

Identification of coronaviruses in farmed wild animals reveals their evolutionary origins in Guangdong, southern China

Xiaohu Wang,¹ Zhaowen Ren,^{1,2,†} Lu Wang,^{1,2} Jing Chen,¹ Pian Zhang,^{1,2} Jin-Ping Chen,³ Xiaofan Chen,^{1,2} Linmiao Li,³ Xuhui Lin,¹ Nanshan Qi,¹ Shengjun Luo,¹ Rong Xiang,¹ Ziguo Yuan,² Jianfeng Zhang,¹ Gang Wang,¹ Min-Hua Sun,¹ Yuan Huang,¹ Yan Hua,⁴ Jiejian Zou,⁴ Fanghui Hou,⁴ Zhong Huang,¹ Shouwen Du,⁵ Hua Xiang,^{1,*} Mingfei Sun,^{1,*} Quan Liu,^{1,6,*} and Ming Liao^{1,2,*}

¹Key Laboratory of Livestock Disease Prevention of Guangdong Province, Scientific Observation and Key Laboratory for prevention and control of Avian Influenza and Other Major Poultry Diseases, Ministry of Agriculture and Rural Affairs, Institute of Animal Health, Guangdong Academy of Agricultural Sciences, Baishigang, Wushan Street, Tianhe District, Guangzhou 510640, China, ²College of Veterinary Medicine, South China Agricultural University, No. 483 Wushan Road, Tianhe District, Guangzhou 510600, China, ³Institute of Zoology, Guangdong Academy of Sciences, No. 105 Xingang West Road, Haizhu District, Guangzhou 510260, China, ⁴Guangdong Provincial Wildlife Rescue Center, No. 139 Yuxi Road, Tianhe District, Guangzhou 510520, China, ⁵Department of Infectious Diseases, The Second Clinical Medical College (Shenzhen People's Hospital) of Jinan University, No. 1017 Dongmen North Road, Luohu District, Shenzhen 518020, China and ⁶School of Life Sciences and Engineering, Foshan University, No. 33 Guangyun Road, Shishan Town, Nanhai District, Foshan 528225, China

[†]<https://orcid.org/0000-0003-3350-8372>

[†]<https://orcid.org/0000-0003-3411-3196>

*Corresponding author: E-mail: xiangh898@163.com; smf7810@126.com; liuquan1973@hotmail.com; mliao@scau.edu.cn

Abstract

Coronavirus infections cause diseases that range from mild to severe in mammals and birds. In this study, we detected coronavirus infections in 748 farmed wild animals of 23 species in Guangdong, southern China, by RT-PCR and metagenomic analysis. We identified four coronaviruses in these wild animals and analysed their evolutionary origins. Coronaviruses detected in *Rhizomys sinensis* were genetically grouped into canine and rodent coronaviruses, which were likely recombinants of canine and rodent coronaviruses. The coronavirus found in *Phasianus colchicus* was a recombinant pheasant coronavirus of turkey coronavirus and infectious bronchitis virus. The coronavirus in *Paguma larvata* had a high nucleotide identity (94.6–98.5 per cent) with a coronavirus of bottlenose dolphin (*Tursiops truncatus*). These findings suggested that the wildlife coronaviruses may have experienced homologous recombination and/or crossed the species barrier, likely resulting in the emergence of new coronaviruses. It is necessary to reduce human–animal interactions by prohibiting the eating and raising of wild animals, which may contribute to preventing the emergence of the next coronavirus pandemic.

Key words: coronavirus; wild animals; *Rhizomys sinensis*; *Phasianus colchicus*; *Paguma larvata*; southern China.

1. Introduction

Coronaviruses (CoVs) are positive-sense single-stranded RNA viruses in the *Coronaviridae* family that can cause the common cold or severe diseases such as pneumonia, bronchiolitis, and gastroenteritis in birds and mammals, including humans (Woo et al. 2012; Alluwaimi et al. 2020). CoVs are comprised of four genera, *Alphacoronavirus* (α -CoV), *Betacoronavirus* (β -CoV), *Gamma-coronavirus* (γ -CoV), and *Deltacoronavirus* (δ -CoV), in which α -CoVs and β -CoVs mainly infect mammals, while γ -CoVs and δ -CoVs primarily infect birds (Woo et al. 2012).

There are at least seven CoVs associated with human diseases, including human coronavirus (HCoV)-HKU1, -OC43, -NL63, -229E, severe acute respiratory syndrome-associated coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), and the recently emerged SARS-CoV-2 (Chen et al. 2020). SARS-CoV and MERS-CoV are considered to have originated from bats and transmitted from palm civets and dromedary camels to

humans (Lu, Wang, and Gao 2015; Cui, Li, and Shi 2019). HCoV-NL63 and -229E may also have originated from bats, whereas HCoV-HKU1 is genetically related to rodent-associated viruses (Corman et al. 2018). HCoV-OC43 is thought to have emerged from cattle (Vijgen et al. 2005). Both bats and pangolins may be potential hosts of SARS-CoV-2 (Gordon et al. 2020; Zhang, Wu, and Zhang 2020; Wacharapluesadee et al. 2021).

The high diversity of coronaviruses identified in bats suggests the origin of many human and animal coronaviruses (Cui, Li, and Shi 2019). In China, a SARS-CoV-like virus has been isolated from palm civets in Guangdong (Song et al. 2005); nearly 2 per cent of 1,465 rodents in Zhejiang have tested positive for α -CoV and β -CoV (Wang et al. 2015). A variant of CoV HKU24 has been found in rats in Qinghai, where *Montifringilla taczanowskii* has been found to be infected with a novel δ -CoV closely related to sparrow δ -CoV ISU42824 (Zhu et al. 2021). Recently, a large-scale surveillance of CoVs in birds in China identified a new γ -CoV, pigeons coronavirus

(Zhuang et al. 2020). These findings help to understand virus origins and evolution of emerging coronaviruses.

The unique replication machinery of CoVs facilitates viral recombination (Lai and Cavanagh 1997; Latinne et al. 2020). In the case of a host coinfecting with more than one CoV strain, the RNA polymerase can jump from the RNA of one strain to that of the other strain, synthesizing a hybrid RNA from both viruses. Recombination may occur not only within CoVs (homologous recombination) but also with different RNA viruses or other organisms (heterologous recombination), by which CoVs can acquire novel biological properties in terms of virulence, host range, and tissue tropism. Therefore, CoVs that are nonpathogenic or have low pathogenicity in the original host may increase their pathogenicity or adapt to different species, leading to rapid spread in the new host (Su et al. 2016).

Southern China is considered to be a centre of CoV diversification (Allen et al. 2017; Latinne et al. 2020). To date, two bat-derived coronaviruses, including SARS-CoV and swine acute diarrhoea syndrome coronavirus, were first identified in Guangdong, southern China (Ksiazek et al. 2003; Zhou et al. 2018). In the present study, we identified four viral members of the *Alphacoronavirus*, *Betacoronavirus*, and *Gammacoronavirus* genera in farmed wild animals in Guangdong, all of which had experienced homologous recombination and/or crossed the species barrier.

2. Materials and methods

2.1 Sample collection

During 17–24 February 2020, we collected wild animal samples from forty-three farms in Guangdong Province (Supplementary Fig. S1). In addition, three dead Malayan pangolins intercepted from Guangdong customs in July 2019 were included in the study (Li et al. 2021). These samples included throat and anal swabs, tissues, and blood specimens. For the tissue samples, the animals were euthanized, followed by autopsy and collection of the trachea, heart, liver, spleen, lung, kidney, gastrointestinal tract, and gallbladder. Blood samples were collected from the wing vein or chicken comb in birds and from the tail vein, ear vein, or jugular vein in mammals. Blood was collected without anticoagulant for serum separation. All samples from the same animal were pooled for further analysis.

The protocols for the animal study were reviewed and approved by the Committee on the Ethics of Animal Experiments of the Institute of Animal Health, Guangdong Academy of Agricultural Science, China.

2.2 Molecular detection of coronaviruses

Samples from individual animals were used to extract viral RNA with the QIAamp Viral RNA Mini kit, and synthesis of cDNA was conducted with a Thermo Scientific RevertAid RT kit (Thermo K1691, USA). All the samples were assessed by RT-PCR with the universal primers of coronaviruses as described elsewhere (Onyuk et al. 2019).

2.3 RNA extraction, library preparation, and sequencing

We further used viral metagenomics to analyse coronavirus infection in these positive samples. The tissue samples were homogenized, followed by slow speed centrifugation to remove the tissue debris, and the viral particles were purified and concentrated by filtration and ultracentrifugation as described elsewhere (Gong et al. 2015). Briefly, 5 g tissue samples were homogenized, centrifuged at $12,000 \times g$ for 15 min at 4°C , and filtered through a $0.22 \mu\text{m}$

filter. The supernatants were transferred into a new tube containing 28 per cent (W/W) sucrose. After centrifugation at $20,000 \times g$ for 2 h at 4°C , precipitates were resuspended and digested with DNase and RNase for 1 h at 37°C .

Viral nucleic acid was extracted using the MagPure Viral DNA/RNA Mini LQ Kit (Magen R6662-02), and genome-wide amplification was conducted with the REPLI-g Cell WGA & WTA kit (Qiagen 150,054). The PCR products were analysed by Life Technologies Qubit 4.0 and 1.5 per cent agarose electrophoresis.

Sequencing libraries were prepared using the NEB Next[®] Ultra[™] DNA Library Prep Kit for Illumina[®] (New England Biolabs, MA, USA) following the manufacturer's recommendations and index codes were added. The library quality was assessed on the Qubit[®] dsDNA HS Assay Kit (Life Technologies, Grand Island, NY) and Agilent 4,200 system (Agilent, Santa Clara, CA). The library was sequenced on an Illumina NovaSeq 6,000, and 150 bp paired-end reads were generated.

2.4 Viral metagenomic analysis

Viral metagenomic analysis was conducted according to the standard procedure by the Magigene Company (Guangzhou, China) as described elsewhere (Liu, Chen, and Chen 2019). Briefly, SOAPnuke software version 1.5.6 (Chen et al. 2018) was used to remove low-quality data to generate clean data, which were further mapped to the ribosomal database and the host reference genome utilizing Burrows–Wheeler Alignment (BWA) software version 0.7.17 (Li and Durbin 2009). Clean reads without ribosome and host sequences were mapped to the virus reference data derived from the GenBank nonredundant nucleotide (NT) database to primarily identify virus reads. Clean reads were *de novo* assembled using MEGAHIT version 1.0 (Li et al. 2016), and CD-HIT version 4.7 (similarity of more than 95 per cent and overlapping area of more than 90 per cent) (Fu et al. 2012) was used to cluster the assembled viral contigs from all samples. Contigs were then classified by BLASTx against the NT database using similarity ≥ 80 per cent, matched length ≥ 500 bp and e-value $\leq 10^{-5}$ (Camacho et al. 2009). Contigs with significant BLASTx hits were confirmed as virus sequences. Viral diversity was evaluated by using the Rhea package for alpha diversity (Lagkouravdos et al. 2017) and the phyloseq package for beta diversity (McMurdie and Holmes 2013).

2.5 Sequence comparison and phylogenetic and recombinant analyses

The genomic sequences were assembled by SeqMan NGen[®], version 7.1 (DNASTAR, Madison, WI) and aligned using MAFFT version 7.487 with the parameter L-INS-I (Katoh et al. 2002). Gaps in incomplete virus genomes were filled by RT-PCR on the individual RNA samples that contained the target virus, and genome termini were determined by using 5'/3' RACE kits (TaKaRa).

We used the Megalign program (Lasergene, version 7.1) to determine the nucleotide and amino acid sequence similarities, and IQ-TREE version 2.1.3 (Minh et al. 2020) to construct a phylogenetic tree by the maximum-likelihood (ML) method; the bootstrap support values were calculated from 1,000 replicates. Partition-Finder 2 (Chernomor, von Haeseler, and Minh 2016; Lanfear et al. 2017) was used to determine the best-fit partitioning scheme and evolution model for genome-wide phylogenetic analysis, and ModelFinder Plus (Kalyaanamoorthy et al. 2017) was used to determine the most suitable evolution model for the phylogenetic tree of the partial coronaviral genome.

We used the recombination detection program (RDP) (version 5.05) with the methods of RDP, GENECONV, BootScan, Maxchi, Chimaera, and 3Seq to evaluate potential recombination events,

respectively (Martin et al. 2021). All analyses were performed with a Bonferroni corrected P value cut-off of 0.01. To further characterize these recombination events, SimPlot software version 3.5.1 was used to infer similarity plots between our sequences and the reference sequences (Lole et al. 1999). The topological discrepancy was evaluated using RELL, KH, SH, ELW, and AU tests in IQ-TREE version 2.1.3 by constraining the topology of the recombinant regions with the topology of the nonrecombinant regions and vice versa (<http://www.iqtree.org/doc/Advanced-Tutorial#tree-topology-tests>). The Bayesian method BEAST version 2.6.6 was used to analyse the potential origin of the virus recombination (Muller, Kistler, and Bedford 2022). We used Bayesian Evaluation of Temporal Signal (BETS) (Duchene et al. 2020) to assess the temporal structure for the genome datasets of RCoVs and CCoVs.

Whole-genome phylogenetic analysis included representative species of the Alphacoronavirus and Betacoronavirus genera, as well as other strains that infect rodents and dogs. For the Gamma-coronavirus genus, representative infectious bronchitis virus (IBV) strains and coronaviruses from birds and aquatic mammals were included as reference sequences. In the recombination analysis, we determined reference sequences from coronaviruses (nucleotide sequence >20,000 nt) in rodents or dogs using RDP software, in which the coronaviruses without recombination events with the identified viruses in this study were excluded. Information on the reference genome retrieved from GenBank and ViPR in this study is listed in Supplementary Table S1.

2.6 Nucleotide sequence accession numbers

All the genome sequences were submitted to GenBank (accession numbers are provided in Supplementary Table S2), and the metagenomic data were deposited into the NCBI sequence archive under accession number PRJNA751997.

3. Results

3.1 Animal sampling

During 17–24 February 2020, we collected tissue samples, pharyngeal or anal swabs from 745 farmed wild animals in fourteen counties across Guangdong Province, southern China (Supplementary Fig. S1), and three dead Malayan pangolins that were intercepted in Guangdong customs in July 2019. In total, these samples were from nine mammals, including an Asian black bear (*Ursus thibetanus*), bamboo rat (*Rhizomys sinensis*), wild boar (*Sus scrofa*), hedgehog (*Erinaceus europaeus*), porcupine (*Hystriacida*), sika deer (*Cervus nippon*), palm civet (*Paguma larvata*), pangolin (*Manis javanica*), and rhesus monkey (*Macaca mulatta*), three amphibians, including tiger frog (*Rana tigrina*), Chinese spiny frog (*Quasipaa spinosa*), and Chinese giant salamander (*Andrias davidianus*), nine reptiles, including a Siamese crocodile (*Crocodylus siamensis*), Chinese cobra (*Naja atra*), Chinese water snake (*Enhydris chinensis*), oriental rat-snake (*Ptyas mucosus*), monocled cobra (*Naja kaouthia*), Chinese pond turtle (*Chinemys reevesii*), red-eared turtle (*Trachemys scripta elegans*), pond slider (*Trachemys scripta*), and a miscellaneous turtle, and two avians, including a ring-necked pheasant (*P. colchicus*) and swan goose (*Anser cygnoides*) (Table 1).

3.2 Detection of coronaviruses

We used universal primers for coronaviruses to screen for possible infections (Onyuk et al. 2019), showing that thirty-six animals may be potentially infected with coronaviruses (Table 1), including the Chinese spiny frog (1/15), oriental rat-snake (1/127), Chinese

Table 1. The farmed wild animals involved in the study in Guangdong province, China.

Animal classification	Animal species	No. of animals	No. of CoV-positive animals ^a
Mammal	Bamboo rat (<i>Rhizomys sinensis</i>)	195	30
	Wild boar (<i>Sus scrofa</i>)	5	0
	Hedgehog (<i>Erinaceus europaeus</i>)	20	0
	Porcupine (<i>Hystriacida</i>)	30	0
	Sika deer (<i>Cervus nippon</i>)	3	0
	Asian black bear (<i>Ursus thibetanus</i>)	24	0
	Palm civets (<i>Paguma larvata</i>)	73	1
	Pangolin (<i>Manis javanica</i>)	3	0
	Rhesus monkey (<i>Macaca mulatta</i>)	10	0
	Subtotal	363	31
Amphibian	Tiger frog (<i>Rana tigrina</i>)	20	0
	Chinese spiny frog (<i>Quasipaa spinosa</i>)	15	1
	Chinese giant salamander (<i>Andrias davidianus</i>)	1	0
	Subtotal	36	1
Reptile	Siamese crocodile (<i>Crocodylus siamensis</i>)	26	0
	Chinese water snake (<i>Enhydris chinensis</i>)	6	0
	Oriental rat-snake (<i>Ptyas mucosus</i>)	127	1
	Chinese cobra (<i>Naja atra</i>)	44	1
	Monocled cobra (<i>Naja kaouthia</i>)	7	0
	Chinese pond turtle (<i>Chinemys reevesii</i>)	6	0
	Red-eared turtle (<i>Trachemys scripta elegans</i>)	6	0
	Pond slider (<i>Trachemys scripta</i>)	3	0
	Miscellaneous turtle	5	0
	Subtotal	230	2
Avian	Ring-necked pheasant (<i>Phasianus colchicus</i>)	99	2
	Swan goose (<i>Anser cygnoides</i>)	20	0
	Subtotal	119	2
Total		748	36

^aCoronavirus infection was tested in farmed wild animals by using a nested RT-PCR with the universal primers.

cobra (1/44), ring-necked pheasant (2/99), bamboo rat (30/175), and masked palm civet (1/73).

3.3 Metagenomic analysis and virome overview

The coronavirus-positive samples and pangolin samples were used for metagenomic analysis, resulting in a total of 1,179.16 Gb of raw nucleotide data (3,930,470,746 valid reads, approximately 150bp in length, Q20 values of 83.11–91.15, and Q30 values of 74.19–84.45). Through data filtering, low-quality data were eliminated, resulting in 694.09 Gb of high-quality nucleotide data (2,758,449,665 clean reads). Subsequently, through assembly, these clean reads were spliced into 2,409,448 contigs.

Furthermore, through the alignment, we discarded the contigs that were classified into prokaryotes, eukaryotes, or those having no significant similarity to any amino acid (aa) sequence in the nonredundant (NR) database, resulting in 41,088 contigs that matched with virus protein sequences (1.71 per cent of the total contigs). The number of virus-associated contigs for different animal species ranged from 117 to 38,898 (Supplementary Table S3). To count the distribution of virus reads, we also determined the abundance statistics of viruses, showing that there were 5,824,828 virus-associated reads, ranging from 85,480 to 4,480,983 in the different animal species (Supplementary Table S3).

These reads were assigned to a wide range of RNA and DNA viruses from forty-six viral families, including twenty-five viral families of mammals, and twenty-one viral families of arthropods and phages (Fig. 1). The greatest viral diversity was observed in bamboo rats (twenty-five virus families), followed by pheasants (seventeen virus families) and palm civets (nine virus families). In contrast, Malayan pangolin, *Quasipaa spinosa* and snake showed the presence of fewer viral families. Viral reads from the families Anelloviridae, Astroviridae, Circoviridae, Coronaviridae, Herpesviridae, Picobirnaviridae, Picornaviridae, Retroviridae, and the subfamily Parvovirinae, were widely distributed in different animal species.

High Shannon diversity indices were revealed in *Rhizomys sinensis* (Samples W1, X6, X4, Z2, X5, X7, and X8 in Supplementary Table S4), *P. colchicus*, and *Quasipaa spinosa*, showing rich viral diversity in these animals. In contrast, a low viral diversity was found in *R. sinensis* (Samples W6, W2, Z4, and W3 in Supplementary Table S4) and *Manis javanica* J3 (Supplementary Table S4). We further calculated beta diversity using Bray Curtis dissimilarity, showing that *R. sinensis* from different regions had different viral abundance, which was different among different individuals in the same region, especially in Shaoguan (Supplementary Fig. S2A). A similar viral abundance was observed in *R. sinensis* from Zhanjiang, Yunfu, Shaoguan, Meizhou, and Heyuan. However, the viral abundance in animals in Guangzhou was different from that in other regions (Supplementary Fig. S2A). There was a different viral abundance among different animal species, with the exception of snakes and pheasants, which had similar viral abundances (Supplementary Fig. S2B). These results showed that there was a certain correlation between the viral abundance in CoV-positive animals and geographic location or animal species.

As we focused on the coronaviruses, 499 (1.22 per cent) contigs, belonging to the family Coronaviridae, were further screened; of these, 454 contigs were from *Rhizomys sinensis*, thirty-one were from pheasants, and five were from palm civets.

3.4 Coronaviruses in *R. sinensis*

Metagenomic analysis generated 222,168 coronaviral reads, accounting for 4.96 per cent of the total mammalian virus families in *R. sinensis* (Fig. 2A). These viral reads were assembled into 485 contigs, of which 397 belonged to the *Alphacoronavirus* genus, and fifty-seven belonged to the *Betacoronavirus* genus. We further obtained the complete or nearly complete genome sequence of three strains of coronavirus from *R. sinensis* using reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) assays.

The first isolate, which was found to belong to the *Betacoronavirus* (β -CoV) genus, had a genome size of 31,157 bp, with a G+C content of approximately 38.7 per cent. Similar to other β -CoVs, the virus had a genome organization and genes characteristic of CoV, including the ORF1ab, haemagglutinin-esterase (HE), spike (S), envelope (E), membrane (M), and nucleocapsid (N) genes.

In addition, ORFs likely coding for accessory proteins, such as nonstructural protein 2a (NS2a), NS4, and NS5, were also found (Fig. 2B).

Sequence comparison showed that the virus had high nucleotide (nt) identities of 92.7–94.76 per cent compared with β -CoVs identified in rodents in China, such as RtBi-CoV/FJ2015 (*Bandicota indica*), Longquan-370 (*Niviventer confucianus*), RtMm-CoV/GD2015 (*Mus musculus*), and RtRn-CoV/YN2013 (*Rattus norvegicus*), but had low identities of 49.7–79.9 per cent compared with β -CoVs identified in other hosts, including HKU14 (rabbit), K37 (dog), and the human coronavirus HKU1, SARS-CoV, and SARS-CoV-2 (Supplementary Table S5). Further comparison of the replicase domains showed >90 per cent similarity with rodent coronaviruses within the *Betacoronavirus* genus (Supplementary Table S6). Therefore, the isolate was classified as rodent coronavirus, and designated RtRs-CoV/GD/2020.

Phylogenetic analysis based on the complete genome sequence showed that RtRs-CoV/GD/2020, together with rodent coronaviruses, such as RtBi-CoV/FJ2015, RtNn-CoV/SAX2015, RtMm-CoV/GD2015, and Longquan-370/708, formed a unique evolutionary clade in lineage A within the *Betacoronavirus* genus (Fig. 2C). However, changes in the topological structure in the phylogenetic position were observed among different genes of the RtRs-CoV/GD/2020 genome in relation to members of the *Betacoronavirus* genus (Supplementary Fig. S3). For the S, M, and N genes, RtRs-CoV/GD/2020 was most closely related to RtNn-CoV/SAX2015; for the E gene, RtRs-CoV/GD/2020 was more closely related to RtBi-CoV/FJ2015; and for haemagglutinin-esterase (HE), RtRs-CoV/GD/2020 was more closely related to Longquan-708. This suggests that recombination may have occurred during their evolution. Recombination analysis by RDP software revealed that RtRs-CoV/GD/2020 was most likely a recombinant of RtBi-CoV/FJ2015, Longquan-708, and RtNn-CoV/SAX2015 (Supplementary Table S7), which was further confirmed by testing a constrained tree (Supplementary Table S8). Simplot analysis showed that RtRs-CoV/GD/2020 contained genome fragments derived from RtBi-CoV/FJ2015 (genome regions 1, 3, 5, 8, and 10), Longquan-708 (genome regions 2, 4, and 6), and RtNn-CoV/SAX2015 (genome regions 7, 9, and 11) (Fig. 3).

BETS revealed a log Bayes factor (BF) of 41 for the RCoV genome dataset, suggesting sufficient temporal signals for Bayesian dating analysis. The evolutionary history of RCoVs was characterized by frequent recombination events, with a recombination rate of approximately 4.97×10^{-5} substitutions per site per year from the Bayesian analysis. RtBi-CoV/FJ2015 was the most recent common ancestor (MRCA) of RtRs-CoV/GD/2020 (Supplementary Fig. S4), and the median estimate of MRCA was 2001 (95 per cent CI, 1991–2017) for RtBi-CoV/FJ2015, 1991 (95 per cent CI, 1977–2015) for Longquan-708, and 1962 (95 per cent CI, 1946–1996) for RtNn-CoV/SAX2015 (Supplementary Fig. S4).

The other two isolates belonged to the *Alphacoronavirus* (α -CoV) genus, with genomes that included 28,823 nt and 28,807 nt, tentatively named X8 and X9, respectively. Their genome organization was similar to those of other α -CoVs, which included the ORF1ab, S, E, ORF3a ~ ORF3c, M, N, and ORF7a-7b genes (Fig. 2B). Sequence comparison showed a high nucleotide identity of 98.4 per cent between X8 and X9 at the complete genome level, and X8 had higher nucleotide identities to canine coronaviruses than transmissible gastroenteritis coronavirus (TGEV) or feline coronaviruses based on the complete genome and the predicted encoding genes (Supplementary Table S9). Phylogenetic analysis based on the complete genome sequence showed that X8 and X9 formed a unique clade with canine coronaviruses in the *Alphacoronavirus*

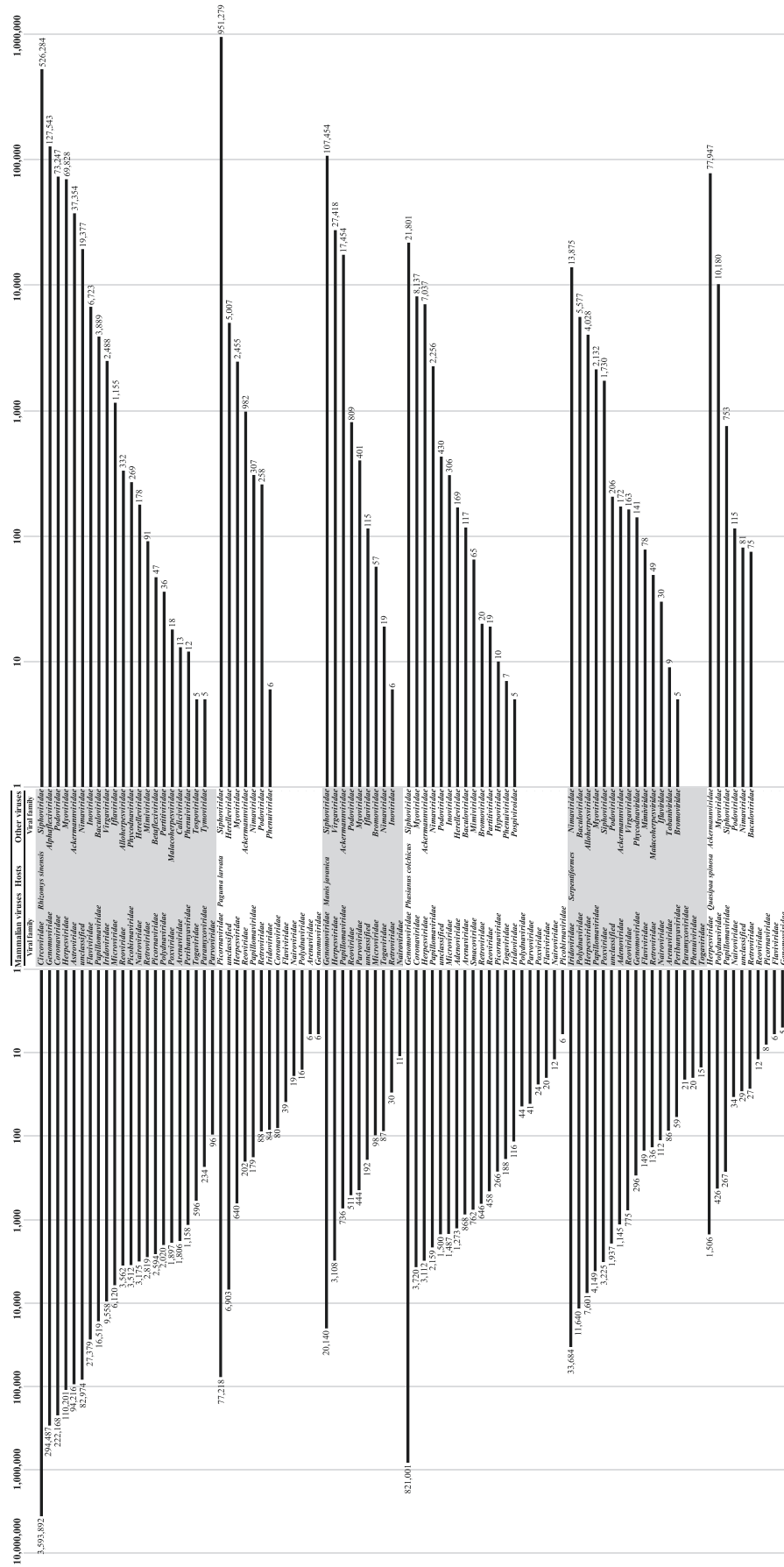


Figure 1. Overview of the diversity and abundance of the viruses classified by viral family in different farmed wild animals. The number of reads annotated to each virus family is shown, with the mammalian virus families on the left and other phage and plant virus families on the right.

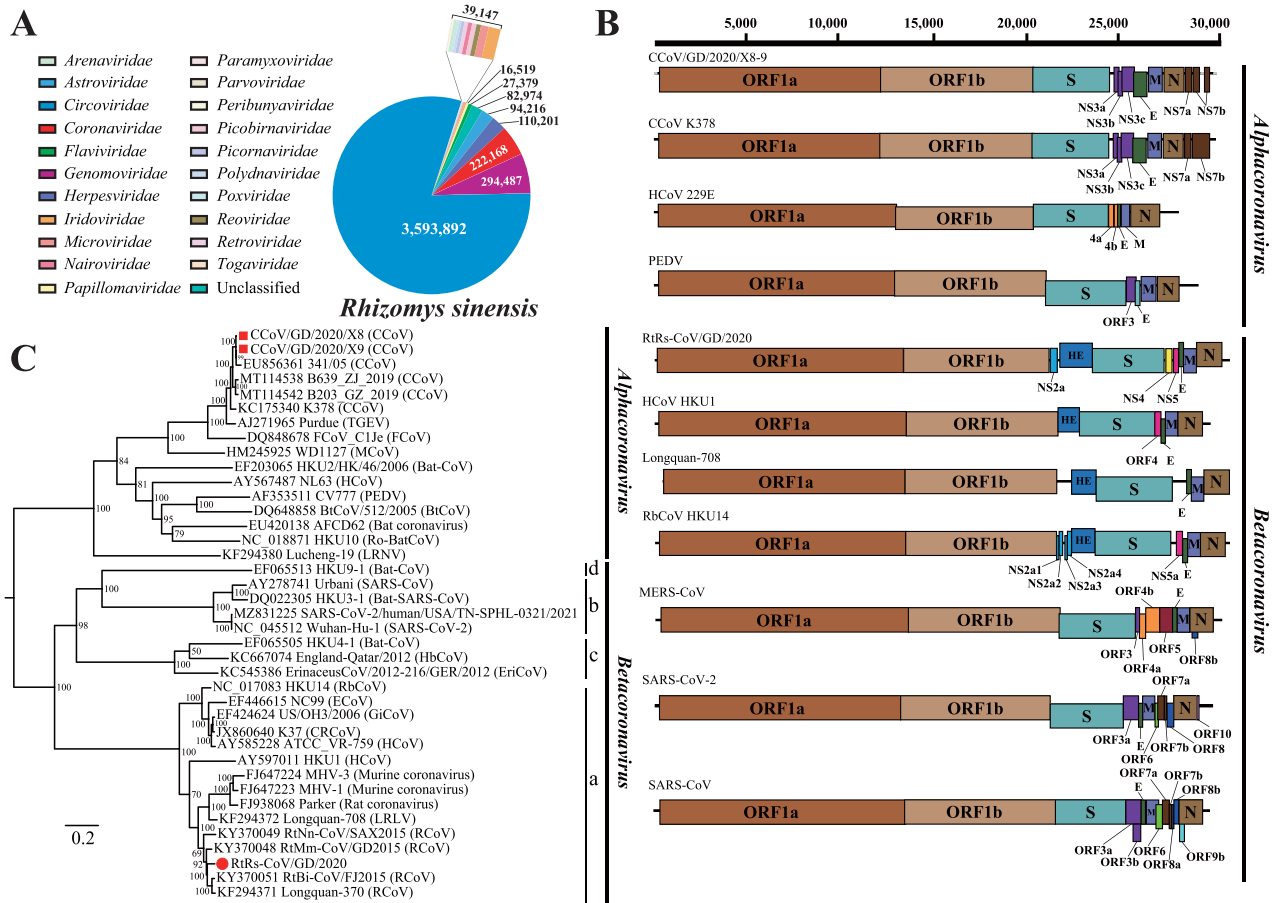


Figure 2. Identification of coronaviruses in *R. sinensis*. (A) Metagenomic analysis of next-generation sequencing of swabs and tissue samples from *R. sinensis*. Only mammalian viruses at the family level are shown. (B) Genomic organization of coronaviruses in *R. sinensis* in comparison with other coronaviruses. The study identified two alphacoronaviruses CCoV/GD/2020/X8 and CCoV/GD/2020/X9, and a betacoronavirus RIRs-CoV/GD/2020. (C) Phylogenetic analysis of the identified coronaviruses by IQ-tree software. The tree was constructed from the complete genome sequences of the included coronaviruses. Bootstrap values expressed as percentages of 1,000 replications are shown at the branch nodes. Circles and squares represent rodent coronavirus and canine coronavirus isolates in *R. sinensis* in this study. The reference virus sequences retrieved from GenBank are included, with names of isolates and their GenBank accession numbers. CCoV, canine coronavirus; HCoV, human coronavirus; PEDV, porcine epidemic diarrhoea virus; RbCoV, rabbit coronavirus; SARS-CoV, severe acute respiratory syndrome coronavirus; MERS-CoV, Middle East respiratory syndrome coronavirus; ORF, open reading frame; HE, haemagglutinin-esterase; S, spike protein; M, membrane protein; N, nucleocapsid protein; NS, nonstructural protein.

genus (Fig. 2C). Both isolated X8 and X9 were further confirmed as canine coronaviruses, as their adenosine diphosphate-ribose-1'-phosphatase (ADRP), nsp5, RdRp, and nsp13-nsp16 had both nucleotide and amino acid identities of more than 90 per cent compared with canine coronaviruses (Supplementary Table S10). The two isolates were classified as canine coronavirus, and designated CCoV/GD/2020/X8 and CCoV/GD/2020/X9, respectively, which was the second time an alphacoronavirus has been found in rodents in addition to Lucheng Rn rat coronavirus. Recombination analysis by RDP showed that CCoV/GD/2020/X8 and CCoV/GD/2020/X9 were probably recombinants of canine coronaviruses 341/05 and B203_GZ_2019 (Supplementary Table S7), which was also confirmed by topology analysis (Supplementary Table S8). Simplot analysis also indicated that CCoV/GD/2020/X8 and CCoV/GD/2020/X9 contained genome fragments derived from 341/05 (regions 1 and 3) and B203_GZ_2019 (region 2) (Supplementary Fig. S5).

The CCoV genome dataset also had sufficient temporal signals for Bayesian dating analysis, with a log BF of 42. Compared with RCoVs, CCoVs have a lower recombination rate of approximately 1.78×10^{-5} substitutions per site per year from the

Bayesian analysis. CCoV 341/05 was the most recent MRCA of CCoV/GD/2020/X8 and X9. The median estimate of the MRCA was 1985 (95 per cent CI: 1970–2000) for 341/05 and 1844 (95 per cent CI: 1830–1978) for B203_GZ_2019 (Supplementary Fig. S4).

3.5 Coronavirus in *P. colchicus*

Our study produced thirty-one coronaviral contigs by metagenomic analysis. These contigs were assembled from 3,720 coronaviral reads, accounting for 0.44 per cent of the total reads of mammalian viral families in pheasants (Fig. 4A). All contigs belong to avian coronavirus within the *Gammacoronavirus* genus, tentatively named pheasant CoV (PhCoV/GD/2020). We attempted to obtain the complete genome sequence by RT-PCR and RACE, but only nine fragments (total approximately 10,642 bp) of the virus, including the complete genes of structural proteins E and M and some nonstructural protein genes (ORF3b, ORF3c, ORF5a, ORF5b, and ORFX/4b), were obtained (Fig. 4B) due to the low viral copy numbers in the collected swab samples.

Sequence comparison showed that most of these fragments had high identities of 84.7–95.5 per cent compared with

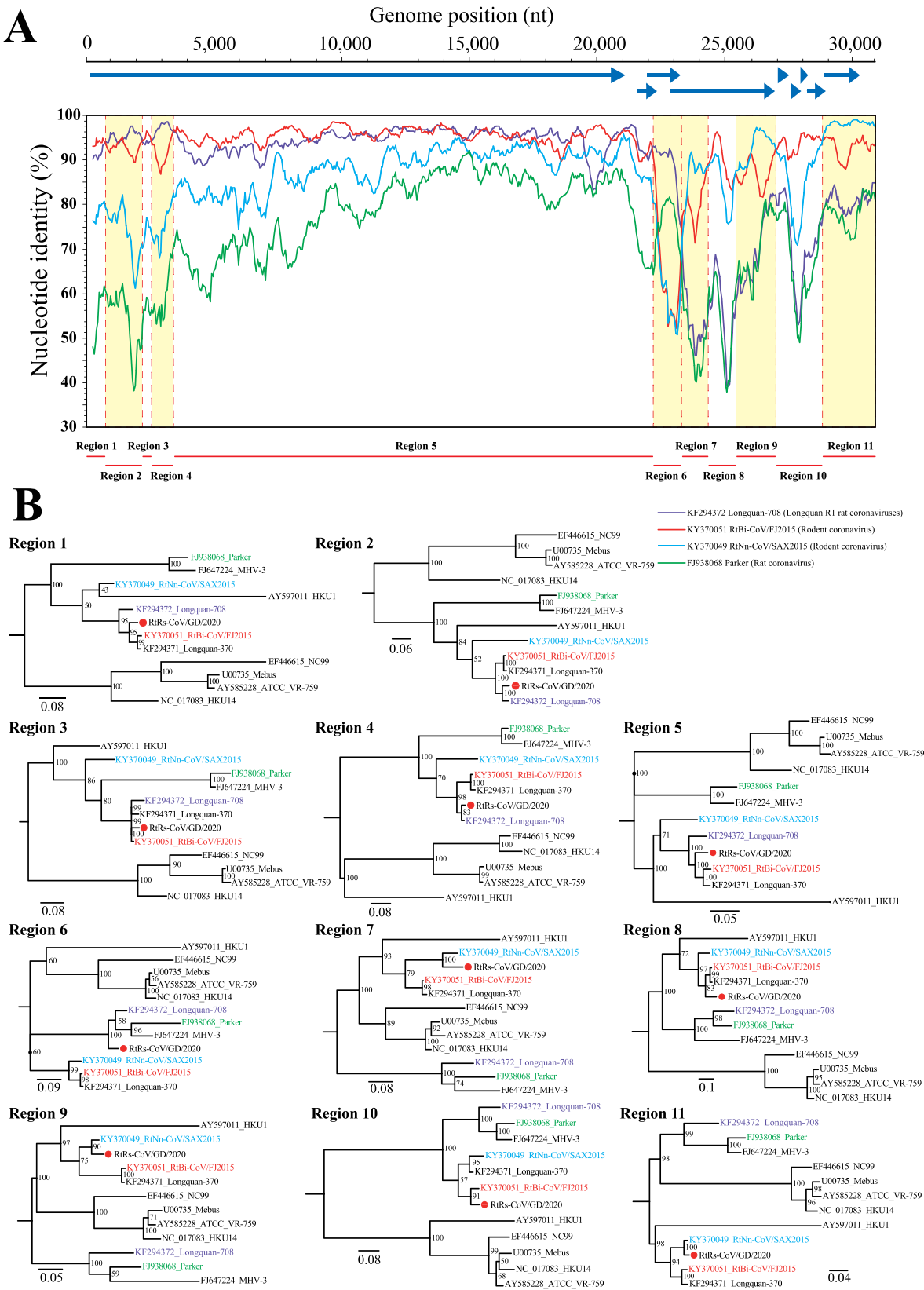


Figure 3. Recombination analysis of the betacoronavirus in *R. sinensis*. (A) Sliding window analysis of changing patterns of sequence similarity between RtRs-CoV/GD/2020 and the reference coronaviruses RtNn-CoV/SAX2015, RiBi-CoV/FJ2015, Longquan-708, and Parker. The potential recombination breakpoints are shown as dashed lines, and regions separated by the breakpoints are alternatively shaded in two different color. These potential breakpoints subdivide the genomes into eleven regions, indicated by the red bars at the bottom of the analysis boxes. The names of the query sequences are shown vertically to the right of the analysis boxes. The similarities to different reference sequences are indicated by different colours. The blue arrows at the top indicate the positions of the ORFs in the alignment. (B) Phylogenetic trees of different genomic regions by IQ-tree software. Branch supports obtained from 1,000 bootstrap replicates are shown.

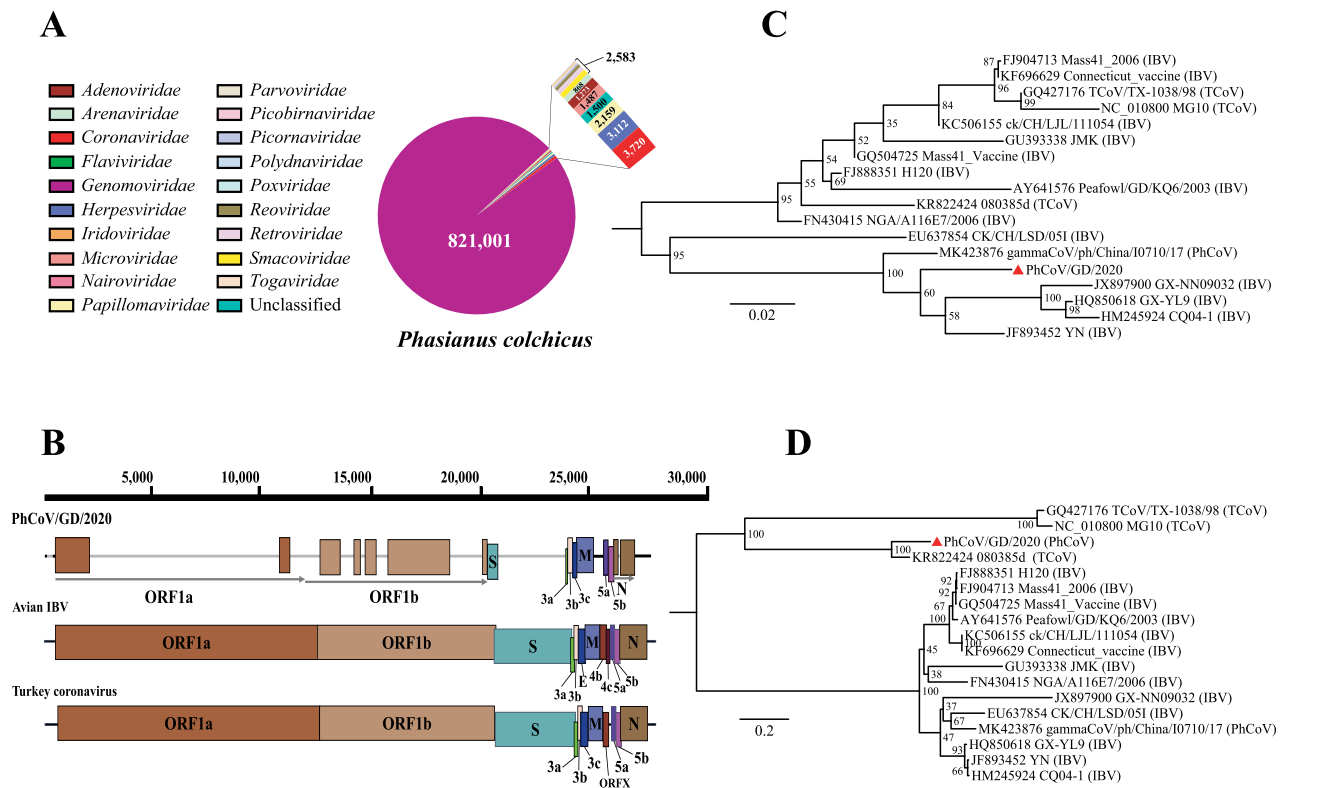


Figure 4. Genome characterization of the coronavirus in *P. colchicus*. (A) Metagenomics analysis of next-generation sequencing of swabs and tissue samples from *P. colchicus*. Only mammalian viruses at the family level are shown. (B) Genomic organization of coronaviruses in *P. colchicus* in comparison with other avian coronaviruses. The study identified a gammacoronavirus designated as PhCoV/GD/2020. The obtained genes are marked with colour rectangles and the other parts that were not obtained in the study are indicated with light grey lines. (C, D) Phylogenetic analysis of identified coronaviruses by IQ-tree software. The trees were constructed based on the nucleotide sequences of the obtained genes (nucleotide) of RdRp (c) and S (d) of PhCoV/GD/2020. Bootstrap values expressed as percentages of 1,000 replications are shown at the branch nodes. The triangle represents PhCoV/GD/2020 in *P. colchicus*. The reference virus sequences retrieved from GenBank are included, with names of isolates and their GenBank accession numbers. IBV, infectious bronchitis virus; TCoV, turkey coronavirus; PhCoV, pheasant coronavirus. ORF, open reading frame; S, spike protein; M, membrane protein; N, nucleocapsid protein.

gammaCoV/ph/China/I0710/17, with the exception of the S and ORF3b genes, which had high identities of 85.3 per cent and 74.9 per cent compared with the European turkey coronavirus 080385d and Chinese infectious bronchitis virus isolate GX-YL9, respectively (Supplementary Table S11). Phylogenetic analysis showed that the partial RNA-dependent RNA polymerase (RdRp) and S genes of the virus were grouped into the clades of IBV and turkey coronavirus (TCoV), respectively (Fig. 4C and D). These results suggested that a recombination event may have occurred in the virus. However, due to the lack of data on the whole genome sequence, we only performed recombination analysis through Simplot software, and further verified it with phylogenetic analysis. The results showed that there was a possibility of recombination events among PhCoV/GD/2020, gammaCoV/ph/China/I0710/17, European turkey coronavirus 080385d and infectious bronchitis virus YN. PhCoV/GD/2020 had genome fragments derived from gammaCoV/ph/China/I0710/17 (genome regions 1, 3, 4, 5, 6, 7, and 9), European turkey coronavirus 080385d (genome region 8), and infectious bronchitis virus YN (genome region 2) (Supplementary Fig. S6). The evaluation of topological differences revealed that the topological structures of genomic regions 2 and 8 were not constrained by other regions (Supplementary Table S8). However, due to the lack of whole genome sequences or complete coding region sequences, we cannot provide more information about these recombination events. It is worth mentioning that, to exclude the possibility of coinfection,

specific PCR was performed to amplify the fragments of region 2 and region 8, and the gene fragments spanning the recombination break site were obtained and verified by sequencing, which suggested that the gamma coronaviruses were not derived from coinfections and that recombination events may have occurred.

3.6 Coronavirus in *P. larvata*

Metagenomic analysis resulted in five coronaviral contigs in *P. larvata*, which were assembled from 80 reads (0.09 per cent of total mammalian virus reads) belonging to the Gammacoronavirus genus (Fig. 5A). Comparison of these contigs with the GenBank database demonstrated that they were closely related to bottlenose dolphin coronavirus HKU22, such as the CF090331 isolate. Unfortunately, we only obtained a partial genome of 2,159 nt, which included the ORF1a (Contig 5 of 304 nt), ORF1b (Contig 2 of 485 and Contig 4 of 282 nt), and ORF10 (Contig 3 of 294 nt) genes, and partial N gene and the 3'-terminal sequences (Contig 1 of 794 nt) targeting the reference virus bottlenose dolphin coronavirus (KF793826) (Fig. 5B). Sequence analysis showed that the isolate had high nt identities of 94.6–98.5 per cent compared with bottlenose dolphin (*Tursiops aduncus*) coronavirus (Supplementary Table S12). Phylogenetic analysis showed that all the contigs of the virus were grouped into the clade of aquatic mammalian CoVs in the γ -CoV genus (Fig. 5C). We named the virus civet coronavirus/GD/2020.

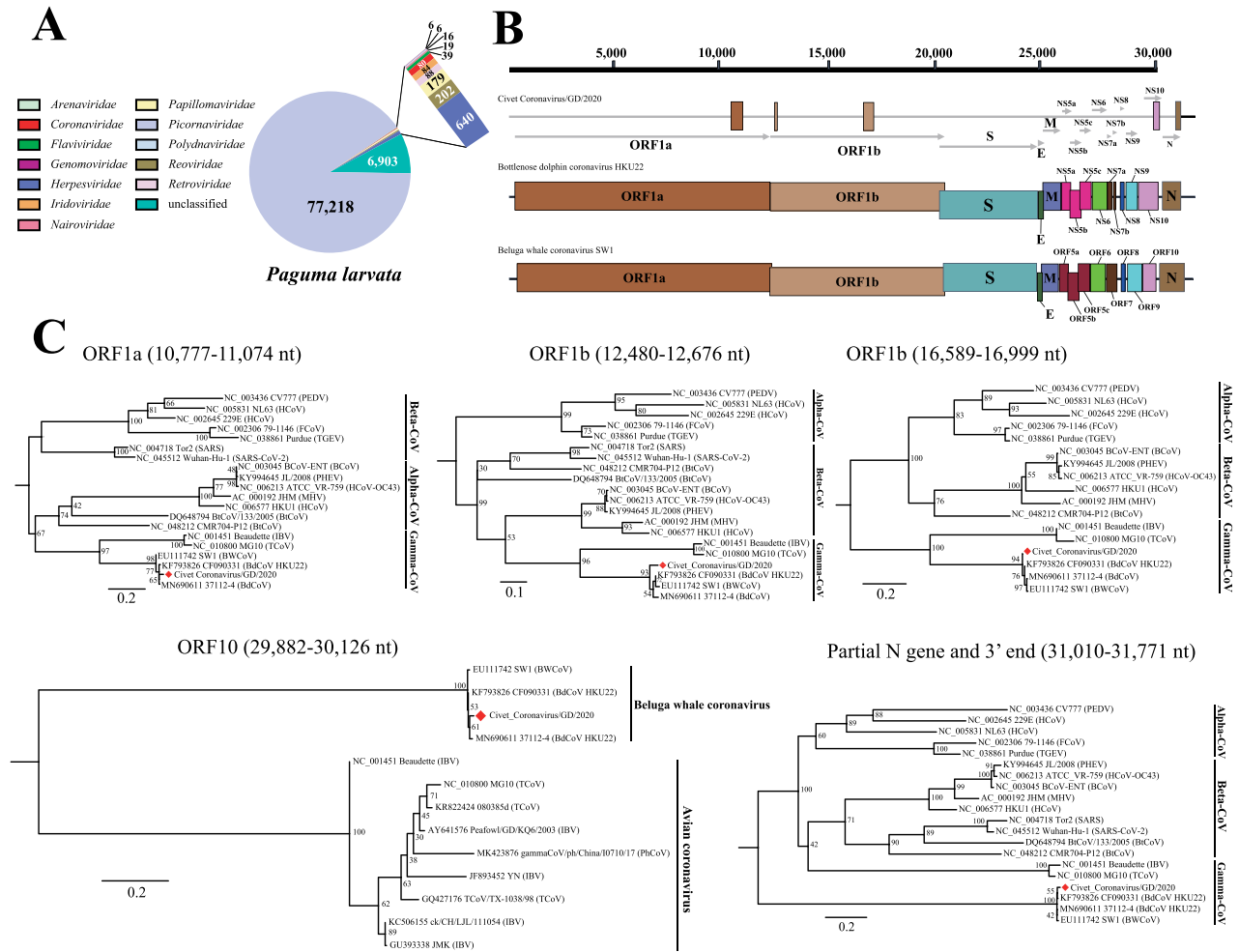


Figure 5. Genome characterization of coronavirus in *P. larvata*.

(A) Metagenomics analysis of next-generation sequencing of swabs and tissue samples from *P. larvata*. Only mammalian viruses at the family level are shown. (B) Genomic organization of coronaviruses in *P. larvata* in comparison with other bottlenose dolphin coronaviruses. The study identified a gammacoronavirus civet designated as CoV/GD/2020. (C) Phylogenetic analysis of identified coronaviruses. The trees were constructed by IQtree software based on the obtained nucleotide sequence of the partial genome of the virus, which included the ORF1a, ORF1b, and ORF10 genes, and partial N gene and 3'-terminal sequences. Bootstrap values expressed as percentages of 1,000 replications are shown at the branch nodes. The red rhomb represents civet CoV/GD/2020 in *P. larvata*. The reference virus sequences retrieved from GenBank are included, with names of isolates and their GenBank accession numbers. BdCoV, bottlenose dolphin coronavirus; BWCoV, beluga whale coronavirus; ORF, open reading frame; S, spike protein; M, membrane protein; N, nucleocapsid protein; NS, nonstructural protein.

4. Discussion

We identified four coronaviruses in farmed wild animals in Guangdong Province, southern China, including one alphacoronavirus, one betacoronavirus, and two gammacoronaviruses. We determined that homologous RNA recombination in CoVs in farmed wild animals occurred with high frequency. For example, the alphacoronaviruses CCoV/GD/2020/X8 and CCoV/GD/2020/X9 in *R. sinensis* were probably recombinants of CCoV 341/05 and B203_GZ_2019, which was the second alphacoronavirus found in rodents. Interestingly, the alphacoronavirus and betacoronavirus in *R. sinensis* were genetically grouped into canine and rat coronaviruses, respectively, and the gammacoronavirus in *P. larvata* was closely related to a bottlenose dolphin coronavirus. These findings suggested that CoVs may generally cross the interspecies barrier in wild animals.

In this study, the collected samples were first screened by RT-PCR, followed by metagenomic analysis. The RT-PCR products were not sequenced, which may generate false-positive results. We found no metagenomic data related to viruses in the family

Coronaviridae in *Quasipaa spinosa*, suggesting false-positive results of the sample.

Coronaviruses are characterized by rapid evolution and changing tissue tropism or host range (Graham and Baric 2010). The replicase RNA-dependent RNA polymerase does not have a good proofreading activity, leading to the incorporation of incorrect nucleotides into the viral genome. Importantly, the coronavirus replication machinery promotes the occurrence of viral recombination events (Muller, Kistler, and Bedford 2022), and many recombinant animal coronaviruses have been found, including CCoV type I (CCoV and feline CoV), CCoV type IIb (CCoV type II and transmissible gastroenteritis virus), feline CoV (FCoV) type II (FCoV and CCoV), and swine CoV (transmissible gastroenteritis virus and porcine epidemic diarrhoea virus). Moreover, all three novel zoonotic human coronaviruses, including SARS-CoV, MERS-CoV, and SARS-CoV-2, have experienced recombination in their evolutionary history (Singh and Yi 2021). Theoretically, as long as there is possible coinfection, all possible recombination may quickly be generated. The spike gene is the most common region

for recombination, resulting in changing tissue tropism or host range (Vakulenko et al. 2021). Our study showed a higher recombination rate in RCoV than in SARS-like coronaviruses (Muller, Kistler, and Bedford 2022), suggesting that RCoV may be a potential interspecies recombination model for betacoronaviruses.

Many interspecies transmission events of coronavirus are associated with emerging animal diseases, including bovine coronavirus, canine coronavirus, feline coronavirus, porcine coronavirus, and transmissible gastroenteritis virus (Vijgen et al. 2006; Lorusso et al. 2008, 2009). Frequent human–wildlife interactions have been considered high-risk factors for the emergence of zoonotic diseases (Daszak, Olival, and Li 2020). It is not a coincidence that SARS and highly pathogenic H5N1 avian influenza have emerged in Guangdong Province where close contact between humans and animals occurs. There is an urgent need to survey high-risk pathogens in wildlife or to close the wildlife markets (Daszak, Olival, and Li 2020). Large percentages of coronaviruses have been detected in bats and rodents at sites where people have close contact and interact with wildlife. The high proportion of coronavirus-positive samples at these human–wildlife interfaces highlights the potential for human exposure to coronaviruses of wildlife origin.

The study had some limitations. Although we attempted to isolate the detected coronaviruses in farmed wild animals, it was unsuccessful. We only obtained partial genome sequences of two coronaviruses from *Ph. colchicus* and *P. larvata* because of the low viral copy numbers in these samples, as farming wild animals has been forbidden in China since 24 February 2020. As it seems highly improbable that transmission from dolphins to civets occurs naturally, the complete genome sequences of the gammacoronavirus identified in civets should be further analysed in future samples and the present data should be used with caution.

In summary, our study identified four viral members of the *Alphacoronavirus*, *Betacoronavirus*, and *Gammacoronavirus* genera in farmed wild animals, and almost all these coronaviruses had undergone homologous recombination and/or crossed the species barrier, likely leading to the emergence of new coronaviruses through human–wildlife interactions (Woo et al. 2009). The emergence of highly pathogenic CoVs in humans highlights the significant threat that CoV spillover poses. The CoVs identified in farmed wild animals in this study may provide important examples for understanding the origin of emerging CoVs in animals and humans. Farming and use of wild animals have been forbidden in China, which may reduce the risk in humans at high risk of contact with wild animals and contribute to preventing the emergence of the next coronavirus pandemic.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

We thank Prof. Zhihong Xu and Weizhen Jiang at the Institute of Animal Health, Guangdong Academy of Agricultural Sciences, Prof. Huijian Hu at the Institute of Animal, Guangdong Academy of Sciences, Prof. Wu Chen at the Guangzhou Zoo, and Director Xianghe Gu at the Guangdong Wildlife Protection Center of China for their helps in sample collection.

Funding

This study was supported by Guangdong Provincial Forestry Department's Provincial Financial Special Fund for Ecological

Forestry Construction-Wildlife Conservation, Special fund for scientific innovation strategy-construction of high level Academy of Agriculture Science-Distinguished Scholar (R2020PY-JC001), the Planning Funds for Science and Technology of Guangdong Province (2020A1515011112 and 2021B1212050021), the Guangdong Province Key Field Research and Development Program (2020B111118001 and 2020B1111320001), the Guangzhou Science Technology and Innovation Commission (201804020080), the Scientific and Technological Research Projects of Foshan (2020001000151) and Guangzhou (202103000008), China.

Conflict of interest: No potential conflicts of interest were disclosed.

References

- Allen, T. et al. (2017) 'Global Hotspots and Correlates of Emerging Zoonotic Diseases', *Nature Communications*, 8: 1124.
- Alluwaimi, A. M. et al. (2020) 'The Coronaviruses of Animals and Birds: Their Zoonosis, Vaccines, and Models for SARS-CoV and SARS-CoV2', *Frontiers in Veterinary Science*, 7: 582287.
- Camacho, C. et al. (2009) 'BLAST+: Architecture and Applications', *BMC Bioinformatics*, 10: 421.
- Chen, B. et al. (2020) 'Overview of Lethal Human Coronaviruses', *Signal Transduction and Targeted Therapy*, 5: 89.
- Chen, Y. et al. (2018) 'SOAPnuke: A MapReduce Acceleration-supported Software for Integrated Quality Control and Preprocessing of High-throughput Sequencing Data', *Gigascience*, 7: 1–6.
- Chernomor, O., von Haeseler, A., and Minh, B. Q. (2016) 'Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices', *Systematic Biology*, 65: 997–1008.
- Corman, V. M. et al. (2018) 'Hosts and Sources of Endemic Human Coronaviruses', *Advances in Virus Research*, 100: 163–88.
- Cui, J., Li, F., and Shi, Z. L. (2019) 'Origin and Evolution of Pathogenic Coronaviruses', *Nature Reviews Microbiology*, 17: 181–92.
- Daszak, P., Olival, K. J., and Li, H. (2020) 'A Strategy to Prevent Future Epidemics Similar to the 2019-nCoV Outbreak', *Biosafety and Health*, 2: 6–8.
- Duchene, S. et al. (2020) 'Bayesian Evaluation of Temporal Signal in Measurably Evolving Populations', *Molecular Biology and Evolution*, 37: 3363–79.
- Fu, L. et al. (2012) 'CD-HIT: Accelerated for Clustering the Next-generation Sequencing Data', *Bioinformatics*, 28: 3150–2.
- Gong, S. et al. (2015) 'Nairobi Sheep Disease Virus RNA in Ixodid Ticks, China, 2013', *Emerging Infectious Diseases*, 21: 718–20.
- Gordon, D. E. et al. (2020) 'A SARS-CoV-2 Protein Interaction Map Reveals Targets for Drug Repurposing', *Nature*, 583: 459–68.
- Graham, R. L., and Baric, R. S. (2010) 'Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-species Transmission', *Journal of Virology*, 84: 3134–46.
- Kalyaanamoorthy, S. et al. (2017) 'ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates', *Nature Methods*, 14: 587–9.
- Katoh, K. et al. (2002) 'MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform', *Nucleic Acids Research*, 30: 3059–66.
- Ksiazek, T. G. et al. (2003) 'A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome', *New England Journal of Medicine*, 348: 1953–66.
- Lagkouvardos, I. et al. (2017) 'Rhea: A Transparent and Modular R Pipeline for Microbial Profiling Based on 16S rRNA Gene Amplicons', *PeerJ*, 5: e2836.
- Lai, M. M., and Cavanagh, D. (1997) 'The Molecular Biology of Coronaviruses', *Advances in Virus Research*, 48: 1–100.

- Lanfear, R. et al. (2017) 'PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses', *Molecular Biology and Evolution*, 34: 772–3.
- Latinne, A. et al. (2020) 'Origin and Cross-species Transmission of Bat Coronaviruses in China', *Nature Communications*, 11: 4235.
- Li, D. et al. (2016) 'MEGAHIT V1.0: A Fast and Scalable Metagenome Assembler Driven by Advanced Methodologies and Community Practices', *Methods*, 102: 3–11.
- Li, H., and Durbin, R. (2009) 'Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform', *Bioinformatics*, 25: 1754–60.
- Li, L. et al. (2021) 'Epidemiological Study of Betacoronaviruses in Captive Malayan Pangolins', *Frontiers in Microbiology*, 12: 657439.
- Liu, P., Chen, W., and Chen, J. P. (2019) 'Viral Metagenomics Revealed Sendai Virus and Coronavirus Infection of Malayan Pangolins (*Manis javanica*)', *Viruses*, 11: 979.
- Lole, K. S. et al. (1999) 'Full-length Human Immunodeficiency Virus Type 1 Genomes from Subtype C-infected Seroconverters in India, with Evidence of Intersubtype Recombination', *Journal of Virology*, 73: 152–60.
- Lorusso, A. et al. (2008) 'Gain, Preservation, and Loss of a Group 1a Coronavirus Accessory Glycoprotein', *Journal of Virology*, 82: 10312–7.
- et al. (2009) 'Molecular Characterization of a Canine Respiratory Coronavirus Strain Detected in Italy', *Virus Research*, 141: 96–100.
- Lu, G., Wang, Q., and Gao, G. F. (2015) 'Bat-to-human: Spike Features Determining 'host jump' of Coronaviruses SARS-CoV, MERS-CoV, and Beyond', *Trends in Microbiology*, 23: 468–78.
- Martin, D. P. et al. (2021) 'RDP5: A Computer Program for Analyzing Recombination In, and Removing Signals of Recombination From, Nucleotide Sequence Datasets', *Virus Evolution*, 7: veaa087.
- McMurdie, P. J., and Holmes, S. (2013) 'Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data', *PLoS One*, 8: e61217.
- Minh, B. Q. et al. (2020) 'IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era', *Molecular Biology and Evolution*, 37: 1530–4.
- Muller, N. F., Kistler, K. E., and Bedford, T. (2022) 'Recombination Patterns in Coronaviruses', *bioRxiv*, 2021: 28.441806.
- Onyuk, S. O. et al. (2019) 'Molecular Detection and Genetic Characterization of Novel RNA Viruses in Wild and Synanthropic Rodents and Shrews in Kenya', *Frontiers in Microbiology*, 10: 2696.
- Singh, D., and Yi, S. V. (2021) 'On the origin and evolution of SARS-CoV-2', *Experimental & Molecular Medicine*, 53: 537–47.
- Song, H. D. et al. (2005) 'Cross-host Evolution of Severe Acute Respiratory Syndrome Coronavirus in Palm Civet and Human', *Proceedings of the National Academy of Sciences of the United States of America*, 102: 2430–5.
- Su, S. et al. (2016) 'Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses', *Trends in Microbiology*, 24: 490–502.
- Vakulenko, Y. et al. (2021) 'Modular Evolution of Coronavirus Genomes', *Viruses*, 13: 1270.
- Vijgen, L. et al. (2005) 'Complete Genomic Sequence of Human Coronavirus OC43: Molecular Clock Analysis Suggests a Relatively Recent Zoonotic Coronavirus Transmission Event', *Journal of Virology*, 79: 1595–604.
- et al. (2006) 'Evolutionary History of the Closely Related Group 2 Coronaviruses: Porcine Hemagglutinating Encephalomyelitis Virus, Bovine Coronavirus, and Human Coronavirus OC43', *Journal of Virology*, 80: 7270–4.
- Wacharapluesadee, S. et al. (2021) 'Evidence for SARS-CoV-2 Related Coronaviruses Circulating in Bats and Pangolins in Southeast Asia', *Nature Communications*, 12: 972.
- Wang, W. et al. (2015) 'Discovery, Diversity and Evolution of Novel Coronaviruses Sampled from Rodents in China', *Virology*, 474: 19–27.
- Woo, P. C. et al. (2009) 'Coronavirus Diversity, Phylogeny and Interspecies Jumping', *Experimental Biology and Medicine*, 234: 1117–27.
- et al. (2012) 'Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavirus', *Journal of Virology*, 86: 3995–4008.
- Zhang, T., Wu, Q., and Zhang, Z. (2020) 'Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak', *Current Biology*, 30: 1578.
- Zhou, P. et al. (2018) 'Fatal Swine Acute Diarrhoea Syndrome Caused by an HKU2-related Coronavirus of Bat Origin', *Nature*, 556: 255–8.
- Zhu, W. et al. (2021) 'Beta- and Novel Delta-coronaviruses are Identified from Wild Animals in the Qinghai-Tibetan Plateau, China', *Virologica Sinica*, 36: 402–11.
- Zhuang, Q. et al. (2020) 'Surveillance and Taxonomic Analysis of the Coronavirus Dominant in Pigeons in China', *Transboundary and Emerging Diseases*. [10.1111/tbed.13541](https://doi.org/10.1111/tbed.13541). Advance online publication.