Molecular diversity and evolution of bat group C betacoronaviruses: origin of the novel human group C betacoronavirus (abridged secondary publication)

SKP Lau *, PCY Woo, BJ Zheng

KEY MESSAGES

1. MERS-CoV is closely related to both Pi-BatCoV HKU5 and Ty-BatCoV HKU4. Pi-BatCoV HKU5 may have a better ability to adapt to new host/environments along with its host living in diverse habitats.
2. A potentially novel lineage C betaCoV that shares genome similarities to MERS-CoV is identified in two bats in Guangdong province, supporting that *Pipistrellus*-related bats are important host of lineage C betaCoVs.
3. Two other novel betaCoVs—SARSr-Rf-BatCoV strains—were identified from greater horseshoe bats. Genome analysis showed that recombination has occurred around ORF8 between SARSr-Rf-BatCoVs and SARSr-Rs-BatCoVs, leading to the generation of civet SARSr-CoVs with ORF8 likely acquired from SARSr-Rf-BatCoVs.
4. Bats in China are important reservoir for diverse betaCoVs including lineage C viruses. This highlights the importance for conservation of these animals and their habitats. Although there is no evidence of direct transmission of CoVs from bats to humans, humans should avoid contact with wild bats.

Hong Kong Med J 2021;27(Suppl 2):S23-7
HMRF project number: 13121102

SKP Lau, PCY Woo, BJ Zheng

Department of Microbiology, The University of Hong Kong

* Principal applicant and corresponding author: skplau@hkucc.hku.hk

Introduction

In 2012, a novel human coronavirus (CoV) in the Middle East region was found to be associated with severe respiratory illness with high mortality rate. This novel CoV was first isolated from the lung tissue of a 60-year-old Saudi Arabian man with fatal acute pneumonia and renal failure. Subsequently, a resident of Qatar with a recent travel history to Saudi Arabia was diagnosed with similar symptoms and detected with the same virus.[1] This novel lineage C betaCoV was named Middle East respiratory syndrome-related coronavirus (MERS-CoV). As of 27 November 2016, of 1832 laboratory-confirmed cases of MERS-CoV infection reported to World Health Organization, there are 649 deaths, accounting for a crude fatality rate of 35%. The MERS-CoV is most closely related to bat lineage C betaCoVs, including *Tylonycteris* bat coronavirus HKU4 (Ty-BatCoV HKU4) and *Pipistrellus* bat coronavirus HKU5 (Pi-BatCoV HKU5) discovered in bats in 2006.[2,3] The virus was subsequently found in dromedary camels, which are likely the immediate source of some human cases. However, it remains unknown where and when the camels have acquired the virus, and whether bats are the ultimate reservoir and origin of MERS-CoV-like viruses, as in the case of SARS-CoV. So far, molecular epidemiology studies on lineage C betaCoVs are scarce, and little is known about how they evolve and diversify. We have reported that CoVs can be transmitted between two bat species of different suborders, implying the natural occurrence of interspecies transmission at any level.[4] Therefore, more intensive surveillance studies for betaCoVs in bats of different species may help understand the origin of the ancestor of MERS-CoV.

Methods

Prospectively collected samples from different bat species in China over a 10-year period were subject to betaCoV detection by RT-PCR. Positive samples were sequenced for complete RNA-dependent RNA polymerase and spike (S) and nucleocapsid (N) genes and were analysed for molecular diversity and evolution. Potential recombinant strains were selected for complete genome sequencing and analysed for possible recombination events. Evolutionary analysis was performed by molecular clock analysis to estimate time of divergence of different lineage C betaCoVs and emergence of MERS-CoV.

Respiratory and alimentary samples from 9866 bats of 54 different bat species captured from

different locations in Hong Kong and mainland China, including Guangdong, Guangxi, and Yunnan provinces, were collected over a 10-year period (2004 to 2014), in collaboration with Agriculture, Fisheries and Conservation Department of the Hong Kong SAR Government, Yunnan Institute of Endemic Diseases Control and Prevention, and Guangdong Entomological Institute. All specimens were immediately placed in viral transport medium before transportation to the laboratory for RNA extraction. Viral RNA was extracted from the bat samples and subjected to RT-PCR for detection of betaCoVs using a 440-bp fragment of RdRp gene. The PCR product of positive samples were sequenced and compared with known sequences of the RdRp genes of betaCoVs in the GenBank database. A preliminary phylogenetic analysis was then performed using these sequences.

Complete RdRp, S and N genes of 13 Ty-BatCoV HKU4 and 15 Pi-BatCoV HKU5 strains as well as the complete genomes of two SARSr-Rf-BatCoVs and the two strains of a novel lineage C betaCoV detected at different time and/or place were amplified and sequenced. The different complete gene sequences were used to construct the phylogenetic trees afterwards. Complete RdRp and N gene sequences were used to estimate the divergence time of Ty-BatCoV HKU4, Pi-BatCoV HKU5, and MERS-CoV. Moreover, the genome sequences of different SARSr-BatCoVs and SARSr-CoVs were subjected to Bootscan analysis in Simplot to detect possible recombination using sliding window approach with civet SARSr-CoV SZ3 as query.

Results

Molecular epidemiology of lineage C betaCoVs in different bat species in China

Among the 9866 bat samples, 267 (2.7%) were positive for betaCoVs by RT-PCR and sequencing of the partial RdRp gene (Table). Phylogenetic analysis showed the presence of Ty-BatCoV HKU4 in 29 alimentary samples from 241 lesser bamboo bat (*Tylonycteris pachypus*) and Pi-BatCoV HKU5 in 58 alimentary samples from 136 Japanese pipistrelle (*Pipistrellus abramus*), all from bats collected in Hong Kong from 2004 to 2014, whereas 19 alimentary samples from lesser bamboo bat collected in mainland China were also positive for Ty-BatCoV HKU4, and four alimentary samples from four bats collected in mainland China were also found to contain a potentially novel lineage C βCoV and a novel SARSr-Rf-BatCoV (Fig). In addition, 52 and 105 alimentary samples were found to contain Ro-BatCoV HKU9 and SARSr-Rs-BatCoV HKU3, respectively. None of the respiratory samples were positive for lineage C betaCoVs, suggesting that these viruses may exhibit enteric tropism. Bats positive for Ty-BatCoV HKU4 and Pi-BatCoV HKU5 were from seven and 13 sampling locations in Hong Kong, respectively, whereas Ty-BatCoV HKU4, the novel lineage C betaCoV and SARSr-Rf-BatCoV were detected in different provinces in mainland China. No obvious disease was observed in bats positive for Ty-BatCoV HKU4 and Pi-BatCoV HKU5. Ty-BatCoV HKU4 was found only in adult bats, whereas Pi-BatCoV HKU5 was found in both adult and juvenile bats.

Molecular diversity, evolution, and potential recombination in lineage C betaCoVs in bats

MERS-CoV was shown to be more closely related to Pi-BatCoV HKU5 than to Ty-BatCoV HKU4 (92.1%-92.3% vs 89.6%-90% identities) in the RdRp gene, but more closely related to Ty-BatCoV HKU4 than to Pi-BatCoV HKU5 in the S gene (66.8%-67.4% vs 63.4%-64.5% identities) and N gene (71.9%-72.3% vs 69.5%-70.5% identities) by comparing the deduced amino acid sequences of the genes. The Ty-BatCoV HKU4 strains and Pi-BatCoV HKU5 strains formed two distinct clusters in the phylogenetic tress of all three genes. Marked sequence polymorphisms were observed in the S gene of Pi-BatCoV HKU5, with up to 12% amino acid differences. Ty-batCoV HKU4, Pi-BatCoV HKU5, and MERS-CoV probably emerged from a common ancestor about 500 years before the epidemic.

Two samples collected from Chinese pipistrelle bats contained a novel lineage C betaCoV species. This novel lineage C betaCoV shares similar genome organisation with other lineage C betaCoVs and possessed 56.3%-78.3%, 78.4%-90.8%, 89.1%-96.4%, 92.0%-97.0%, 84.7%-97.3%, 76.9%-89.2%, and 84.8%-96.7% amino acid identities to other lineage C betaCoVs in the seven conserved replicase domains: ADRP, nsp5 (3CLpro), nsp12 (RdRp), nsp13 (Hel), nsp14 (ExoN), nsp15 (NendoU), and nsp16 (O-MT), respectively.

In addition, two novel betaCoVs—SARSr-Rf-BatCoV strains—were identified from greater horseshoe bats in mainland China. Their genomes shared 88.2% nucleotide identities to the genomes of SARSr-Rs-BatCoV HKU3 and 93% nucleotide identities to the genomes of human/civet SARSr-CoVs. The sequence identity between the ORF8 of SARSr-Rf-BatCoVs and human/civet SARSr-CoVs (80.4%-81.3% aa identity) was exceptionally high and was much higher than that between human/civet SARSr-CoVs and other SARSr-BatCoVs (23.2%-37.3% aa identity). Phylogenetic and sliding window analyses suggested potential recombination between SARSr-BatCoVs from different bat host species, in which the ORF8 of civet and human SARSr-CoV would have originated from SARSr-Rf-BatCoVs. Several recombination breakpoints were observed, and ORF8 was located between two breakpoints in particular.

TABLE. Detection of betacoronaviruses (betaCoVs) in bats by RT-PCR

Bat species	Common name	No. of bats tested	No. (%) of bats positive for betaCoVs in alimentary samples	BetaCoV species*
Megachiroptera				
Pteropodidae				
Cynopterus sphinx	Short-nosed fruit bat	130	0 (0)	-
Eonycteris spelaea	Cave nectar bat	6	0 (0)	-
Megaerops ecaudatus	Tailless fruit bat	1	0 (0)	-
Rousettus leschenaulti	Leschenault's rousette	741	52 (7.0)	Ro-BatCoV HKU9 (n_1=1; n_2=51)
Microchiroptera				
Emballonuridae				
Taphozous melanopogon	Black-bearded tomb bat	25	0 (0)	-
Taphozous sp		19	0 (0)	-
Hipposideridae				
Aselliscus stoliczkanus	Stoliczka's trident bat	48	0 (0)	-
Hipposideros armiger	Himalayan leaf-nosed bat	365	0 (0)	-
Hipposideros larvatus	Intermediate roundleaf bat	142	0 (0)	-
Hipposideros pomona	Pomona leaf-nosed bat	1001	0 (0)	-
Hipposideros pratti	Pratt's roundleaf bat	19	0 (0)	-
Megadermatidae				
Megaderma lyra	Greater false vampire bat	1	0 (0)	-
Rhinolophidae				
Coelops frithi	East Asian tailless leaf-nosed bat	9	0 (0)	-
Rhinolophus affinus	Intermediate horseshoe bat	670	0 (0)	-
Rhinolophus ferrumequinum	Greater horseshoe bat	97	2 (2.1)	SARSr-Rf-BatCoV (n_1=0; n_2=2)
Rhinolophus luctus	Woolly horseshoe bat	63	0 (0)	-
Rhinolophus macrotis	Big-eared horseshoe bat	18	0 (0)	-
Rhinolophus osgoodi	Osgood's horseshoe bat	1	0 (0)	-
Rhinolophus pearsoni	Pearson's horseshoe bat	24	0 (0)	-
Rhinolophus pusillus	Least horseshoe bat	419	0 (0)	-
Rhinolophus rex	King horseshoe bat	2	0 (0)	-
Rhinolophus siamensis	Thai horseshoe bat	3	0 (0)	-
Rhinolophus sinicus	Chinese horseshoe bat	2430	105 (4.3)	SARSr-Rs-BatCoV HKU3 (n_1=99; n_2=6)
Rhinolophus stheno	Lesser brown horseshoe bat	34	0 (0)	-
Rhinolophus thomasi	Thomas's horseshoe bat	1	0 (0)	-
Rhinolophus sp.		22	0 (0)	-
Vespertilionidae				
Eptesicus sp.		1	0 (0)	-
Hypsugo pulveratus	Chinese pipistrelle	11	2 (1.8)	novel lineage C betaCoV (n_1=0; n_2=2)
Ia io	Great evening bat	20	0 (0)	-
Miniopterus fuliginosus	Eastern bent-winged bat	90	0 (0)	-
Miniopterus magnater	Greater bent-winged bat	29	0 (0)	-
Miniopterus pusillus	Lesser bent-winged bat	541	0 (0)	-
Miniopterus schreibersii	Common bent-winged bat	1016	0 (0)	-

* n_1= No. of bats positive for betaCoVs identified in Hong Kong; n_2= No. of bats positive for betaCoVs identified in mainland China.

TABLE. (cont'd)

Bat species	Common name	No. of bats tested	No. (%) of bats positive for betaCoVs in alimentary samples	BetaCoV species*
Myotis adversus	Large-footed mouse-eared bat	4	0 (0)	-
Myotis altarium	Szechwan myotis	15	0 (0)	-
Myotis chinensis	Chinese myotis	224	0 (0)	-
Myotis daubentoniid	Daubenton's bat	98	0 (0)	-
Myotis fimbriatus	Fringed long-footed myotis	6	0 (0)	-
Myotis formosus	Hodgson's bat	1	0 (0)	-
Myotis horsfieldii	Horsfield's bat	7	0 (0)	-
Myotis longipes	Kashmir cave bat	5	0 (0)	-
Myotis muricola	Whiskered myotis	4	0 (0)	-
Myotis pequinius	Beijing mouse-eared bat	29	0 (0)	-
Myotis ricketti	Rickett's big-footed bat	451	0 (0)	-
Myotis sp.		1	0 (0)	-
Nyctalus noctula	Brown noctule	56	0 (0)	-
Nyctalus plancyi	Chinese noctule	1	0 (0)	-
Pipistrellus abramus	Japanese pipistrelle	314	58 (18.5)	Pi-BatCoV HKU5 (n_1=58; n_2=0)
Pipistrellus minus	Wroughton's pipistrelle	3	0 (0)	-
Pipistrellus pipistrellus	Common pipistrelle	6	0 (0)	-
Pipistrellus tenuis	Least pipistrelle	11	0 (0)	-
Scotophilus kuhlii	Lesser yellow bat	220	0 (0)	-
Tylonycteris pachypus	Lesser bamboo bat	306	48 (15.7)	Ty-BatCoV HKU4 (n_1=29; n_2=19)
Tylonycteris robustula	Greater bamboo bat	105	0 (0)	-

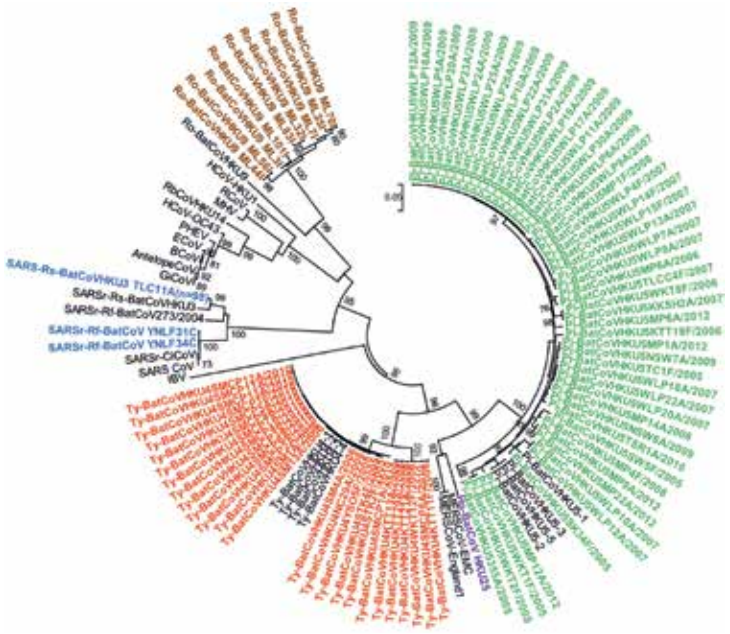

FIG. Phylogenetic analysis of nucleotide sequences of the 311-bp fragment of RNA-dependent RNA polymerase (RdRp) of the betacoronaviruses (betaCoVs) identified in bats. The tree is constructed by neighbour-joining method and Jukes-Cantor model with bootstrap values calculated from 1000 trees by MEGA 6.0. Only bootstrap values >70% are shown. Scale bar indicates 20 nucleotide substitutions per site.

Discussion

In this extensive epidemiological study, betaCoVs were detected in 267 of 9866 bats, giving a detection rate of 2.7%. At least five different betaCoV species are circulating in specific bat species including Ty-Bat CoV HKU4, Pi-BatCoV HKU5, a novel lineage C betaCoV, SARSr-BatCoVs, and Ro-BatCoV HKU9. This supports bats are important reservoirs for betaCoVs. Ty-BatCoV HKU4 and Pi-BatCoV HKU5 were highly prevalent among lesser bamboo bats and Japanese pipistrelle in Hong Kong, respectively, with detection rates of 21% to 24% in their alimentary samples. Pi-BatCoV HKU5 may have a better ability to adapt to new host/environments, given its high divergence in the S gene sequences. This may be explained by the diverse habitat of its host Japanese pipistrelle, which would favour the evolution of the viral S protein, in particular the receptor binding domain, and allow efficient interspecies transmission to other animals or human. Using molecular clock analysis, we showed that Ty-BatCoV HKU4, Pi-BatCoV HKU5, and MERS-CoV have diverged at least centuries ago from their common ancestor. These two bat lineage C betaCoVs are different from

SARSr-CoV, which diverged between civet and bat strains only several years before the SARS epidemic. Therefore, they are not likely the direct ancestor of MERS-CoV. However, another novel lineage C betaCoV species was detected in two bat samples collected from Chinese pipistrelle. Although the genome of this novel CoV only possessed about 73% nucleotide identities to that of human/camel MERS-CoV, this finding further supports that *Pipistrellus*-related bats are likely important host of lineage C betaCoVs. To understand how these bat lineage C betaCoVs may have evolved leading to the emergence of MERS-CoV, more epidemiological studies on bats and other animals are warranted.

In addition to the lineage C betaCoVs, two novel betaCoVs—SARSr-Rf-BatCoV strains—were identified from greater horseshoe bats. Their genomes only shared 93% nucleotide identities to that of human/civet SARSr-CoVs, but the ORF8 of these two bat CoVs was highly similar to that of civet SARSr-CoV among all SARSr-BatCoVs. Based on the results from phylogenetic analysis and the identification of potential recombination sites between SARS-Rs-BatCoVs and SARS-Rf-BatCoVs around the ORF8 region, civet SARSr-CoV SZ3 is likely to have originated from genetic recombination between SARS-Rs-BatCoVs and SARS-Rf-BatCoVs from different horseshoe bat species with the ORF8 acquired from SARSr-Rf-BatCoVs. Further studies should be sought to understand how the evolution of ORF8 may have played a role during interspecies transmission of SARSr-CoV.

The present results are important for future research on the emergence of CoVs in humans. It has provided clues on the animal origins and the evolutionary pathways of MERS-CoV and SARS-CoV. We have also discovered a potentially novel lineage C betaCoV closely related to MERS-CoV. Hence, future studies may focus on Pipistrellus and Rhinolophus or related bats to identify yet closer ancestors of MERS-CoV and SARS-CoV. This will help better understand how CoVs may emerge in humans and predict the next possible epidemic. Our findings support a diversity of betaCoVs among bats in China. Therefore, it is important to conserve these animals and their habitats. Although there is no evidence of direct transmission of CoVs from bats to humans, humans should avoid contact with wild bats to prevent transmission of other viruses such as rabies.

Funding

This study was supported by the Health and Medical Research Fund, Food and Health Bureau, Hong Kong SAR Government (#13121102). The full report is available from the Health and Medical Research Fund website (https://rfs1.fhb.gov.hk/index.html).

Disclosure

The results of this research have been previously published in:

1. Lau SK, Li KS, Tsang AK, et al. Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. J Virol 2013;87:8638-50.
2. Lau SK, Feng Y, Chen H, et al. Severe acute respiratory syndrome (SARS) coronavirus ORF8 protein is acquired from SARS-related coronavirus from greater horseshoe bats through recombination. J Virol 2015;89:10532-47.
3. Lau SKP, Zhang L, Luk HKH, et al. Receptor usage of a novel bat lineage C betacoronavirus reveals evolution of Middle East respiratory syndrome-related coronavirus spike proteins for human dipeptidyl peptidase 4 binding. J Infect Dis 2018;218:197-207.

References

1. Butler D. SARS veterans tackle coronavirus. Nature 2012;490:20.
2. Woo PC, Wang M, Lau SK, et al. Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. J Virol 2007;81:1574-85.
3. Woo PC, Lau SK, Li KS, Tsang AK, Yuen KY. Genetic relatedness of the novel human group C betacoronavirus to Tylonycteris bat coronavirus HKU4 and Pipistrellus bat coronavirus HKU5. Emerg Microbes Infect 2012;11:e35.
4. Lau SK, Li KS, Tsang AK, et al. Recent transmission of a novel alphacoronaivurs, bat coronavirus HKU10, from Leschenault's rousettes to pomona leaf-nosed bats: first evidence of interspecies transmission of coronavirus between bats of different suborders. J Virol 2012;86:11906-18.