MINIREVIEW

Coronavirus Diversity, Phylogeny and Interspecies Jumping

PATRICK C. Y. WOO,*,†,‡,[1] SUSANNA K. P. LAU,*,†,‡,[1] YI HUANG,‡ AND KWOK-YUNG YUEN*,†,‡,[2]

*State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong; †Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong; and ‡Department of Microbiology, The University of Hong Kong, Hong Kong

The SARS epidemic has boosted interest in research on coronavirus biodiversity and genomics. Before 2003, there were only 10 coronaviruses with complete genomes available. After the SARS epidemic, up to December 2008, there was an addition of 16 coronaviruses with complete genomes sequenced. These include two human coronaviruses (human coronavirus NL63 and human coronavirus HKU1), 10 other mammalian coronaviruses [bat SARS coronavirus, bat coronavirus (bat-CoV) HKU2, bat-CoV HKU4, bat-CoV HKU5, bat-CoV HKU8, bat-CoV HKU9, bat-CoV 512/2005, bat-CoV 1A, equine coronavirus, and beluga whale coronavirus] and four avian coronaviruses (turkey coronavirus, bulbul coronavirus HKU11, thrush coronavirus HKU12, and munia coronavirus HKU13). Two novel subgroups in group 2 coronavirus (groups 2c and 2d) and two novel subgroups in group 3 coronavirus (groups 3b and 3c) have been proposed. The diversity of coronaviruses is a result of the infidelity of RNA-dependent RNA polymerase, high frequency of homologous RNA recombination, and the large genomes of coronaviruses. Among all hosts, the diversity of coronaviruses is most evidenced in bats and birds, which may be a result of their species diversity, ability to fly, environmental pressures, and habits of roosting and flocking. The present evidence supports that bat coronaviruses are the gene pools of group 1 and 2 coronaviruses, whereas bird coronaviruses are the gene pools of group 3 coronaviruses. With the increasing number of coronaviruses, more and more closely related coronaviruses from distantly related animals have been observed, which were results of recent interspecies jumping and may be the cause of disastrous outbreaks of zoonotic diseases. Exp Biol Med 234:1117–1127, 2009

Key words: coronavirus; genome; diversity; phylogeny; interspecies jumping

Introduction

Among the 7800 "coronavirus" papers found by MEDLINE search, almost half of them were published after the SARS epidemic, which has boosted interest in all directions of coronavirus research, most notably, coronavirus biodiversity and genomics (1). Infectious bronchitis virus (IBV), the first coronavirus discovered, was isolated from chicken embryos in 1937 (2). This was followed by mouse hepatitis virus (MHV) and other mammalian coronaviruses in the 1940s (3, 4). The two human coronaviruses, human coronavirus 229E (HCoV-229E) and human coronavirus OC43 (HCoV-OC43), were discovered in the 1960s (5, 6). Before 2003, there were only 10 coronaviruses with complete genomes available, with two human coronaviruses (HCoV-229E and HCoV-OC43), seven other mammalian coronaviruses [MHV, bovine coronavirus (BCoV), porcine hemagglutinating encephalomyelitis virus (PHEV), transmissible gastroenteritis virus (TGEV), porcine epidemic diarrhea virus (PEDV), porcine respiratory coronavirus (PRCV), and feline coronavirus (FCoV)] and one avian coronavirus (IBV) (Table 1, Fig.

This work was partly supported by the Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Disease for the Department of Health of the Hong Kong Special Administrative Region of China; the Research Grant Council Grant (780709), University Development Fund, Outstanding Young Researcher Award, The University of Hong Kong.

[1] These authors contributed equally to this paper.

[2] To whom correspondence should be addressed at State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong. E-mail: hkumicro@hkucc.hku.hk

DOI: 10.3181/0903-MR-94
1535-3702/09/23410-1117$15.00

Table 1. Genomic Features of Coronaviruses with Complete Genomes Available

Coronaviruses[a]	Genome features					
	Size (bases)	G + C content	TRS	No. of nsp in ORF1ab	No. of PL^{pro}	ORF downstream to N
Group 1a						
PEDV	28033	0.42	CUAAAC	16	2	1
TGEV	28586	0.38	CUAAAC[b]	16	2	1
FCoV	29355	0.38	CUAAAC	16	2	2
Group 1b						
HCoV-229E	27317	0.38	CUAAAC	16	2	—
HCoV-NL63	27553	0.34	CUAAAC	16	2	—
Bat-CoV 512/2005	28203	0.40	CUAAAC	16	2	1
Bat-CoV HKU2	27165	0.39	CUAAAC	16	2	1
Bat-CoV HKU8	28773	0.42	CUAAAC	16	2	1
Bat-CoV 1A	28326	0.38	CUAAAC	16	2	—
Group 2a						
HCoV-OC43	30738	0.37	CUAAAC[b]	16	2	—
BCoV	31028	0.37	CUAAAC[b]	16	2	—
PHEV	30480	0.37	CUAAAC[b]	16	2	—
ECoV	30992	0.37	CUAAAC[b]	16	2	—
MHV	31357	0.42	CUAAAC[b]	16	2	—
HCoV-HKU1	29926	0.32	CUAAAC[b]	16	2	—
Group 2b						
SARS-CoV	29751	0.41	ACGAAC	16	1	—
Bat-SARS-CoV HKU3	29728	0.41	ACGAAC	16	1	—
Group 2c						
Bat-CoV HKU4	30286	0.38	ACGAAC	16	1	—
Bat-CoV HKU5	30488	0.43	ACGAAC	16	1	—
Group 2d						
Bat-CoV HKU9	29114	0.41	ACGAAC	16	1	2
Group 3a						
IBV	27608	0.38	CUUAACAA	15	1	—
TCoV	27657	0.38	CUUAACAA	15	1	—
Group 3b						
SW1	31686	0.39	AAACA	15	1	—
Group 3c						
BuCoV HKU11	26476	0.39	ACACCA	15	1	3
ThCoV HKU12	26396	0.38	ACACCA	15	1	3
MuCoV HKU13	26552	0.43	ACACCA	15	1	3

[a] HCoV-229E, human coronavirus 229E; PEDV, porcine epidemic diarrhea virus; TGEV, porcine transmissible gastroenteritis virus; HCoV-NL63, human coronavirus NL63; FCoV, feline coronavirus; bat-CoV 512/2005, bat coronavirus 512/2005; bat-CoV HKU2, bat coronavirus HKU2; bat-CoV 1A, bat coronavirus 1A; bat-CoV HKU8, bat coronavirus HKU8; HCoV-HKU1, human coronavirus HKU1; HCoV-OC43, human coronavirus OC43; MHV, mouse hepatitis virus; BCoV, bovine coronavirus; PHEV, porcine hemagglutinating encephalomyelitis virus; ECoV, equine coronavirus; SARS-CoV, SARS coronavirus; bat-SARS-CoV HKU3; bat SARS coronavirus HKU3; bat-CoV HKU4, bat coronavirus HKU4; bat-CoV HKU5, bat coronavirus HKU5; bat-CoV HKU9, bat coronavirus HKU9; IBV, infectious bronchitis virus; TCoV, turkey coronavirus; SW1, beluga whale coronavirus; BuCoV HKU11, Bulbul coronavirus HKU11; ThCoV HKU12, Thrush coronavirus HKU12; MuCoV HKU13, Munia coronavirus HKU13.
[b] Internal ribosomal entry site is employed for orf3b of TGEV and E of group 2a coronaviruses.

1a). These coronaviruses were classified into three groups, with groups 1 and 2 comprising the nine mammalian coronaviruses and group 3 the avian coronavirus (Fig. 1a) (7–9).

After the SARS epidemic, up to December 2008, there was an addition of 16 coronaviruses with complete genomes sequenced. These include two globally distributed human coronaviruses, human coronavirus NL63 (HCoV-NL63) and human coronavirus HKU1 (HCoV-HKU1) (10–26); 10 other mammalian coronaviruses, bat SARS coronavirus (bat-SARS-CoV), bat coronavirus (bat-CoV) HKU2, bat-CoV HKU4, bat-CoV HKU5, bat-CoV HKU8, bat-CoV HKU9, bat-CoV 512/2005, bat-CoV 1A, equine coronavirus, and beluga whale coronavirus (SW1) (27–35); and four avian coronaviruses, turkey coronavirus (TCoV), bulbul coronavirus HKU11, (BuCoV HKU11), thrush coronavirus HKU12 (ThCoV HKU12), and munia coronavirus HKU13 (MuCoV HKU13) (Table 1, Fig. 1b) (36, 37). Moreover, two novel subgroups in group 2 coronavirus (groups 2c and 2d) and two novel subgroups in group 3 coronavirus (groups 3b and 3c) have been proposed (33, 37). Recently, the Coronavirus Study Group of the International Committee

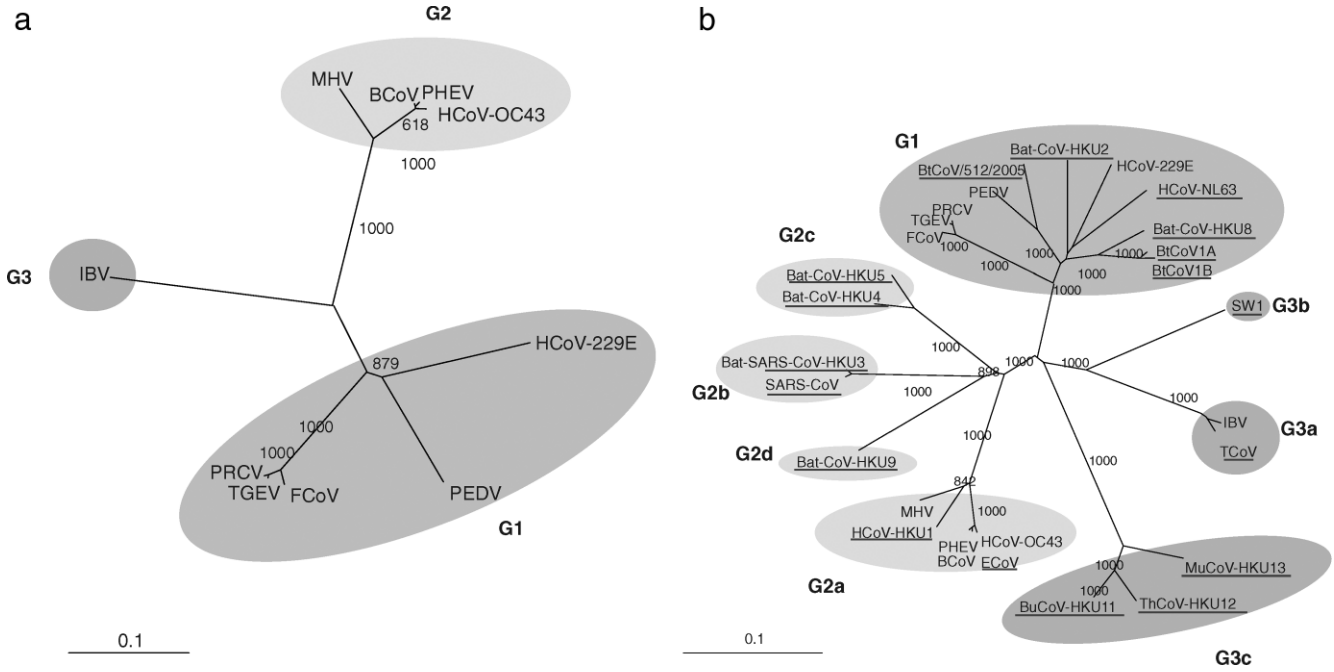

Figure 1. Phylogenetic analysis of RNA-dependent RNA polymerases (Pol) of the 10 coronaviruses with complete genome sequences available before SARS (panel A), and that of all coronaviruses with complete genome sequences available by the end of 2008 (panel B). The trees were constructed by neighbor joining method using Kimura's two-parameter correction and bootstrap values calculated from 1000 trees. 948 and 958 amino acid positions in Pol were included in the two analyses, respectively. The scale bars indicate the estimated number of substitutions per 10 amino acids. HCoV-229E, human coronavirus 229E (NC_002645); PEDV, porcine epidemic diarrhea virus (NC_003436); TGEV, porcine transmissible gastroenteritis virus (NC_002306); FCoV, feline coronavirus (AY994055); PRCV, porcine respiratory coronavirus (DQ811787); HCoV-NL63, human coronavirus NL63 (NC_005831); bat-CoV-HKU2 (EF203064), HKU4 (NC_009019), HKU5 (NC_009020), HKU8 (NC_010438), HKU9 (NC_009021), 1A (NC_010437), 1B (NC_010436), 512/2005 (NC_009657); HCoV-HKU1, human coronavirus HKU1 (NC_006577), HCoV-OC43, human coronavirus OC43 (NC_005147); MHV, mouse hepatitis virus (NC_006852); BCoV, bovine coronavirus (NC_003045); PHEV, porcine hemagglutinating encephalomyelitis virus (NC_007732); ECoV, equine coronavirus (NC_010327); SARS-CoV, SARS coronavirus (NC_004718); bat-SARS-CoV-HKU3, bat-SARS coronavirus HKU3 (NC_009694); IBV, infectious bronchitis virus (NC_001451); TCoV, turkey coronavirus (NC_010800); SW1, beluga whale coronavirus (NC_010646); BuCoV-HKU11, Bulbul coronavirus HKU11 (NC_011548); ThCoV-HKU12, Thrush coronavirus HKU12 (NC_011549); MuCoV-HKU13, Munia coronavirus HKU13 (NC_011550). A color version of this figure is available in the online journal.

for Taxonomy of Viruses (ICTV) has proposed three genera, *Alphacoronavirus*, *Betacoronavirus*, and *Gammacoronavirus*, to replace the traditional groups 1, 2, and 3 coronaviruses (http://talk.ictvonline.org/cfs-filesystemfile.ashx/__key/CommunityServer.Components.PostAttachments/00.00.00.06.26/2008.085_2D00_122V.01.Coronaviridae.pdf).

The diversity of coronaviruses is a result of three major reasons. First, the infidelity of RNA-dependent RNA polymerase of coronaviruses makes their mutation rates in the order of one per 1000 to 10000 nucleotides replicated, which makes them especially plastic (38, 39). Second, as a result of their unique random template switching during RNA replication, thought to be mediated by a "copy-choice" mechanism, coronaviruses have a high frequency of homologous RNA recombination (40, 41). Third, as coronaviruses possess the largest genomes (26.4–31.7 kb) among all known RNA viruses, it has given this family of virus extra plasticity in accommodating and modifying genes. These three factors have not only led to the generation of a diversity of strains and genotypes of one coronavirus species, but also to new species which are able to adapt to new hosts and ecological niches, sometimes causing major zoonotic outbreaks with disastrous consequences (42). As a result of the numerous coronaviruses discovered and genomes sequenced in the past few years, our understanding of the diversity, genomics, and phylogeny of coronavirus has greatly improved. In this article, we review the recent work by us and others on coronavirus diversity and genomics, with an emphasis on phylogeny and interspecies jumping.

Group 1 Coronaviruses (*Alphacoronavirus*)

Among the three groups of coronaviruses, the phylogeny of group 1 coronaviruses is the least well understood. Although it has been proposed that group 1 coronaviruses can be subdivided into groups 1a and 1b based on phylogenetic clustering of group 1a coronaviruses and >90% overall genome identity among the members of this subgroup (Fig. 1b), no additional genomic evidence, such as gene contents, transcription regulatory sequence (TRS) or other unique genomic features, as in the subgroups in groups 2 and 3 coronaviruses, as described below, support such a subclassification. For the group 1b coronaviruses, in addition to

the lack of common genomic features, there is no phylogenetic clustering (Fig. 1b). Therefore, the group 1b coronaviruses are in fact "non–group 1a" coronaviruses, rather than having common features that make them a distinct lineage. In the recent proposal of the Coronavirus Study Group of the ICTV (http://talk.ictvonline.org/cfs-filesystemfile.ashx/_key/CommunityServer.Components.PostAttachments/00.00.00.06.26/2008.085_2D00_122V.01.Coronaviridae.pdf), *Geselavirus* was proposed to be the name given to group 1a coronavirus. Although the genomes of all the members of this subgroup contain one (NS7a) or two (NS7a and 7b) ORFs downstream to N, hence the name *Geselavirus*, which stands for "gene seven last," the genomes of some group 1b coronaviruses, such as bat-CoV HKU8, also contain ORF downstream to N.

Although the present sub-classification of group 1 coronaviruses into groups 1a and 1b may not be ideal, the best documented example of generation of coronavirus species through homologous recombination is present in group 1a coronavirus, which is the generation of FCoV [also called feline infectious peritonitis virus (FIPV) in some publications] type II strains by double recombination between FCoV (FIPV) type I strains and canine coronavirus (CCoV). It was originally observed that the sequence of S in type II FCoV was closely related to that of CCoV (43, 44) but the sequence downstream of E in type II FCoV was closely related to that of type I FCoV (45, 46). This suggests that there may have been a homologous RNA recombination event between the 3' ends of the genomes of CCoV and type I FCoV, giving rise to a type II FCoV genome. Further analysis by multiple alignments pinpointed the site of recombination to a region in the E gene. A few years later, Herrewegh et al. further discovered an additional recombination region in the *pol* gene, and they concluded that type II FCoV in fact originated from two recombination events between genomes of CCoV and type I FCoV (47).

Group 2 Coronaviruses (*Betacoronavirus*)

Among the three groups of coronaviruses, the greatest improvement in our understanding in coronavirus phylogeny lies in group 2 coronaviruses. Before the discovery of SARS-CoV, group 2 coronaviruses were considered to include one lineage, with all members possessing haemagglutinin esterase genes and two papain-like proteases ($PL1^{pro}$ and $PL2^{pro}$) in nsp3 of ORF1ab (Fig. 2). When SARS-CoV was first identified and its genome sequenced, it was proposed that it constituted a fourth group of coronavirus (48, 49). However, after more extensive analyses of the amino-terminal domain of S of SARS-CoV, it was observed that 19 out of the 20 cysteine residues were spatially conserved with those of the consensus sequence for group 2 coronaviruses (50). In contrast, only five of the cysteine residues were spatially conserved with those of the consensus sequences in group 1 and group 3 coronaviruses (50). Furthermore, using both genomic and proteomic approaches, it was confirmed that SARS-CoV is probably an early split-off from the group 2 coronavirus lineage (51). Therefore, SARS-CoV was subsequently classified as group 2b coronaviruses and the historical group 2 coronaviruses were classified as group 2a coronaviruses. In 2005, we and others described the discovery of SARS-CoV–like viruses from at least four species of horseshoe bats in Hong Kong (*Rhinolophus sinicus*) and mainland China (*Rhinolophus ferrumequinum, Rhinolophus macrotis, and Rhinolophus pearsoni*) (31, 34). These bat SARS-CoV were closely related to SARS-CoV found in humans and civets, with the *pol*, helicase, and N genes possessing more than 95% amino acid similarity with the corresponding ones in SARS-CoV from humans and civets. The greatest difference between the genomes of bat SARS-CoV and human and civet SARS-CoV lay in the sars3, sars8, and S genes, with amino acid identities as low as 33% between the sars8 gene in bat SARS-CoV and those in human and civet SARS-CoV. These three genes were also the three genes that showed the greatest variations in the various genomes of human and civet SARS-CoV (31). Despite the finding of bat-SARS-CoV, its S protein only shared 79–80% amino acid identity to that of SARS-CoV, suggesting that SARS-CoV may have acquired a distinct S protein that has allowed interspecies transmission. In our previous studies, another novel group 1 coronavirus, bat-CoV-HKU2, was also found in Chinese horseshoe bats, the same bat species that harbors bat-SARS-CoV (29). Since co-infection of the same bat species by two different coronaviruses may have allowed the opportunities for recombination, the genome sequences of bat-CoV-HKU2 were compared to SARS-CoV–like viruses to reveal possible recombination events. Bat-CoV HKU2 was found to possess a unique spike protein evolutionarily distinct from the rest of the genome. Its spike protein, sharing similar deletions with other group 2 coronaviruses in its C-terminus, also contained a 15-amino acid peptide homologous to a corresponding peptide within the receptor binding motif of SARS-CoV spike protein, which was absent in other coronaviruses, except bat-SARS-CoV. Although no recombination events could be identified, the results suggest a common evolutionary origin in the spike proteins between bat-CoV HKU2 (a group 1 coronavirus) and bat-SARS-CoV and SARS-CoV (group 2 coronaviruses). It is also noteworthy that at least one member of group 1b, HCoV-NL63, also uses ACE2 as the receptor for cell entry, as in the case of SARS-CoV, though the site of binding on ACE2 is different (52, 53).

In 2006 and 2007, we proposed two additional subgroups of group 2 coronaviruses: group 2c and group 2d (33). These two subgroups form two unique lineages, most closely related to, but distinct from group 2a and group 2b coronaviruses. In addition to phylogenetic evidence, there is also clear-cut evidence from gene contents and other genomic features that four subgroups exist in group 2 coronaviruses. For the gene contents of the genomes of

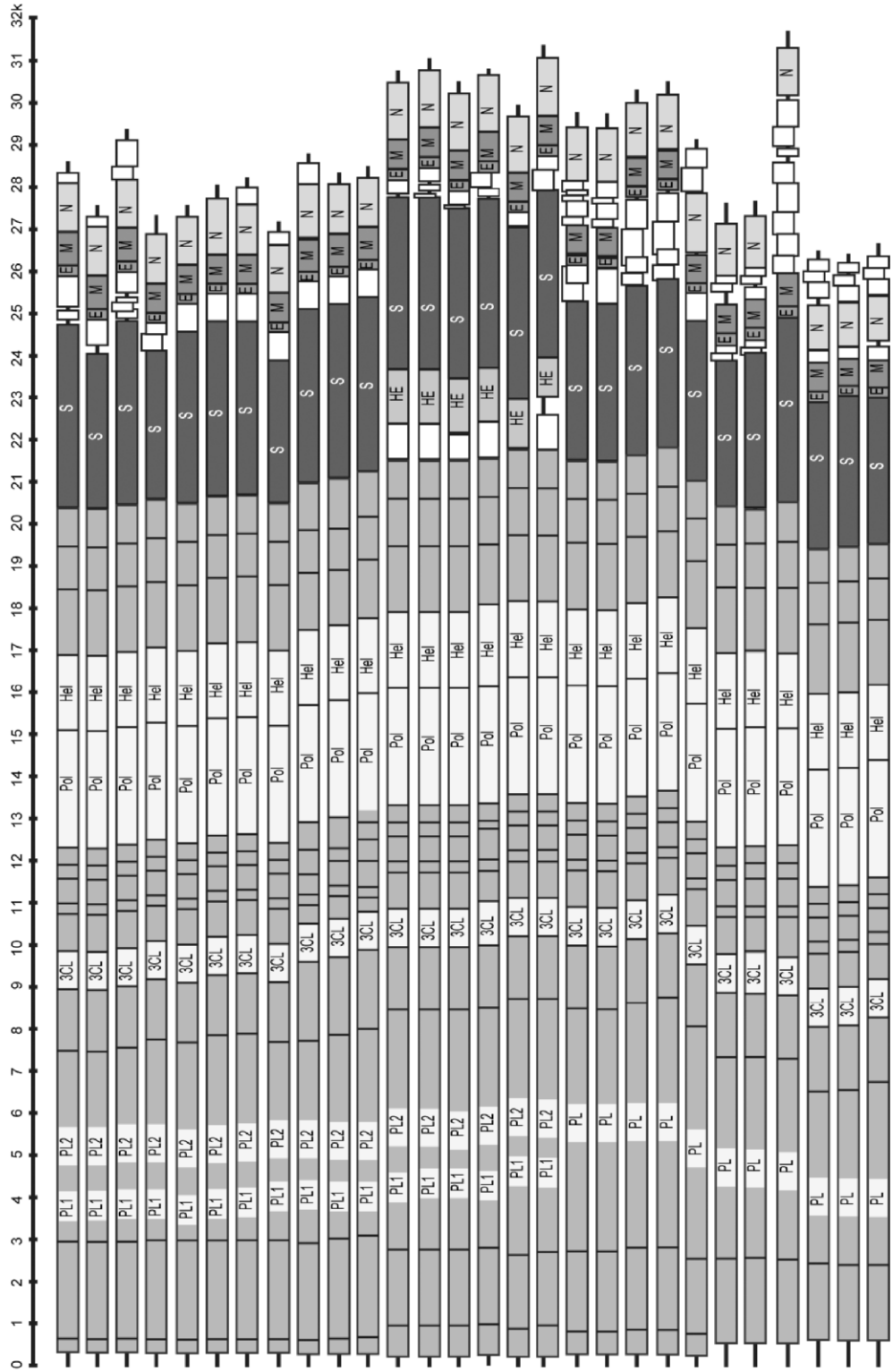

Figure 2. Genome organizations of members in different groups of coronaviruses. PL1, papain-like protease 1; PL2, papain-like protease 2; PL, papain-like protease; 3CL, chymotrypsin-like protease; Pol, RNA-dependent RNA polymerase; Hel, helicase; HE, haemagglutinin esterase; S, spike; E, envelope; M, membrane; N, nucleocapsid. TGEV, porcine transmissible gastroenteritis virus (NC_002306); PRCV, porcine respiratory coronavirus (DQ811787); FCoV, feline coronavirus (AY994055); HCoV-229E, human coronavirus 229E (NC_002645); HCoV-NL63, human coronavirus NL63 (NC_005831); PEDV, porcine epidemic diarrhea virus (NC_003436); BtCoV/512/2005, bat coronavirus 512/2005 (NC_009657); bat-CoV-HKU2, bat coronavirus HKU2 (EF203064); bat-CoV-HKU8, bat coronavirus HKU8 (NC_010438); BtCoV1A, bat coronavirus 1A (NC_010437); BtCoV1B, bat coronavirus 1B (NC_010436); HCoV-OC43, human coronavirus OC43 (NC_005147); BCoV, bovine coronavirus (NC_003045); PHEV, porcine hemagglutinating encephalomyelitis virus (NC_007732); HCoV-HKU1, human coronavirus HKU1 (NC_006577); MHV, mouse hepatitis virus (NC_006852); ECoV, equine coronavirus (NC_010327); human SARS-CoV, human SARS coronavirus (NC_004718); bat SARS-CoV, bat SARS coronavirus (NC_009694); bat-CoV-HKU4, bat coronavirus HKU4 (NC_009019); bat-CoV-HKU5, bat coronavirus HKU5 (NC_009020); bat-CoV-HKU9, bat coronavirus HKU9 (NC_009021); IBV, infectious bronchitis virus (NC_001451); TCoV, turkey coronavirus (NC_010800); SW1, beluga whale coronavirus (NC_010646); BuCoV HKU11, Bulbul coronavirus HKU11 (NC_011548); ThCoV-HKU12, Thrush coronavirus HKU12 (NC_011549); MuCoV-HKU13, Munia coronavirus HKU13 (NC_011550). A color version of this figure is available in the online journal.

group 2a coronaviruses, they possess $PL1^{pro}$ and $PL2^{pro}$ in nsp3 of ORF1ab, but group 2b, 2c, and 2d coronaviruses only possess one PL^{pro}, which is homologous to $PL2^{pro}$. Furthermore, the genomes of group 2a, but not those of group 2b, 2c, and 2d coronaviruses, encode haemagglutinin esterase. For group 2b coronaviruses, their genomes, but not those of group 2a, 2c, and 2d coronaviruses, contain several small ORFs between the M and N genes. As for group 2d coronaviruses, their genomes, but not those of group 2a, 2b, and 2c coronaviruses, contain two ORFs downstream to the N gene. As for the TRS, the sequence for the TRS of group 2a coronaviruses is CUAAAC and that of group 2b, 2c, and 2d coronaviruses is ACGAAC (33, 54–56). For the E gene, TRS is present in group 2b, 2c, and 2d, but not group 2a, coronaviruses, in which an internal ribosomal entry site is used for their translation (33, 48, 49, 57). The genomes of group 2a, 2b, and 2c contain clear-cut overlapped bulged stem-loop and pseudoknot structures at the 3′ untranslated region and immediately downstream to N. On the other hand, whether the genomes of group 2d coronaviruses possess similar bulged stem-loop and pseudoknot structures is controversial. Obviously, the genome of bat CoV-HKU9, the only member of group 2d coronaviruses identified so far, does not possess the classical bulged stem-loop and pseudoknot structures immediately downstream to N that were present in the genomes of group 2a, 2b, and 2c coronaviruses (33). Although it has been suggested that a candidate pseudoknot structure could be present at 1073 bases downstream to N and a predicted candidate bulged stem-loop can be found upstream to it (58), the part occupied by the predicted candidate bulged stem-loop belongs to the putative coding region of NS7b, which is probably an ORF that is expressed because of the presence of TRS.

Extensive homologous and heterologous recombination events have been documented in both human and animal group 2 coronaviruses, which has led to the generation of various genotypes and strains within a coronavirus species, as well as acquisition of new genes from other non-coronavirus RNA donors. Among the coronaviruses, MHV is one of the most extensively studied examples of homologous recombination in coronaviruses, and is also the coronavirus in which homologous recombination was first observed. Over 20 years ago, Lai et al. first observed homologous recombination as a result of mixed infection of DBT cells with MHV strains A59 and JHM (59). Genome analysis showed that the recombinant strain contained sequences from both parents and one crossover site. Subsequently, homologous recombination in MHV was further observed in tissue culture (59, 60) and experimentally infected animals (61). In the 1990s, it was found that as much as 25% of MHV were recombinants (62, 63). Furthermore, in vitro studies have shown variations in both sites and rates of recombination, with the S gene having a frequency threefold that of the *pol* gene (60, 63). As for human coronavirus, the most studied example was HCoV-HKU1. In our study on complete genome sequencing and phylogenetic studies of 22 strains of HCoV-HKU1, extensive recombination in different parts of the genomes was observed, which has led to the generation of three genotypes, A, B, and C, of HCoV-HKU1 (64). The two most notable examples were observed in a stretch of 143 nucleotides near the 3' end of nsp6, where recombination between genotypes B and C led to generation of genotype A, and in another stretch of 29 nucleotides near the 3′ end of nsp16, where recombination between genotypes A and B led to generation of genotype C. This represented the first example of recombination in human coronavirus and was also the first report to describe a distribution of the recombination spots in the entire genome of field isolates of a coronavirus. As for the acquisition of new genes from non-coronavirus RNA donors by heterologous recombination, the most notable example is the HE gene from influenza C virus (65, 66). The presence of HE genes in group 2a, but not other group 2, coronaviruses suggested that the recombination had probably occurred in the ancestor of group 2a coronaviruses, after diverging from the ancestor of other group 2 coronaviruses.

Group 3 Coronaviruses (*Gammacoronavirus*)

Dramatic improvement in our understanding of the diversity and phylogeny, and potential interspecies jumping, of group 3 coronaviruses occurred in the last year. Since its discovery in 1937, IBV has been the only species of group 3 coronavirus for over 50 years. In the last decade of the last century and the first few years of the 21st century, a few IBV-like viruses, including TCoV, have been described in various species of birds, with some of their genomes sequenced (36, 67–69). The sizes, G + C contents, and genome organizations of their genomes were similar, indicating that they probably have diverged from the same ancestor recently. This 70 years of quiescence was broken by two discoveries in 2008—first, the report on SW1 from a beluga whale, with the largest coronavirus genome; and second, the discovery of a novel subgroup of coronavirus from birds of different families, with the smallest coronavirus genomes (Table 1) (28, 37).

SW1 was discovered from the liver tissue of a dead beluga whale (28). It was the first reported group 3 mammalian coronavirus with complete genome sequence and was phylogenetically distantly related to IBV. Uniquely, eight ORFs, occupying a 4105-base region, were observed between M and N, giving rise to the largest reported coronavirus genome. We propose that this lineage should be group 3b coronavirus, whereas the IBV and IBV-like viruses should be group 3a coronaviruses.

The novel subgroup of avian coronaviruses, group 3c coronavirus, that we recently described consisted of at least three members (BuCoV HKU11, ThCoV HKU12, and MuCoV HKU13), infecting at least three different families of birds (bulbuls, thrushes, and munias) (37). These coronaviruses were distantly related to IBV and SW1. Most

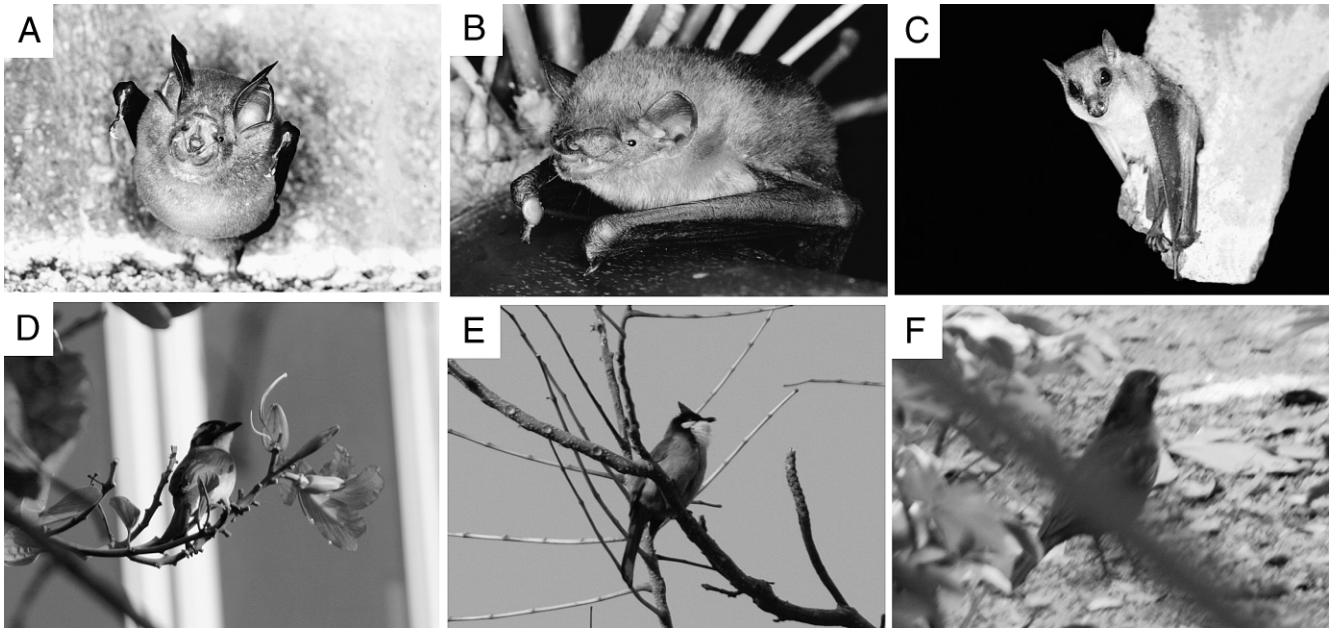

Figure 3. Examples of bats and birds in Hong Kong from which novel coronaviruses were discovered. Chinese horseshoe bat (*Rhinolophus sinicus*) (panel A), from which bat-SARS-CoV and bat-CoV HKU2 were discovered; Lesser bamboo bat (*Tylonycteris pachypus*) (panel B), from which bat-CoV HKU4 was discovered; Leschenault's rousette (*Rousettus lechenaulti*) (panel C), from which bat-CoV HKU9 was discovered; Chinese Bulbul (*Pycnonotus sinensis*) (panel D) and Red-whiskered Bulbul (*Pycnonotus jocosus*) (panel E), from which BuCoV HKU11 was discovered; and Blackbird (*Turdus merula*) (panel F), from which ThCoV HKU12 was discovered. A color version of this figure is available in the online journal.

interestingly, these three avian coronaviruses were also clustered with a coronavirus recently discovered in the Asian leopard cat (ALC-CoV), for which the complete genome sequence was not available (70). From the sequences of the gene fragments available, it was observed that ALC-CoV probably also employed the same putative TRS, NS6 is also present between M and N, and a stem-loop II motif (s2m), a conserved RNA element downstream to N and upstream to the polyA tail, is also present. This represents the hitherto closest relationship between mammalian and avian coronaviruses, as the puffinosis virus, a group 2a coronavirus that had been found in birds, was considered as a contaminating MHV as a result of its passage in mouse brains (71). Complete genome sequencing of ALC-CoV and comparative genomics studies may reveal the secret behind interclass jumping in coronaviruses.

Bat Coronaviruses as Gene Pool for Group 1 and Group 2 Coronaviruses and Avian Coronaviruses as Gene Pool for Group 3 Coronaviruses

The discovery of bat-SARS-CoV has marked the beginning of the race of coronavirus hunting in bats (31, 34). Among the 23 group 1 and group 2 coronaviruses with complete genome sequence available, 9 (39%) were from bats (Fig. 3). Furthermore, bats were also the hosts of 103 (GenBank taxonomy data in Feb. 2009) additional coronaviruses, discovered in Asia, Europe, America, and Africa, although complete genome sequences were still not available (32, 35, 72–74). As for group 3 coronaviruses, they have been exclusively found in birds, with the exception of SW1 from the beluga whale and ALC-CoV from Asian Leopard cats (28, 70). As the race of coronavirus hunting in birds has just begun, we speculate that there are still many unrecognized coronaviruses in birds (Fig. 3). This diversity of coronaviruses in bats and birds could be related to the unique properties of these two groups of animals (75, 76). First, the diversity of bats and birds themselves is huge. Bats account for more than 20% of the 4800 mammalian species recorded in the world. For example, although Hong Kong is an urbanized, subtropical city, it has extensive natural areas with more than 50 different species of terrestrial mammals, with 40% of the species being bats. As for birds, this class contains around 10,000 species, making them the most diverse tetrapod vertebrates; and in Hong Kong, there are more than 460 different species of birds. This diversity of bats and birds would potentially provide a large number of different cell types for different coronaviruses. This is in line with the genus specificity for different bat and bird coronaviruses. For example, bat-SARS-CoV was found in *Rhinolophus* bats, bat-CoV HKU4 in *Tylonycteris* bats, bat-CoV HKU5 in *Pipistrellus* bats, bat-CoV HKU9 in *Rousettus* bats, BuCoV HKU11 in *Pycnonotus* birds, ThCoV HKU12 in *Turdus* birds, and MuCoV HKU13 in *Lonchura* birds. Second, the ability to fly has given bats and birds the opportunity to go almost anywhere, free from obstacles faced by land-based mammals. Bats have been found at altitudes as high as 5000 m, and some birds can fly for over

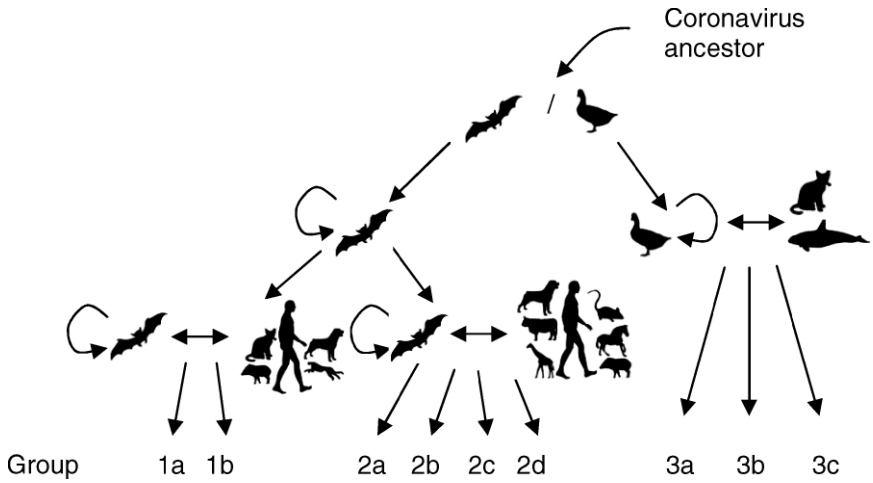

Figure 4. A model of coronavirus evolution. Coronaviruses in bats are the hypothesized gene pool of group 1 and group 2 coronaviruses and coronaviruses in birds are the hypothesized gene pool of group 3 coronaviruses.

10,000 km in their journeys of long-distance non-stop migration. This ability of bats and birds would have allowed possible exchange of viruses and/or their genetic materials with different kinds of living organisms. Third, the different environmental pressures such as food, climate, shelter, and predators would have provided different selective pressures on parasitisation of different coronaviruses in different species of bats and birds. Fourth, the habit of roosting in bats and flocking in birds results in a large number of bats and birds to gather together. This would have also facilitated exchange of viruses among individual bats and birds.

The huge diversity of coronaviruses in bats and birds has made them excellent gene pools for groups 1 and 2 coronaviruses and group 3 coronaviruses, respectively (Fig. 4). It has been proposed that bat coronaviruses were the gene pools of all three groups of coronaviruses (77). However, it seems that there is no evidence supporting this hypothesis because more than 100 bat coronaviruses have been discovered and still none of them belonged to group 3. Instead, the present evidence supports that bat coronaviruses are the gene pools of groups 1 and 2 coronaviruses, whereas bird coronaviruses are the gene pools of group 3 coronaviruses. We speculate that the ancestor of the present coronaviruses infected a bat and it jumped from the bat to a bird, or alternatively, it infected a bird and it jumped from the bird to a bat, evolving dichotomously. On the one hand, the bat coronavirus jumped to another species of bat, giving rise to the group 1 and group 2 coronaviruses, evolving dichotomously. These bat coronaviruses in turn jumped to other bat species and other mammals, including humans, with each interspecies jumping evolving dichotomously. On the other hand, the bird coronavirus jumped to other species of birds, and occasionally to some specific mammalian species, such as whale and Asian Leopard cat, with each interspecies jumping evolving dichotomously, giving rise to the group 3 coronaviruses. The properties of bats and birds mentioned above have facilitated the generation of a huge diversity of bat and bird coronaviruses as well as dissemination to other animals.

Concluding Remarks

In the past six years of the 21st century, we have witnessed a drastic increase in the number of coronaviruses discovered and coronavirus genomes being sequenced. With this increase in the number of coronavirus species and genomes, we are starting to appreciate the diversity of coronaviruses. Databases for efficient sequence retrieval and the ever-improving bioinformatics tools have further enabled us to start to understand the phylogeny of coronaviruses and perform additional genomic analyses (78, 79). With the increasing number of coronaviruses, more and more closely related coronaviruses from distantly related animals have been observed. Examples included FCoV and CCoV in group 1a; MHV and rat coronavirus or HCoV-OC43, BCoV, and PHEV in group 2a; bat, civet SARS-CoV, and human SARS-CoV in group 2b; IBV and TCoV in group 3a; and the Asian Leopard cat coronavirus and the novel avian coronaviruses in group 3c. These were results of recent interspecies jumping and may be the cause of disastrous outbreaks of zoonotic diseases. Detailed analysis of their genomes, particularly the S protein sequences and structures, as well as the receptors for the individual coronaviruses, will enable rational design of experiments to understand the secret behind interspecies jumping at the molecular level.

We are grateful to Chung-Tong Shek for providing the photos of bats and Rex K. H. Au-Yeung for providing the photos of birds.

1. Cheng VC, Lau SK, Woo PC, Yuen KY. Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. Clin Microbiol Rev 20:660–694, 2007.

2. Beaudette FR, Hudson CB. Cultivation of the virus of infectious bronchitis. J Am Vet Med Assoc 90:51–58, 1937.
3. Cheever FS, Daniels JB, Pappenheimer AM, Bailey OT. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. J Exp Med 90:181–194, 1949.
4. Doyle LP, Hutchings LM. A transmissible gastroenteritis virus in pigs. J Am Vet Assoc 108:257–259, 1946.
5. Tyrrell DA, Bynoe ML. Cultivation of viruses from a high proportion of patients with colds. Lancet 1:76–77, 1966.
6. Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med 121:190–193, 1966.
7. Pensaert MB, Debouck P, Reynolds DJ. An immunoelectron microscopic and immunofluorescent study on the antigenic relationship between the coronavirus-like agent, CV 777, and several coronaviruses. Arch Virol 68:45–52, 1981.
8. Bradburne AF. Antigenic relationships amongst coronaviruses. Arch Gesamte Virusforsch 31:352–364, 1970.
9. McIntosh K, Kapikian AZ, Hardison KA, Hartley JW, Chanock RM. Antigenic relationships among the coronaviruses of man and between human and animal coronaviruses. J Immunol 102:1109–1118, 1969.
10. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, Wertheim-van Dillen PM, Kaandorp J, Spaargaren J, Berkhout B. Identification of a new human coronavirus. Nat Med 10: 368–373, 2004.
11. Koetz A, Nilsson P, Linden M, van der Hoek L, Ripa T. Detection of human coronavirus NL63, human metapneumovirus and respiratory syncytial virus in children with respiratory tract infections in south-west Sweden. Clin Microbiol Infect 12:1089–1096, 2006.
12. Vabret A, Mourez T, Dina J, van der Hoek L, Gouarin S, Petitjean J, Brouard J, Freymuth F. Human coronavirus NL63, France. Emerg Infect Dis 11:1225–1229, 2005.
13. Suzuki A, Okamoto M, Ohmi A, Watanabe O, Miyabayashi S, Nishimura H. Detection of human coronavirus-NL63 in children in Japan. Pediatr Infect Dis J 24:645–646, 2005.
14. Chiu SS, Chan KH, Chu KW, Kwan SW, Guan Y, Poon LL, Peiris JS. Human coronavirus NL63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong, China. Clin Infect Dis 40:1721–1729, 2005.
15. Moes E, Vijgen L, Keyaerts E, Zlateva K, Li S, Maes P, Pyrc K, Berkhout B, van der Hoek L, Van Ranst M. A novel pancoronavirus RT-PCR assay: frequent detection of human coronavirus NL63 in children hospitalized with respiratory tract infections in Belgium. BMC Infect Dis 5:6, 2005.
16. Arden KE, Nissen MD, Sloots TP, Mackay IM. New human coronavirus, HCoV-NL63, associated with severe lower respiratory tract disease in Australia. J Med Virol 75:455–462, 2005.
17. Ebihara T, Endo R, Ma X, Ishiguro N, Kikuta H. Detection of human coronavirus NL63 in young children with bronchiolitis. J Med Virol 75: 463–465, 2005.
18. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BH, Poon RW, Cai JJ, Luk WK, Poon LL, Wong SS, Guan Y, Peiris JS, Yuen KY. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 79:884–895, 2005.
19. Woo PC, Lau SK, Tsoi HW, Huang Y, Poon RW, Chu CM, Lee RA, Luk WK, Wong GK, Wong BH, Cheng VC, Tang BS, Wu AK, Yung RW, Chen H, Guan Y, Chan KH, Yuen KY. Clinical and molecular epidemiological features of coronavirus HKU1–associated community-acquired pneumonia. J Infect Dis 192:1898–1907, 2005.
20. Lau SK, Woo PC, Yip CC, Tse H, Tsoi HW, Cheng VC, Lee P, Tang BS, Cheung CH, Lee RA, So LY, Lau YL, Chan KH, Yuen KY. Coronavirus HKU1 and other coronavirus infections in Hong Kong. J Clin Microbiol 44:2063–2071, 2006.
21. Esper F, Weibel C, Ferguson D, Landry ML, Kahn JS. Coronavirus HKU1 infection in the United States. Emerg Infect Dis 12:775–779, 2006.
22. Vabret A, Dina J, Gouarin S, Petitjean J, Corbet S, Freymuth F. Detection of the new human coronavirus HKU1: a report of 6 cases. Clin Infect Dis 42:634–639, 2006.
23. Sloots TP, McErlean P, Speicher DJ, Arden KE, Nissen MD, Mackay IM. Evidence of human coronavirus HKU1 and human bocavirus in Australian children. J Clin Virol 35:99–102, 2006.
24. Gerna G, Percivalle E, Sarasini A, Campanini G, Piralla A, Rovida F, Genini E, Marchi A, Baldanti F. Human respiratory coronavirus HKU1 versus other coronavirus infections in Italian hospitalised patients. J Clin Virol 38:244–250, 2007.
25. Bosis S, Esposito S, Niesters HG, Tremolati E, Pas S, Principi N, Osterhaus AD. Coronavirus HKU1 in an Italian pre-term infant with bronchiolitis. J Clin Virol 38:251–253, 2007.
26. Kupfer B, Simon A, Jonassen CM, Viazov S, Ditt V, Tillmann RL, Muller A, Matz B, Schildgen O. Two cases of severe obstructive pneumonia associated with an HKU1-like coronavirus. Eur J Med Res 12:134–138, 2007.
27. Zhang J, Guy JS, Snijder EJ, Denniston DA, Timoney PJ, Balasuriya UB. Genomic characterization of equine coronavirus. Virology 369:92–104, 2007.
28. Mihindukulasuriya KA, Wu G, St Leger J, Nordhausen RW, Wang D. Identification of a novel coronavirus from a beluga whale by using a panviral microarray. J Virol 82:5084–5088, 2008.
29. Lau SK, Woo PC, Li KS, Huang Y, Wang M, Lam CS, Xu H, Guo R, Chan KH, Zheng BJ, Yuen KY. Complete genome sequence of bat coronavirus HKU2 from Chinese horseshoe bats revealed a much smaller spike gene with a different evolutionary lineage from the rest of the genome. Virology 367:428–439, 2007.
30. Chu DK, Peiris JS, Chen H, Guan Y, Poon LL. Genomic characterizations of bat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in Miniopterus bats. J Gen Virol 89:1282–1287, 2008.
31. Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, Wong SS, Leung SY, Chan KH, Yuen KY. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci U S A 102:14040–14045, 2005.
32. Woo PC, Lau SK, Li KS, Poon RW, Wong BH, Tsoi HW, Yip BC, Huang Y, Chan KH, Yuen KY. Molecular diversity of coronaviruses in bats. Virology 351:180–187, 2006.
33. 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. 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, 2007.
34. 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. Bats are natural reservoirs of SARS-like coronaviruses. Science 310:676–679, 2005.
35. Tang XC, Zhang JX, Zhang SY, Wang P, Fan XH, Li LF, Li G, Dong BQ, Liu W, Cheung CL, Xu KM, Song WJ, Vijaykrishna D, Poon LL, Peiris JS, Smith GJ, Chen H, Guan Y. Prevalence and genetic diversity of coronaviruses in bats from China. J Virol 80:7481–7490, 2006.
36. Gomaa MH, Barta JR, Ojkic D, Yoo D. Complete genomic sequence of turkey coronavirus. Virus Res 135:237–246, 2008.
37. Woo PC, Lau SK, Lam CS, Lai KK, Huang Y, Lee P, Luk GS, Dyrting KC, Chan KH, Yuen KY. Comparative analysis of complete genome sequences of three avian coronaviruses reveals a novel group 3c coronavirus. J Virol 83:908–917, 2009.
38. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 9:267–276, 2008.
39. Jenkins GM, Rambaut A, Pybus OG, Holmes EC. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J Mol Evol 54:156–165, 2002.

40. Lai MM. RNA recombination in animal and plant viruses. Microbiol Rev 56:61–79, 1992.
41. Pasternak AO, Spaan WJ, Snijder EJ. Nidovirus transcription: how to make sense...? J Gen Virol 87:1403–1421, 2006.
42. Woo PC, Lau SK, Yuen KY. Infectious diseases emerging from Chinese wet-markets: zoonotic origins of severe respiratory viral infections. Curr Opin Infect Dis 19:401–407, 2006.
43. Wesseling JG, Vennema H, Godeke GJ, Horzinek MC, Rottier PJ. Nucleotide sequence and expression of the spike (S) gene of canine coronavirus and comparison with the S proteins of feline and porcine coronaviruses. J Gen Virol 75:1789–1794, 1994.
44. Motokawa K, Hohdatsu T, Aizawa C, Koyama H, Hashimoto H. Molecular cloning and sequence determination of the peplomer protein gene of feline infectious peritonitis virus type I. Arch Virol 140:469–480, 1995.
45. Motokawa K, Hohdatsu T, Hashimoto H, Koyama H. Comparison of the amino acid sequence and phylogenetic analysis of the peplomer, integral membrane and nucleocapsid proteins of feline, canine and porcine coronaviruses. Microbiol Immunol 40:425–433, 1996.
46. Herrewegh AA, Vennema H, Horzinek MC, Rottier PJ, de Groot RJ. The molecular genetics of feline coronaviruses: comparative sequence analysis of the ORF7a/7b transcription unit of different biotypes. Virology 212:622–631, 1995.
47. Herrewegh AA, Smeenk I, Horzinek MC, Rottier PJ, de Groot RJ. Feline coronavirus type II strains 79–1683 and 79–1146 originate from a double recombination between feline coronavirus type I and canine coronavirus. J Virol 72:4508–4514, 1998.
48. Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Coughlin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JM, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Krajden M, Petric M, Skowronski DM, Upton C, Roper RL. The genome sequence of the SARS-associated coronavirus. Science 300:1399–1404, 2003.
49. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300:1394–1399, 2003.
50. Eickmann M, Becker S, Klenk HD, Doerr HW, Stadler K, Censini S, Guidotti S, Masignani V, Scarselli M, Mora M, Donati C, Han JH, Song HC, Abrignani S, Covacci A, Rappuoli R. Phylogeny of the SARS coronavirus. Science 302:1504–1505, 2003.
51. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE. 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, 2003.
52. Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci U S A 102:7988–7993, 2005.
53. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426:450–454, 2003.
54. Hussain S, Pan J, Chen Y, Yang Y, Xu J, Peng Y, Wu Y, Li Z, Zhu Y, Tien P, Guo D. Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. J Virol 79:5288–5295, 2005.
55. Makino S, Joo M. Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. J Virol 67:3304–3311, 1993.
56. Lai MM, Cavanagh D. The molecular biology of coronaviruses. Adv Virus Res 48:1–100, 1997.
57. Jendrach M, Thiel V, Siddell S. Characterization of an internal ribosome entry site within mRNA 5 of murine hepatitis virus. Arch Virol 144:921–933, 1999.
58. Zust R, Miller TB, Goebel SJ, Thiel V, Masters PS. Genetic interactions between an essential 3' cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the mouse coronavirus genome. J Virol 82:1214–1228, 2008.
59. Lai MM, Baric RS, Makino S, Keck JG, Egbert J, Leibowitz JL, Stohlman SA. Recombination between nonsegmented RNA genomes of murine coronaviruses. J Virol 56:449–456, 1985.
60. Makino S, Keck JG, Stohlman SA, Lai MM. High-frequency RNA recombination of murine coronaviruses. J Virol 57:729–737, 1986.
61. Keck JG, Matsushima GK, Makino S, Fleming JO, Vannier DM, Stohlman SA, Lai MM. In vivo RNA-RNA recombination of coronavirus in mouse brain. J Virol 62:1810–1813, 1988.
62. Fu K, Baric RS. Map locations of mouse hepatitis virus temperature-sensitive mutants: confirmation of variable rates of recombination. J Virol 68:7458–7466, 1994.
63. Fu K, Baric RS. Evidence for variable rates of recombination in the MHV genome. Virology 189:88–102, 1992.
64. Woo PC, Lau SK, Yip CC, Huang Y, Tsoi HW, Chan KH, Yuen KY. Comparative analysis of 22 coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in coronavirus HKU1. J Virol 80:7136–7145, 2006.
65. Zeng Q, Langereis MA, van Vliet AL, Huizinga EG, de Groot RJ. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. Proc Natl Acad Sci U S A 105: 9065–9069, 2008.
66. Luytjes W, Bredenbeek PJ, Noten AF, Horzinek MC, Spaan WJ. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. Virology 166:415–422, 1988.
67. Jonassen CM, Kofstad T, Larsen IL, Lovland A, Handeland K, Follestad A, Lillehaug A. Molecular identification and characterization of novel coronaviruses infecting graylag geese (*Anser anser*), feral pigeons (*Columbia livia*) and mallards (*Anas platyrhynchos*). J Gen Virol 86:1597–1607, 2005.
68. Liu S, Chen J, Chen J, Kong X, Shao Y, Han Z, Feng L, Cai X, Gu S, Liu M. Isolation of avian infectious bronchitis coronavirus from domestic peafowl (*Pavo cristatus*) and teal (*Anas*). J Gen Virol 86:719–725, 2005.
69. Gough RE, Cox WJ, Winkler CE, Sharp MW, Spackman D. Isolation and identification of infectious bronchitis virus from pheasants. Vet Rec 138:208–209, 1996.
70. Dong BQ, Liu W, Fan XH, Vijaykrishna D, Tang XC, Gao F, Li LF, Li GJ, Zhang JX, Yang LQ, Poon LL, Zhang SY, Peiris JS, Smith GJ, Chen H, Guan Y. Detection of a novel and highly divergent coronavirus from asian leopard cats and Chinese ferret badgers in Southern China. J Virol 81:6920–6926, 2007.
71. Nuttall PA, Harrap KA. Isolation of a coronavirus during studies on puffinosis, a disease of the Manx shearwater (*Puffinus puffinus*). Arch Virol 73:1–13, 1982.
72. Gloza-Rausch F, Ipsen A, Seebens A, Gottsche M, Panning M, Felix Drexler J, Petersen N, Annan A, Grywna K, Muller M, Pfefferle S, Drosten C. Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. Emerg Infect Dis 14:626–631, 2008.
73. Carrington CV, Foster JE, Zhu HC, Zhang JX, Smith GJ, Thompson N, Auguste AJ, Ramkissoon V, Adesiyun AA, Guan Y. Detection and

phylogenetic analysis of group 1 coronaviruses in South American bats. Emerg Infect Dis 14:1890–1893, 2008.

74. Poon LL, Chu DK, Chan KH, Wong OK, Ellis TM, Leung YH, Lau SK, Woo PC, Suen KY, Yuen KY, Guan Y, Peiris JS. Identification of a novel coronavirus in bats. J Virol 79:2001–2009, 2005.
75. Wong S, Lau S, Woo P, Yuen KY. Bats as a continuing source of emerging infections in humans. Rev Med Virol 17:67–91, 2007.
76. Cavanagh D. Coronaviruses in poultry and other birds. Avian Pathol 34:439–448, 2005.
77. Vijaykrishna D, Smith GJ, Zhang JX, Peiris JS, Chen H, Guan Y. Evolutionary insights into the ecology of coronaviruses. J Virol 81: 4012–4020, 2007.
78. Huang Y, Lau SK, Woo PC, Yuen KY. CoVDB: a comprehensive database for comparative analysis of coronavirus genes and genomes. Nucleic Acids Res 36:D504–D511, 2008.
79. Woo PC, Wong BH, Huang Y, Lau SK, Yuen KY. Cytosine deamination and selection of CpG suppressed clones are the two major independent biological forces that shape codon usage bias in coronaviruses. Virology 369:431–442, 2007.