JVI Journals.ASM.org

Rooting the Phylogenetic Tree of Middle East Respiratory Syndrome Coronavirus by Characterization of a Conspecific Virus from an African Bat

Victor Max Corman,[a,e] Ndapewa Laudika Ithete,[b] Leigh Rosanne Richards,[c] M. Corrie Schoeman,[d] Wolfgang Preiser,[b] Christian Drosten,[a] Jan Felix Drexler[a]*

Institute of Virology, University of Bonn Medical Centre, Bonn, Germany[a]; Division of Medical Virology, University of Stellenbosch/National Health Laboratory Service, Tygerberg, South Africa[b]; Durban Natural Science Museum, Durban, South Africa[c]; School of Life Sciences, University of KwaZulu Natal, Durban, South Africa[d]; German Centre for Infection Research (DZIF), Partner Site Bonn-Cologne, Bonn, Germany[e]

ABSTRACT

The emerging Middle East respiratory syndrome coronavirus (MERS-CoV) causes lethal respiratory infections mainly on the Arabian Peninsula. The evolutionary origins of MERS-CoV are unknown. We determined the full genome sequence of a CoV directly from fecal material obtained from a South African *Neoromicia capensis* bat (NeoCoV). NeoCoV shared essential details of genome architecture with MERS-CoV. Eighty-five percent of the NeoCoV genome was identical to MERS-CoV at the nucleotide level. Based on taxonomic criteria, NeoCoV and MERS-CoV belonged to one viral species. The presence of a genetically divergent S1 subunit within the NeoCoV *spike* gene indicated that intraspike recombination events may have been involved in the emergence of MERS-CoV. NeoCoV constitutes a sister taxon of MERS-CoV, placing the MERS-CoV root between a recently described virus from African camels and all other viruses. This suggests a higher level of viral diversity in camels than in humans. Together with serologic evidence for widespread MERS-CoV infection in camelids sampled up to 20 years ago in Africa and the Arabian Peninsula, the genetic data indicate that camels act as sources of virus for humans rather than vice versa. The majority of camels on the Arabian Peninsula is imported from the Greater Horn of Africa, where several *Neoromicia* species occur. The acquisition of MERS-CoV by camels from bats might have taken place in sub-Saharan Africa. Camelids may represent mixing vessels for MERS-CoV and other mammalian CoVs.

IMPORTANCE

It is unclear how, when, and where the highly pathogenic MERS-CoV emerged. We characterized the full genome of an African bat virus closely related to MERS-CoV and show that human, camel, and bat viruses belong to the same viral species. The bat virus roots the phylogenetic tree of MERS-CoV, providing evidence for an evolution of MERS-CoV in camels that preceded that in humans. The revised tree suggests that humans are infected by camels rather than vice versa. Although MERS-CoV cases occur mainly on the Arabian Peninsula, the data from this study together with serologic and molecular investigations of African camels indicate that the initial host switch from bats may have taken place in Africa. The emergence of MERS-CoV likely involved exchanges of genetic elements between different viral ancestors. These exchanges may have taken place either in bat ancestors or in camels acting as mixing vessels for viruses from different hosts.

Human coronaviruses (HCoVs) belong to the genera *Alphacoronavirus* and *Betacoronavirus* within the subfamily *Coronavirinae*. Betacoronaviruses are further divided into four genetic clades, termed clades a to d (1). The genetic diversity of CoVs in bats exceeds that known for any other host, which is compatible with bats being the major reservoir of mammalian CoVs (2).

In 2002 to 2003, an emerging HCoV, termed severe acute respiratory syndrome coronavirus (SARS-CoV), caused a pandemic involving about 8,000 cases, about 10% of whom died. SARS-CoV belonged to *Betacoronavirus* clade b. The evolutionary origins of SARS-CoV involved bat hosts, possibly with civets as intermediate hosts and the source of human infection (3, 4).

In 2012, a novel HCoV, termed Middle East respiratory syndrome CoV (MERS-CoV), was detected in a patient with a fatal respiratory infection in Saudi Arabia (5). As of 11 June 2014, MERS-CoV infection has been diagnosed in 699 patients, mainly from the Arabian Peninsula, with a case fatality rate potentially exceeding the rate observed during the SARS pandemic (6). Because camels on the Arabian Peninsula show high rates of neutralizing antibodies against MERS-CoV and harbor viruses that are genetically highly related to those from human cases, these animals are considered to constitute the source of human infections (7–9). High rates of antibodies against MERS-CoV were recently found in African camels, and a MERS-CoV strain was detected in an Egyptian camel likely imported from Sudan (10–12).

MERS-CoV belongs to *Betacoronavirus* clade c (13, 14). The

Received 23 May 2014 Accepted 12 July 2014
Published ahead of print 16 July 2014
Editor: S. Perlman

Address correspondence Christian Drosten, drosten@virology-bonn.de, or Jan Felix Drexler, drexler@virology-bonn.de.

* Present address: Jan Felix Drexler, Department of Viroscience, Erasmus Medical Centre, Rotterdam, The Netherlands.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.01498-14.

doi:10.1128/JVI.01498-14

prototype clade c betacoronaviruses, termed HKU4 and HKU5, were detected in bats (15). HKU4 and HKU5 form two separate species in genetic sister relationship to MERS-CoV (13). Distinct clade c betacoronaviruses putatively representing another clade c *Betacoronavirus* species (Erinaceus CoV [EriCoV]) were recently described in hedgehogs (16). We and others characterized small genomic sequence fragments of bat CoVs (BtCoVs) that were closely related to MERS-CoV and suggested that MERS-CoV ancestors may have evolved in bats (17–19). Because these sequence fragments encompassed only a few hundred nucleotides from a single gene, the *RNA-dependent RNA polymerase (RdRp)* gene, the evolutionary relationship of these bat CoVs with MERS-CoV could not be conclusively defined (20, 21). A bat virus most likely representing a potential MERS-CoV ancestor was detected in a *Neoromicia capensis* bat from South Africa (22).

In this study, we characterized the full genome of the *Neoromicia* bat CoV, referred to here as NeoCoV. We determined that NeoCoV belongs to the same viral species defined by MERS-CoV based upon established taxonomic criteria (1, 21). Analysis of the NeoCoV genome pointed toward nonrecent recombination events within the MERS-CoV species. The bat virus roots the phylogenetic tree of MERS-CoV and shows that MERS-CoV evolution in camelids likely preceded that in humans.

MATERIALS AND METHODS

Sample processing and full-genome sequencing. A fecal specimen from a *Neoromicia* bat from South Africa was sampled and tested positive for CoVs, as described previously (22). To obtain the full sequence of the NeoCoV genome, 70 heminested reverse transcription-PCR (RT-PCR) assays were developed (primer sequences and PCR conditions are listed in Table S1 in the supplemental material). These assays were designed to amplify about 800 overlapping base pairs of all known MERS-CoV sequences.

Genomic fragments that could not be amplified by these assays were connected by bridging RT-PCR using NeoCoV-specific primers (available upon request) and sequenced by dye terminator chemistry. Determination of genome ends was done by using a rapid amplification of cDNA ends kit (Roche, Penzberg, Germany).

Genomic analyses. Nucleotide and amino acid sequences of predicted open reading frames (ORFs) and the full genome of NeoCoV and related betacoronaviruses were aligned by using MAFFT (23). The pairwise identities of the genome and all ORFs and predicted proteins of NeoCoV were calculated by using MEGA5 (24). Similarity plots of CoV clade c genomes were generated by using SSE v1.1 (25), using a sliding window of 400 and a step size of 40 nucleotides (nt). Phylogenetic analyses of predicted ORFs were done by using MrBayes v3.1 (26), using a GTR+G+I nucleotide or a WAG amino acid substitution model and 2,000,000 generations sampled every 100 steps. Trees were annotated by using the latter 75% of all trees in TreeAnnotator v1.5 and visualized with FigTree v1.4 from the BEAST package (27).

RESULTS AND DISCUSSION

The NeoCoV-positive bat was identified as a female *Neoromicia capensis* (shown in Fig. 1) based on size (forearm length of 35 mm and mass of 5.5 g) and dental and cranial characteristics (28). Typing was confirmed by characterization of the cytochrome *b* and cytochrome oxidase I (COI) genes (GenBank accession numbers KJ756000 and KJ756001), allowing definite species identification as *N. capensis*.

The full NeoCoV genome sequence was obtained directly from fecal material stored in RNAlater by using the panel of PCR assays developed for this study as well as NeoCoV-specific primers. This

FIG 1 *Neoromicia capensis* bat. The absence of a tiny upper premolar separates it from similarly sized *Pipistrellus* and *Hypsugo* bats. The presence of an occipital helmet separates it from *Neoromicia zuluensis*, the species to which it was assigned based on preliminary morphological criteria.

suggests an applicability of this panel for characterizations of genetically diversified MERS-CoVs in further studies. Figure 2A shows a graphical representation of the NeoCoV genome (GenBank accession number KC869678). This genome contained 30,100 nt excluding the poly(A) tail, with a G/C content of 40%. This was comparable to MERS-CoV strains, which range in size from 30,100 to 30,107 nt and have a G/C content of 41%. The number and order of NeoCoV open reading frames (ORFs) were identical to those of MERS-CoV in the order *ORF1ab-spike-ORF3-ORF4ab-ORF5-envelope (E)-membrane (M)-nucleocapsid (N)-ORF8b*. As in MERS-CoV, a ribosomal frameshift site and 16 nonstructural protein (NSP) domains within *ORF1ab* were predicted. Table 1 provides information on the size and genomic location of these NSP domains.

Figure 2A and Table 2 provide details on the predicted ORFs, transcription regulatory sequences (TRSs), and their genomic localizations. In analogy to MERS-CoV, eight putative TRSs with the conserved TRS core motif of clade c betacoronaviruses, AAC GAA, preceded predicted ORFs. The predicted leader TRS core sequence of NeoCoV (TTAACGAACT) and the predicted body TRS core sequences of NeoCoV were completely identical to those of MERS-CoVs. This included TRS core sequences preceding the *E* (AAAACGAACT) and *N* (TTAACGAATC) genes showing minor sequence differences, as observed previously for MERS-CoV (13). As in MERS-CoV, no separate body TRSs preceding the predicted AUG codons of *ORF4b* and *ORF8b* were detected.

Amino acid sequence identity in seven concatenated NSP domains has been established by the International Committee on Taxonomy of Viruses (ICTV) for CoV species demarcation (1, 21). As shown in Table 3, the amino acid sequence identity of these translated domains of NeoCoV was 97.2 to 97.4% compared to MERS-CoV strains. Because this exceeded the 90% threshold defined to separate CoV species (1), NeoCoV and MERS-CoV be-

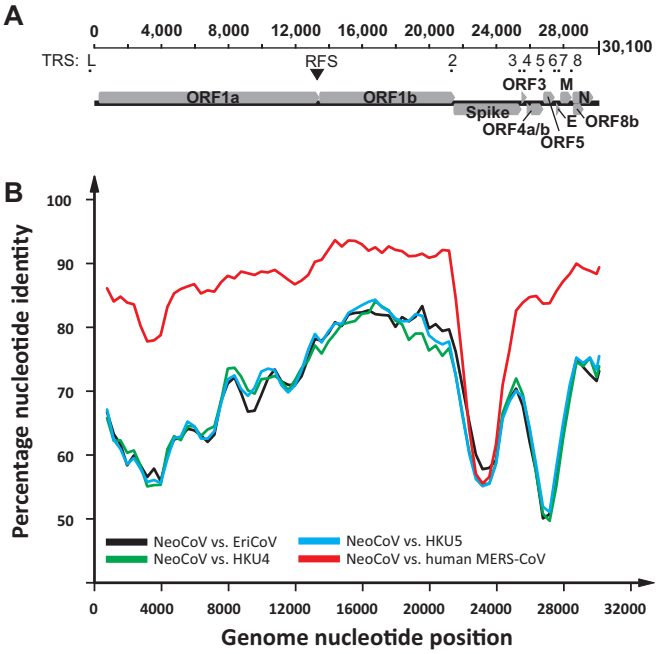


FIG 2 Genome organization of NeoCoV and sequence identity compared to other clade c betacoronaviruses. (A) Genome organization of NeoCoV. The NeoCoV genome is represented by a black line; ORFs are indicated by gray arrows. The ribosomal frameshift site (RFS) is marked with an arrowhead. The locations of transcription regulatory core sequences (TRSs) following the leader (L) are marked by labeled dots and numbered in their order of appearance from the genomic 5′ terminus. (B) Genomic sequence identity between NeoCoV and other clade c betacoronaviruses. Plots were generated by using SSE version 1.1 (25). The graph representing the comparison of the phylogenetically basal camel virus NRCE-HKU205 and NeoCoV is not shown due to a total overlap in the curve resulting from the comparison between NeoCoV and human MERS-CoV.

longed to the same species. The established *Betacoronavirus* clade c species HKU4 and HKU5 and hedgehog CoV shared 85.3 to 88.7% amino acid sequence identity with MERS-CoVs from humans and camels as well as NeoCoV, substantiating the classification of MERS-CoV as a separate species.

To analyze the relationships of NeoCoV with MERS-CoV and other clade c betacoronaviruses beyond the domains used for species delineation, full-genome comparisons were made. NeoCoV shared 85.5% to 85.6% overall nucleotide identity with MERS-CoVs from humans and camels. Nucleotide identity with other clade c betacoronaviruses was considerably lower, at 25.5 to 51.5%. Figure 2B shows that the nucleotide identity between NeoCoV and MERS-CoV decreased in the genomic region encoding the spike glycoprotein. Sequence identity was lower toward the 5′ end of the *spike* ORF than toward its 3′ end. The translated *spike* ORF of NeoCoV showed 64.3 to 64.6% identity with MERS-CoV and 60.5 to 63.6% identity with other clade c betacoronaviruses. Table 2 shows that the amino acid sequence identities between NeoCoV and MERS-CoV were higher for all other ORFs. The genes encoding the structural proteins E, M, and N showed high levels of sequence identity between NeoCoV and MERS-CoV strains, up to 89.0% (E), 94.5% (M) and 91.7% (N). The functions of MERS-CoV ORF3, ORF4ab, ORF5, and ORF8b are poorly understood. Sequence identities between these ORFs from NeoCoV and those from MERS-CoV ranged between 76.5% (ORF3) and 88.4% (ORF5). These identity levels exceed those between any other clade c betacoronavirus and MERS-CoV by 2- to 3-fold. This includes ORF4a, which acts as an interferon antagonist, presumably by interaction with double-stranded RNA in MERS-CoV (29). The 23 amino acid (aa) positions suggested to form a double-stranded RNA-binding domain in the MERS-CoV 4a protein (29) were completely conserved in NeoCoV, suggesting that the predicted NeoCoV protein might have a homologous function.

To confirm the genetic relationships suggested by sequence distance comparisons, Bayesian phylogenies of major ORFs were reconstructed. As shown in Fig. 3A, NeoCoV clustered with a basal sister relationship to a clade containing MERS-CoV from humans and camels in all ORFs except the *spike* ORF. In the *spike* ORF, NeoCoV clustered with European hedgehog CoVs and a *Nycteris* bat CoV from Ghana. To find reasons for the variant tree topology in the *spike* ORF, Bayesian phylogenetic reconstructions and sequence distance analyses were done on two different *spike* ORF data sets. One data set represented subunit S1, which contains the receptor-binding domain (RBD). Another data set represented subunit S2, involved in the fusion of viral and cellular membranes. Figure 3B shows that in the phylogenetic tree of the S1 subunit, NeoCoV clustered distantly from MERS-CoV at the same position as that observed in the tree based on the full *spike* sequence. Accordingly, MERS-CoV and NeoCoV showed only 46.0% amino acid sequence identity in the S1 subunit. In the S2 subunit, NeoCoV shared a monophyletic origin with MERS-CoV in a basal sister relationship to MERS-CoVs from humans and camels, similar to the phylogenetic reconstructions in all other ORFs. Monophyly correlated with a higher degree of amino acid sequence

TABLE 1 Prediction of the putative pp1a/pp1ab cleavage sites of NeoCoV based on sequence comparison with MERS-CoV strain EMC/2012

NSP	1st amino acid residue-last amino acid residue[a]	Protein size (no. of amino acids)	Putative functional domain(s)[b]
NSP1	Met^{1}-Gly^{193}	193	
NSP2	Asp^{194}-Gly^{857}	664	
NSP3	Ala^{858}-Gly^{2743}	1,886	ADRP, PL2pro
NSP4	Ala^{2744}-Gln^{3250}	507	
NSP5	Ser^{3251}-Gln^{3556}	306	3CLpro
NSP6	Ser^{3557}-Gln^{3848}	292	
NSP7	Ser^{3849}-Gln^{3931}	83	
NSP8	Ala^{3932}-Gln^{4130}	198	Primase
NSP9	Asn^{4131}-Gln^{4240}	110	
NSP10	Ala^{4241}-Gln^{4380}	140	
NSP11	Ser^{4381}-Leu^{4394}	14	Short peptide at the end of *ORF1a*
NSP12	Ser^{4381}-Gln^{5314}	933	RdRp
NSP13	Ala^{5315}-Gln^{5912}	598	HEL, NTPase
NSP14	Ser^{5913}-Gln^{6436}	524	ExoN, NMT
NSP15	Gly^{6437}-Gln^{6779}	343	NendoU
NSP16	Ala^{6780}-Arg^{7082}	303	OMT

[a] Superscript numbers indicate positions in polyprotein pp1a/pp1ab or positions in the available sequence with the supposition of a ribosomal frameshift based on the conserved slippery sequence (UUUAAAC) of coronaviruses.
[b] ADRP, ADP-ribose 1-phosphatase; PL2pro, papain-like protease 2; 3CLpro, coronavirus NSP5 protease; Hel, helicase; NTPase, nucleoside triphosphatase; ExoN, exoribonuclease; NMT, N7 methyltransferase; NendoU, endoribonuclease; OMT, 2′-*O*-methyltransferase.

TABLE 2 Coding potential, putative transcription regulatory sequences, and sequence comparison with prototype clade c betacoronaviruses

ORF	nt positions (start-end)	No. of amino acids	Sequence[a]	% amino acid identity[b] of BtCoV/Neoromicia/PML-PHE1/RSA/2011 with:			
				MERS-CoV[c]	HKU4[d]	HKU5[e]	EriCoV[f]
ORF 1ab	281–21528	7,082	$_{00057}$GATCTTAACGAACTTAAA$_{74}$	92.7	73.7	76.0	73.9
Spike	21470–25504	1,344	$_{21412}$CAGATTAACGAACTTGTA$_{21429}$	64.3–64.6	60.5–60.8	61.5	63.6
ORF3	25519–25830	103	$_{25501}$CTAATTAACGAACTTCCA$_{25518}$	76.5–78.4	40.7–44.0	47.9	27.3–28.4
ORF4a	25839–26168	109	$_{25823}$TTAATTAACGAACTCTAT$_{25840}$	87.0–88.0	37.4–38.3	41.9–42.9	40.7
ORF4b	26044–26820	258		83.7–85.4	28.6–29.4	26.8–26.8	39.0
ORF5	26827–27501	224	$_{26813}$GATTTTAACGAACTATGG$_{26830}$	87.1–88.4	47.1	54.8–55.7	51.3
E	27577–27825	82	$_{27563}$ATGGAAAACGAACTATGT$_{27580}$	89.0	73.2–74.4	72.0	76.8
M	27840–28499	219	$_{27818}$GGGTTTAACGAACTCCTT$_{27835}$	93.6–94.5	81.7–82.2	82.6–83.1	80.7–81.2
N	28557–29801	414	$_{28528}$GATCTTAACGAATCTTAA$_{28545}$	91.3–91.7	76.5–76.8	76.0	72.7–73.0
ORF8b	28603–29202	199		81.1–83.9	48.7–50.8	52.6–55.8	45.4–46.4

[a] Underlined type indicates conserved nucleotides of the putative leader TRS core sequence. Subscripted numbers indicate positions in the BtCoV/Neoromicia/PML-PHE1/RSA/2011 genome.
[b] Calculated with MEGA5 (24) using a pairwise deletion option.
[c] GenBank accession numbers JX869059, KC164505, KC776174, KF186567, KF192507, KF600612, KF600620, and KJ477102.
[d] GenBank accession numbers EF065506, EF065507, EF065508, and DQ648794.
[e] GenBank accession numbers EF065505, EF065509, EF065510, EF065511, and EF065512.
[f] GenBank accession numbers KC545383 and KC545386.

identity (87.2%) between MERS-CoV and NeoCoV in the S2 subunit.

These data suggested that the human-pathogenic MERS-CoV variant might be the result of nonrecent recombination events involving as-yet-unknown partners. Typical recombination breakpoints in CoV genomes encompass the *spike* gene (2). Intraspike recombination between the S1 and S2 subunits has been hypothesized to be the major mechanism involved in the emergence of SARS-CoV from bat and civet ancestors (30). The location of RBDs at either the N or the C termini of the S1 subunits of HCoV-229E and mouse hepatitis virus (MHV) has been interpreted as further evidence for the interchangeability of *spike* subunits (30).

The different S1 subunit suggested that NeoCoV was not the direct ancestor of MERS-CoV. Of note, CoVs are mostly associated with chiropteran hosts on the genus level (2), and MERS-CoV was shown to infect cells from vespertilionid bats (31). According to this principle, MERS-CoV variants carrying a *spike* gene closely related to human-pathogenic MERS-CoV may exist in bats belonging to the family *Vespertilionidae* and specifically the genus *Neoromicia*. This scenario parallels the *spike* gene diversity found in bat CoV ancestors of SARS-CoV. All ancestral bat SARS-related CoVs described since 2005 have had highly diversified *spike* genes, which differed from human SARS-CoV by about 20% of their amino acid sequences (4). Only recently, a bat virus carrying a *spike* gene related to human SARS-CoV and capable of using the SARS-CoV receptor molecule ACE2 was found (3). In agreement with the principle that chiropteran hosts can harbor closely related CoVs, the bat CoV carrying the human SARS-CoV-related *spike* gene occurred in *Rhinolophus sinicus*, which is the same bat species yielding multiple SARS-related CoV lineages carrying divergent *spike* genes since 2005 (2).

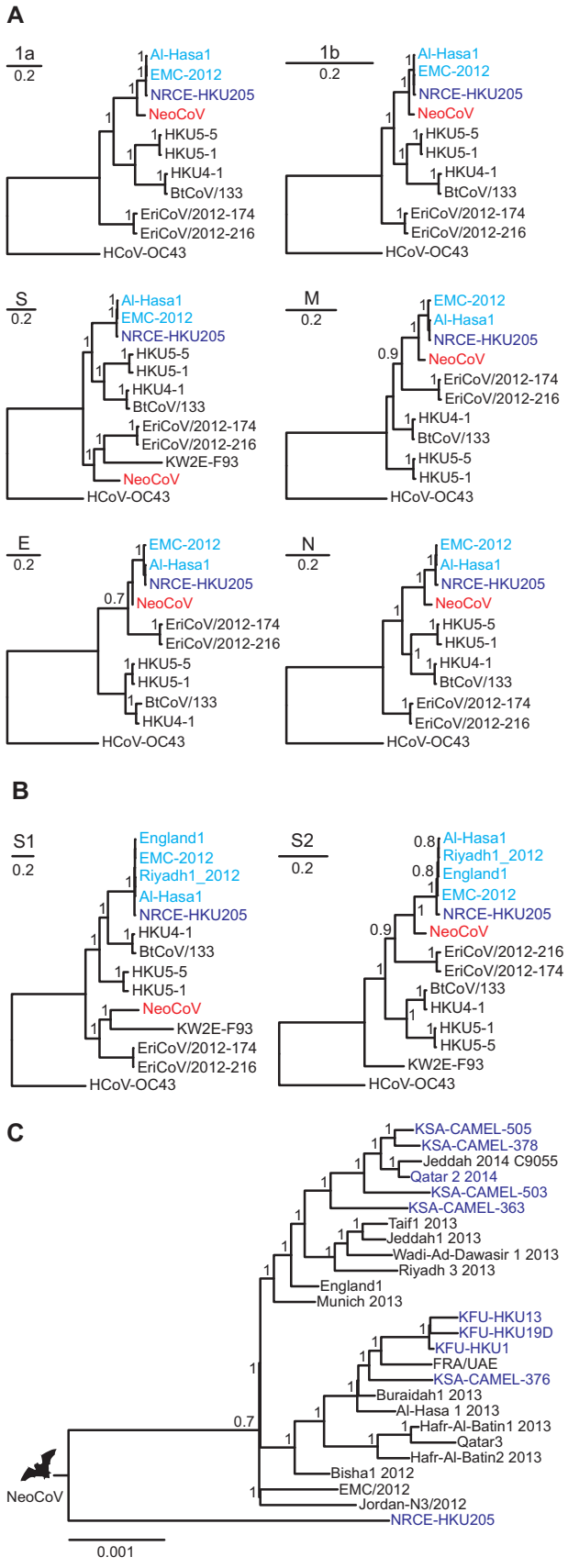
TABLE 3 Comparison of amino acid identities of seven conserved replicase domains of NeoCoV for species classification

BtCoV/Neoromicia/PML-PHE1/RSA/2011 domain	% amino acid identity with[a]:			
	MERS-CoV[b]	HKU4[c]	HKU5[d]	EriCoV[e]
ADRP	90.6	59.1–60.4	63.1	67.5
NSP5 (3CLpro)	96.4–97.1	80.4–80.7	82.4–83.0	78.4–79.1
NSP12 (RdRp)	98.6–98.7	89.9–90.0	92.5	89.1–89.5
NSP13 (Hel, NTPase)	98.5–98.7	91.8–92.3	94.5	91.0–91.1
NSP14 (ExoN, NMT)	97.9–98.3	85.7–86.6	91.8–92.0	89.1–89.5
NSP15 (NendoU)	93.6–94.5	77.2–77.8	81.9–82.2	81.9–82.5
NSP16 (OMT)	96.4	83.4	87.1	86.8–87.8
Concatenated domains	97.2–97.4	85.3–85.4	88.7	86.5

[a] Calculated with MEGA5 (24) using a pairwise deletion option.
[b] Including GenBank accession numbers JX869059, KC164505, KC776174, KF186567, KF192507, KF600612, KF600620, and KJ477102.
[c] Including GenBank accession numbers EF065506, EF065507, EF065508, and DQ648794.
[d] Including GenBank accession numbers EF065505, EF065509, EF065510, EF065511, and EF065512.
[e] Including GenBank accession numbers KC545383 and KC545386.

Because NeoCoV clustered with a basal sister relationship to MERS-CoV in all ORFs and because NeoCoV and MERS-CoV belonged to one virus species, the bat virus can be used to infer the root of the phylogenetic tree of MERS-CoV. Figure 3C shows a Bayesian phylogenetic reconstruction of all available MERS-CoV full genomes from camels and representative MERS-CoV full genomes from humans, rooted by NeoCoV. The MERS-CoV detected in an African camel, termed NRCE-HKU205 (10), clustered with high statistical support in basal sister relationship to MERS-CoVs from humans and camels from the Arabian Peninsula. Despite the phylogenetic clustering of NRCE-HKU205 in an intermediate position between NeoCoV and MERS-CoVs from the Arabian Peninsula, the maximum nucleotide distance of NeoCoV and NRCE-HKU205 did not differ from the maximum nucleotide distance of NeoCoV and the other MERS-CoVs, at 14.5% and 14.4 to 14.5%, respectively. The maximum nucleotide distance within MERS-CoVs from camels was 0.6%. This was slightly higher than the 0.4% maximum distance within MERS-CoVs from humans, although only 11 near-full-length genomes are available for MERS-CoVs from camels, compared to 38 near-full-length genomes for MERS-CoVs from humans.

The absence of more diversified MERS-CoV sequences in humans, particularly outside the Arabian Peninsula, cannot be fully



excluded. However, the phylogenetic position of the outlier camel MERS-CoV and the slightly higher level of genetic diversity in camels suggest that the evolution of MERS-CoV in camels preceded that in humans and that camels represent donors of viruses for humans rather than vice versa. These genetic data are corroborated by the existence of specific antibodies against MERS-CoV in up to 90% of camelids from the Arabian Peninsula and Africa for at least 20 years (7, 9, 11, 12). On the contrary, sera from children with respiratory disease sampled in 2010 to 2011 in Saudi Arabia (32) and sera from blood donors and slaughterhouse workers sampled in Saudi Arabia in 2012 (33) showed no evidence of antibodies against MERS-CoV.

Interestingly, the camel yielding the genetic outlier MERS-CoV was likely imported from Sudan into Egypt (10). The majority of camels in the Arabian Peninsula are imported from countries in the Greater Horn of Africa, such as Somalia, Sudan, and Kenya (10, 11). We and others have recently shown that MERS-CoV-neutralizing antibodies occur frequently in camels from eastern Africa (11, 12). A hypothetical scenario might thus imply a spillover of viruses from bats to camels in the Greater Horn of Africa. Of note, 10 out of 11 *Neoromicia* species listed by the International Union for Conservation of Nature (http://www.iucnredlist.org/), including *N. capensis*, are common in this region. An alternative scenario implying an exchange of viruses between *Neoromicia* bats and camels on the Arabian Peninsula is unlikely, because *Neoromicia* bats are not known to occur in this region.

The deep branches leading to NeoCoV in phylogenetic reconstructions and the differences observed in the *spike* gene of this bat virus suggest that this putative host switch may have occurred nonrecently. Although NeoCoV formed one viral species together with MERS-CoVs from humans and camels, unidirectional host switching events from bat hosts to other mammalian hosts such as camels and different evolutionary histories in these hosts can be assumed. The bat, camel, and human hosts of this CoV species are thus unlikely to fulfill the population criteria required for coalescent dating. Additionally, dating of CoV branches deeper than

FIG 3 Bayesian phylogenies of clade c betacoronaviruses, including NeoCoV. (A) Phylogenies of *ORF1a, ORF1b*, and ORFs coding for structural proteins. (B) Phylogenies of the S1 and S2 subunits, corresponding to amino acid positions 1 to 747 and 748 to 1353, respectively, of MERS-CoV strain EMC/2012. NeoCoV is shown in red, camel MERS-CoV is shown in blue, and human MERS-CoV is shown in cyan. HCoV-OC43 was used as an outgroup. (C) Phylogeny of MERS-CoV full genomes. MERS-CoVs obtained from humans are shown in black, and MERS-CoVs from camels are shown in blue. NeoCoV was used for rooting the tree. For all trees, statistical support of grouping from Bayesian posterior probabilities is shown at deep nodes. Only values above 0.7 are shown. The bar represents genetic distance. GenBank accession numbers are KJ477102 for NRCE-HKU205, KJ156881 for Wadi-Ad-Dawasir 1 2013, JX869059 for EMC/2012, KJ650296 for KFU-HKU19D, KC776174 for Jordan-N3/2012, KJ650297 for KFU-HKU1, KJ156910 for Hafr-Al-Batin2 2013, KF600613 for Riyadh 3 2013, KF186567 for Al-Hasa 1 2013, KC164505 for England1, KF961221 for Qatar3, KJ713299 for KSA-CAMEL-376, KJ156949 for Taif1 2013, KJ556336 for Jeddah1 2013, KJ713297 for KSA-CAMEL-503, KJ713295 for KSA-CAMEL-505, KF192507 for Munich 2013, KJ650098 for Qatar 2 2014, KF745068 for FRA/UAE, KF600630 for Buraidah1 2013, KJ650295 for KFU-HKU13, KF600628 for Hafr-Al-Batin1 2013, KJ713298 for KSA-CAMEL-363, KJ713296 for KSA-CAMEL-378, KF600620 for Bisha1 2012, KC869678 for NeoCoV, NC_005147 for HCoV-OC43, EF065512 for HKU5-5, NC_009020 for HKU5-1, NC_008315 for BtCoV/133, NC_009019 for HKU4-1, KC545386 for EriCoV/2012-216, KC545383 for EriCoV/2012-174, and KM027259 for Jeddah 2014 C9055.

0.05 substitutions per site, such as those leading to the predicted ancestor of NeoCoV and MERS-CoV, was found to be highly unreliable and may greatly underestimate the true evolutionary history of CoVs (34). Therefore, and because of the evidence for recombination detected in the NeoCoV *spike* gene, no molecular dating of the projected ancestor shared by NeoCoV and other MERS-CoVs was conducted.

The putative ancient recombination events giving rise to MERS-CoV may have taken place in two candidate hosts to be explored. The first and more likely option is vespertilionid bats. It would be highly relevant to fully characterize additional bat viruses from Africa and the Arabian Peninsula. Alternatively, camelids may represent a putative mixing vessel, similar to the role of swine in influenza A viruses (35). Sera from camelids should be tested for antibodies against the NeoCoV S1 subunit to gather evidence for infection with this CoV lineage. In parallel, camelids should be screened to identify genetically diversified viruses potentially related to their putative bat ancestors. The putative role of camelids as recipients of CoVs from other mammalian hosts is supported by the occurrence of viruses in camelids that are closely related to human CoV-229E and bovine CoV (36, 37).

ACKNOWLEDGMENTS

We thank Monika Eschbach-Bludau, Tobias Bleicker, and Sebastian Brünink (Institute of Virology, Bonn) for excellent technical assistance.

This study was supported by funds from European Union FP7 projects EMPERIE (contract number 223498) and ANTIGONE (contract number 278976) granted to C.D. and by funds from the National Health Laboratory Service Research Trust, the South African National Research Foundation, and the German Research Foundation (project numbers KR1293/13-1 and KR1293/9-1) granted to W.P.

REFERENCES

1. **de Groot RJ, Baker SC, Baric R, Enjuanes L, Gorbalenya AE, Holmes KV, Perlman S, Poon L, Rottier PJM, Talbot PJ, Woo PCY, Ziebuhr J.** 2012. Family *Coronaviridae*, p 806–820. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy: classification and nomenclature of viruses. Ninth report of the International Committee on Taxonomy of Viruses. Academic Press, London, United Kingdom.
2. **Drexler JF, Corman VM, Drosten C.** 2014. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. Antiviral Res. **101:**45–56. http://dx.doi.org/10.1016/j.antiviral.2013.10.013.
3. **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. http://dx.doi.org/10.1038/nature12711.
4. **Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S, Wang LF.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science **310:**676–679. http://dx.doi.org/10.1126/science.1118391.
5. **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. http://dx.doi.org/10.1056/NEJMoa1211721.
6. **WHO.** 11 June 2014. MERS-CoV update 16: summary and literature update as of 11 June 2014. WHO, Geneva, Switzerland. http://www.who.int/csr/don/2014_06_16_mers/en/.
7. **Meyer B, Muller MA, Corman VM, Reusken CB, Ritz D, Godeke GJ, Lattwein E, Kallies S, Siemens A, van Beek J, Drexler JF, Muth D, Bosch BJ, Wernery U, Koopmans MP, Wernery R, Drosten C.** 2014. Antibodies against MERS coronavirus in dromedary camels, United Arab Emirates, 2003 and 2013. Emerg. Infect. Dis. **20:**552–559. http://dx.doi.org/10.3201/eid2004.131746.
8. **Memish ZA, Cotten M, Meyer B, Watson SJ, Alsahafi AJ, Al Rabeeah AA, Corman VM, Sieberg A, Makhdoom HQ, Assiri A, Al Asmari M, Aldabbagh S, Bosch BJ, Beer M, Müller MA, Kellam P, Drosten C.** 2014. Human infection with MERS coronavirus after exposure to infected camels, Saudi Arabia, 2013. Emerg. Infect. Dis. **20:**1012–1015. http://dx.doi.org/10.3201/eid2006.140402.
9. **Alagaili AN, Briese T, Mishra N, Kapoor V, Sameroff SC, de Wit E, Munster VJ, Hensley LE, Zalmout IS, Kapoor A, Epstein JH, Karesh WB, Daszak P, Mohammed OB, Lipkin WI.** 2014. Middle East respiratory syndrome coronavirus infection in dromedary camels in Saudi Arabia. mBio **5**(2):e00884-14. http://dx.doi.org/10.1128/mBio.00884-14.
10. **Chu DK, Poon LL, Gomaa MM, Shehata MM, Perera RA, Abu Zeid D, El Rifay AS, Siu LY, Guan Y, Webby RJ, Ali MA, Peiris M, Kayali G.** 2014. MERS coronaviruses in dromedary camels, Egypt. Emerg. Infect. Dis. **20:**1049–1053. http://dx.doi.org/10.3201/eid2006.140299.
11. **Corman VM, Jores J, Meyer B, Younan M, Liljander A, Said MY, Gluecks I, Lattwein E, Bosch BJ, Drexler JF, Bornstein S, Drosten C, Müller MA.** MERS coronavirus antibody prevalence in dromedary camels, Kenya, 1992-2013. Emerg. Infect. Dis. **20:**1319–1322. http://dx.doi.org/10.3201/eid2008.140596.
12. **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 MPG.** Geographic distribution of MERS coronavirus among dromedary camels, Africa. Emerg. Infect. Dis. **20:**1370–1374. http://dx.doi.org/10.3201/eid2008.140590.
13. **van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, Osterhaus AD, Haagmans BL, Gorbalenya AE, Snijder EJ, Fouchier RA.** 2012. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. mBio **3**(6):e00473-12. http://dx.doi.org/10.1128/mBio.00473-12.
14. **de Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, Fouchier RA, Galiano M, Gorbalenya AE, Memish ZA, Perlman S, Poon LL, Snijder EJ, Stephens GM, Woo PC, 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. http://dx.doi.org/10.1128/JVI.01244-13.
15. **Woo PC, Wang M, Lau SK, Xu H, Poon RW, Guo R, Wong BH, Gao K, Tsoi HW, Huang Y, Li KS, Lam CS, 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. http://dx.doi.org/10.1128/JVI.02182-06.
16. **Corman VM, Kallies R, Philipps H, Gopner G, Muller MA, Eckerle I, Brunink S, Drosten C, Drexler JF.** 2014. Characterization of a novel betacoronavirus related to Middle East respiratory syndrome coronavirus in European hedgehogs. J. Virol. **88:**717–724. http://dx.doi.org/10.1128/JVI.01600-13.
17. **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 EK, Lina PH, 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. http://dx.doi.org/10.3201/eid1903.121503.
18. **Reusken CB, Lina PH, 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. http://dx.doi.org/10.1089/vbz.2009.0173.
19. **De Benedictis P, Marciano S, Scaravelli D, Priori P, Zecchin B, Capua I, Monne I, Cattoli G.** 2014. Alpha and lineage C betaCoV infections in Italian bats. Virus Genes **48:**366–371. http://dx.doi.org/10.1007/s11262-013-1008-x.
20. **Gorbalenya AE, Snijder EJ, Spaan WJ.** 2004. Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. J. Virol. **78:**7863–7866. http://dx.doi.org/10.1128/JVI.78.15.7863-7866.2004.
21. **Lauber C, Gorbalenya AE.** 2012. Partitioning the genetic diversity of a virus family: approach and evaluation through a case study of picornaviruses. J. Virol. **86:**3890–3904. http://dx.doi.org/10.1128/JVI.07173-11.
22. **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. http://dx.doi.org/10.3201/eid1910.130946.
23. **Katoh K, Misawa K, Kuma K, Miyata T.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. **30:**3059–3066. http://dx.doi.org/10.1093/nar/gkf436.

24. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.
25. **Simmonds P.** 2012. SSE: a nucleotide and amino acid sequence analysis platform. BMC Res. Notes **5:**50. http://dx.doi.org/10.1186/1756-0500-5-50.
26. **Ronquist F, Huelsenbeck JP.** 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:**1572–1574. http://dx.doi.org/10.1093/bioinformatics/btg180.
27. **Drummond AJ, Suchard MA, Xie D, Rambaut A.** 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. **29:**1969–1973. http://dx.doi.org/10.1093/molbev/mss075.
28. **Monadjem A, Taylor PJ, Cotterill W, Schoeman MC.** 2010. Bats of southern and central Africa: a biogeographic and taxonomic synthesis. Wits University Press, Johannesburg, South Africa.
29. **Niemeyer D, Zillinger T, Muth D, Zielecki F, Horvath G, Suliman T, Barchet W, Weber F, Drosten C, Muller MA.** 2013. Middle East respiratory syndrome coronavirus accessory protein 4a is a type I interferon antagonist. J. Virol. **87:**12489–12495. http://dx.doi.org/10.1128/JVI.01845-13.
30. **Graham RL, Baric RS.** 2010. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J. Virol. **84:**3134–3146. http://dx.doi.org/10.1128/JVI.01394-09.
31. **Bosch BJ, Raj VS, Haagmans BL.** 2013. Spiking the MERS-coronavirus receptor. Cell Res. **23:**1069–1070. http://dx.doi.org/10.1038/cr.2013.108.
32. **Gierer S, Hofmann-Winkler H, Albuali WH, Bertram S, Al-Rubaish AM, Yousef AA, Al-Nafaie AN, Al-Ali AK, Obeid OE, Alkharsah KR, Pohlmann S.** 2013. Lack of MERS coronavirus neutralizing antibodies in humans, eastern province, Saudi Arabia. Emerg. Infect. Dis. **19:**2034–2036. http://dx.doi.org/10.3201/eid1912.130701.
33. **Aburizaiza AS, Mattes FM, Azhar EI, Hassan AM, Memish ZA, Muth D, Meyer B, Lattwein E, Muller MA, Drosten C.** 2014. Investigation of anti-Middle East respiratory syndrome antibodies in blood donors and slaughterhouse workers in Jeddah and Makkah, Saudi Arabia, fall 2012. J. Infect. Dis. **209:**243–246. http://dx.doi.org/10.1093/infdis/jit589.
34. **Wertheim JO, Chu DK, Peiris JS, Kosakovsky Pond SL, Poon LL.** 2013. A case for the ancient origin of coronaviruses. J. Virol. **87:**7039–7045. http://dx.doi.org/10.1128/JVI.03273-12.
35. **Ma W, Lager KM, Vincent AL, Janke BH, Gramer MR, Richt JA.** 2009. The role of swine in the generation of novel influenza viruses. Zoonoses Public Health **56:**326–337. http://dx.doi.org/10.1111/j.1863-2378.2008.01217.x.
36. **Crossley BM, Mock RE, Callison SA, Hietala SK.** 2012. Identification and characterization of a novel alpaca respiratory coronavirus most closely related to the human coronavirus 229E. Viruses **4:**3689–3700. http://dx.doi.org/10.3390/v4123689.
37. **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. http://dx.doi.org/10.1016/j.virol.2007.03.035.