoVs existing in natural bat hosts, and some of them have the potential to spill over by sharing the same host receptor as MERS-CoV. In order to understand the evolutionary history of MERS-CoV and to prevent future emerging infectious diseases, it is critical to take extensive and long-term surveillances for those that are more closely related to MERS-CoV than those currently discovered, with particular focus on Africa and the Middle East.

MATERIALS AND METHODS

Ethics statement. All sampling procedures were performed by veterinarians with approval from the Animal Ethics Committee of the Wuhan Institute of Virology (WIVH5210201). The study was conducted in accordance with the Guide for the Care and Use of Wild Mammals in Research of the People's Republic of China.

Sample collection. The animal surveillance program was conducted from the summer of 2012 to the autumn of 2015 in three southern provinces (Guangdong, Guangxi, and Sichuan) of China. Alimentary specimens were collected using procedures described previously (18). All specimens were collected using disposable swabs and protective gloves that were changed between samples for preventing cross contamination and were immediately placed in viral transport medium. All samples were kept in liquid nitrogen or dry ice before transportation to the laboratory for RNA extraction.

RNA extraction. Viral RNA was extracted from the alimentary specimens using a High Pure viral RNA kit (Roche Diagnostics GmbH, Mannheim, Germany). RNA was extracted from 200 μl of sample, used as the template for RT-PCR, eluted in 50 μl of RNase-free water provided by the viral RNA extraction kit, and stored at −80°C.

RT-PCR for CoV detection. Two pairs of previously described conserved nested primers (1st round, CoV-FWD3, 5′-GGTTGGGAYTAYCCHAARTGTGA-3′; CoV-RVS3, 5′-CCATCATCASWYRAATCATCATA-3′; 2nd round, CoV-FWD4/Bat, 5′-GAYTAYCCHAARTGTGAYAGAGC-3′; CoV-RVS3, 5′-CCATCATCASWYRAATCATCATA-3′) for amplifying a 440-bp fragment of the CoV *RdRp* gene were used for initial CoV screening (19). The first round of screening PCR was performed with a Superscript III one-step RT-PCR with Platinum *Taq*

kit (Invitrogen, San Diego, CA). The second round of screening PCR was performed with a Platinum *Taq* DNA polymerase kit (Invitrogen, San Diego, CA).

Host species identification. Host species were identified first morphologically and then by amplifying and sequencing the NADH dehydrogenase subunit 1 (*ND1*) gene and/or the cytochrome *b* (*CytB*) gene (20, 21). The sequence data, excluding primer sequences, were assembled, and BLASTN (default parameters) was used to identify host species based on the most closely related sequences with the highest query coverage and a minimum identity of 95%.

Sequencing of full-length genomes. Two novel MERS-related CoV strains, named BtCoV/li/GD/2013-845 and BtCoV/li/GD/2014-422, were selected for amplification of their complete genomes through gene walking with primers based on the sequences of known MERS-related CoVs as previously described (primers provided upon request) (22). The 5′ and 3′ ends of the viral genomes were confirmed by rapid amplification of cDNA ends (RACE) using a 5′/3′ RACE kit (Clontech, Heidelberg, Germany). Sequences were assembled and manually edited to produce final sequences of the viral genome with Lasergene SeqMan (v7; DNASTAR, Madison, WI).

Genome analysis. Putative open reading frames (ORFs) were predicted using NCBI's ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/), with a minimal ORF length of 150 nt, and then inspected manually. The nucleotide sequences of the genomes and the deduced amino acid sequences of the ORFs were preliminarily compared to those of other CoV full-length genome sequences available from EMBOSS Needle (http://www.ebi.ac.uk). Sequencing alignment were performed in MEGA (v6.0) with the ClustalW method (23, 24). The similarity plot (window size, 400 bp; step size, 40 bp) was generated with Simplot (v3.5.7) (25). Pairwise comparison of the complete genome and putative proteins was calculated using MEGA 6.0 (24) with a previously described method (26). The TRSs and nonstructural protein (nsp1 to nsp16) prediction was done according to described methods (27). Phylogenetic trees were constructed using the maximum likelihood method set as previously described (5), with 100 bootstrap replications in MEGA (v6.0) (24). Protein family analysis was performed using PFAM (28) and InterProScan (29). The genome recombination bootscan plot (window size, 1,500 bp; step size, 300 bp) was first scanned using the Recombination Detection Program (RDP; v4.94 Beta) (30) and then generated using Simplot (v3.5.7) (25).

Estimation of synonymous and nonsynonymous substitution rates. The corrected Nei-Gojobori method (Jukes-Cantor) in the KaKs Calculator (v2.0) (31) and the Ka/Ks calculation tool were used to calculate the number of synonymous substitutions per synonymous site (K_s) and the number of nonsynonymous substitutions per nonsynonymous site (K_a) for each coding region between each pair of strains.

Cell lines. The HEK293T (human embryonic kidney) and Tb1-Lu (*Tadarida brasiliensis* bat lung) cell lines were obtained from the ATCC (www.atcc.org). These cell lines were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 1% penicillin-streptomycin (Life Technologies Inc., Grand Island, NY).

Protein expression and purification. The RBDs of the spike proteins of MERS-CoV (residues 367 to 588; GenBank accession no. AFS88936), HKU4 (residues 372 to 593; GenBank accession no. ABN10839), BtCoV/li/GD/2014-422 (residues 357 to 592), and HKU5 (residues 375 to 586; GenBank accession no. ABN10875) were expressed and purified as previously described (32, 33). Briefly, CoV RBDs containing an N-terminal honey bee melittin signal peptide and a C-terminal human IgG4 Fc tag were expressed in insect cells using the Bac-to-Bac expression system (Life Technologies Inc.), secreted into cell culture medium, and purified sequentially on HiTrap protein G HP columns and Superdex 200 gel filtration columns (GE Healthcare, Pittsburgh, PA).

The ectodomains of human DPP4 (residues 39 to 766; GenBank accession no. NP_001926) and *Pipistrellus* bat DPP4 (residues 36 to 760; GenBank accession no. KC249974), containing an N-terminal human CD5 signal peptide and an N-terminal His tag, were expressed and purified using the same procedure as that for CoV RBDs, except they were purified sequentially on HiTrap nickel-chelating HP columns and Superdex 200 gel filtration columns (GE Healthcare) (17).

AlphaScreen protein-protein binding assay. Interactions between CoV RBDs and human or bat DPP4 were measured using AlphaScreen, as previously described (34, 35). Briefly, each of the Fc-tagged CoV RBDs at a final concentration of 3 nM was mixed with 100 nM His_8-tagged hDPP4 or bDPP4 in half of an AreaPlate (PerkinElmer, Waltham, MA) for 1 h at room temperature as recommended by the manufacturer. AlphaScreen nickel chelate donor beads and AlphaScreen protein A acceptor beads (PerkinElmer) were added to the mixture at a final concentration of 5 μg/ml. The mixtures were incubated at room temperature for 1 h and protected from light. The assay plates were read using an EnSpire plate reader (PerkinElmer).

Dot blot hybridization assay. Dot blot hybridization assays were performed as previously described (36). Briefly, 2 μl of each 50 μg/ml Fc-tagged CoV RBD was dotted onto a nitrocellulose membrane. The membranes were dried completely and blocked with 5% skim milk at 37°C for 1 h. A sample of 20 μg/ml His_8-tagged hDPP4 or bDPP4 was preincubated at 37°C for 1 h, added to the membrane, and incubated at 4°C overnight. The membrane was washed five times with phosphate-buffered saline with 0.1% Tween 20 (PBST), incubated with anti-His_6 mouse monoclonal IgG1 horseradish peroxidase-conjugated antibody (Santa Cruz Biotechnology, Santa Cruz, CA) at 37°C for 2 h, and washed five times with PBST. Finally, the bound proteins were detected using ECL plus (GE Healthcare).

BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry into human and bat cells. Retroviruses pseudotyped with MERS-CoV spike, BtCoV/li/GD/2014-422 spike, or no spike (mock) were used to infect Tb1Lu cells *trans*-expressing hDPP4, bDPP4, or no DPP4 in 96-well plates. Measurements of infection and luciferase activity were conducted as described previously (17).

Neutralization of BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry by anti-hDPP4 antibodies. Tb1Lu cells expressing hDPP4, bDPP4, or no DPP4 were preincubated with 0 or 10 μg/ml goat anti-hDPP4 polyclonal antibody (R&D Systems) at 37°C for 1 h and then infected by equal amounts of MERS-CoV or BtCoV/li/GD/2014-422 spike-pseudotyped retroviruses. Measurements of infection and luciferase activity were conducted as described previously (17).

Neutralization of BtCoV/li/GD/2014-422 spike-mediated pseudovirus entry by anti-MERS-CoV-spike antibodies. The mouse serum and monoclonal antibody against MERS-CoV-spike were kindly provided by Lanying Du at New York Blood Center. MERS-CoV or BtCoV/li/GD/2014-422 spike-pseudotyped retrovirus was preincubated with mouse serum or monoclonal antibody against MERS-CoV-spike, as previously described (37), for 10 min at 37°C. The mixture was used to infect Tb1Lu cells expressing hDPP4, bDPP4, or no DPP4. Measurements of infection and luciferase activity were conducted as described previously (17).

Accession number(s). The nucleotide sequences of the complete genomes of the lineage C BetaCoVs obtained in this study have been submitted to GenBank under accession numbers MG021451 and MG021452. The GenBank accession numbers of the reference complete genomes are the following: MERS-CoV, JX869059, KC164505, and KJ713298; bat MERS-related CoV, KJ473821, KC869678, and NC_034440; HKU4-CoV, NC_009019, EF065506–EF065508, and KJ473822; HKU5-CoV, NC_009020, EF065510–EF065512, and KJ473820; hedgehog coronavirus, KC545383 and KC545386; HCoV-OC43, AY391777; murine hepatitis virus, AY700211; severe acute respiratory syndrome virus, NC_004718.

ACKNOWLEDGMENTS

We thank Lanying Du for the mouse antiserum and monoclonal antibody against MERS-CoV-spike. This study was jointly funded by the Scientific and Technological Basis Special Project (2013FY113500) from the Ministry of Science and Technology of the People's Republic of China to Z.-L.S., a USAID Emerging Pandemic Threats PREDICT project grant (Cooperative Agreement no. AID-OAA-A-14-00102), NIH grants (R01AI110964 to Z.-L.S.; R01AI089728 and R01AI110700 to F.L.), and the National Natural Science Foundation of China (31727901) to Z.-L.S.

REFERENCES

1. Masters PS, Perlman S. 2013. Coronaviridae, p 825–858. *In* Knipe DM, Howley PM (ed), Fields virology, 6th ed, vol 2. Lippincott Williams & Wilkins, Philadelphia, PA.
2. Lai MMC. 1996. Recombination in large RNA viruses: coronaviruses. Semin Virol 7:381–388. https://doi.org/10.1006/smvy.1996.0046.
3. Li F. 2016. Structure, function, and evolution of coronavirus spike proteins. Annu Rev Virol 3:237–261. https://doi.org/10.1146/annurev-virology-110615-042301.
4. de Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, Fouchier RAM, Galiano M, Gorbalenya AE, Memish ZA, Perlman S, Poon LLM, Snijder EJ, Stephens GM, Woo PCY, Zaki AM, Zambon M, Ziebuhr J. 2013. Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the coronavirus study group. J Virol 87:7790–7792. https://doi.org/10.1128/JVI.01244-13.
5. Lau SKP, Li KSM, Tsang AKL, Lam CSF, Ahmed S, Chen HL, Chan KH, Woo PCY, 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.
6. Ithete NL, Stoffberg S, Corman VM, Cottontail VM, Richards LR, Schoeman MC, Drosten C, Drexler JF, Preiser W. 2013. Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. Emerg Infect Dis 19:1697–1699. https://doi.org/10.3201/eid1910.130946.
7. Corman VM, Ithete NL, Richards LR, Schoeman MC, Preiser W, Drosten C, Drexler JF. 2014. Rooting the phylogenetic tree of Middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. J Virol 88:11297–11303. https://doi.org/10.1128/JVI.01498-14.
8. Yang L, Wu ZQ, Ren XW, Yang F, Zhang JP, He GM, Dong J, Sun LL, Zhu YF, Zhang SY, Jin Q. 2014. MERS-related betacoronavirus in Vespertilio superans bats, China. Emerg Infect Dis 20:1260–1262. https://doi.org/10.3201/eid2007.140318.
9. Anthony SJ, Gilardi K, Menachery VD, Goldstein T, Ssebide B, Mbabazi R, Navarrete-Macias I, Liang E, Wells H, Hicks A, Petrosov A, Byarugaba DK, Debbink K, Dinnon KH, Scobey T, Randell SH, Yount BL, Cranfield M, Johnson CK, Baric R, Lipkin WI, Mazet JK. 2017. Further evidence for bats as the evolutionary source of Middle East respiratory syndrome coronavirus. mBio 8:e00373-17. https://doi.org/10.1128/mBio.00373-17.
10. Lelli D, Papetti A, Sabelli C, Rosti E, Moreno A, Boniotti MB. 2013. Detection of coronaviruses in bats of various species in Italy. Viruses 5:2679–2689. https://doi.org/10.3390/v5112679.
11. Wacharapluesadee S, Sintunawa C, Kaewpom T, Khongnomnan K, Olival KJ, Epstein JH, Rodpan A, Sangsri P, Intarut N, Chindamporn A, Suksawa K, Hemachudha T. 2013. Group C betacoronavirus in bat guano fertilizer, Thailand. Emerg Infect Dis 19:1349–1351. https://doi.org/10.3201/eid1908.130119.
12. Reusken C, Lina PHC, Pielaat A, de Vries A, Dam-Deisz C, Adema J, Drexler JF, Drosten C, Kooi EA. 2010. Circulation of group 2 coronaviruses in a bat species common to urban areas in Western Europe. Vector Borne Zoonotic Dis 10:785–791. https://doi.org/10.1089/vbz.2009.0173.
13. Annan A, Baldwin HJ, Corman VM, Klose SM, Owusu M, Nkrumah EE, Badu EK, Anti P, Agbenyega O, Meyer B, Oppong S, Sarkodie YA, Kalko EKV, Lina PHC, Godlevska EV, Reusken C, Seebens A, Gloza-Rausch F, Vallo P, Tschapka M, Drosten C, Drexler JF. 2013. Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. Emerg Infect Dis 19:456–459. https://doi.org/10.3201/eid1903.121503.
14. Anthony SJ, Ojeda-Flores R, Rico-Chavez O, Navarrete-Macias I, Zambrana-Torrelio CM, Rostal MK, Epstein JH, Tipps T, Liang E, Sanchez-Leon M, Sotomayor-Bonilla J, Aguirre AA, Avila-Flores R, Medellin RA, Goldstein T, Suzan G, Daszak P, Lipkin WI. 2013. Coronaviruses in bats from Mexico. J Gen Virol 94:1028–1038. https://doi.org/10.1099/vir.0.049759-0.
15. Falcon A, Vazquez-Moron S, Casas I, Aznar C, Ruiz G, Pozo F, Perez-Brena P, Juste J, Ibanez C, Garin I, Aihartza J, Echevarria JE. 2011. Detection of alpha and betacoronaviruses in multiple Iberian bat species. Arch Virol 156:1883–1890. https://doi.org/10.1007/s00705-011-1057-1.
16. 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. https://doi.org/10.1038/nature12005.
17. 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. https://doi.org/10.1073/pnas.1405889111.
18. Li WD, Shi ZL, Yu M, Ren WZ, Smith C, Epstein JH, Wang HZ, Crameri G, Hu ZH, Zhang HJ, Zhang JH, McEachern J, Field H, Daszak P, Eaton BT, Zhang SY, Wang LF. 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science 310:676–679. https://doi.org/10.1126/science.1118391.
19. de Souza Luna K, 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.
20. Irwin DM, Kocher TD, Wilson AC. 1991. Evolution of the cytochrome b gene of mammals. J Mol Evol 32:128–144. https://doi.org/10.1007/BF02515385.
21. Mayer F, von Helversen O. 2001. Cryptic diversity in European bats. Proc Biol Sci 268:1825–1832. https://doi.org/10.1098/rspb.2001.1744.
22. 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.
23. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. https://doi.org/10.1093/bioinformatics/btm404.
24. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.
25. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 73:152–160.
26. Woo PCY, Wang M, Lau SKP, Xu HF, Poon RWS, Guo RT, Wong BHL, Gao K, Tsoi HW, Huang Y, Li KSM, Lam CSF, Chan KH, Zheng BJ, Yuen KY. 2007. Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. J Virol 81:1574–1585. https://doi.org/10.1128/JVI.02182-06.
27. van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, Osterhaus A, Haagmans BL, Gorbalenya AE, Snijder EJ, Fouchier RAM. 2012. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. mBio 3:e00473-12. https://doi.org/10.1128/mBio.00473-12.
28. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.
29. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SC, Wu CH, Xenarios I, Yeh LS, Young SY, Mitchell AL. 2017. InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 4:D190–D199. https://doi.org/10.1093/nar/gkw1107.
30. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol 1:vev003. https://doi.org/10.1093/ve/vev003.
31. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. 2010. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics 8:77–80. https://doi.org/10.1016/S1672-0229(10)60008-3.
32. Chen YQ, Rajashankar KR, Yang Y, Agnihothram SS, Liu C, Lin YL, Baric RS, Li F. 2013. Crystal structure of the receptor-binding domain from newly emerged Middle East respiratory syndrome coronavirus. J Virol 87:10777–10783.
33. Li F, Li WH, Farzan M, Harrison SC. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309:1864–1868.
34. Du LY, Zhao GY, Yang Y, Qiu HJ, Wang LL, Kou ZH, Tao XR, Yu H, Sun SH, Tseng CTK, Jiang SB, Li F, Zhou YS. 2014. A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in Middle East respiratory syndrome coronavirus spike protein. J Virol 88:7045–7053.
35. Cassel JA, Blass BE, Reitz AB, Pawlyk AC. 2010. Development of a novel nonradiometric assay for nucleic acid binding to TDP-43 suitable for high-throughput screening using AlphaScreen technology. J Biomol Screen 15:1099–1106.
36. Peng GQ, Sun DW, Rajashankar KR, Qian ZH, Holmes KV, Li F. 2011. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. Proc Natl Acad Sci U S A 108:10696–10701.
37. Du LY, Tai WB, Yang Y, Zhao GY, Zhu Q, Sun SH, Liu C, Tao XR, Tseng CTK, Perlman S, Jiang SB, Zhou YS, Li F. 2016. Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. Nat Commun 7:13473.
38. Zumla A, Hui DS, Perlman S. 2015. Middle East respiratory syndrome. Lancet 386:995–1007.
39. Thabah A, Li G, Wang YN, Liang B, Hu KL, Zhang SY, Jones G. 2007. Diet, echolocation calls, and phylogenetic affinities of the great evening bat (Ia IO; Vespertilionidae): another carnivorous bat. J Mammal 88:728–735.
40. Lau SKP, Zhang L, Luk HKH, Xiong L, Peng X, Li KSM, He X, Zhao PS, Fan RYY, Wong ACP, Ahmed SS, Cai JP, Chan JFW, Sun Y, Jin D, Chen H, Lau TCK, Kok RKH, Li W, Yuen KY, Woo PCY. 16 January 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 https://doi.org/10.1093/infdis/jiy018.