AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

# Diversity of Dromedary Camel Coronavirus HKU23 in African Camels Revealed Multiple Recombination Events among Closely Related Betacoronaviruses of the Subgenus Embecovirus

**Ray T. Y. So,[a] Daniel K. W. Chu,[a] Eve Miguel,[c] Ranawaka A. P. M. Perera,[a] Jamiu O. Oladipo,[a,d] Ouafaa Fassi-Fihri,[e] Gelagay Aylet,[f] Ronald L. W. Ko,[a] Ziqi Zhou,[a] Mo-Sheung Cheng,[a] Sulyman A. Kuranga,[d] François L. Roger,[b,g] Veronique Chevalier,[b,h] Richard J. Webby,[i] Patrick C. Y. Woo,[j] Leo L. M. Poon,[a] Malik Peiris[a]**

[a]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, Republic of China

[b]Animal, Santé, Territoires, Risques et Ecosystèmes, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Institut National de la Recherche Agronomique, Université de Montpellier, Montpellier, France

[c]MIVEGEC Maladies Infectieuses et Vecteurs: Ecologie, Génétique, Evolution et Contrôle, IRD L'Institut de Recherche pour le Développement, CNRS Centre National de Recherche Scientifique, Universitè de Montpellier, Montpellier, France

[d]Department of Surgery, Faculty of Clinical Sciences, University of Ilorin, Ilorin, Nigeria

[e]Institut Agronomique et Vétérinaire, Hassan II Université, Rabat, Morocco

[f]Pan African Veterinary Center of the African Union (AU-PANVAC), Debre Zeit, Ethiopia

[g]Kasetsart University, Bangkok, Thailand

[h]Institut Pasteur du Cambodge, Phnom Penh, Cambodia

[i]St. Jude Children's Research Hospital, Memphis, Tennessee, USA

[j]Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, Republic of China

**ABSTRACT** Genetic recombination has frequently been observed in coronaviruses. Here, we sequenced multiple complete genomes of dromedary camel coronavirus HKU23 (DcCoV-HKU23) from Nigeria, Morocco, and Ethiopia and identified several genomic positions indicative of cross-species virus recombination events among other betacoronaviruses of the subgenus Embecovirus (clade A beta-CoVs). Recombinant fragments of a rabbit coronavirus (RbCoV-HKU14) were identified at the hemagglutinin esterase gene position. Homolog fragments of a rodent CoV were also observed at 8.9-kDa open reading frame 4a at the 3′ end of the spike gene. The patterns of recombination differed geographically across the African region, highlighting a mosaic structure of DcCoV-HKU23 genomes circulating in dromedaries. Our results highlighted active recombination of coronaviruses circulating in dromedaries and are also relevant to the emergence and evolution of other betacoronaviruses, including Middle East respiratory syndrome coronavirus (MERS-CoV).

**IMPORTANCE** Genetic recombination is often demonstrated in coronaviruses and can result in host range expansion or alteration in tissue tropism. Here, we showed interspecies events of recombination of an endemic dromedary camel coronavirus, HKU23, with other clade A betacoronaviruses. Our results supported the possibility that the zoonotic pathogen MERS-CoV, which also cocirculates in the same camel species, may have undergone similar recombination events facilitating its emergence or may do so in its future evolution.

**KEYWORDS** coronaviruses, betacoronaviruses, recombination, phylogeny, dromedary camels, evolution

Emerging infectious disease outbreaks usually arise by interspecies jumps of viruses between animal species, sometimes including humans. Coronaviruses (CoVs) have repeatedly made species jumps between animal species (e.g., swine acute diarrhea

**Citation** So RTY, Chu DKW, Miguel E, Perera RAPM, Oladipo JO, Fassi-Fihri O, Aylet G, Ko RLW, Zhou Z, Cheng M-S, Kuranga SA, Roger FL, Chevalier V, Webby RJ, Woo PCY, Poon LLM, Peiris M. 2019. Diversity of dromedary camel coronavirus HKU23 in African camels revealed multiple recombination events among closely related betacoronaviruses of the subgenus Embecovirus. J Virol 93:e01236-19. https://doi.org/10.1128/JVI.01236-19.

**Editor** Julie K. Pfeiffer, University of Texas Southwestern Medical Center



Address correspondence to Malik Peiris, malik@hku.hk.

**Received** 29 July 2019
**Accepted** 11 September 2019
**Accepted manuscript posted online** 18 September 2019
**Published** 13 November 2019

syndrome [SADS] coronavirus from bats to swine) (1) and from animals to humans (2). Two human coronaviruses (HCoVs), 229E and OC43, now endemic in the human population, emerged from camels and bovines, respectively, within the past few hundred years (2, 3). Severe acute respiratory syndrome (SARS) coronavirus emerged from bats via intermediate mammalian hosts in live game animal markets in Guangdong to spread to over 25 countries across 5 continents, causing illness in almost 8,000 people and leading to almost 800 deaths (4–6). The ability of coronaviruses to make interspecies jumps is facilitated by complex virus-host interactions. The high frequency of virus genetic recombination represents one mechanism by which the virus adapts to a new host.

Virus genetic recombination is frequently observed in coronaviruses and other positive-sense RNA viruses. Murine hepatitis virus (MHV), a clade A beta-CoV, is a well-studied example of homologous recombination, with up to 25% of its progeny in infected cells being demonstrated to be recombinants (7). High frequency of recombination is believed to be contributed by the large genome size, the intrinsic template-switching property of the viral RNA-dependent RNA polymerase (RdRp) during replication, and the abundance of subgenomic RNA strands for template switching (8, 9). The role of RdRp in RNA recombination has been shown in poliovirus, where a single amino acid residue mutation in the RdRp of poliovirus can result in a decrease in RNA recombination frequency (10). The activity of exoribonucleases (ExoN) in replicase nonstructural protein 14 (nsp 14) of CoVs that constitutes the proofreading activity of genome replication has been suggested to be a potential regulator of RNA recombination (11). The presence of group-specific genes in CoVs is assumed to be a result of heterologous recombination, which involves exchange of nonhomologous viral or cellular RNAs. The hemagglutinin esterase (HE) gene that is expressed only in clade A beta-CoVs is believed to have been acquired from influenza C virus through such heterologous recombination (12).

In 2012, a novel respiratory pathogen, Middle East respiratory syndrome coronavirus (MERS-CoV), was isolated from a patient with severe respiratory illness in Jeddah, Saudi Arabia (13). It was a zoonotic virus found in dromedary camels which occasionally transmits to human (14–16). MERS-CoV is enzootic in dromedary camels in the Middle East as well as in Africa, with the greatest virus diversity found in Africa (17). Circulation of multiple lineages of clade B MERS-CoV in dromedary camels eventually resulted in the emergence of a recombinant lineage 5 virus that caused major outbreaks during 2015 within both Saudi Arabia and South Korea, following introduction of the virus by a returning traveler (18). Recent studies have shown that two other coronaviruses, alphacoronavirus dromedary camel coronavirus 229E (DcCoV-229E) and beta-CoV dromedary camel coronavirus HKU23 (DcCoV-HKU23), cocirculate in dromedaries in Saudi Arabia (18, 19). The cocirculation of at least three coronaviruses within camels provides an opportunity for the emergence of novel emerging infections via recombination. It is therefore important to look for evidence of recombination between coronaviruses cocirculating in dromedary camels because such recombination might have contributed to the emergence of MERS-CoV and might contribute to the future emergence of viruses of zoonotic and epidemic potential. Here, we report the genetic diversity of DcCoV-HKU23 in the African camel population and identify several recombination events that have taken place with clade A beta-CoVs, including recombination with bovine coronavirus (BCoV) and with more-distant species such as rabbit coronavirus (RbCoVHKU14) and rodent coronavirus (RodentCoV). We carried out our studies in West Africa (Nigeria), East Africa (Ethiopia), and North Africa (Morocco) because over 70% of the global population of dromedaries are found in Africa and this is where the greatest diversity of these dromedary coronaviruses is likely to be found and where MERS-CoV emerged.

## RESULTS

**Screening of DcCoV-HKU23 in African camels by RT-qPCR and microneutralization assay.** Nasal swab samples of dromedary camels in Nigeria ($n = 2,529$), Mo-

**TABLE 1** Screening of DcCoV-HKU23 in dromedary camels in Nigeria, Ethiopia, and Morocco

| Country | Sampling yr(s) | No. of positive sample results/no. tested (% positive) | $C_T$ range | Field collection reference |
|---|---|---|---|---|
| Nigeria | 2015, 2016 | Total: 55/2,529 (2.2)<br>Young camels: 8/194 (4.1)<br>Adult camels: 47/2,335 (2.0) | 18.8–36.9 | 28 |
| Morocco | 2015, 2016 | Total: 7/1,569[a] (0.45)<br>Young camels: 6/584 (1.0)<br>Adult camels: 1/577 (0.17) | 26.5–33.4 | 29 |
| Ethiopia | 2015 | Total: 9/621[a] (1.4)<br>Young camels: 1/136 (0.74)<br>Adult camels: 5/314 (1.6) | 23.2–30.2 | 29 |

[a]Age information was not available for all sampled camels.

rocco ($n = 1,569$), and Ethiopia ($n = 621$) were tested for coronaviruses using the pan-CoV reverse transcription-PCR (RT-PCR) method and identified by sequencing PCR amplicons (Table 1). (20). The overall prevalence of HKU23 viruses at each location ranged from 0.4% of 1,569 samples tested in Morocco to 2.2% of 2,529 collected in Nigeria (Table 1). A DcCoV-HKU23-specific quantitative real-time reverse transcription-PCR (RT-qPCR) assay was subsequently performed to identify samples with a high viral RNA copy number for whole-genome sequencing.

In Morocco, the rate of DcCoV-HKU23 RNA positivity in young camels aged ≤2 years ($n = 584$) was 1.0% and was not significantly different from that seen in adults ($n = 577$), with 0.17% positive (Fisher's exact test, $P = 0.124$). In Nigeria, the rate of DcCoV-HKU23 RNA positivity in young camels was 4.1% ($n = 194$) compared to 2.0% in adults ($n = 2,335$) (Fisher's exact test, $P = 0.0674$). In Ethiopia, young camels ($n = 136$) had 1 positive swab whereas adults ($n = 314$) had 5 positive swabs, which also did not represent significant differences (Fisher's exact test, $P = 0.673$). Swab specimens were collected during October to April, with virus detection occurring in most months (Table 2).

To study the seroprevalence of Dc-CoV-HKU23 in African camels, dromedary sera were also collected from a subset of camels on the same sampling occasions and were tested by microneutralization assay. High seroprevalence was detected in dromedary camels in all three countries, with seroprevalences of 92% of 150 sera in Nigeria, 91% of 100 sera in Ethiopia, and 79% of 100 sera in Morocco (Table 3). A lower seropositive rate was observed in younger (≤2 years) than in older Moroccan camels from abattoirs (48% versus 92%; Fisher's exact test, $P = 0.0036$) and farms (76% versus 100%; Fisher's exact test, $P = 0.0223$). There were no marked differences in the seroprevalence levels measured for young and old camels in the Nigerian abattoir or in abattoirs or farms in Ethiopia.

Cross-neutralizing antibody responses of DcCoV-HKU23 and BCoV were evaluated by testing camel sera with high, medium, and low levels of and no neutralizing

**TABLE 2** Monthly Dc-CoV-HKU23 RNA prevalence in camels in Africa and Saudi Arabia

| | No. of positive sample results/no. tested (% positive) | | | |
|---|---|---|---|---|
| | 2014–2015 | | 2015–2016 | |
| Mo | Morocco | Ethiopia | Nigeria | Morocco |
| May–September | ND[a] | ND | ND | ND |
| October | ND | ND | 1/526 (0.2) | ND |
| November | ND | ND | 15/739 (2.0) | ND |
| December | ND | ND | 1/35 (2.9) | ND |
| January | ND | 2/120 (1.7) | 12/531 (2.3) | 0/349 (0) |
| February | 0/195 (0) | 7/501 (1.4) | 26/698 (3.7) | ND |
| March | 0/186 (0) | ND | ND | 5/385 (1.3) |
| April | ND | ND | ND | 2/453 (0.4) |

[a]ND, no data.

**TABLE 3** Seropositive rates of Dc-CoV-HKU23 in camel sera from Nigeria, Ethiopia, and Morocco determined by microneutralization assay

| Region | Location type | Camel age group[a] | No. of positive sample results/no. tested (% positive) | Median titer |
|---|---|---|---|---|
| Nigeria | Abattoir | Young | 12/16 (75) | 1:40 |
| | | Adult | 123/134 (92) | 1:20 |
| Ethiopia | Abattoir | Young | 22/24 (92) | 1:40 |
| | | Adult | 22/26 (85) | 1:40 |
| | Farm | Young | 23/25 (92) | 1:40 |
| | | Adult | 24/25 (96) | 1:40 |
| Morocco | Abattoir | Young | 12/25 (48) | <1:10 |
| | | Adult | 23/25 (92) | 1:20 |
| | Farm | Young | 19/25 (76) | 1:20 |
| | | Adult | 25/25 (100) | 1:160 |

[a]Young, age ≤2 years; Adult, age >2 years.

DcCoV-HKU23 titers by neutralization tests with strain BCoV-Mebus. There was a significant correlation between the titers of DcCoV-HKU23 and BCoV, suggesting likely serological cross-reactivity between the two viruses (Fig. 1).

**Evolutionary divergence and genetic diversity of DcCoV-HKU23.** Full genomes of DcCoV-HKU23 were obtained from four swab samples collected in Nigeria (NV1010, NV1092, NV1097, and NV1385) and from one sample each collected in Morocco (CAC2586) and Ethiopia (CAC1019). The African virus genomes were found to be closely related, with the number of pairwise base substitutions per site being below 0.0270 (Table 4). These full genomes were compared with those previously reported from Saudi Arabia (18). DcCoV-HKU23 isolates collected in the African region differed from those collected in Saudi Arabia by a range of 0.0223 to 0.0270 pairwise base substitutions per site, comparable to the divergence observed among the regions in Africa. Compared to other closely related species within clade A beta-CoVs, they differed from BCoV by a range of 0.0249 to 0.0300 pairwise base substitutions per site and from HCoV-OC43 by a range of 0.0445 to 0.0468 pairwise base substitutions per site. The species most closely related to HCoV-OC43 remains BCoV rather than DcCoV-HKU23.

The genetic diversity of DcCoV-HKU23 across Africa and the Middle East was studied on the basis of the ORF1ab (open reading frame 1ab) gene by generating a distance plot using SSE version 1.3 (21) (Fig. 2). As recombination has previously been shown to increase progressively from the 5′ to the 3′ end of the genome (22), ORF1ab was

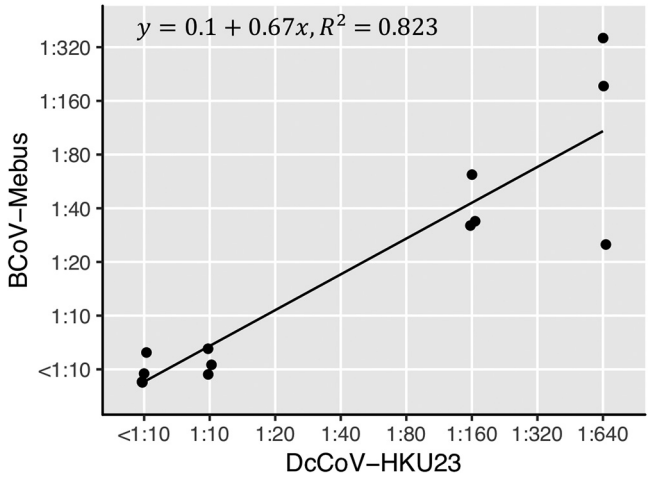


**FIG 1** Scatterplot showing camel sera ($n = 13$) with different neutralizing titers against DcCoV HKU23 tested for cross-neutralization against BCoV.

**TABLE 4** Estimates of evolutionary divergence between bovine CoV and human CoV-OC43 and the complete genome sequences of DcCoV-HKU23 strains identified in Africa (Nigeria, Ethiopia, and Morocco) and Saudi Arabia[a]

| Host | Country | Strain (GenBank accession no.) | Genome size (bp) | %GC | Pairwise evolutionary divergence (no. of base substitutions per site) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | NV1010 | NV1092 | NV1097 | NV1385 | CAC1019 | CAC2586 | 368F | 362F | 265F | Ry123 | DB2 | OC43 |
| Camel | Nigeria | DcCoV-HKU23-NV1010 | 30,780 | 36.90 | | | | | | | | | | | | |
| | | DcCoV-HKU23-NV1092 | 30,799 | 36.90 | 0.0001 | | | | | | | | | | | |
| | | DcCoV-HKU23-NV1097 | 31,075 | 37.00 | 0.0167 | 0.0168 | | | | | | | | | | |
| | | DcCoV-HKU23-NV1385 | 30,798 | 36.90 | 0.0004 | 0.0004 | 0.0168 | | | | | | | | | |
| | Ethiopia | DcCoV-HKU23-CAC1019 | 31,021 | 36.90 | 0.0206 | 0.0207 | 0.0246 | 0.0206 | | | | | | | | |
| | Morocco | DcCoV-HKU23-CAC2586 | 31,062 | 37.00 | 0.0191 | 0.0191 | 0.0129 | 0.0191 | 0.0270 | | | | | | | |
| | Saudi Arabia | DcCoV-HKU23-368F (KF906251.1) | 31,052 | 37.00 | 0.0263 | 0.0265 | 0.0231 | 0.0263 | 0.0223 | 0.0225 | | | | | | |
| | | DcCoV HKU23-362F (KF906250.1) | 31,052 | 37.00 | 0.0263 | 0.0265 | 0.0231 | 0.0263 | 0.0223 | 0.0225 | 0.0000 | | | | | |
| | | DcCoV-HKU23-265F (KF906249.1) | 31,052 | 37.00 | 0.0264 | 0.0266 | 0.0235 | 0.0264 | 0.0227 | 0.0229 | 0.0017 | 0.0017 | | | | |
| | | DcCoV-HKU23-Ry123 (KT368891.1) | 31,041 | 37.00 | 0.0269 | 0.0270 | 0.0237 | 0.0268 | 0.0229 | 0.0231 | 0.0017 | 0.0017 | 0.0026 | | | |
| Bovine | | BCoV-DB2 (DQ811784.2) | 31007 | 37.10 | 0.0299 | 0.0300 | 0.0288 | 0.0299 | 0.0249 | 0.0279 | 0.0197 | 0.0197 | 0.0200 | 0.0204 | | |
| Human | | HCoV-OC43 (AY391777.1) | 30738 | 36.80 | 0.0466 | 0.0467 | 0.0468 | 0.0468 | 0.0445 | 0.0463 | 0.0420 | 0.0420 | 0.0421 | 0.0425 | 0.0333 | |

[a]Analyses were conducted using the Tamura-Nei model and MEGA7.

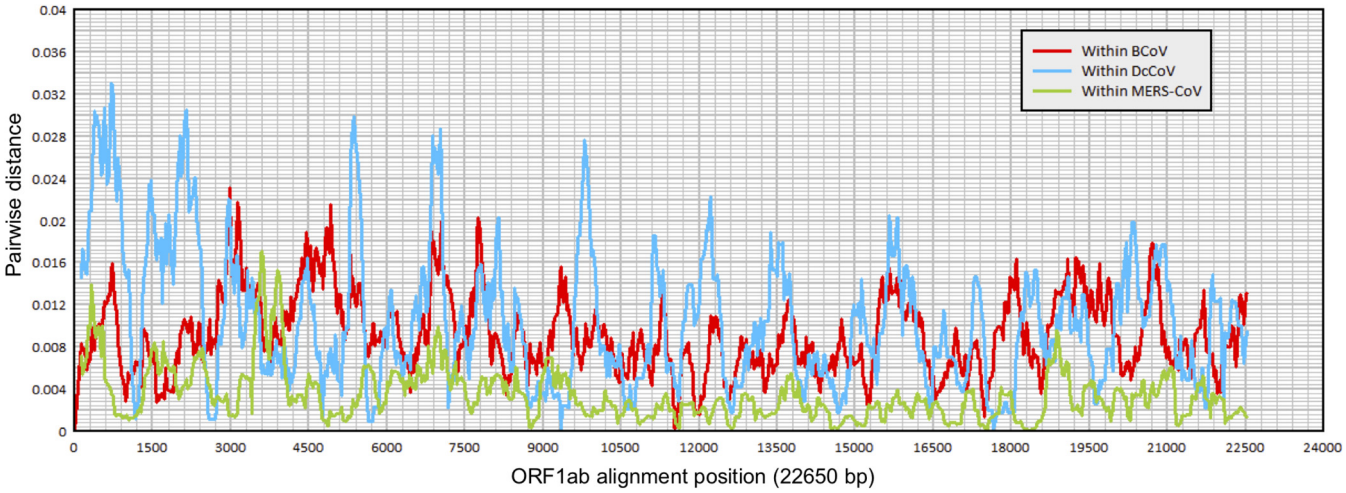


**FIG 2** Comparison of the levels of genetic diversity of the ORF1ab gene of DcCoV-HKU23, BCoV, and MERS-CoV. Sequence distance plots were generated using SSE version 1.3, a sliding window of 250 nucleotides, and a step size of 25 nucleotides. Sequences were obtained from the GenBank database, and closely related sequences with a pairwise distance of <0.001 were excluded from the analysis. A total of 10 DcCoV-HKU23 sequences, 28 BCoV sequences, and 88 MERS-CoV sequences were included in the analysis.

selected to enable study of the genetic diversity with minimal confounding by recombination. Along the position of ORF1ab gene, a mean pairwise distance of about 0.01 was observed within the 6 DcCoV-HKU23 sequences from Africa and the 4 reference sequences available from Saudi Arabia. The observed diversity of DcCoV-HKU23 was comparable to that of BCoV, suggesting that the two viruses were introduced into their animal hosts at similar points in time. Another CoV circulating in the same camel populations, MERS-CoV, was included in the analysis and showed a diversity level of about 0.004 by the same analysis (Fig. 2), which was lower than that determined for DcCoV-HKU23.

**Phylogenetic analysis of DcCoV-HKU23 with BCoV and HCoV-OC43 sequences.** To infer the phylogenetic relationship of the newly identified African DcCoV-HKU23 strain to previously reported DcCoV-HKU23 strains from Saudi Arabia and to bovine coronaviruses which are closely related, phylogenetic trees based on the complete coding sequences of the RdRp gene (2,783 nucleotides [nt]), the spike gene (4,101 nt), and the nucleocapsid gene (1,347 nt) were constructed. In addition to the 6 full genomes, 4 more virus sequences with complete RdRp, S, and N genes representing DcCoV-HKU23 isolates from Ethiopia (CAC1320 and CAC1452) and Morocco (CAC2505 and CAC2753) were obtained and included in the analysis. Using the genotyping nomenclature previously described for HCoV-OC43 and BCoV (23, 24) as a reference point, the sequences of DcCoV-HKU23 examined in this study were mapped into the 3 main subclusters of BCoV, namely, C1, C2, and C3, where C1 contains BCoVs from the Americas, C2 contains BCoVs from Europe, and C3 contains the prototype BCoV (Fig. 3). In the phylogenetic analysis of the RdRp gene, all the African and Saudi DcCoV-HKU23 sequences clustered within clade C2, which includes BCoVs from Europe. African DcCoV-HKU23 sequences did not form a monophyletic clade with the Saudi Arabia strain; instead, these sequences were scattered within clade C2, suggesting multiple ancestral origins of DcCoV-HKU23 across different geographic areas.

The analysis of the spike gene showed that the 8 African and Saudi DcCoV-HKU23 sequences clustered together and grouped into a clade distinct from BCoV, which we designated clade $C_{Outlier}$, and that that clade has a basal phylogenetic relationship to the BCoV clades. The phylogeny of the viruses was geographically structured such that most of the sequences within the $C_{Outlier}$ clade grouped into subclades of Saudi Arabia, Ethiopia, and Western/Northern Africa (Nigeria and Morocco) viruses. The phylogenetic tree of the spike gene resembles the region-dependent diversity of MERS-CoV as observed in these camels, in which viruses from Africa and the Middle East were

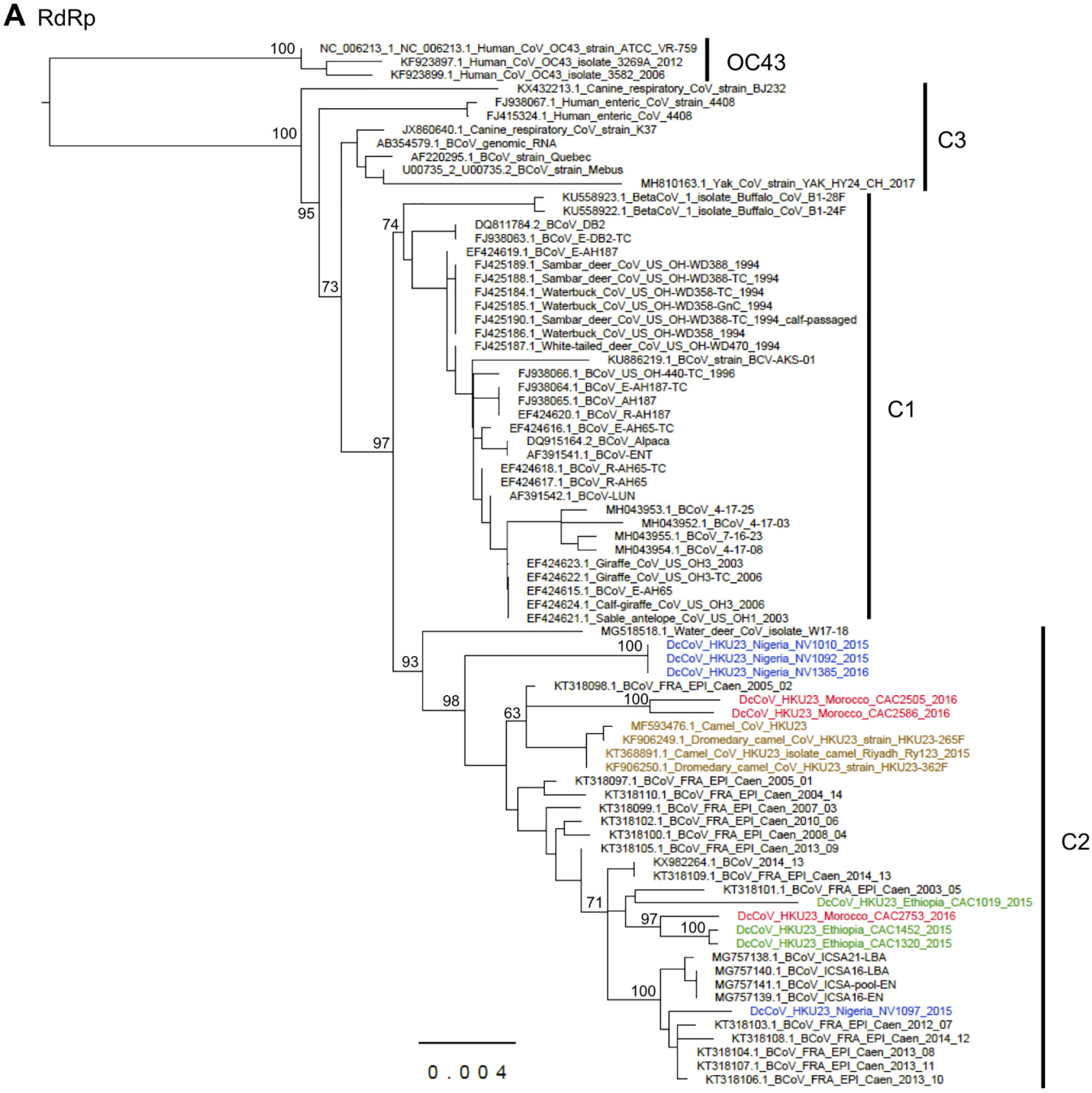


**FIG 3** Phylogenetic analysis of (A) RdRp, (B) spike, and (C) nucleocapsid genes of DcCoV-HKU23 collected in Nigeria (blue), Morocco (red), and Ethiopia (green). Reference DcCoV-HKU23 sequences from Saudi Arabia are indicated in brown. The alignment of each gene was manually trimmed to obtain alignments of 2,769 nucleotides (nt) for the RdRp gene, 4,137 nt for the spike gene, and 1,347 nt for the nucleocapsid gene. The tree was constructed by the maximum likelihood method using IQ-Tree, with the best-fit model automatically selected by ModelFinder. Nodes indicate bootstrap values calculated using ultrafast bootstrap with 1,000 replicates. Trees were midpoint rooted.

grouped into two separate clades (17). However, two sequences from Morocco (CAC2505 and CAC2753) were distinct from the other sequences and fell into clade C2 of BCoVs, sharing a common ancestor with a cluster of BCoVs from France.

The phylogeny of the N gene of DcCoV-HKU23 was more diverse, with virus sequences being distributed in BCoV clade C3, which included 4 Nigerian sequences (NV1010, NV1092, NV1097, and NV1385), one Moroccan sequence (CAC2586), and two sequences from Ethiopia (CAC1320). These sequences clustered together monophyl-

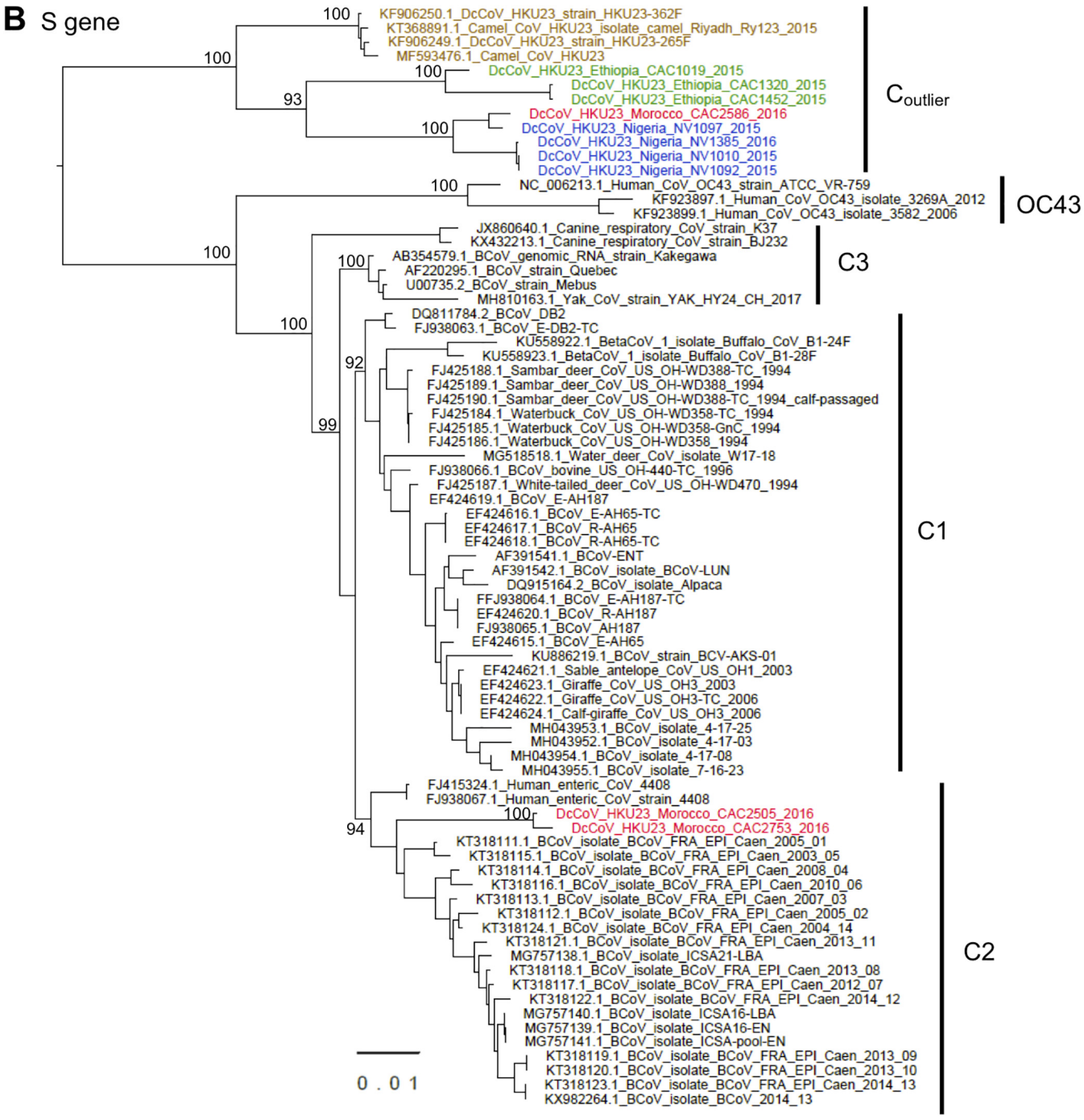


**FIG 3** (Continued)

etically and were related to human enteric coronavirus strain 4408. The other 3 sequences from Ethiopian isolate CAC1019 and Moroccan isolates CAC2753 and CAC2505 were grouped within BCoV clade C2 together with the strains from Saudi Arabia.

Combining the clade classifications of these three gene regions, there were 3 circulating genotypes of DcCoV-HKU23, viz., C2/C2/C2; C2/$C_{Outlier}$/C2, and C2/$C_{Outlier}$/C3 (Table 5), suggesting that multiple genetic recombination events had occurred in the past. This contrasts with BCoV, which does not appear to exhibit such genetic instability (24). However, there is a lack of BCoV sequence data from Africa. Recombinant

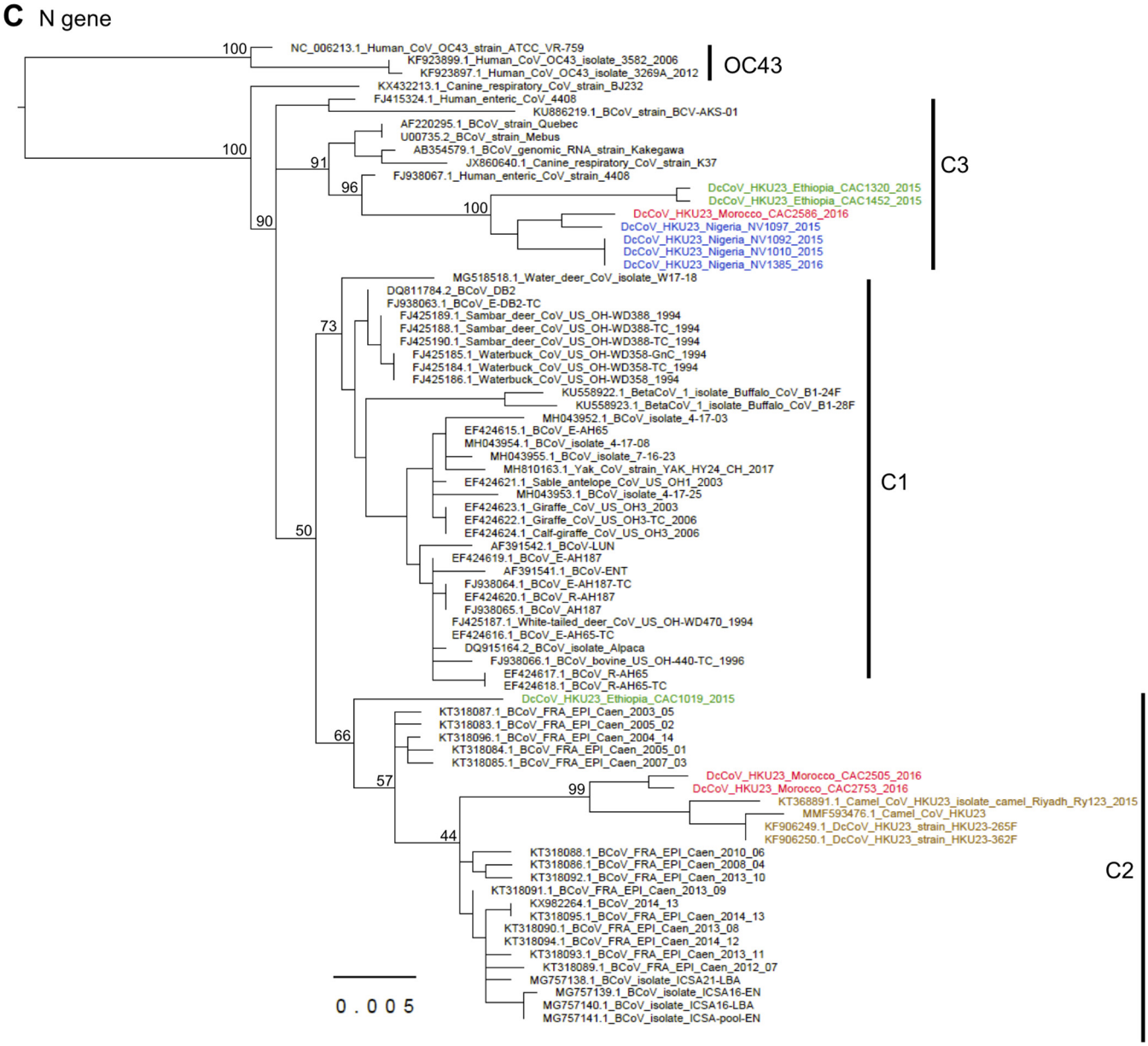


**FIG 3** (Continued)

genotypes C2/$C_{Outlier}$/C2 and C2/$C_{Outlier}$/C3 were observed in DcCoV-HKU23 across the African region without a distinct geographic pattern. Genotype C2/$C_{Outlier}$/C2 was observed in one sample from Saudi Arabia and in one from Ethiopia, while C2/$C_{Outlier}$/C3 was observed in four samples from Nigeria, one from Ethiopia, and one from Morocco. BCoV genotype C2/C2/C2 was observed in two Moroccan strains (CAC2753 and CAC2505), possibly suggesting a direct spillover of BCoV genotype C2 to the camel population.

**Distinct genetic region upstream of NS5a among DcCoV-HKU23 strains.** The genomic organization of African DcCoV-HKU23 is almost identical to those of Saudi Arabia DcCoV-HKU23 and BCoV, except for a 400-nt region between the S gene and NS5a that was found to be highly divergent between DcCoV-HKU23 and other clade A beta-CoV strains (Fig. 4). In BCoV, this region contains 2 ORFs (ORF4a and ORF4b) that encode 4.9-kDa and 4.8-kDa nonstructural proteins, respectively. The absence of ORF4a and ORF4b in HCoV-OC43 suggested that they are not essential for viral replication (25).

**TABLE 5** Summary of the genotypes of DcCoV-HKU23 identified in this study[a]

| Virus strain | Country of origin | Sample | RdRp | S gene | N gene | Genotype |
|---|---|---|---|---|---|---|
| DcCoV-HKU23 | Nigeria | NV1010 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | | NV1092 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | | NV1097 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | | NV1385 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | Ethiopia | CAC1019 | C2 | C outlier | C2 | C2/C outlier/C2 |
| | | CAC1320 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | | CAC1452 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | Morocco | CAC2505 | C2 | C2 | C2 | C2/C2/C2 |
| | | CAC2586 | C2 | C outlier | C3 | C2/C outlier/C3 |
| | | CAC2753 | C2 | C2 | C2 | C2/C2/C2 |
| | Saudi Arabia | 362F | C2 | C outlier | C2 | C2/C outlier/C2 |
| BCoV | Europe | BCoV/FRA | C2 | C2 | C2 | C2/C2/C2 |
| | Americas | BCoV ENT | C1 | C1 | C1 | C1/C1/C1 |

[a]RdRp, RNA-dependent RNA polymerase; S, spike; N, nucleocapsid.

Pairwise comparisons of this region among all DcCoV-HKU23 strains with BCoV-DB2, RbCoV HKU14, and HCoV-OC43 revealed nonsense mutations in both ORF4a and ORF4b in DcCoV-HKU23 strains from Saudi Arabia and Africa that have resulted in premature stop codons and a truncated protein. A 200-nt deletion was observed in NV1385 after the premature stop codon of ORF4a. Similar though not identical patterns of deletion were also found in RbCoV-HKU14 and HCoV-OC43 that resulted in more-truncated protein sequences. Although these protein sequences differ, the nucleotide sequences of DcCoV-HKU23/362F, DcCoV-HKU23/CAC1019, RbCoV-HKU14, and HCoV-OC43 in fact share high pairwise similarity to the BCoV sequences that contain full-length sequence of both ORF4a and ORF4b. Interestingly, DcCoV/NV1097 and

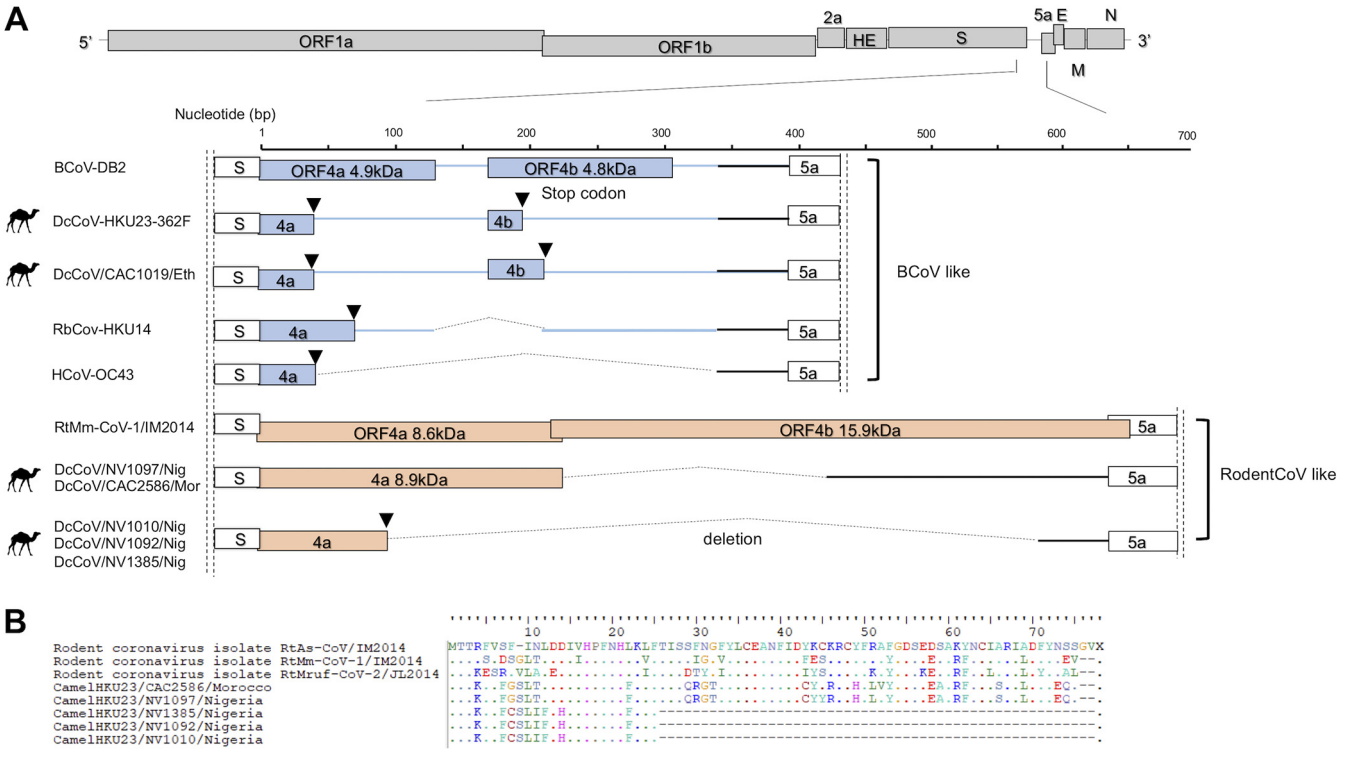


**FIG 4** (A) Genomic organization of DcCoV-HKU23 and other clade A beta-CoVs of the region between the spike gene and NS5a. Distinct ORF patterns were found in DcCoV-HKU23 in this region. Stop codons in the ORFs are labeled with a black triangle. Horizontal dotted lines indicated regions of deletion. (B) The amino acid sequence alignment of the 8.9-kDa ORF4a of rodent coronaviruses (RtAs-CoV/IM2014 [GenBank accession no. KY370044, RtMm-CoV-1/IM2014 [accession no. KY370052], and RtMruf-CoV-2/JL2014 [accession no. KY370046]) and DcCoV-HKU23.

DcCoV/CAC2586, among other DcCoVs, contain a distinct ORF4a sequence that encodes an 8.9-kDa nonstructural protein in this region. Homology analysis of this protein was performed by BLAST searches, and the results mapped to another 8.6-kDa nonstructural protein encoded by a rodent CoV (RtMm-CoV-1/IM2014 (GenBank accession no. KY370052.1) with about 60% amino acid similarity (Fig. 4B). BCoV ORF4a and ORF4b have previously been suggested to be counterparts of the 11-kDa nonstructural protein in mouse hepatitis virus (MHV) (26). The discovery of this rodent-like ORF4a encoded in DcCoV-HKU23 illustrated the possibility that homologous recombination with rodent coronaviruses had occurred, highlighting its distinct evolutionary history compared to BCoV.

**Recombination analysis with other clade A beta-CoVs.** To study the possibility that cross-species recombination had occurred, bootscan analysis of full genomes of DcCoV-HKU23 isolates collected in Africa with other clade A beta-CoVs was performed using Simplot, version 3.5.1. Multiple-sequence alignment of the 6 African DcCoV-HKU23 strains with RbCoV-HKU14, porcine hemagglutinating encephalomyelitis virus (PHEV), BCoV-DB2, EqCoV-NC99, RodentCoV-IM2014, DcCoV-HKU23 from Saudi Arabia, and HCoV-OC43 was performed. Using the Nigerian DcCoV-HKU23 strain as the query, recombination signals were observed with BCoV-DB2 at the NS2a gene (positions 21901 to 22600), with RbCoV-HKU14 at the hemagglutinin esterase (HE) gene (positions 22601 to 23600), and with the RodentCoV-IM2014 at the regions of ORF4a and ORF4b and the NS5a gene (positions 27901 to 28800) (Fig. 5A). Phylogenetic analysis of the BCoV signal region showed that BCoV-DB2 clustered with the group of DcCoV-HKU23 from Africa (Fig. 5B). The signal at the NS2a gene extended the region, showing the mixing of BCoV with DcCoV-HKU23 in addition to the RdRp and N genes. The tree of the RbCoV signal region showed that RbCoV-HKU14 had changed its phylogenetic position and was linked to the cluster of DcCoV-HKU23 isolates from Nigeria and Ethiopia. This clustering suggests that a recombination event had taken place between a common ancestor of DcCoV-HKU23 sequences from Nigeria or Ethiopia and RbCoV-HKU14 or a RbCoV-like virus. The presence of RodentCoV signal at the region of ORF4a and ORF4b supported the idea of a homologous recombinant resulting from recombination of West African DcCoV-HKU23 with a RodentCoV sequence in the genetic organization analysis. A tree plotted on the basis of positions 27901 to 28800 showed that the DcCoV-HKU23 strains were split into two separate evolutionary pathways (illustrated as "BCoV-like" and "RodentCoV-like" in Fig. 4A) in which West African DcCoV-HKU23 strains were clustered in an outgroup with RodentCoV-IM2014 whereas East African DcCoV-HKU23 strains were clustered with BCoV.

## DISCUSSION

Our data provide enhanced understanding of the diversity and circulation of an endemic DcCoV-HKU23 in African dromedaries in East Africa (Ethiopia), West Africa (Nigeria), and North Africa (Morocco) in comparison with viruses in Saudi Arabia and of the evolutionary relationships between HKU23 and BCoV, an important pathogen of cattle (27). In this study, HKU23 viral RNA was detected in 2.2% of dromedary nasal swabs tested in Nigeria, 0.5% in Morocco, and 1.4% in Ethiopia. The rates were higher than those seen in a previous study performed with year-around sampling of camel nasal swabs in Saudi Arabia, where a virus detection rate of 0.2% was reported (18). We detected DcCoV-HKU23 in most sampling months from October to April, which were the only periods of the year that we investigated, suggesting that there was no clear seasonality of virus activity. While cocirculation in Saudi Arabian camels of three CoVs, namely, MERS-CoV, DcCoV-229E, and DcCoV-HKU23, was reported, similar virus circulation was also observed in Nigeria, with positive rates of 2.2% for MERS-CoV (28) and 1.0% for DcCoV-229E (data not published). The serological prevalence rates of DcCoV-HKU23 antibodies in camels were 92% in Nigeria, 91% in Ethiopia, and 79% in Morocco, suggesting widespread circulation of this or an antigenically related virus over a broad geographical area across Africa in a manner comparable with the circulation of MERS-CoV (29, 30). Younger camels had lower seropositive rates than adult camels in

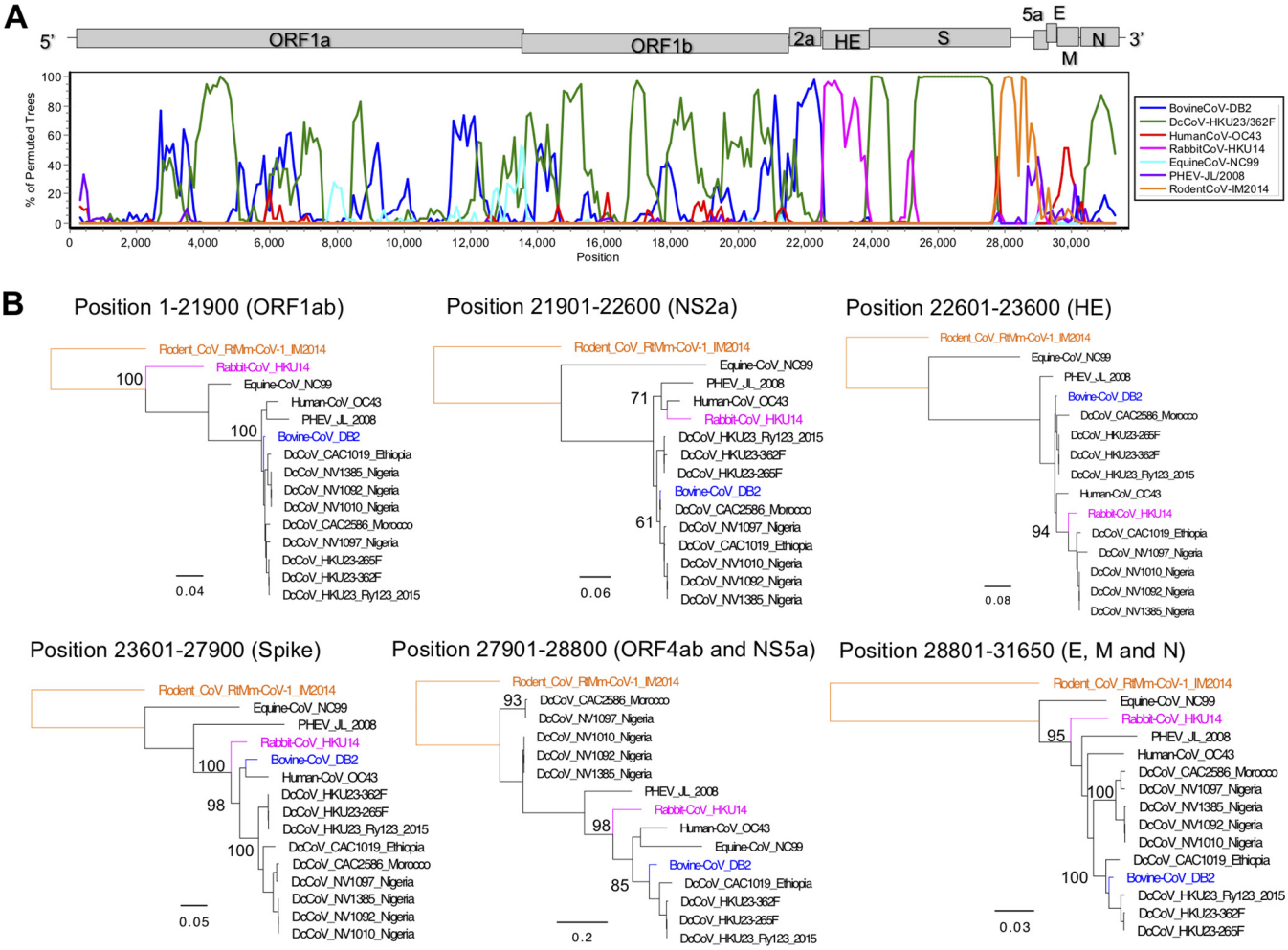


**FIG 5** (A) Recombination analysis of the genomes of DcCoV-HKU23, BCoV-DB2, PHEV, HCoV-OC43, RbCoV-HKU14, equine CoV-NC99 (EqCoV-NC99), and rodent CoV (RtMm-CoV-1/IM2014). Bootscan analysis was performed by Simplot, version 3.5.1, using a 50% consensus sequence of DcCoV-HKU23 collected in Nigeria with the genotype C2/$C_{outlier}$/C3 as the query. (B) Phylogenetic trees from representative regions were constructed by the maximum likelihood method using IQ-Tree, version 1.6.8. Trees were midpoint rooted. The CoVs (accession numbers) used in this analysis were as follows: DcCoV-HKU23/362F (KF906250.1), BCoV-DB2 (DQ811784.2), PHEV (KY994645), HCoV-OC43 (AY391777.1), RbCoV-HKU14 (JN874559), equine CoV-NC99 (EF446615), and rodent CoV-IM2014 (KY370052).

Morocco, but no age-related differences were observed in Nigeria and Ethiopia. Differences in DcCoV-HKU23 seroprevalence rates may be attributable to variations in husbandry practices, cohabitant animal hosts, or climatic factors between these countries. Similar seroprevalence rates were observed in camels in abattoirs and farms, suggesting that virus circulation had already been established at the farm or herd level and did not solely reflect amplification in the camel marketing chain. Extensive cross-neutralization of DcCoV-HKU23 and BCoV was detected, with a trend to higher titers for DcCoV-HKU23. The high (92% to 97%) amino acid sequence identity of the spike proteins of DcCoV-HKU23 and BCoV very likely contributed the cross-neutralization seen between the two viruses.

The full-genome sequences obtained from DcCoV-HKU23 isolates collected across Africa and Saudi Arabia allowed us to study the genetic diversity of this virus in camel populations. The distance plot generated using the ORF1ab gene can, to a modest degree, enable evaluation of the diversity within DcCoV-HKU23 sequences due to random mutations, with a lesser effect contributed by recombination. The observed diversity of DcCoV-HKU23 was comparable to that of BCoV, suggesting that the two viruses have been established in their hosts for similar periods of time. It is of interest that DcCoV-HKU23 has a much higher level of diversity than MERS-CoV. At present,

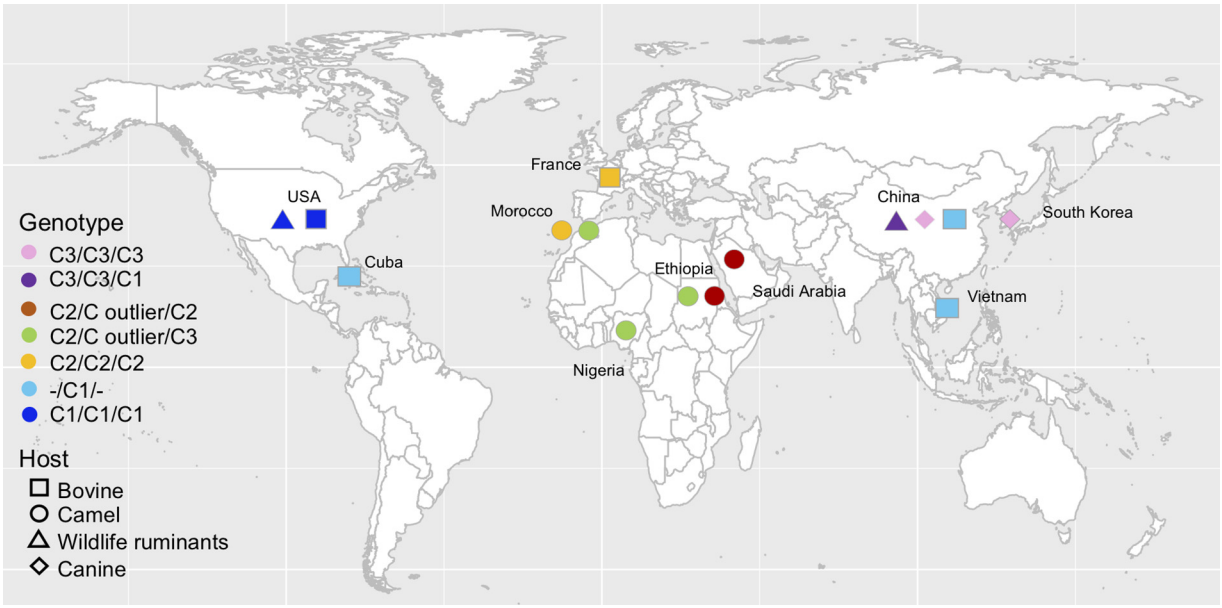


**FIG 6** The geographic distribution of different genotypes of BCoV/HKU23-like viruses in camels, bovines, and other species. The map was drawn using R software.

there are three CoVs (camel CoV 229E, DcCoV-HKU23, and MERS-CoV) cocirculating in dromedary camels; the narrower range of genetic diversity of MERS-CoV possibly indicates a more recent introduction into camels or a purifying selection event in its more recent evolutionary history.

Phylogenetic analysis of the full genomes of DcCoV-HKU23 in Africa with other published BCoV sequences available at GenBank identified incongruent topologies in the phylogenetic trees of the RdRp, S, and N genes, indicating events of recombination. With reference to the BCoV genotyping method described previously, the DcCoV-HKU23 strains obtained in this study were classified into 3 genotypes: C2/C2/C2, C2/$C_{Outlier}$/C2, and C2/$C_{Outlier}$/C3. The $C_{Outlier}$ clade has no BCoV sequences and is a uniquely DcCoV-HKU23 clade. Two DcCoV-HKU23 samples from Morocco (CAC2505 and CAC2753) showed a C2/C2/C2 genotype, which represents a nonrecombinant variant of BCoV clade 2, suggesting a BCoV evolutionary origin and the possibility that a BCoV spillover into the Moroccan camel population had taken place. This is the first time that a nonrecombinant BCoV variant has been detected in camels. Previously, BCoV was detected in a wide range of ungulate hosts, including bovines, waterbuck, sambar deer, white-tailed deer, alpaca, giraffe, stable antelope, buffalo, and yak (31–34). The expansion of BCoV to camel hosts further illustrates its ability to cross species and to infect other similar ungulate hosts. Recently, surveillance of BCoV has been expanded into regions of East Asia, including China (35), South Korea (36), and Vietnam (37), and into the Caribbean (Cuba) (38). The spike genes of BCoV strains from these regions were all phylogenetically clustered into C1. The currently known geographic distribution of different genotypes of BCoV/HKU23-like viruses in camels, bovines, and other species is summarized in Fig. 6.

The phylogeny of the S gene of DcCoV exhibited region-dependent diversity such as was also noted in BCoV and MERS-CoV. The C1 and C2 genotypes of BCoV correspond to the American-Asian and European clusters, respectively. In this study, the C2 genotype of the two Morocco strains clustered with BCoVs detected in France, suggesting a divergence from the common ancestor of genotype C2 BCoVs. Other DcCoVs identified in this study, as well as those previously identified in Saudi Arabia, are recombinants, because their S gene showed an outgroup topology with respect to BCoVs and HCoV-OC43, indicating that DcCoV-HKU23 acquired its S gene through recombination with an ancestor yet to be identified. However, we cannot infer whether

the recombination occurred prior to or after the introduction of the virus to camels from the present data. The phylogenies of the N genes of DcCoV-HKU23 also differed among the 8 recombinant sequences with the $C_{Outlier}$ genotype in the S gene. A total of 7 DcCoV-HKU23 sequences (NV1010, NV1092, NV1097, NV1385, CAC1320, CAC1452, and CAC2586) were grouped into C3 in the N gene tree, indicating a close relatedness to the prototype BCoV in the N gene. DcCoV-HKU23/CAC1019 from Ethiopia and DcCoV-HKU23 from Saudi Arabia were grouped into the C2 European clade. These C2 genotype sequences suggest the possibility that an event of recombination between the N gene of a nonrecombinant BCoV strain and a recombinant DcCoV-HKU23 strain (as the backbone) had taken place. Overall, DcCoV-HKU23 exhibited broad diversity that contrasted with the genetic stability observed in BCoV. However, a limitation in the analysis was the lack of BCoV sequence data in Africa. With more BCoV genetic data sets from Africa, the recombination events of DcCoV-HKU23 and BCoV may be more clearly resolved.

Cross-species recombination of DcCoV-HKU23 was also observed with other clade A beta-CoV species that involved a rabbit-CoV HKU14-like virus. Rabbit-CoV14 is a virus initially discovered through surveillance in China (39), but similar viruses may be present in a much wider geographic region. Bootscan analysis of DcCoV-HKU23 strains from Nigeria and Ethiopia showed a recombination signal with RbCoV-HKU14-like virus at positions 22601 to 23600, harboring the NS2a and HE gene. The HE gene in clade A beta-CoVs has been suggested to have been acquired as a consequence of heterologous recombination with influenza C virus. The recombination identified here was located in a similar region, suggesting a possible recombination hot spot. The phylogeny of the signal region showed that the RbCoV sequence was linked to the DcCoV-HKU23 sequences from Ethiopia and Nigeria at a basal position, suggesting that the recombination event might have occurred with the ancestral sequence from these two regions. The HE protein encoded by clade A beta-CoVs plays a role in receptor binding through its receptor-binding or receptor-destroying activity with respect to glycan components (40). An example of the functional properties of HE has been provided in the adaptation of HCoV-OC43 and HKU1 to infect humans, when the HE lectin domain was progressively lost through accumulated mutations (41). The diversity of the HE genes present in parental and recombinant DcCoVs deserves further research to characterize the glycoconjugates targeted by the receptor-destroying activity and to study how such virion-glycan interactions may contribute to its host tropism.

We further identified multiple mutations in ORF4a and ORF4b located downstream between the S and NS5a genes, providing insight into the evolutionary origin of clade A beta-CoVs. The deletion patterns of ORF4a and ORF4b observed in DcCoV-HKU23, RbCoV-HKU14, and HCoV-OC43 revealed stepwise deletions among these sequences. While these patterns may suggest a BCoV origin of these sequences, it is also possible that an ancestral virus that infected multiple hosts and bovines preferentially retained those ORFs. Nonsense mutations and deletions of these ORFs in DcCoV-HKU23, RbCoV-HKU14, and HCoV-OC43 supported the contention that this region may not contain essential genetic sequences and that the loss of such genetic information will not impair virus fitness in dromedaries. In fact, ORF4a and ORF4b in BCoV had previously been suggested to represent vestiges of an 11-kDa protein encoded by mouse hepatitis virus (MHV) resulting from a nonsense mutation in the middle of ORF4 (26). The region between the S and E genes may suggest a mouse or murine CoV origin of this region. As additional evidence, we also observed another 8.9-kDa protein encoded by ORF4a in DcCoV-HKU23 identified in this study which mapped to a similar protein in rodent CoV with 60% amino acid similarity. Bootscan analysis showed that DcCoV sequences from isolates collected in Morocco (CAC2586) and Nigeria (NV1010, NV1092, NV1097, and NV1385) were phylogenetically classified in an outgroup with rodent CoV at positions 27901 to 28800, suggesting a possible homologous recombination. Together, these sequences illustrated multiple origins of clade A beta-CoVs from rodent-like species. Results of recent surveillance of rodent species have identified many more novel CoV species, including ChRCoV-HKU24, Longquan Aa mouse coronavirus (LAMV),

Longquan RI rat coronavirus (LRLV), and rodent CoV (42–44). These sequences are phylogenetically positioned at the deep branch rooting members of clade A beta-CoVs. The discovery of sequence remnants of rodent CoV in DcCoV-HKU23 further supports the involvement of rodent-like species in the evolutionary history of clade A beta-CoVs (42).

The occurrence of the recombination events identified in this study required the coinfection of two or more different CoVs in the same host and same cell, and this requires the parental viruses to have been cocirculating in the same geographic region. The wide host range and geographic range of BCoV-like viruses could potentially facilitate recombination with other coronaviruses. Recombination events that led to the emergence of SARS-like CoV likely occurred in its natural reservoirs in bats, after which the virus spilled over to intermediate hosts such as civets, raccoons, dogs, and humans. Since rodents harbor the highest diversity of clade A beta-CoVs (44), one may speculate that it is possible for recombination events to occur in rodents, with the recombinant virus subsequently spilling over to other mammalian hosts if it has competitive advantages over preexisting strains. Thus, one may speculate that the recombination of DcCoV-HKU23 with RabbitCoV-HKU14 and RodentCoV-IM2014 might have occurred in a rodent species.

Limitations of this study included the lack of availability of rectal swabs to evaluate the tissue tropism associated with DcCoV-HKU23 infection. Although the virus seems not to cause significant disease, as it was detected in apparently healthy camels in abattoirs, it is unclear whether such infections are limited to the upper respiratory mucosa, similarly to MERS-CoV, or whether they spread more systemically. Specimens from sites of the body other than the upper respiratory tract may provide information on the tropism of the virus. The lack of year-round sampling precludes making conclusions on the seasonality of virus activity. On the other hand, the strength of the study is in its range of sampling across East, West, and North Africa, which allows an understanding of the virus diversity across a large geographic region. The lack of BCoV sequences from Africa precludes a more definitive analysis of the origins of the different genotypes of HKU23.

In conclusion, the study showed a mosaic structure of DcCoV-HKU23 that is likely to have been contributed by several recombination events among clade A beta-CoVs. Among the three identified DcCoVs that circulate in dromedary camels, MERS-CoV has so far demonstrated intraspecies recombination, while DcCoV-HKU23, in addition, demonstrated interspecies recombination. Our study highlighted the importance of studying recombination of CoVs to understand the evolutionary history and cross-species transmission of coronaviruses in dromedaries.

## MATERIALS AND METHODS

**Sample collection.** Nasal swabs and sera were collected from dromedary camels sampled in Nigeria, Morocco, and Ethiopia in previous studies of MERS-CoV in 2015 and 2016 (17, 29). Camel nasal swabs were collected from a camel abattoir in Kano, Nigeria ($n = 2,529$), in 2015 and 2016 (17, 28) and from dromedary herds and abattoirs in Morocco ($n = 1,569$) in 2015 and 2016 and in Ethiopia ($n = 621$) in 2015 (29) (Table 1). The camels from Morocco and Ethiopia were mostly raised for meat, milk production, or transport. Camel sera were concurrently collected from the abattoir in Nigeria ($n = 150$) and from abattoirs and farms in both Ethiopia ($n = 100$) and Morocco ($n = 100$). Sampled camels were 1 month to 20 years of age (median age, 3 years).

**DcCoV-HKU23 detection and genome sequencing.** Total nucleic acid was extracted from swab samples using an EasyMAG (bioMérieux, France) system. RNA was reverse transcribed into cDNA with random hexamers using a PrimeScript RT reagent kit (Perfect Real Time) (TaKaRa, Japan), according to the manufacturer's protocol. cDNAs were screened for DcCoV-HKU23 using a broad-range pancoronavirus nested PCR assay designed to detect known and unknown CoVs targeting the consensus region of the RNA-dependent RNA polymerase (RdRP) gene (20). RT-PCR-positive amplicons were purified using ExoSap-IT reagent (USB, USA) and were subjected to Sanger sequencing to identify the identity of the CoV. Samples with DcCoV-HKU23-like virus sequences were identified and subjected to viral load quantification using a quantitative reverse transcription-PCR (RT-qPCR) assay. Oligonucleotide sequences were designed to target the N genes of both DcCoV-HKU23 and bovine CoV (forward primer, 5′-GTCA ATACCCCGGCTGAC-3′; probe, 5′-6-carboxyfluorescein (FAM)-TCGGGACCCAAGTAGCGATGAGGC-black hole quencher (BHQ)-3′; reverse primer, 5′-AACCCTGAGGGAGTACCG-3′). RT-qPCR was performed using TaqMan Fast Virus 1-Step master mix (Thermo Fisher Scientific, USA), with the following cycling protocol:

5 min at 50°C for reverse transcription, followed by 20 s at 95°C and 40 cycles of 3 s at 95°C and 30 s at 60°C. Samples with low cycle threshold ($C_T$) values were selected for full-genome sequencing. Reverse transcription was performed with HKU23-specific primers targeting different regions of the genome to generate cDNAs, which were subsequently amplified by PCR with primers designed to generate overlapping amplicons that can cover the whole genome. The primer sequences are available upon request. PCR amplicons from each sample were pooled for next-generation sequencing and processed with a Nextera XT library preparation kit following the protocol provided by the manufacturer. Sequencing was performed using an Illumina MiSeq instrument with read lengths of paired ends of approximately 300 bp. The raw sequence reads generated were mapped to a reference DcCoV-HKU23 genome (GenBank accession no. KF906250.1) using Burrows-Wheeler alignment (BWA) (45). The sequence of the target virus was generated by taking the majority consensus of the mapped reads with sequencing coverage at each position more than 100 times.

**Genomic and phylogenetic analysis.** Opening reading frames (ORFs) of the virus genome encoding proteins were predicted using ORFfinder (NIH, USA). Full genomes of DcCoV-HKU23 were aligned with previously determined sequences from samples collected in Saudi Arabia and with bovine CoV and human CoV OC43 sequences by the use of MAFFT (multiple alignment using fast Fourier transform). Gaps and poorly aligned regions in the alignment were manually edited. Pairwise genetic distances were calculated using MEGA 7 (46). Phylogenetic analysis of DcCoV-HKU23 was performed using the maximum likelihood method and IQ-Tree, version 1.6.8 (47).

Recombination analysis was performed using Simplot version 3.5.1 (48). Bootscan analysis for a recombination event was performed on an alignment of the genome sequences as described above, with a 50% consensus sequence representing 4 DcCoV-HKU23 samples collected in Nigeria with the same C2/$C_{Outlier}$/C3 genotype (NV1010, NV1092, NV1097, and NV1385) as the query sequence. A sliding window of 600 nucleotides and a step size of 100 nucleotides were used as the scanning settings.

**Microneutralization assays.** Heat-inactivated (56°C for 30 min) camel sera were first diluted 1:10 and then subjected to serial 2-fold dilutions and mixed with equal volumes of virus at a dose of 200 50% tissue culture infective doses ($TCID_{50}$) of DcCoV-HKU23 isolate 368F (49). After 1 h of incubation at 37°C, 35 $\mu$l of the virus-serum mixture was added in quadruplicate to HRT-18G cell monolayers in 96-well microtiter plates. After 1 h of adsorption, the virus-serum mixture was removed and replaced with 150 $\mu$l of virus growth medium in each well. The plates were incubated for 5 days at 37°C in 5% $CO_2$ in a humidified incubator. Cytopathic effect was observed at day 5 postinoculation. The highest serum dilution protecting ≥50% of the replicate wells was denoted the neutralizing antibody titer. A virus back-titration of the input virus was included in each batch of tests.

**Data availability.** The complete genomes and segment sequences generated in this study are available in GenBank under accession numbers MN514962 to MN514979.

## ACKNOWLEDGMENTS

The study was supported by research grants from NIH contract HHSN272201400006C (to M.P. and R.J.W.); a Commissioned Grant from the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region (to M.P.); and the Theme-based Research Scheme (Project no. T11/707/15), University Grant Committee, Hong Kong (to P.C.Y.W.).

## REFERENCES

1. 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 XL, 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.
2. Vijgen L, Keyaerts E, Moes E, Thoelen I, Wollants E, Lemey P, Vandamme AM, Van Ranst M. 2005. Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. J Virol 79:1595–1604. https://doi.org/10.1128/JVI.79.3.1595-1604.2005.
3. Corman VM, Eckerle I, Memish ZA, Liljander AM, Dijkman R, Jonsdottir H, Juma Ngeiywa KJ, Kamau E, Younan M, Al Masri M, Assiri A, Gluecks I, Musa BE, Meyer B, Muller MA, Hilali M, Bornstein S, Wernery U, Thiel V, Jores J, Drexler JF, Drosten C. 2016. Link of a ubiquitous human coronavirus to dromedary camels. Proc Natl Acad Sci U S A 113:9864–9869. https://doi.org/10.1073/pnas.1604472113.
4. Peiris JS, Yuen KY, Osterhaus AD, Stohr K. 2003. The severe acute respiratory syndrome. N Engl J Med 349:2431–2441. https://doi.org/10.1056/NEJMra032498.
5. Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, Wong SS, Leung SY, Chan KH, Yuen KY. 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci U S A 102:14040–14045. https://doi.org/10.1073/pnas.0506735102.
6. 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. https://doi.org/10.1038/nature12711.
7. Makino S, Keck JG, Stohlman SA, Lai MM. 1986. High-frequency RNA recombination of murine coronaviruses. J Virol 57:729–737.
8. Cheng CP, Nagy PD. 2003. Mechanism of RNA recombination in carmo- and tombusviruses: evidence for template switching by the RNA-dependent RNA polymerase in vitro. J Virol 77:12033–12047. https://doi.org/10.1128/jvi.77.22.12033-12047.2003.
9. Fu K, Baric RS. 1992. Evidence for variable rates of recombination in the MHV genome. Virology 189:88–102. https://doi.org/10.1016/0042-6822(92)90684-h.
10. Kempf BJ, Peersen OB, Barton DJ. 2016. Poliovirus polymerase Leu420 facilitates RNA recombination and ribavirin resistance. J Virol 90:8410–8421. https://doi.org/10.1128/JVI.00078-16.
11. Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS. 2011. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. RNA Biol 8:270–279. https://doi.org/10.4161/rna.8.2.15013.
12. Luytjes W, Bredenbeek PJ, Noten AF, Horzinek MC, Spaan WJ. 1988.

Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. Virology 166:415–422. https://doi.org/10.1016/0042-6822(88)90512-0.
13. 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.
14. Perera RA, Wang P, Gomaa MR, El-Shesheny R, Kandeil A, Bagato O, Siu LY, Shehata MM, Kayed AS, Moatasim Y, Li M, Poon LL, Guan Y, Webby RJ, Ali MA, Peiris JS, Kayali G. 2013. Seroepidemiology for MERS coronavirus using microneutralisation and pseudoparticle virus neutralisation assays reveal a high prevalence of antibody in dromedary camels in Egypt, June 2013. Euro Surveill 18:pii=20574. https://doi.org/10.2807/1560-7917.es2013.18.36.20574.
15. Reusken CBEM, Haagmans BL, Müller MA, Gutierrez C, Godeke G-J, Meyer B, Muth D, Raj VS, Smits-De Vries L, Corman VM, Drexler J-F, Smits SL, El Tahir YE, De Sousa R, van Beek J, Nowotny N, van Maanen K, Hidalgo-Hermoso E, Bosch B-J, Rottier P, Osterhaus A, Gortázar-Schmidt C, Drosten C, Koopmans MPG. 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.
16. Haagmans BL, Al Dhahiry SH, Reusken CB, Raj VS, Galiano M, Myers R, Godeke GJ, Jonges M, Farag E, Diab A, Ghobashy H, Alhajri F, Al-Thani M, Al-Marri SA, Al Romaihi HE, Al Khal A, Bermingham A, Osterhaus AD, AlHajri MM, Koopmans MP. 2014. Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. Lancet Infect Dis 14:140–145. https://doi.org/10.1016/S1473-3099(13)70690-X.
17. Chu DKW, Hui KPY, Perera R, Miguel E, Niemeyer D, Zhao J, Channappanavar R, Dudas G, Oladipo JO, Traore A, Fassi-Fihri O, Ali A, Demissie GF, Muth D, Chan MCW, Nicholls JM, Meyerholz DK, Kuranga SA, Mamo G, Zhou Z, So RTY, Hemida MG, Webby RJ, Roger F, Rambaut A, Poon LLM, Perlman S, Drosten C, Chevalier V, Peiris M. 5 March 2018, posting date. MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.1718769115.
18. Sabir JS, Lam TT, Ahmed MM, Li L, Shen Y, Abo-Aba SE, Qureshi MI, Abu-Zeid M, Zhang Y, Khiyami MA, Alharbi NS, Hajrah NH, Sabir MJ, Mutwakil MH, Kabli SA, Alsulaimany FA, Obaid AY, Zhou B, Smith DK, Holmes EC, Zhu H, Guan Y. 2016. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. Science 351:81–84. https://doi.org/10.1126/science.aac8608.
19. 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.
20. Poon LL, Chu DK, Chan KH, Wong OK, Ellis TM, Leung YH, Lau SK, Woo PC, Suen KY, Yuen KY, Guan Y, Peiris JS. 2005. Identification of a novel coronavirus in bats. J Virol 79:2001–2009. https://doi.org/10.1128/JVI.79.4.2001-2009.2005.
21. Simmonds P. 2012. SSE: a nucleotide and amino acid sequence analysis platform. BMC Res Notes 5:50. https://doi.org/10.1186/1756-0500-5-50.
22. Fu K, Baric RS. 1994. Map locations of mouse hepatitis virus temperature-sensitive mutants: confirmation of variable rates of recombination. J Virol 68:7458–7466.
23. Lau SK, Lee P, Tsang AK, Yip CC, Tse H, Lee RA, So LY, Lau YL, Chan KH, Woo PC, Yuen KY. 2011. Molecular epidemiology of human coronavirus OC43 reveals evolution of different genotypes over time and recent emergence of a novel genotype due to natural recombination. J Virol 85:11325–11337. https://doi.org/10.1128/JVI.05512-11.
24. Kin N, Miszczak F, Diancourt L, Caro V, Moutou F, Vabret A, Ar Gouilh M. 2016. Comparative molecular epidemiology of two closely related coronaviruses, bovine coronavirus (BCoV) and human coronavirus OC43 (HCoV-OC43), reveals a different evolutionary pattern. Infect Genet Evol 40:186–191. https://doi.org/10.1016/j.meegid.2016.03.006.
25. Mounir S, Talbot PJ. 1993. Human coronavirus OC43 RNA 4 lacks two open reading frames located downstream of the S gene of bovine coronavirus. Virology 192:355–360. https://doi.org/10.1006/viro.1993.1043.
26. Abraham S, Kienzle TE, Lapps WE, Brian DA. 1990. Sequence and expression analysis of potential nonstructural proteins of 4.9, 4.8, 12.7, and 9.5 kDa encoded between the spike and membrane protein genes of the bovine coronavirus. Virology 177:488–495. https://doi.org/10.1016/0042-6822(90)90513-q.
27. Saif LJ. 2010. Bovine respiratory coronavirus. Vet Clin North Am Food Anim Pract 26:349–364. https://doi.org/10.1016/j.cvfa.2010.04.005.
28. So RT, Perera RA, Oladipo JO, Chu DK, Kuranga SA, Chan KH, Lau EH, Cheng SM, Poon LL, Webby RJ, Peiris M. 23 August 2018, posting date. Lack of serological evidence of Middle East respiratory syndrome coronavirus infection in virus exposed camel abattoir workers in Nigeria, 2016. Euro Surveill. https://doi.org/10.2807/1560-7917.ES.2018.23.32.1800175.
29. Miguel E, Chevalier V, Ayelet G, Ben Bencheikh MN, Boussini H, Chu DK, El Berbri I, Fassi-Fihri O, Faye B, Fekadu G, Grosbois V, Ng BC, Perera RA, So TY, Traore A, Roger F, Peiris M. 30 March 2017, posting date. Risk factors for MERS coronavirus infection in dromedary camels in Burkina Faso, Ethiopia, and Morocco, 2015. Euro Surveill. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30498.
30. Reusken CB, Messadi L, Feyisa A, Ularamu H, Godeke GJ, Danmarwa A, Dawo F, Jemli M, Melaku S, Shamaki D, Woma Y, Wungak Y, Gebremedhin EZ, Zutt I, Bosch BJ, Haagmans BL, Koopmans MP. 2014. Geographic distribution of MERS coronavirus among dromedary camels, Africa. Emerg Infect Dis 20:1370–1374. https://doi.org/10.3201/eid2008.140590.
31. Hasoksuz M, Alekseev K, Vlasova A, Zhang X, Spiro D, Halpin R, Wang S, Ghedin E, Saif LJ. 2007. Biologic, antigenic, and full-length genomic characterization of a bovine-like coronavirus isolated from a giraffe. J Virol 81:4981–4990. https://doi.org/10.1128/JVI.02361-06.
32. Jin L, Cebra CK, Baker RJ, Mattson DE, Cohen SA, Alvarado DE, Rohrmann GF. 2007. Analysis of the genome sequence of an alpaca coronavirus. Virology 365:198–203. https://doi.org/10.1016/j.virol.2007.03.035.
33. Decaro N, Martella V, Elia G, Campolo M, Mari V, Desario C, Lucente MS, Lorusso A, Greco G, Corrente M, Tempesta M, Buonavoglia C. 2008. Biological and genetic analysis of a bovine-like coronavirus isolated from water buffalo (Bubalus bubalis) calves. Virology 370:213–222. https://doi.org/10.1016/j.virol.2007.08.031.
34. Alekseev KP, Vlasova AN, Jung K, Hasoksuz M, Zhang X, Halpin R, Wang S, Ghedin E, Spiro D, Saif LJ. 2008. Bovine-like coronaviruses isolated from four species of captive wild ruminants are homologous to bovine coronaviruses, based on complete genomic sequences. J Virol 82:12422–12431. https://doi.org/10.1128/JVI.01586-08.
35. Keha A, Xue L, Yan S, Yue H, Tang C. 2019. Prevalence of a novel bovine coronavirus strain with a recombinant hemagglutinin/esterase gene in dairy calves in China. Transbound Emerg Dis 66:1971. https://doi.org/10.1111/tbed.13228.
36. Jeong JH, Kim GY, Yoon SS, Park SJ, Kim YJ, Sung CM, Shin SS, Lee BJ, Kang MI, Park NY, Koh HB, Cho KO. 2005. Molecular analysis of S gene of spike glycoprotein of winter dysentery bovine coronavirus circulated in Korea during 2002–2003. Virus Res 108:207–212. https://doi.org/10.1016/j.virusres.2004.07.003.
37. Shin J, Tark D, Le VP, Choe S, Cha RM, Park GN, Cho IS, Nga BTT, Lan NT, An DJ. 2019. Genetic characterization of bovine coronavirus in Vietnam. Virus Genes 55:415–420. https://doi.org/10.1007/s11262-019-01647-1.
38. Martínez N, Brandão PE, de Souza SP, Barrera M, Santana N, de Arce HD, Pérez LJ. 2012. Molecular and phylogenetic analysis of bovine coronavirus based on the spike glycoprotein gene. Infect Genet Evol 12:1870–1878. https://doi.org/10.1016/j.meegid.2012.05.007.
39. Lau SK, Woo PC, Yip CC, Fan RY, Huang Y, Wang M, Guo R, Lam CS, Tsang AK, Lai KK, Chan KH, Che XY, Zheng BJ, Yuen KY. 2012. Isolation and characterization of a novel betacoronavirus subgroup A coronavirus, rabbit coronavirus HKU14, from domestic rabbits. J Virol 86:5481–5496. https://doi.org/10.1128/JVI.06927-11.
40. Schultze B, Wahn K, Klenk HD, Herrler G. 1991. Isolated HE-protein from hemagglutinating encephalomyelitis virus and bovine coronavirus has receptor-destroying and receptor-binding activity. Virology 180:221–228. https://doi.org/10.1016/0042-6822(91)90026-8.
41. Bakkers MJ, Lang Y, Feitsma LJ, Hulswit RJ, de Poot SA, van Vliet AL, Margine I, de Groot-Mijnes JD, van Kuppeveld FJ, Langereis MA, Huizinga EG, de Groot RJ. 2017. Betacoronavirus adaptation to humans involved progressive loss of hemagglutinin-esterase lectin activity. Cell Host Microbe 21:356–366. https://doi.org/10.1016/j.chom.2017.02.008.
42. Lau SK, Woo PC, Li KS, Tsang AK, Fan RY, Luk HK, Cai JP, Chan KH, Zheng BJ, Wang M, Yuen KY. 2015. Discovery of a novel coronavirus, China Rattus coronavirus HKU24, from Norway rats supports the murine origin of betacoronavirus 1 and has implications for the ancestor of betacoronavirus lineage A. J Virol 89:3076–3092. https://doi.org/10.1128/JVI.02420-14.
43. Wang W, Lin XD, Guo WP, Zhou RH, Wang MR, Wang CQ, Ge S, Mei SH, Li MH, Shi M, Holmes EC, Zhang YZ. 2015. Discovery, diversity and

evolution of novel coronaviruses sampled from rodents in China. Virology 474:19–27. https://doi.org/10.1016/j.virol.2014.10.017.
44. Wu Z, Lu L, Du J, Yang L, Ren X, Liu B, Jiang J, Yang J, Dong J, Sun L, Zhu Y, Li Y, Zheng D, Zhang C, Su H, Zheng Y, Zhou H, Zhu G, Li H, Chmura A, Yang F, Daszak P, Wang J, Liu Q, Jin Q. 2018. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. Microbiome 6:178. https://doi.org/10.1186/s40168-018-0554-9.
45. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
46. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.
47. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300.
48. 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.
49. Woo PC, Lau SK, Fan RY, Lau CC, Wong EY, Joseph S, Tsang AK, Wernery R, Yip CC, Tsang CC, Wernery U, Yuen KY. 17 May 2016, posting date. Isolation and characterization of dromedary camel coronavirus UAE-HKU23 from dromedaries of the Middle East: minimal serological cross-reactivity between MERS coronavirus and dromedary camel coronavirus UAE-HKU23. Int J Mol Sci https://doi.org/10.3390/ijms17050691.