Journal of Virology, Jan. 2009, p. 908–917 Vol. 83, No. 2
0022-538X/09/$08.00+0 doi:10.1128/JVI.01977-08

Comparative Analysis of Complete Genome Sequences of Three Avian Coronaviruses Reveals a Novel Group 3c Coronavirus▽

Patrick C. Y. Woo,[1,2,3,4]† Susanna K. P. Lau,[1,2,3,4]† Carol S. F. Lam,[4] Kenneth K. Y. Lai,[4] Yi Huang,[4] Paul Lee,[4] Geraldine S. M. Luk,[5] Kitman C. Dyrting,[5] Kwok-Hung Chan,[4] and Kwok-Yung Yuen[1,2,3,4]*

State Key Laboratory of Emerging Infectious Diseases,[1] Research Centre of Infection and Immunology,[2] Carol Yu Centre for Infection,[3] and Department of Microbiology,[4] The University of Hong Kong, and Agriculture, Fisheries, and Conservation Department,[5] Hong Kong

Received 18 September 2008/Accepted 20 October 2008

In this territory-wide molecular epidemiology study of coronaviruses (CoVs) in Hong Kong involving 1,541 dead wild birds, three novel CoVs were identified in three different bird families (bulbul CoV HKU11 [BuCoV HKU11], thrush CoV HKU12 [ThCoV HKU12], and munia CoV HKU13 [MuCoV HKU13]). Four complete genomes of the three novel CoVs were sequenced. Their genomes (26,396 to 26,552 bases) represent the smallest known CoV genomes. In phylogenetic trees constructed using chymotrypsin-like protease (3CL^{pro}), RNA-dependent RNA polymerase (Pol), helicase, spike, and nucleocapsid proteins, BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 formed a cluster distantly related to infectious bronchitis virus and turkey CoV (group 3a CoVs). For helicase, spike, and nucleocapsid, they were also clustered with a CoV recently discovered in Asian leopard cats, for which the complete genome sequence was not available. The 3CL^{pro}, Pol, helicase, and nucleocapsid of the three CoVs possessed higher amino acid identities to those of group 3a CoVs than to those of group 1 and group 2 CoVs. Unique genomic features distinguishing them from other group 3 CoVs include a distinct transcription regulatory sequence and coding potential for small open reading frames. Based on these results, we propose a novel CoV subgroup, group 3c, to describe this distinct subgroup of CoVs under the group 3 CoVs. Avian CoVs are genetically more diverse than previously thought and may be closely related to some newly identified mammalian CoVs. Further studies would be important to delineate whether the Asian leopard cat CoV was a result of interspecies jumping from birds, a situation analogous to that of bat and civet severe acute respiratory syndrome CoVs.

Coronaviruses (CoVs) are found in a wide variety of animals in which they can cause respiratory, enteric, hepatic, and neurological diseases of varying severity. Based on genotypic and serological characterization, CoVs have been divided into three distinct groups (3, 21, 50). As a result of the unique mechanism of viral replication, CoVs have a high frequency of recombination (21). Their tendency for recombination and high mutation rates may allow them to adapt to new hosts and ecological niches (17, 47).

The recent severe acute respiratory syndrome (SARS) epidemic, the discovery of SARS-CoV, and the identification of SARS-CoV-like viruses from Himalayan palm civets and a raccoon dog from wildlife markets in China have boosted interest in the discovery of novel CoVs in both humans and animals (5, 15, 29, 32, 35, 36, 45). For human CoVs (HCoVs), a novel group 1 HCoV, HCoV-NL63, was reported independently by two groups in 2004 (11, 40). In 2005, we also described the discovery, complete genome sequence, clinical features, and molecular epidemiology of another novel group 2 HCoV, HCoV-HKU1 (23, 41, 42, 46). As for animal CoVs, we and others have described the discovery of SARS-CoV-like viruses in horseshoe bats in Hong Kong Special Administrative Region (HKSAR) and other provinces of China (22, 26). Based on these findings, we conducted molecular surveillance studies to examine the diversity of CoVs in bats of our locality, as well as of the Guangdong province of southern China where the SARS epidemic originated and wet markets and game food restaurants serving bat dishes are commonly found. In these studies, at least nine other novel CoVs were discovered, including two novel subgroups of CoVs, groups 2c and 2d (24, 33, 43, 48). Other groups have also conducted molecular surveillance studies in bats and other animals, and additional novel CoVs were discovered and complete genomes sequenced (4, 6, 7, 8, 9, 12, 13, 14, 16, 20, 27, 28, 39, 49). Recently, beluga whale CoV (SW1), a novel CoV most closely related to infectious bronchitis virus (IBV), was discovered in a dead whale (30).

Birds are the reservoir of major emerging viruses, most notably, avian influenza viruses (25). Due to their ability to fly over long distances, birds have the potential to disseminate these emerging viruses efficiently. As for CoVs, the number of known CoVs in birds is relatively small in comparison to the number in bats. Therefore, we hypothesized that previously unrecognized CoVs may be present in birds. To test this hypothesis, we carried out a territory-wide molecular epidemiology study of CoVs in dead wild birds in HKSAR. In this study, three previously undescribed CoVs (bulbul CoV HKU11 [BuCoV HKU11], thrush CoV HKU12 [ThCoV HKU12], and munia CoV HKU13 [MuCoV HKU13]), which form a unique group of CoV distantly related to IBV, were discovered. In

* Corresponding author. Mailing address: State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong. Phone: (852) 28554892. Fax: (852) 28551241. E-mail: hkumicro@hkucc.hku.hk.

† P. C. Y. Woo and S. K. P. Lau contributed equally to the manuscript.

▽ Published ahead of print on 29 October 2008.

addition, we sequenced two complete genomes of BuCoV HKU11 and one complete genome each of ThCoV HKU12 and MuCoV HKU13. Based on the results of the present study, we propose a novel subgroup, group 3c, in the group 3 CoVs.

MATERIALS AND METHODS

Dead wild bird surveillance and sample collection. The Department of Agriculture, Fisheries, and Conservation (AFCD), HKSAR, provided access to samples collected from various locations in HKSAR over a 7-month period (December 2006 to June 2007) as part of the AFCD avian influenza surveillance program on dead wild birds. Tracheal and cloacal swabs were collected from these birds by the Tai Lung Veterinary Laboratory, AFCD, using procedures described previously (10). A total of 1,548 samples from 1,541 dead wild birds of 77 different species in 32 families were tested.

RNA extraction. Viral RNA was extracted from the tracheal and cloacal swabs by using an RNeasy Mini spin column (QIAgen, Hilden, Germany). The RNA was eluted in 50 μl of RNase-free water and was used as the template for reverse transcription-PCR (RT-PCR).

RT-PCR of the *pol*

TABLE 1. Bird species and associated CoVs in the present surveillance study

Bird family name	Scientific name	Common name	No. of birds tested	No. (%) of birds positive for CoVs	CoV
Accipitridae			5	0 (0)	
Alcedinidae			1	0 (0)	
Anatidae			1	0 (0)	
Ardeidae			11	0 (0)	
Cacatuidae			5	0 (0)	
Chloropseidae			1	0 (0)	
Columbidae			253	0 (0)	
Corvidae			19	0 (0)	
Cuculidae			3	0 (0)	
Emberizidae			4	0 (0)	
Estrildidae	*Lonchura atricapilla*	Chestnut munia	10	0 (0)	
	Lonchura punctulata	Scaly-breasted munia	35	1 (2.9)	MuCoV HKU13
	Lonchura striata	White-rumped munia	82	1 (1.2)	MuCoV HKU13
Fringillidae			3	0 (0)	
Hirundinidae			1	0 (0)	
Motacillidae			5	0 (0)	
Muscicapidae			30	0 (0)	
Nectariniidae			6	0 (0)	
Passeridae			85	0 (0)	
Phalacrocoracidae			1	0 (0)	
Phasianidae			4	0 (0)	
Phylloscopidae			1	0 (0)	
Podicipedidae			1	0 (0)	
Psittacidae			6	0 (0)	
Pycnonotidae	*Hemixos castanonotus*	Chestnut bulbul	19	0 (0)	
	Pycnonotus jocosus	Red-whiskered bulbul	178	5 (2.8)	BuCoV HKU11
	Pycnonotus melanicterus	Black-crested bulbul	1	0 (0)	
	Pycnonotus sinensis	Chinese bulbul	242	10 (4.1)	BuCoV HKU11
Rallidae			7	0 (0)	
Recurvirostridae			4	0 (0)	
Scolopacidae			1	0 (0)	
Strigidae			1	0 (0)	
Sturnidae			15	0 (0)	
Sylviidae			1	0 (0)	
Timaliidae			14	0 (0)	
Turdidae	*Myiophonus caeruleus*	Blue whistling thrush	4	0 (0)	
	Monticola solitarius	Blue rock thrush	2	0 (0)	
	Turdus cardis	Japanese thrush	30	0 (0)	
	Turdus hortulorum	Gray-backed thrush	221	3 (1.4)	ThCoV HKU12
	Turdus obscurus	Eyebrowed thrush	1	0 (0)	
	Turdus kessleri	White-backed thrush	1	0 (0)	
	Turdus merula	Blackbird	16	1 (6.3)	ThCoV HKU12
	Turdus pallidus	Pale thrush	66	0 (0)	
	Zoothera dauma	Scaly thrush	22	0 (0)	
Zosteropidae			122	0 (0)	

HCoV-HKU1, bovine CoV, HCoV-OC43, mouse hepatitis virus [MHV], porcine hemagglutinating encephalomyelitis virus, SARS-CoV, and bat SARS-CoV to 52 bases in IBV). Alternatively, the S in these genomes may be using other potential imperfect TRS's located closer to the corresponding AUG (Table 3). In the genomes of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13, one ORF (NS6) between M and N and three ORFs (NS7a, -7b, and -7c) downstream from N, which encode putative nonstructural proteins, were observed. Among these four ORFs, only NS7b was preceded by a TRS. The absence of a TRS preceding NS6 and the relatively high (0.500) K_a/K_s ratio (the ratio between the number of nonsynonymous substitutions per nonsynonymous site and the number of synonymous substitutions per synonymous site; an index of the action of selective forces) of NS6 in BuCoV HKU11 (data not shown) implied that this ORF may not be expressed.

BLAST search revealed no amino acid similarities between these four putative nonstructural proteins and other known proteins, and no functional domain was identified by PFAM and InterProScan. TMHMM and TMpred analyses showed one putative transmembrane domain in NS7b of BuCoV HKU11 (residues 42 to 64 by TMHMM and 41 to 62 by TMpred analysis),

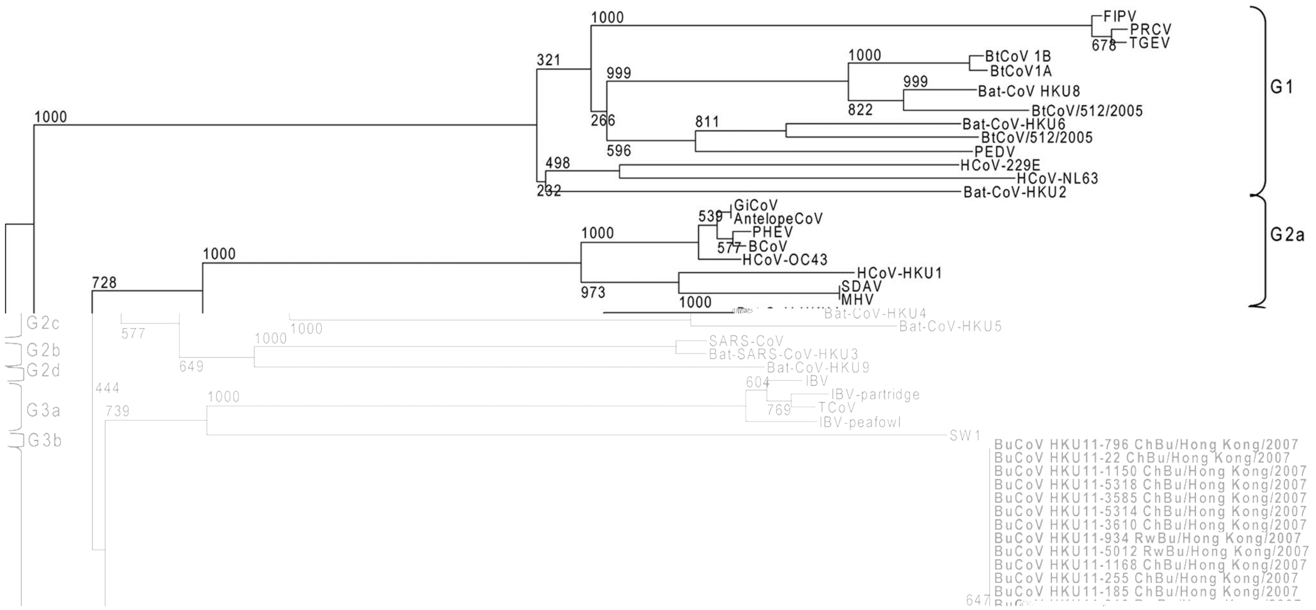

FIG. 1. Phylogenetic analysis of amino acid sequences of the 393-bp fragment (excluding primer sequences) of RNA-dependent RNA Pol of CoVs identified from dead wild birds in the present study. The tree was constructed by the neighbor joining method using Kimura's two-parameter correction and bootstrap values calculated from 1,000 trees. The scale bar indicates the estimated number of substitutions per 50 amino acids. The three novel CoVs identified in the present study are shown in bold. The four genomes completely sequenced are highlighted in gray. HCoV-229E (NC_002645); PEDV, porcine epidemic diarrhea virus (NC_003436); TGEV, porcine transmissible gastroenteritis virus (NC_002306); FIPV, feline infectious peritonitis virus (AY994055); PRCV, porcine respiratory CoV (DQ811787); HCoV-NL63 (NC_005831); bat-CoV HKU2 (EF203064); bat-CoV HKU4 (NC_009019); bat-CoV HKU5 (NC_009020); bat-CoV HKU6 (DQ249224); bat-CoV HKU7 (DQ249226); bat-CoV HKU8 (NC_010438); bat-CoV HKU9 (NC_009021) BtCoV 1A, bat CoV 1A (NC_010437); BtCoV 1B, bat CoV 1B 1B (NC_010436); BtCoV/512/2005, bat CoV 512/2005 (NC_009657); HCoV-HKU1 (NC_006577); HCoV-OC43 (NC_005147); MHV (NC_006852); BCoV, bovine CoV (NC_003045); SDAV, rat sialodacryoadenitis CoV (AF124990); AntelopeCoV, sable antelope CoV (EF424621); GiCoV, giraffe CoV (EF424622); PHEV, porcine hemagglutinating encephalomyelitis virus (NC_007732); SARS-CoV (NC_004718); bat-SARS-CoV HKU3 (NC_009694); IBV (NC_001451); IBV-partridge, partridge CoV (AY646283); IBV-peafowl, peafowl CoV (AY641576); TCoV (NC_010800); ALC, Asian leopard cat CoV Guangxi/F230/2006 (EF584908); SW1 (NC_010646). BuCoV HKU11 (ChBu, chinese Bulbul, and RwBu, red-whiskered bulbul); ThCoV HKU12 (BlTh, blackbird, and GbTh, gray-backed thrush); MuCoV HKU13 (WrMu, white-rumped munia, and SbMu, scaly-breasted munia).

ThCoV HKU12 (residues 43 to 65 by TMHMM and 41 to 62 by TMpred analysis), and MuCoV HKU13 (residues 43 to 65 by TMHMM and 46 to 65 by TMpred analysis). TMHMM and TMpred analyses showed two putative transmembrane domains in NS7c of BuCoV HKU11 (residues 2 to 19 and 24 to 46 by TMHMM and 1 to 17 and 31 to 47 by TMpred analysis), one putative transmembrane domain in NS7c of ThCoV HKU12 (residues 10 to 32 by TMHMM and 10 to 30 by TMpred analysis) and two putative transmembrane domains in NS7c of MuCoV HKU13 (residues 2 to 19 and 29 to 48 by TMHMM and 1 to 19 and 31 to 47 by TMpred analysis). Each of the genomes of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 contains a stem-loop II motif (s2m) (bases 26262 to 26304, 26175 to 26217, and 26390 to 26432, respectively) as a conserved RNA element downstream from N and upstream from the poly(A) tail, similar to those in IBV, TCoV, bat SARS-CoV, and SARS-CoV, as well as other in CoVs discovered in Asian leopard cats, graylag geese, feral pigeons, and mallards but without complete genomes available (Fig. 3) (13, 20, 34).

Phylogenetic analyses. The phylogenetic trees constructed using the amino acid sequences of the $3CL^{pro}$, Pol, helicase, S, and N proteins of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 and other CoVs are shown in Fig. 4, and the corresponding pairwise amino acid identities are shown in Table 2. For all five gene products, BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 possessed higher amino acid identities to each other than to any other known CoVs with complete genomes available (Table 2). In all five trees, BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 were clustered (Fig. 4). For the Hel, S, and N genes, BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 were also clustered with a CoV recently discovered in Asian leopard cats (9) for which the sequences of these genes were available (Fig. 4). There were 14.5 to 18%, 37.4 to 47%, and 25.1 to 29.3% base differences between the helicase, S, and N genes of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 and that of the Asian leopard cat CoV. Notably, a short fragment (184 bases) of *pol* in MuCoV HKU13 pos-

TABLE 2. Comparison of genomic features and amino acid identities between BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 and other CoVs[a]

CoV	Genome features		Pairwise amino acid identity (%)														
			BuCoV HKU11					ThCoV HKU12					MuCoV HKU13				
	Size (bases)	G+C content	$3CL^{pro}$	Pol	Hel	S	N	$3CL^{pro}$	Pol	Hel	S	N	$3CL^{pro}$	Pol	Hel	S	N
Group 1a																	
PEDV	28033	0.42	36.7	49.1	47.8	35.5	21.3	35.8	48.8	48.1	35.8	21.6	36.4	49.6	47.9	37.4	22.4
TGEV	28586	0.38	33.4	49.5	50.9	33.5	23.5	33.1	49.5	50.7	34.5	22.5	34.1	49.8	50.1	34.0	23.1
FIPV	29355	0.38	34.9	49.6	50.6	34.5	23.7	35.3	49.3	50.4	34.8	22.4	34.7	50.4	49.8	34.2	22.9
Group 1b																	
HCoV-229E	27317	0.38	33.8	48.8	49.7	41.4	19.0	34.2	48.8	48.8	40.8	21.4	34.5	49.6	49.7	40.5	20.3
HCoV-NL63	27553	0.34	35.6	49.0	49.5	36.5	21.2	34.6	49.2	49.3	36.0	22.3	35.7	49.6	49.3	36.9	21.6
Bat-CoV HKU2	27165	0.39	35.1	50.1	50.3	23.8	21.9	34.4	49.9	50.6	25.9	20.8	33.4	50.1	50.2	24.1	21.6
BtCoV 1A	28326	0.38	33.8	49.2	51.5	34.2	21.8	34.2	48.6	51.0	33.6	23.9	34.1	49.5	51.1	34.4	23.6
BtCoV 1B	28476	0.39	32.8	48.7	50.5	33.2	23.2	33.9	48.4	51.2	33.7	23.8	34.1	49.2	50.7	34.1	24.1
Bat-CoV HKU8	28773	0.42	32.8	49.4	49.1	34.4	21.4	33.1	49.3	48.8	35.0	19.2	32.8	49.6	49.1	34.9	20.9
Group 2a																	
HCoV-OC43	30738	0.37	38.1	51.3	48.3	20.9	22.0	38.1	51.3	47.8	23.9	19.2	37.5	52.1	49.3	24.8	21.0
BCoV	31028	0.37	38.5	51.4	48.4	24.8	21.7	38.4	51.3	47.9	22.6	19.4	37.8	52.3	49.4	25.4	20.8
PHEV	30480	0.37	38.5	51.4	48.3	25.5	21.9	38.4	51.3	47.8	24.3	20.7	37.8	52.1	49.3	25.2	19.6
AntelopeCoV	30995	0.37	38.5	51.5	48.4	25.0	21.7	38.4	51.4	47.9	25.6	19.4	37.8	52.4	49.4	25.7	20.8
GiCoV	30979	0.37	38.5	51.5	48.4	25.1	21.7	38.4	51.4	47.9	25.6	19.4	37.8	52.4	49.4	25.8	20.8
MHV	31357	0.42	37.0	51.5	47.6	25.5	23.7	38.3	51.3	47.5	25.2	22.8	36.2	52.8	49.1	24.9	24.2
HCoV-HKU1	29926	0.32	36.5	51.4	47.9	25.2	23.8	38.2	50.9	47.8	24.4	23.0	35.6	52.1	49.3	24.3	23.6
Group 2b																	
SARS-CoV	29751	0.41	35.1	51.4	51.5	25.8	23.7	35.1	51.0	50.5	26.6	23.4	35.1	51.2	51.9	26.3	23.6
Bat-SARS-CoV HKU3	29728	0.41	35.1	51.4	51.6	27.0	23.2	32.9	51.3	50.7	25.5	22.8	34.8	50.9	51.9	25.5	23.0
Group 2c																	
Bat-CoV HKU4	30286	0.38	33.6	51.3	50.8	25.2	24.7	33.3	51.2	49.7	25.4	22.9	35.0	50.6	50.0	23.9	25.1
Bat-CoV HKU5	30488	0.43	33.6	50.3	50.8	25.8	23.0	32.6	50.3	49.5	25.6	23.6	35.0	50.3	50.0	22.9	24.4
Group 2d																	
Bat-CoV HKU9	29114	0.41	35.1	52.6	51.2	26.1	23.2	35.8	52.5	50.4	25.7	23.1	35.1	51.5	51.4	26.4	22.8
Group 3a																	
IBV	27608	0.38	43.6	54.2	56.4	27.0	30.2	43.8	54.4	56.6	28.7	29.0	43.5	54.7	56.1	28.2	30.4
TCoV	27657	0.38	43.3	54.0	57.4	28.6	30.2	44.1	54.7	57.5	30.0	30.0	43.1	54.6	56.8	29.7	29.6
Group 3b																	
SW1	31686	0.39	40.3	52.3	52.8	26.2	31.4	39.0	52.8	52.8	24.7	30.6	37.1	52.5	52.8	25.9	30.1
Group 3c																	
BuCoV HKU11	26476	0.39						88.3	92.5	96.7	46.3	74.6	79.8	88.4	93.6	68.1	72.5
ThCoV HKU12	26396	0.38	88.3	92.5	96.7	46.3	74.6						80.5	86.9	93.1	48.1	75.4
MuCoV HKU13	26552	0.43	79.8	88.4	93.6	68.1	72.5	80.5	86.9	93.1	48.1	75.4					

[a] Comparison of genomic features of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 and other CoVs with complete genome sequences available and of amino acid identities between the predicted $3CL^{pro}$, RNA-dependent RNA Pol, helicase (Hel), S, and N proteins of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 and the corresponding proteins of other CoVs. PEDV, porcine epidemic diarrhea virus; TGEV, porcine transmissible gastroenteritis virus; FIPV, feline infectious peritonitis virus; BtCoV 1A, bat CoV 1A; BtCoV 1B, bat CoV 1B; BCoV, bovine CoV; AntelopeCoV, sable antelope CoV; GiCoV, giraffe CoV; PHEV, porcine hemagglutinating encephalomyelitis virus.

sessed 92.5% nucleotide identity to a recently isolated CoV from a parrot (14). However, no sequence from other regions of the latter was available for further analysis. For $3CL^{pro}$, Pol, helicase, and N, BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 possessed higher amino acid identities to the homologous gene products in IBV and TCoV than to those of group 1 and group 2 CoVs (Table 2). Based both on phylogenetic tree analyses and amino acid differences, we propose a novel subgroup, group 3c, under the group 3 CoVs to describe this distinct subgroup of CoVs which includes BuCoV HKU11, ThCoV HKU12, MuCoV HKU13, and probably, the recently discovered CoV from Asian leopard cats for which a complete genome sequence is not available (9). Interestingly, the S proteins of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 possessed higher amino acid identities to those of group 1 CoVs than those of IBV and TCoV. This is analogous to a phenomenon we recently described in a group 1 CoV, bat CoV HKU2, in which S is not closely related to those of any known CoVs (Fig. 4) (24).

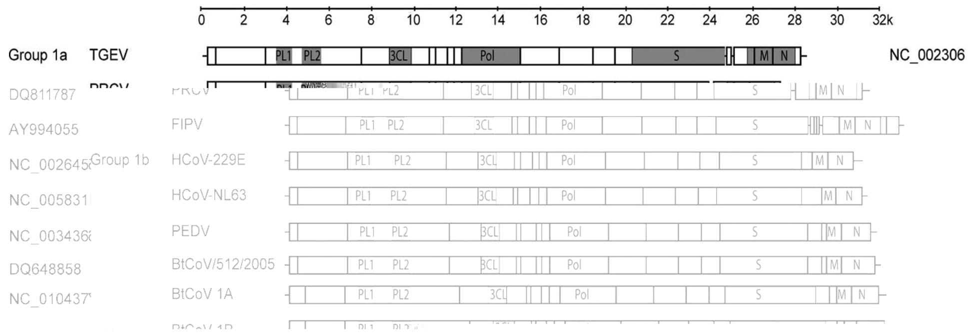

FIG. 2. Genome organizations of BuCoV HKU11, ThCoV HKU12, MuCoV HKU13, and representative CoVs from each group. Papain-like proteases (PL1, PL2, and PL), 3CL^{pro}, RNA-dependent RNA Pol, hemagglutinin esterase (HE), spike (S), envelope (E), membrane (M), and nucleocapsid (N) are represented by gray boxes. Virus name abbreviations may be found in the Fig. 1 legend and the text.

DISCUSSION

In this territory-wide surveillance study of dead wild birds for CoVs, we identified a novel putative subgroup, 3c, of CoVs. BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 formed three distinct branches within this putative subgroup 3c lineage in all five phylogenetic trees analyzed (Fig. 4). Moreover, all 15 strains of BuCoV HKU11 were found in bulbuls, the four strains of ThCoV HKU12 were found in thrushes, and the two strains of MuCoV HKU13 were found in munias. These data support the idea that BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 are three separate novel CoV species infecting three different families of birds. As the three novel CoVs possess the same genome organization and share the same putative TRS, we speculate that they originated from the same ancestor, which subsequently diverged to adapt to different hosts and ecological niches. Based on phylogenetic tree analyses, the three novel avian CoVs formed a unique cluster, most closely related to but distinct from other group 3 CoVs. These three novel avian CoVs were not cultivable using six different cell lines and specific-pathogen-free embryonated chicken eggs (data not shown). BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 of this new proposed subgroup possessed genomic features different from those of other group 3 CoVs (Table 4). For the coding potential of the genomes, IBV and TCoV but not SW1, BuCoV HKU11, ThCoV HKU12, or MuCoV HKU13, possess NS3a and -3b between S and E. IBV, TCoV, SW1, and the three novel putative subgroup 3c CoVs possess two, three, eight, and one small ORF between M and N, respectively. BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 but not IBV, TCoV, or SW1 possess NS7a, -7b, and -7c downstream from N. As for the TRS, the sequences for the TRS's of IBV and TCoV are CUUAACAA, that of SW1 is AAACA, and those of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 are ACACCA.

Group 3 CoVs are found in both birds and mammals and probably contain at least three distinct lineages. Historically,

TABLE 3. Coding potential and putative transcription regulatory sequences of the genomes of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13

CoV	ORF	Start–end (nucleotide position)	No. of nucleotides	No. of amino acids	Frame	Putative TRS	
						Nucleotide position in genome(s)	TRS sequence(s) (distance in bases to AUG)
BuCoV HKU11	1ab	607–19394	18,788	6,262	+1, +3	72	ACACCA(529)AUG
	S	19376–22867	3,492	1,163	+2	19230 or 19320 or 19368	

group 1 and 2 CoVs were found in mammals and group 3 CoVs were found in birds. Although puffinosis virus, a group 2a CoV, had been found in birds, its passage in mouse brains had raised the suspicion that it could be a contaminant of MHV (31). Recently, SW1, with a genome most closely related to IBV, was discovered from a whale (30). Interestingly, BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 were also clustered with a CoV recently described in Asian leopard cats (9), for which a complete genome sequence is not available, in phylogenetic trees constructed using Hel, S, and N. This implied that this lineage of CoVs may be present not only in birds but also in some mammals. The clustering of IBV, TCoV, and IBV-like viruses into one lineage, SW1 as a second lineage, and BuCoV HKU11, ThCoV HKU12, MuCoV HKU13, and the Asian leopard cat CoV as a third lineage implies that group 3 CoVs probably contain at least three subgroups. Although the complete genome sequence is not available for the Asian leopard cat CoV, the same putative TRS for the ORFs available, the presence of NS6 between M and N, and the presence of s2m imply that it may possess a genome organization similar to those of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13 (Fig. 2 and 3). Moreover, the presence of s2m in the Asian leopard cat CoV genome could also suggest that bat SARS-CoV and SARS-CoV may well have acquired their s2m's by recombining with a group 3 mammalian CoV rather than an avian CoV (38). Of note is that 13 of the 21 dead wild

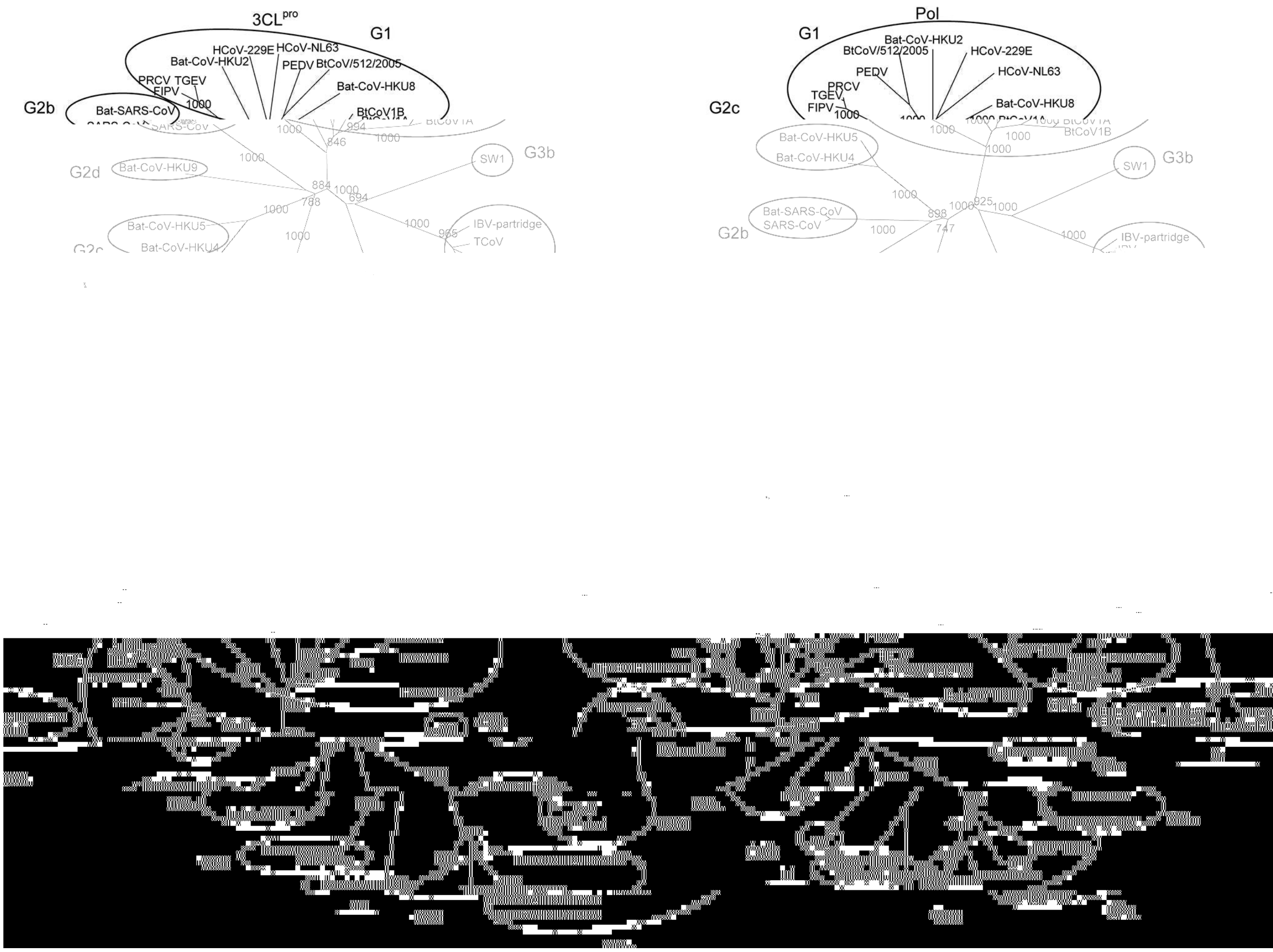

FIG. 4. Phylogenetic analyses of 3CL^{pro}, RNA-dependent RNA Pol, helicase (Hel), S, and N proteins of BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13. The trees were constructed by using the neighbor joining method using Kimura's two-parameter correction and bootstrap values calculated from 1,000 trees. Three hundred eight, 958, 609, 1,735, and 592 amino acid positions in 3CL^{pro}, Pol, helicase, S, and N, respectively, were included in the analyses. The scale bar indicates the estimated number of substitutions per 10 amino acids. PEDV, porcine epidemic diarrhea virus; TGEV, porcine transmissible gastroenteritis virus; FIPV, feline infectious peritonitis virus; BtCoV 1A, bat CoV 1A; BtCoV 1B, bat CoV 1B; BtCoV/512/2005, bat CoV 512/2005; BCoV, bovine CoV; AntelopeCoV, sable antelope CoV; GiCoV, giraffe CoV; IBV-partridge, partridge CoV; IBV-peafowl, peafowl CoV; PHEV, porcine hemagglutinating encephalomyelitis virus; ALC, Asian leopard cat CoV Guangxi/F230/2006.

birds were collected in urban areas of HKSAR. It would be of great importance to determine whether the Asian leopard cat CoV was a result of interspecies jumping from birds, a situation analogous to that of bat SARS-CoV and civet SARS-CoV. White-rumped Munias are commonly released or are found as wild birds around human dwellings, and bulbuls and thrushes are also resident birds found in most habitats in Hong Kong. If group 3c CoV can really adapt to infect mammals such as Asian leopard cats, a common wild mammal in southern China, the mixing of such birds and mammals in wildlife markets may provide the correct environment for interspecies jumping and could subsequently pose the risk of further genetic changes toward adapting to the human host, as in the case of SARS (44).

TABLE 4. Comparison of characteristics in the genomes of group 3a, group 3b, and group 3c CoVs

Group (CoV[s])	Coding potential			TRS sequence
	Small ORFs between S and E	Small ORFs between M and N	Small ORFs downstream from N[a]	
3a (IBV, TCoV)	NS3a, -3b	NS5a, -5b (IBV), ORFx, NS5a, -5b (TCoV)		CUUAACAA
3b (SW1)		NS5a, -5b, -5c, -6, -7, -8, -9, -10		AAACA
3c		NS6	NS7a, -7b, -7c	ACACCA

[a] Newly discovered CoVs from geese, ducks, and pigeons contain small ORFs (orfx and orfy) downstream from N.

Similar to bats, birds also contain a wide diversity of CoVs. Before the SARS epidemic in 2003, 19 (2 human, 13 mammalian, and 4 avian) CoVs were known. After the SARS epidemic, two novel HCoVs, HCoV-NL63 and HCoV-HKU1, were discovered (11, 40, 42). In the most-recent 4 years, at least 10 previously unrecognized CoVs from bats have been described in Hong Kong and mainland China (22, 23, 26, 33, 39, 43, 48). In addition to the generation of a large number of CoV species, recombination has also resulted in different genotypes in certain CoV species, most notably the three genotypes in HCoV-HKU1 (47). In the present study and others, a wide diversity of CoVs are also observed in birds. The wide diversity of CoVs in bats and birds is probably a result of both a higher mutation rate of RNA viruses due to the infidelity of their Pols and a higher chance of recombination due to their unique replication mechanism. Further molecular epidemiological studies in bats and birds of other countries, as well as in other animals, and complete genome sequencing will shed more light on the diversity of CoVs and their evolutionary histories.

ACKNOWLEDGMENTS

We thank York Y. N. Chow, Secretary of Health, Welfare, and Food, HKSAR, The Peoples' Republic of China; and Virology and Biotechnology Laboratory staff of the Tai Lung Veterinary Laboratory of the Agriculture, Fisheries, and Conservation Department of the HKSAR.

The views expressed in this paper are those of the authors only and may not represent the opinion of the Agriculture, Fisheries, and Conservation Department of the government of the HKSAR.

We are grateful for the generous support of Hui Hoy and Hui Ming in the genomic sequencing platform. This work is partly supported by a Research Grant Council grant; a University Development Fund and Outstanding Young Researcher Award from The University of Hong Kong; The Tung Wah Group of Hospitals Fund for Research in Infectious Diseases; the HKSAR Research Fund for the Control of Infectious Diseases of the Health, Welfare, and Food Bureau; and the Shaw Foundation.

REFERENCES

1. **Apweiler, R., T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, L. Cerutti, F. Corpet, M. D. Croning, R. Durbin, L. Falquet, W. Fleischmann, J. Gouzy, H. Hermjakob, N. Hulo, I. Jonassen, D. Kahn, A. Kanapin, Y. Karavidopoulou, R. Lopez, B. Marx, N. J. Mulder, T. M. Oinn, M. Pagni, F. Servant, C. J. Sigrist, and E. M. Zdobnov.** 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res. **29:**37–40.
2. **Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall, and E. L. Sonnhammer.** 2002. The Pfam protein families database. Nucleic Acids Res. **30:**276–280.
3. **Brian, D. A., and R. S. Baric.** 2005. Coronavirus genome structure and replication. Curr. Top. Microbiol. Immunol. **287:**1–30.
4. **Cao, J., C. C. Wu, and T. L. Lin.** 2008. Complete nucleotide sequence of polyprotein gene 1 and genome organization of turkey coronavirus. Virus Res. **136:**43–49.
5. **Cheng, V. C., S. K. Lau, P. C. Woo, and K. Y. Yuen.** 2007. Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. Clin. Microbiol. Rev. **20:**660–694.
6. **Chu, D. K., J. S. Peiris, H. Chen, Y. Guan, and L. L. Poon.** 2008. Genomic characterizations of bat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in Miniopterus bats. J. Gen. Virol. **89:**1282–1287.
7. **Circella, E., A. Camarda, V. Martella, G. Bruni, A. Lavazza, and C. Buonavoglia.** 2007. Coronavirus associated with an enteric syndrome on a quail farm. Avian Pathol. **36:**251–258.
8. **Dominguez, S. R., T. J. O'Shea, L. M. Oko, and K. V. Holmes.** 2007. Detection of group 1 coronaviruses in bats in North America. Emerg. Infect. Dis. **13:**1295–1300.
9. **Dong, B. Q., W. Liu, X. H. Fan, D. Vijaykrishna, X. C. Tang, F. Gao, L. F. Li, G. J. Li, J. X. Zhang, L. Q. Yang, L. L. Poon, S. Y. Zhang, J. S. Peiris, G. J. Smith, H. Chen, and Y. Guan.** 2007. Detection of a novel and highly divergent coronavirus from Asian leopard cats and Chinese ferret badgers in Southern China. J. Virol. **81:**6920–6926.
10. **Ellis, T. M., R. B. Bousfield, L. A. Bissett, K. C. Dyrting, G. S. Luk, S. T. Tsim, K. Sturm-Ramirez, R. G. Webster, Y. Guan, and J. S. Peiris.** 2004. Investigation of outbreaks of highly pathogenic H5N1 avian influenza in waterfowl and wild birds in Hong Kong in late 2002. Avian Pathol. **33:**492–505.
11. **Fouchier, R. A., N. G. Hartwig, T. M. Bestebroer, B. Niemeyer, J. C. de Jong, J. H. Simon, and A. D. Osterhaus.** 2004. A previously undescribed coronavirus associated with respiratory disease in humans. Proc. Natl. Acad. Sci. USA **101:**6212–6216.
12. **Gloza-Rausch, F., A. Ipsen, A. Seebens, M. Göttsche, M. Panning, J. Felix Drexler, N. Petersen, A. Annan, K. Grywna, M. Müller, S. Pfefferle, and C. Drosten.** 2008. Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. Emerg. Infect. Dis. **14:**626–631.
13. **Gomaa, M. H., J. R. Barta, D. Ojkic, and D. Yoo.** 2008. Complete genomic sequence of turkey coronavirus. Virus Res. **135:**237–246.
14. **Gough, R. E., S. E. Drury, F. Culver, P. Britton, and D. Cavanagh.** 2006. Isolation of a coronavirus from a green-cheeked Amazon parrot (Amazon viridigenalis Cassin) Avian Pathol. **35:**122–126.
15. **Guan, Y., B. J. Zheng, Y. Q. He, X. L. Liu, Z. X. Zhuang, C. L-23 F11o1Tf15.ff95J21.8373-1g**

Roper. 2003. The genome sequence of the SARS-associated coronavirus. Science **300:**1399–1404.

30. **Mihindukulasuriya, K. A., G. Wu, J. St. Leger, R. W. Nordhausen, and D. Wang.** 2008. Identification of a novel coronavirus from a beluga whale by using a panviral microarray. J. Virol. **82:**5084–5088.
31. **Nuttall, P. A., and K. A. Harrap.** 1982. Isolation of a coronavirus during studies on puffinosis, a disease of the Manx shearwater (*Puffinus puffinus*). Arch. Virol. **73:**1–13.
32. **Peiris, J. S., S. T. Lai, L. L. Poon, Y. Guan, L. Y. Yam, W. Lim, J. Nicholls, W. K. Yee, W. W. Yan, M. T. Cheung, V. C. Cheng, K. H. Chan, D. N. Tsang, R. W. Yung, T. K. Ng, and K. Y. Yuen.** 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet **361:**1319–1325.
33. **Poon, L. L., D. K. Chu, K. H. Chan, O. K. Wong, T. M. Ellis, Y. H. Leung, S. K. Lau, P. C. Woo, K. Y. Suen, K. Y. Yuen, Y. Guan, and J. S. Peiris.** 2005. Identification of a novel coronavirus in bats. J. Virol. **79:**2001–2009.
34. **Robertson, M. P., H. Igel, R. Baertsch, D. Haussler, M. Ares, Jr., and W. G. Scott.** 2005. The structure of a rigorously conserved RNA element within the SARS virus genome. PLoS Biol. **3:**e5.
35. **Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science **300:**1394–1399.
36. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Rozanov, W. J. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. **331:**991–1004.
37. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. **6:**175–182.
38. **Stavrinides, J., and D. S. Guttman.** 2004. Mosaic evolution of the severe acute respiratory syndrome coronavirus. J. Virol. **78:**76–82.
39. **Tang, X. C., J. X. Zhang, S. Y. Zhang, P. Wang, X. H. Fan, L. F. Li, G. Li, B. Q. Dong, W. Liu, C. L. Cheung, K. M. Xu, W. J. Song, D. Vijaykrishna, L. L. Poon, J. S. Peiris, G. J. Smith, H. Chen, and Y. Guan.** 2006. Prevalence and genetic diversity of coronaviruses in bats from China. J. Virol. **80:**7481–7490.
40. **van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout.** 2004. Identification of a new human coronavirus. Nat. Med. **10:**368–373.
41. **Woo, P. C., Y. Huang, S. K. Lau, H. W. Tsoi, and K. Y. Yuen.** 2005. In silico analysis of ORF1ab in coronavirus HKU1 genome reveals a unique putative cleavage site of coronavirus HKU1 3C-like protease. Microbiol. Immunol. **49:**899–908.
42. **Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris, and K. Y. Yuen.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. **79:**884–895.
43. **Woo, P. C., S. K. Lau, K. S. Li, R. W. Poon, B. H. Wong, H. W. Tsoi, B. C. Yip, Y. Huang, K. H. Chan, and K. Y. Yuen.** 2006. Molecular diversity of coronaviruses in bats. Virology **351:**180–187.
44. **Woo, P. C., S. K. Lau, and K. Y. Yuen.** 2006. Infectious diseases emerging from Chinese wet-markets: zoonotic origins of severe respiratory viral infections. Curr. Opin. Infect. Dis. **19:**401–407.
45. **Woo, P. C., S. K. Lau, H. W. Tsoi, K. H. Chan, B. H. Wong, X. Y. Che, V. K. Tam, S. C. Tam, V. C. Cheng, I. F. Hung, S. S. Wong, B. J. Zheng, Y. Guan, and K. Y. Yuen.** 2004. Relative rates of non-pneumonic SARS coronavirus infection and SARS coronavirus pneumonia. Lancet **363:**841–845.
46. **Woo, P. C., S. K. Lau, H. W. Tsoi, Y. Huang, R. W. Poon, C. M. Chu, R. A. Lee, W. K. Luk, G. K. Wong, B. H. Wong, V. C. Cheng, B. S. Tang, A. K. Wu, R. W. Yung, H. Chen, Y. Guan, K. H. Chan, and K. Y. Yuen.** 2005. Clinical and molecular epidemiological features of coronavirus HKU1-associated community-acquired pneumonia. J. Infect. Dis. **192:**1898–1907.
47. **Woo, P. C., S. K. Lau, C. C. Yip, Y. Huang, H. W. Tsoi, K. H. Chan, and K. Y. Yuen.** 2006. Comparative analysis of 22 coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in coronavirus HKU1. J. Virol. **80:**7136–7145.
48. **Woo, P. C., M. Wang, S. K. Lau, H. Xu, R. W. Poon, R. Guo, B. H. Wong, K. Gao, H. W. Tsoi, Y. Huang, K. S. Li, C. S. Lam, K. H. Chan, B. J. Zheng, and K. Y. Yuen.** 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.
49. **Zhang, J., J. S. Guy, E. J. Snijder, D. A. Denniston, P. J. Timoney, and U. B. Balasuriya.** 2007. Genomic characterization of equine coronavirus. Virology **369:**92–104.
50. **Ziebuhr, J.** 2004. Molecular biology of severe acute respiratory syndrome coronavirus. Curr. Opin. Microbiol. **7:**412–419.