Virologica Sinica 38 (2023) 868–876

Contents lists available at ScienceDirect

Virologica Sinica

journal homepage: www.keaipublishing.com/en/journals/virologica-sinica
www.virosin.org

Research Article

Pangolin HKU4-related coronaviruses found in greater bamboo bats from southern China

Min Guo [a,1], Kai Zhao [b,1], Xingwen Peng [a], Xiangyang He [a], Jin Deng [a], Bo Wang [c], Xinglou Yang [b,d,*], Libiao Zhang [a,*]

[a] Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Guangdong Public Laboratory of Wild Animal Conservation and Utilization, Institute of Zoology, Guangdong Academy of Sciences, Guangzhou, 510260, China
[b] Yunnan Key Laboratory of Biodiversity Information, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650023, China
[c] Department of Biomedical Sciences and Pathobiology, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24060, USA
[d] Hubei Jiangxia Lab, Wuhan, 430071, China

ARTICLE INFO

Keywords:
Coronavirus
Pangolin HKU4r-CoV
Bat HKU4r-CoV
Tylonycteris robustula
Spillover risk

ABSTRACT

Coronavirus (CoV) spillover originating from game animals, particularly pangolins, is currently a significant concern. Meanwhile, vigilance is urgently needed for coronaviruses carried by bats, which are known as natural reservoirs of many coronaviruses. In this study, we collected 729 anal swabs of 20 different bat species from nine locations in Yunnan and Guangdong provinces, southern China, in 2016 and 2017, and described the molecular characteristics and genetic diversity of alphacoronaviruses (αCoVs) and betacoronaviruses (βCoVs) found in these bats. Using RT-PCR, we identified 58 (8.0%) bat CoVs in nine bat species from six locations. Furthermore, using the Illumina platform, we obtained two representative full-length genomes of the bat CoVs, namely TyRo-CoV-162275 and TyRo-CoV-162269. Sequence analysis showed that TyRo-CoV-162275 shared the highest identity with Malayan pangolin (*Manis javanica*) HKU4-related coronaviruses (MjHKU4r-CoVs) from Guangxi Province, whereas TyRo-CoV-162269 was closely related to HKU33-CoV discovered in a greater bamboo bat (*Tylonycteris robustula*) from Guizhou Province. Notably, TyRo-CoV-162275 has a putative furin protease cleavage site in its S protein and is likely to utilize human dipeptidyl peptidase-4 (hDPP4) as a cell-entry receptor, similar to MERS-CoV. To the best of our knowledge, this is the first report of a bat HKU4r-CoV strain containing a furin protease cleavage site. These findings expand our understanding of coronavirus geographic and host distributions.

1. Introduction

Zoonotic viruses carried by wild animals threaten human health, and growing evidences suggest that game animals are intermediate hosts for zoonotic viral spillover. However, the transmission route of zoonotic viruses from hosts to intermediate hosts and then to humans is largely unknown (He et al., 2022). In the past few decades, bats have been recognised as natural hosts of various zoonotic viruses, such as coronaviruses (CoVs), and bat-borne viruses directly or indirectly spilling into human society have caused multiple devastating outbreaks (Drexler et al., 2010; Li et al., 2005; Wu et al., 2016). For example, in the outbreak of severe acute respiratory syndrome (SARS) in 2003, some intermediate host game animals, such as the civet (*Paguma larvata*), were once thought to be the culprits. However, with the discovery of a large number of SARS-CoV-1 related CoVs (SARSr-CoV) and SARS natural host libraries in bats, it was confirmed that the spillover of SARSr-CoV from bats was the evolutionary origin of these outbreaks, and civets were the culprit intermediate host (Hu et al., 2017). Zoonotic spillovers events include Middle East respiratory syndrome (MERS) outbreaks in the Middle East in 2012, swine acute diarrhoea syndrome (SADS) outbreaks in Guangdong in 2018, and the coronavirus disease (COVID-19) pandemic in late 2019. Bat-borne Merbecoviruses (HKU4r-CoV and HKU5r-CoV), Rhinacoviruses (HKU2r-CoV), and Sarbecoviruses (RaTG-13) are associated with these outbreaks (Zaki et al., 2012; Zhou et al., 2018, 2020). Thus, long-term epidemiological surveys of natural reservoirs that have intimate interactions with humans and livestock are the key to preventing the next pandemic.

Bat coronaviruses, which are closely related to the coronaviruses that cause these outbreaks, have been reported in bats from southern China,

* Corresponding authors.
E-mail addresses: yangxinglou@mail.kiz.ac.cn (X. Yang), zhanglb@giz.gd.cn (L. Zhang).
[1] Min Guo and Kai Zhao contributed equally to this work.

https://doi.org/10.1016/j.virs.2023.11.003
Received 10 April 2023; Accepted 10 November 2023
Available online 13 November 2023

bringing significant attention to these hot spots (Cui et al., 2019). The suitable geographical environment and climatic conditions in southern China have created an abundance of bat species, providing the greatest possibility for genetic diversity and evolution of coronaviruses in bat habitats in this region (Fan et al., 2019). Among the provinces in southern China, the top four in terms of bat species richness are Yunnan, Guangdong, Guangxi, and Guizhou, with 75, 51, 51, and 51 bat species, respectively (Yang, 2016). The highest diversity of bat-borne coronaviruses in China was found in Yunnan and Guangdong (Fan et al., 2019; Zhu et al., 2023). Recently, SARS-CoV-2-related CoVs identified in several pangolin species in Guangdong and Guangxi provinces have shown the ability to exploit human angiotensin-converting enzyme 2 (hACE2) as a receptor for viral infection and may share common ancestry with RaTG-13, BANAL-236 and other SARS-CoV-2 related CoVs found in bats from China and Laos (Lam et al., 2020; Shi et al., 2022; Xiao et al., 2020). Surprisingly, the Merbecovirus MjHKU4r-CoV found in pangolins from Guangxi has also been shown to use the human dipeptidyl peptidase-4 (hDPP4) receptor to achieve infection. However, the discovered MjHKU4r-CoV has only 67.9% and 86.8% genomic identity with its relatives MERS-CoV and HKU4-CoV, respectively (Chen et al., 2023). In this study, we investigated coronavirus diversity in bats from Yunnan and Guangdong provinces to enrich our knowledge of the current distribution and evolution of coronaviruses in bats.

2. Materials and methods

2.1. Sample collection

A total of 729 specimens from 20 different bat species belonging to 11 genera and six families were collected from nine different locations in the Yunnan and Guangdong provinces of southern China. Each bat was kept in a clean cloth bag after being captured. Separate fecal samples were collected from individual bats using anal swabs. After sampling, the anal swab was immediately placed in a 2 mL tube prefilled with 1 mL viral transport media (VTM), snap-frozen in liquid nitrogen and then stored under −80 °C until used.

2.2. Bat species identification and coronavirus detection

Each bat species was identified by analysing the external morphological characteristics and mitochondrial *Cyt-b* gene sequences (using patagium samples of bats). For viral detection, tubes containing anal swab samples were thawed (on ice) and homogenised (2.5 m/s, 1 min). All the tubes were then divided into quintuplets according to the bat species, merged and kept on ice. Total RNA was extracted from the samples using a high-purity viral RNA kit (Roche, USA) in a biosafety cabinet. Broadly reactive consensus hemi-nested reverse transcription PCR (RT-PCR) was used for the detection of the RNA-dependent RNA polymerase (*RdRp*) gene conserved fragment (440 bp) of coronavirus using the Superscript III one-step RT-PCR system with Platinum high-fidelity Taq polymerase (Invitrogen, USA). The primers were: CoV-FWD3 (5′-GGTTGGGAYTAYVVHAARTGTGA-3′), CoV-RVS3 (5′-CCATCATCASWYRAATCATCATA-3′), and CoV-FWD4 (5′-GAYTAYCCHAARTGTGAYAGAGC-3′). Positive amplicons were sequenced by Sangon Biotech Co. (Guangzhou, China). Coronavirus *RdRp* sequences were deposited in the GenBank database (accession numbers: MW600658–MW600715). If a merged sample was positive for coronavirus, the individual samples were assayed separately by repeating the above steps to verify that the individual samples were positive for coronavirus.

2.3. Amplification of coronavirus genome

Two individual coronavirus-positive samples were subjected to virome sequencing using an Illumina NovaSeq 6000 sequencer (OriginGene Co., Shanghai, China). Raw reads were subjected to quality control using SOAPnuke (1.5.6) to filter out low-quality and adaptor-contaminated reads. Megahit (http://i.cs.hku.hk/∼alse/hkubrg/projects/idba_ud/) was used for the assembly of high-quality reads. Obtained contigs were then aligned to a cohort of published HKU4r-CoV genomes (NCBI ID: KJ473822.1, MK720944.1, EF065505.1, MH002337.1, and NC_009019.1). After gap filling and completeness of 5′ and 3′ rapid amplification of cDNA ends (RACE), two complete coronavirus genomes were made available in the GenBank database (accession numbers ON745165 and ON745166). MetaGeneMark (http://exon.gatech.edu/meta_gmhmmp.cgi) was used to search for coding sequences (CDSs) in the complete genome.

2.4. Phylogenetic analysis

The *RdRp* gene sequences of bat coronaviruses were assembled using the Seqman program in the DNAstar software, version 8.1.3. Pairwise alignments of the assembled sequences were performed using BioEdit software version 7.0.5. Maximum likelihood phylogenetic trees of the partial *RdRp* (387 bp) were constructed based on alignments using MEGA7 version 7.0.21, with 1000 bootstrap replicates. The optimal model (GTR + G + I) for tree construction was determined using J model test version 2.1.7. Simplot version 3.5.1 was used for recombination and similarity analyses. Neighbour-joining phylogenetic trees of the complete genome, spike protein nucleotide acids, and amino acids were also constructed based on the alignments using MEGA7 version 7.0.21, with 1000 bootstrap replicates. All trees were edited using FigTree version 1.4.4.

2.5. Genomic mining

The identities of the genomic sequences were analysed using Simplot version 3.5.1. Prediction of furin cleavage sites in the *S* proteins of bat coronavirus was performed using the ProP-1.0 Server (https://services.healthtech.dtu.dk/service.php?ProP-1.0). The structure model of spike protein was predicted using the SWISS-MODEL online tool (https://swissmodel.expasy.org) using the structure of the related PDB protein (PDB: 5X5C) as a template.

3. Results

3.1. Sampling

Of the 729 bats sampled, 409 were from Yunnan Province and 320 were from Guangdong Province. These bats included two species from Pteropodidae (*Eonycteris spelaea*, and *Rousettus leschenaulti*), one from Megadermatidae (*Megaderma lyra*), four from Hipposideridae (*Aselliscus stoliczkanus*, *Hipposideros armiger*, *H. larvatus*, and *H. pomona*), five from Rhinolophidae (*Rhinolophus affinis*, *R. pearsonii*, *R. pusillus*, *R. siamensis*, and *R. sinicus*), one from Molossidae (*Tadarida plicata*), and seven from Vespertilionidae (*Miniopterus pusillus*, *Myotis chinensis*, *M. laniger*, *M. pilosus*, *Ia io*, *Tylonycteris pachypus*, *T. robustula*) (Table 1).

3.2. Coronavirus detection

Through RT-PCR, CoV RNA was detected in 58 (8.0%) specimens from nine different bat species: *T. robustula* (TyRo, 28.95%), *I. io* (IaIo, 16.67%), *M. chinensis* (MyCh, 13.33%), *M. pilosus* (MyPi 2.17%), *M. laniger* (MyLa, 20.00%), *R. sinicus* (RhSi, 9.57%), *R. pearsoni* (RhPe, 2.43%), *R. affinis* (RhAf, 44.83%), *A. stoliczkanus* (AsSt, 2.70%) (Fig. 1). The detection rates of bat CoVs ranged from 0.0% to 31.1% per site in nine of the sampling locations, and the positive rates at each sampling site, ordered from high to low, were as follows: Jingning (JN, 31.1%), Maoming (MM, 21.6%), Menghai (MH, 8.9%), Conghua (CH, 8.2%), Lufeng (LF, 6%), and Yingde (YD, 3%). CoV RNA was not detected in bat samples from Jinghong (JH), Mengla (ML), and Huidong (HD) (Fig. 1). The positive detection rates in individual bat colonies were: 66.7% (LF,

Table 1
Bat species tested for coronavirus in this study.

Bat species	Pos/total	Sampling site (month/year,[a] positive)	CoV clade(s) [cluster]/(no. positive)
Family Pteropodidae			
Eonycteris spelaea	0/90	ML (8/2016, 8/2017);	
Rousettus leschenaulti	0/38	LF (8/2017); ML (8/2016, 8/2017)	
Famliy Megadermatidae			
Megaderma lyra	0/2	JH (8/2016)	
Family Hipposideridae			
Aselliscus stoliczkanus	1/37	JH (8/2016); JN (8/2017[a])	Decacovirus [HKU10]/(1)
Hipposideros armiger	0/34	JN (8/2016); LF (8/2016, 8/2017); MM (8/2017); CH (9/2017); HD (9/2017); YD (9/2017)	
Hipposideros larvatus	0/13	YD (9/2017)	
Hipposideros pomona	0/1	CH (9/2017)	
Family Rhinolophidae			
Rhinolophus affinis	26/58	JN (8/2016, 8/2017[a]); JH (8/2016); MM (8/2017[a]); CH (9/2017[a])	Rhinacovirus [HKU2/SADSr]/(23); Sarbecovirus [SARSr]/(1); Decacovirus [HKU10]/(2)
Rhinolophus pearsoni	1/41	MM (8/2017[a]); YD (9/2017)	Rhinacovirus [HKU2/SADSr]/(1)
Rhinolophus pusillus	0/5	JN (8/2016); JH (8/2016); MM (8/2017); YD (9/2017)	
Rhinolophus siamensis	0/2	YD (9/2017)	
Rhinolophus sinicus	11/116	JN (8/2016[a], 8/2017); JH (8/2016); LF (8/2017); CH (9/2017); YD (9/2017); HD (9/2017)	Sarbecovirus [SARSr]/(11)
Family Molossidae			
Tadarida plicata	0/8	LF (8/2016)	
Familiy Vespertilionidae			
Miniopterus pusillus	0/27	HD (8/2017)	
Myotis chinensis	4/30	JH (8/2016); HD (9/2017); YD (9/2017[a])	Minunacovirus [1B]/(1); HKU6r-CoV [HKU6]/(3)
Myotis laniger	1/5	JN (8/2016, 2017[a]); YD (9/2017)	HKU6r-CoV [HKU6]/(1)
Myotis pilosus	2/92	YD (9/2017); LF (8/2016[a], 8/2017); HD (9/2017)	HKU6r-CoV [HKU6]/(2)
Ia io	1/6	JN (8/2017); LF (8/2016[a])	HKU6r-CoV [HKU6]/(1)
Tylonycteris pachypus	0/86	MH (8/2016)	
Tylonycteris robustula	11/38	MH (8/2016[a])	Merbecovirus [MjHKU4r-CoV]/(8); Nyctacovirus [HKU33]/(3)

Notes: Conghua (CH), Huidong (HD), Maoming (MM) and Yingde (YD) sampling sites are in Guangdong Province; Jingning (JN), Jinghong (JH), Lufeng (LF), Menghai (MH), and Mengla (ML) sampling sites are in Yunnan Province.
[a] Denotes the sampling time when coronavirus was detected.

MyPi), 57.7% (JN, RhAf), 50% (JN, MyLa), 50% (MM, RhAf), 35.5% (JN, RhSi), 33.3% (LF, IoIa), 28.9% (MH, TyRo), 28.6% (CH, RhAf), 18.2% (YD, MyCh), 4.8% (JN, AsSt), and 4.8% (MM, RhPe) (Fig. 1).

3.3. Geo-evolutionary analysis based on RdRp fragments

The coronavirus *RdRp* region is an RNA polymerase fragment widely used to identify coronaviruses. Phylogenetic analysis of the nucleotide sequences of a conserved partial *RdRp* gene segment (387 bp) showed that the 38 samples were members of the *αCoV* genus, and 20 samples belonged to the *βCoV* genus. All sequences could be divided into seven clades, five of which belonged to the *αCoV* genus (HKU6r-CoV, Decacovirus, Nyctacovirus, Minuacovirus, and Rhinacovirus clade), and two clades belonged to the *βCoV* genus (Merbecovirus and Sarbecovirus) (Fig. 2). The percentage of nucleotide identity within each clade was calculated and is shown in Fig. 2.

CoVs from the Rhinacovirus clade (green, Fig. 2) were divided into four branches, where species and geographical restriction characteristics were distinguishable between the branches. CoVs in branch one were from *R. affinis* in CH, CoVs in branches two and three were from *R. affinis* in MM, and CoVs in branch four were from *R. pearsonii* in MM. We detected one CoV in *M. chinensis* from YD in the Minunacovirus clade (orange, Fig. 2). The detected CoVs in the Nyctacovirus clade were from MH's *T. robustula* and were close to HKU33-CoV, which was found in *T. robustula* from Guizhou Province (purple, Fig. 2). CoV sequence 172044 was detected in *A. stoliczkanus* from JN and was related to HKU10, while other CoVs detected in the Decacovirus clade were related to HuB2013, which was found in Hubei's *R. ferrumequinum* (yellow, Fig. 2). The CoVs detected in the HKU6r-CoV clade appeared to be of poor species and geographical restrictions, whereas the CoVs detected in *M. pilosus*, *I. io*, *M. chinensis*, and *M. laniger* were from LF, YD, or JN (light blue, Fig. 2). The CoVs detected in *R. sinicus* from JN and *R. affinis* from MM in the Sarbecovirus clade were related to SARSr-CoV in *Rhinolophus* bats from YN (red, Fig. 2). Interestingly, the eight CoVs found in MH's *T. robustula* were related to MjHKU4r-CoV in pangolins and belonged to the Merbecovirus clade with MERS-CoV, HKU4-CoV, and HKU5-CoV (dark blue, Fig. 2).

3.4. Phylogenetic analysis of complete genomes of bat coronaviruses TyRo-CoV-162275 and TyRo-CoV-162269

To further confirm the relationship between the two clades of viruses detected in this study of Merbecovirus and Nyctacovirus, which were less reported in known viruses, we selected the HKU4r-CoV related sample (TyRo-CoV-162275) from the Merbecovirus clade and the HKU33r-CoV related sample (TyRo-CoV-162269) from the Nyctacovirus clade for whole genome sequencing.

TyRo-CoV-162275 showed a typical Merbecovirus genome structure and contained all ten viral open reading frames (ORFs) present in the subgenus Merbecovirus, and similarity plot analysis showed that TyRo-CoV-162275 exhibited the highest genome nucleotide identity with bat MjHKU4r-CoVs (Fig. 3A and C). Phylogenetic analysis of the *S*, *ORF1ab*, *E*, *M*, and *N* genes and the whole genome sequences further confirmed that the TyRo-CoV-162275 detected in MH's *T. robustula* was closely related (nucleotide similarity between 93.2% and 93.9%) to the P251T and MjHKU4r-CoVs found in pangolins in Guangxi, while the TyRo-CoV-162269 detected in the same *T. robustula* population was closely related (nucleotide similarity up to 94.1%) to the HKU33-CoV detected in Guizhou *T. robustula* (Fig. 3B and Supplementary Fig. S2). Recombination analysis revealed no recombination events in the TyRo-CoV-162275 genomic sequences. Furthermore, similarity plot analysis showed that TyRo-CoV-162269 exhibited the highest genomic nucleotide identity to

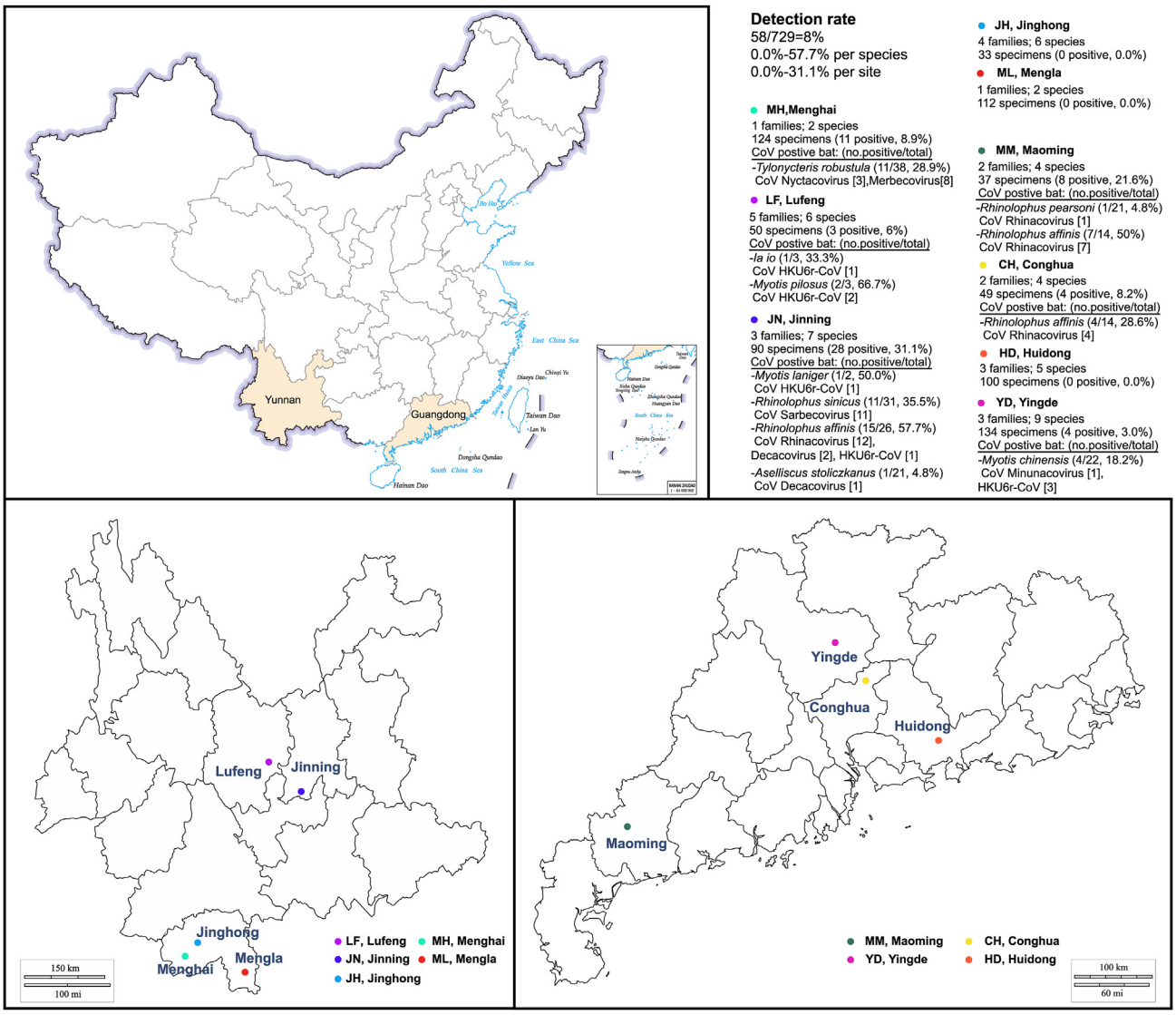

Fig. 1. Maps showing bat surveillance locations in Yunnan and Guangdong. The solid dots represent the nine locations where bat anal swabs were collected.

bat HKU33, and recombination analysis showed no recombination events in the TyRo-CoV-162269-genome sequence (Supplementary Fig. S1). Of the five conserved replicase domains of ORF1ab that were used for CoV species classification, TyRo-CoV-162275 shared 99% amino acid identity with MjHKU4r-CoVs from pangolins and 97% identity with *Tylonycteris* bat coronavirus HKU4-CoV, with nucleotide identities of 96% and 90%, respectively, indicating that TyRo-CoV-162275 is a bat MjHKU4r-CoV-like virus belonging to the species *Tylonycteris* bat coronavirus HKU4 and the subgenus Merbecovirus (Fig. 3D).

3.5. Receptor usage and S protein cleavage site analysis of TyRo-CoV-162275

Receptor binding and *S* protein cleavage are two barriers to Merbecovirus access for infecting human cells. MERS-CoV uses human dipeptidyl peptidase-4 (hDPP4) as a cell entry receptor, and cell membrane fusion is mediated by the proteolytic cleavage of its S protein by host proteases, a process that can be enhanced by furin proteases. Given that MERS-CoV, bat HKU4-CoV, and pangolin MjHKU4r-CoV use hDPP4 as a cell-entry receptor, we investigated whether bat TyRo-CoV-162275 also uses hDPP4 for entry.

The complete spike nucleotide and amino acid phylogenetic trees of TyRo-CoV-162275, HKU4-CoV, HKU5-CoV, MERS-CoV, and MjHKU4r-CoVs showed that TyRo-CoV-162275 was closer to the pangolin HKU4r-CoVs than to the bat HKU4-CoV (Fig. 4A and B). There are 16 key residues in the receptor-binding motif (RBM) of the MERS-CoV *S* protein that determine its binding to hDPP4. Bat TyRo-CoV-162275 shared 7 of these 16 key residues, suggesting that it may also use hDPP4 as a cell-entry receptor. In comparison, pangolin MjHKU4r-CoVs and bat HKU4-CoV, which use hDPP4 as a receptor, shared 8 and 7 out of 16 contact residues with MERS-CoV respectively, whereas HKU5-CoV, which shared only 1 residue, did not bind to hDPP4 (Fig. 4D). TyRo-CoV-162275 shared 10 of the 16 DPP4 key residues, suggesting that it may also use pangolin DPP4 as a cell-entry receptor (Fig. 4D).

In addition, we predicted a potential furin protease cleavage site (RQQR) in the *S* protein subdomain 1 (SD1) of the bat TyRo-CoV-162275 genome, which was similar to that in the MERS-CoV and pangolin MjHKU4r-CoV genomes (Fig. 4C, and D). Considering that such a cleavage site was not found in any of the publicly available bat HKU4r-CoV genomes, this is the first report of bat HKU4r-CoV containing a furin protease cleavage site.

4. Discussion

Bats play important roles in the emergence of zoonotic viruses. Most bat-borne viruses have been discovered through passive origination studies during disease outbreak investigations. During the past decades,

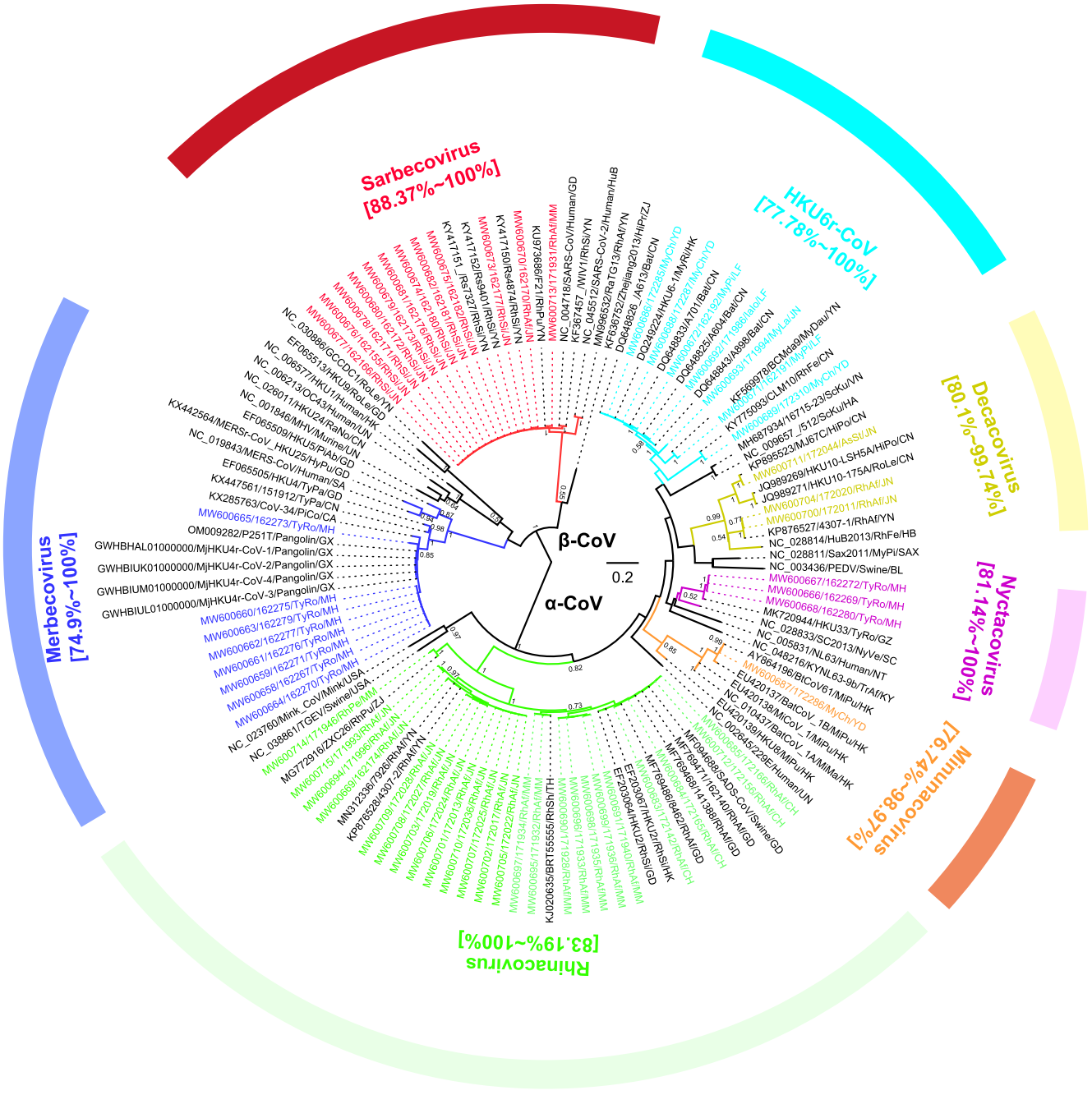

Fig. 2. Phylogenetic tree of the partial RNA-dependent RNA polymerase (*RdRp*) gene (387 bp) of coronavirus (CoV) strains found in bats. All CoVs detected are divided into seven clades, five of which belong to *αCoV* genus: HKU6r-CoV (light blue), Decacovirus (yellow), Nyctacovirus (purple), Minunacovirus (orange) and Rhinacovirus (green) clade, and two clades in *βCoV* genus: Merbecovirus (dark blue) and Sarbecovirus (red) clade. The scale bars represent 0.2 substitutions per nucleotide position.

several coronavirus disease outbreaks, such as SARS-CoV in 2002–2003, MERS-CoV in 2012, and SARS-CoV-2 in 2019, had severe impacts on human health and the world economy. To date, approximately 5000 bat coronavirus sequences have been reported, and bats are recognised as the major natural reservoirs of alpha- and beta-coronaviruses. Recent studies have shown that approximately 66,280 people are infected with SARSr-CoVs annually in Southeast Asia (Sanchez et al., 2022). We are currently in the era of proactive coronavirus discovery to fight future bat coronaviruses. We conducted a bat coronavirus survey in southern China, which has a wide distribution and rich diversity of bats. Fifty-eight coronavirus sequences have been identified, including HKU10r-CoV, HKU6r-CoV, SARSr-CoV, HKU2r-CoV, and MERSr-CoV. Most importantly, a cluster of HKU4r-CoVs was detected in greater bamboo bats, providing new evidence that pangolins are the HKU4r-CoVs intermediate reservoir (Chen et al., 2023).

Rhinacoviruses were detected in two species, *R. affinis* and *R. pearsoni*, from three locations in Yunnan (JN and CH) and Guangdong (MM) (Table 1). These sequences are in branch one and are closely related to HKU2-CoV and SADS-CoV, which were detected in *Rhinolophus* bats and swine from the Pan-Pearl River Delta region; the CoVs detected in this branch were all from *R. affinis* in CH. Unlike branch one, the CoVs detected in branch two all came from MM's *R. affinis*; meanwhile, CoVs in branch three also came from *R. affinis* in MM, but this branch was closer to the HKU2r-CoV evolution detected in *R. shameli* in Thailand. The CoVs

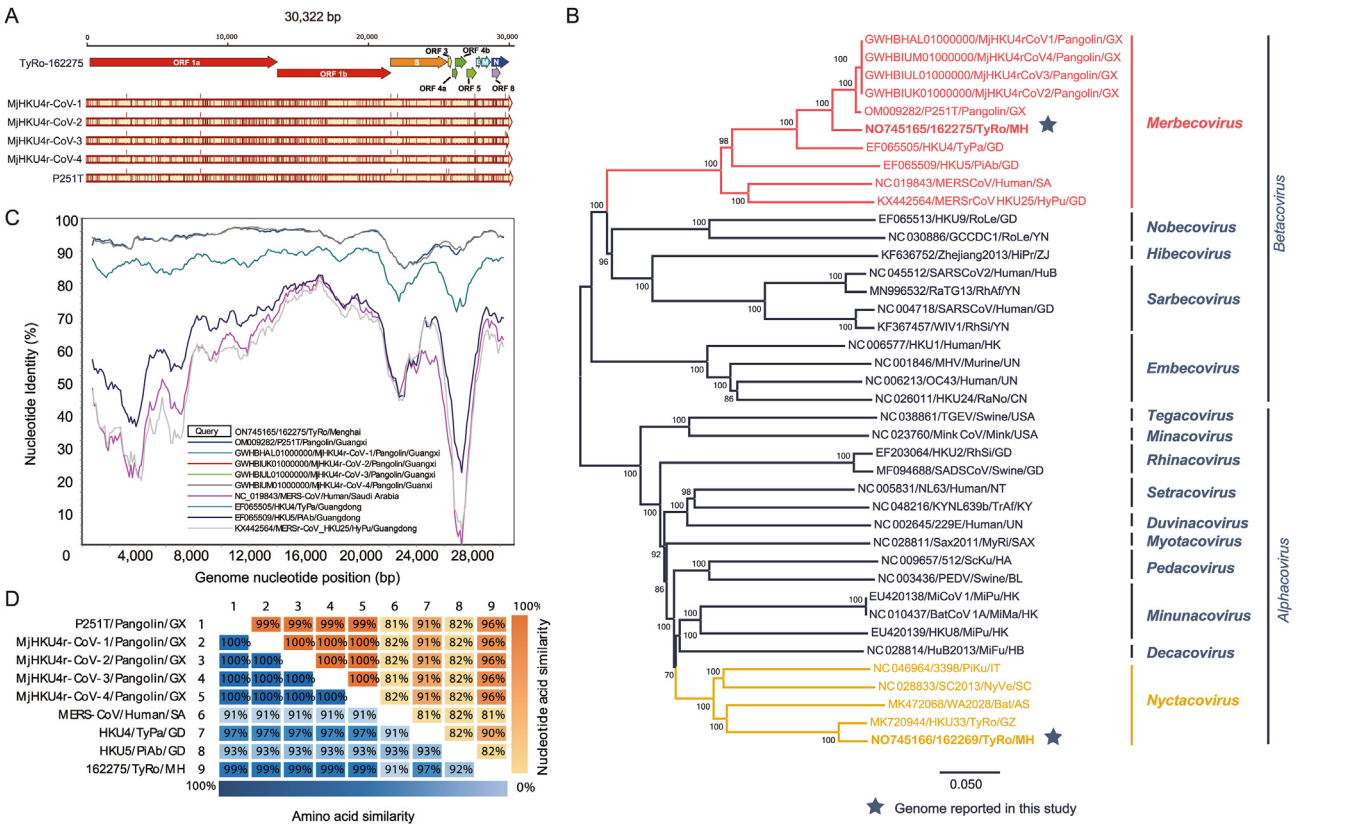

Fig. 3. Genomic phylogenetic analysis of bat coronaviruses TyRo-CoV-162275 and TyRo-CoV-162269. **A** Genome structure of TyRo-CoV-162275. **B** Phylogenetic tree based on the nucleotide sequences of the complete genomes of representative αCoV and βCoV. Merbecoviruses are shown in red, and Nyctacoviruses are shown in yellow. Shaded colours represent different subgenera of coronaviruses. The scale bars represent 0.05 substitutions per nucleotide position. **C** Similarity plot based on the full-length genome sequences. TyRo-CoV-162275 was used as a query sequence, Tylonycteris bat CoV HKU4-1, Pipistrellus bat CoV HKU5-1, pangolin CoV MjHKU4r-CoV-1–4, P251T, and human MERS-CoV were used as reference sequences. **D** Full spike amino acid and nucleotide similarity between P251T, MjHKU4r-CoV-1–4, MERS-CoV, HKU4-CoV, HKU5-CoV, and TyRo- CoV-162275.

detected in branches three and four came from *R. affinis* of JN and *R. pearsonii* of MM and were similar to the HKU2r-CoV detected in *R. affinis* in Yunnan Province and *R. pusillus* in Zhejiang Province (Fig. 2). SADS-CoV (also known as SeACoV and PEAV), isolated in 2017, caused outbreaks of severe watery diarrhoea in suckling piglets, with a mortality rate of up to 90% in several commercial pig farms in Guangdong Province, China (Zhou et al., 2018). Virological, epidemiological, evolutionary, and experimental evidences have shown that SADS-CoV is related to the bat coronavirus HKU2r-CoV (Zhou et al., 2018), which was found in Hong Kong, Guangdong, in a previous surveillance (Lau et al., 2007). Another interesting point is that locations MM in Guangdong Province and JN in Yunnan Province are separated by Guangxi Province; the direct distance between the two sampling locations is approximately 900 km (Fig. 1). Similar to SARSr-CoVs in *R. sinicus* and *R. affinis* from JN, RhAf-CoV-171931 also belongs to the Sarbecovirus clade but was sampled at MM. It shares 99.74% nucleotide identity with strain KU973686/F21/RhPu/YN, which was found in the Yunnan Province (Wang et al., 2017). This suggests a high exchange of CoV strains between Yunnan and Guangdong provinces via bat migration or intermediate hosts.

The highest positive rate and CoV diversity were observed in *R. affinis* at several of our sites; *R. affinis* usually roosts in the same cave as other *Rhinolophus* species, which are subsequently found to harbour the same bat CoV. For example, *R. affinis* harboured CoVs similar to those of *R. pearsoni* in the Rhinacovirus clade, *A. stoliczkanus* in the Decacovirus clade, and *R. sinicus* in the Sarbecovirus clade. Co-roosting of these bats in an enclosed cave environment may facilitate viral inter-species transmission. The bat species captured from MM (*R. pearsoni* and *R. affinis*) carried similar viruses that clustered in clade one, *A. stoliczkanus* and *R. affinis* from JN carried similar viruses clustered in the Decacovirus clade, and *M. pilosus* and *I. io* from LF carried similar viruses clustered in the HKU6r-CoV clade. This supports results from previous studies conducted in China, Spain, South America, and Thailand, where different bat species sampled within the same location carried genetically related viruses (Carrington et al., 2008; Langeland and Moore, 1990; Tang et al., 2006; Wacharapluesadee et al., 2015). Furthermore, HKU10r-CoV, HKU6r-CoV, SARSr-CoV, and HKU2r-CoV are carried by multiple but related host species (from the same genus or family), suggesting the host specificity of these CoVs. This confirms the inference of previous studies on bat hosts and geographic variations of bat-carrying coronaviruses in China, and bats from closely related taxa are prone to transmit similar coronaviruses compared to those from different geographies (Latinne et al., 2020; Yang et al., 2023).

Here, we report a branch of *T. robustula* HKU4r-CoVs in Yunnan that is evolutionarily related to bat HKU4-CoV in the Pan-Pearl River Delta region and pangolin MjHKU4-CoV in Guangxi. Receptor usage and spike protein cleavage site analysis of one of eight *T. robustula* HKU4r-CoV (TyRo-CoV-162275) suggested that *T. robustula* HKU4r-CoV may use hDPP4 as an entry receptor, and this was the first reported bat HKU4r-CoV to possess a furin protease cleavage site. Analysis of key residues in the hDPP4 receptor-binding region showed that bat TyRo-CoV-162275 shared 10 key sites with pangolin MjHKU4r-CoV. Although the hypothesis that hDPP4 could be used as a receptor by TyRo-CoV-162275 could not be directly supported due to a shortage of receptor utilisation experiments, previous studies on pangolin MjHKU4r-CoV showed that hDPP4 could be utilised in cases where MjHKU4r-CoV shared only eight key sites with MERS-CoV (Chen et al., 2023). Although no HKU4-CoV

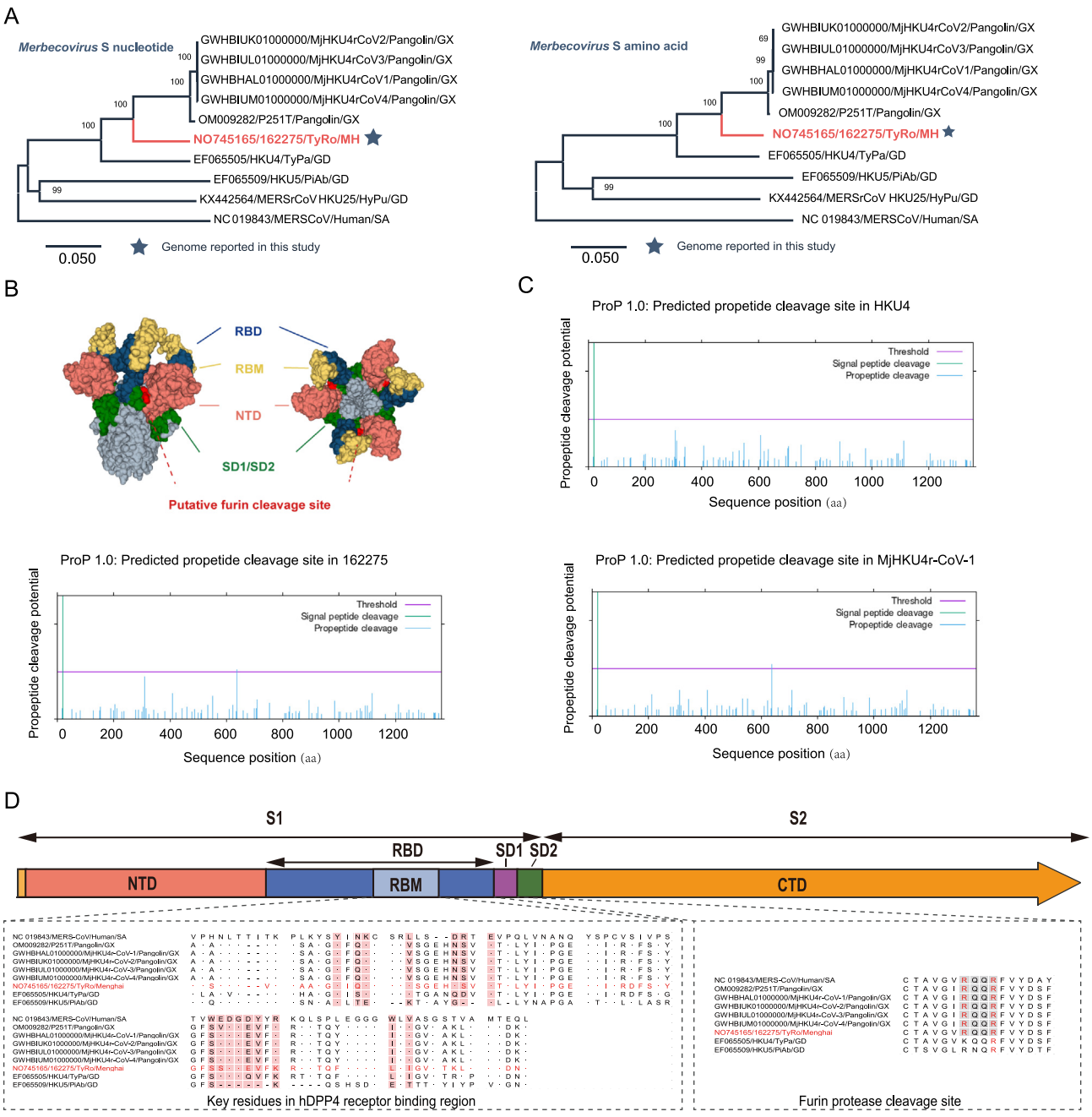

Fig. 4. Receptor usage and spike protein cleavage site analysis of TyRo-CoV-162275. **A** Phylogenetic tree based on the complete spike nucleotide and amino acid similarity between P251T, MjHKU4r-CoV-1-4, MERS-CoV, HKU4-CoV, HKU5-CoV, MERSr-CoV HKU25 and TyRo-CoV-162275. **B** Structure model for TyRo-CoV-162275 spike protein predicted based on the PDB: 5X5C structure as a template using the SWISS-MODEL online tool (swissmodel.expasy.org). The predicted furin cleavage site is shown in red. **C** Prediction of furin cleavage sites in the spike proteins of bat HKU4-CoV, pangolin CoV MjHKU4r-CoV-1 and TyRo-CoV-162275 was carried out using ProP-1.0 Server (https://services.healthtech.dtu.dk/service.php?ProP-1.0), using the furin-specific prediction as the default. **D** Schematic representation of the spike protein. S1 and S2 subunits are indicated, as well as four domains within S1, including the N-terminal domain (NTD), RBD, subdomain 1 (SD1), and subdomain 2 (SD2). RBM alignment is shown at the bottom left, and the 16 key residues at the surface between MERS-CoV RBM and human DPP4 are pink-shaded. Predicted furin cleavage sites of TyRo-CoV-162275 and the corresponding sites in other Merbecoviruses are shown at the bottom right. The key arginine sites are shown in red, and the predicted furin cleavage sites are grey-shaded.

infection has been reported in humans, previous studies have shown that pangolin-carried MjHKU4r-CoV is at risk of spilling into humans via hDPP4 (Chen et al., 2023). A previous study on the evolutionary analysis of coronaviruses in bats and pangolins also indicated gene flow between bat and pangolin coronaviruses (Yang et al., 2023). Enhanced surveillance of the natural hosts of bat HKU4r-CoV and serological investigation of bat habitat populations will help prevent a potential pandemic caused by the next bat coronavirus spillover triggered by HKU4r-CoV. Further studies on pangolin MjHKU4r-CoVs will also help clarify their potential spillover risk, although this experiment has not yet involved studies related to cellular and virulence levels.

Although samples of *R. sinicus*, *M. pilosus*, and *M. chinensis* were obtained from various locations, samples positive for CoV were only found at specific locations. The prevalence of the virus in host species varies from population to population. In addition, the prevalence of coronavirus has changed during the same season in different years. Hu et al. (2017)

conducted a 5-year longitudinal surveillance of bats SARSr-CoVs in Kunming, China. They collected 602 samples from one site in Kunming in April, May, July, September, and October (Hu et al., 2017). In their study, the positivity rate of SARSr-CoVs in *R. sinicus* varied in the same season and in different years. In addition, 4 out of 54 *R. sinicus* samples (obtained in May 2012) tested positive for SARSr-CoVs, but all *R. sinicus* samples collected in May 2014 and May 2015 were negative. This finding suggests that in addition to the bat host and geography, other factors also affect CoV transmission, making understanding the mechanism regulating CoV transmission elusive.

5. Conclusions

In conclusion, our findings demonstrate the greater diversity and ecological complexity of bat CoVs in southern China, Yunnan and Guangdong provinces than previously appreciated. We found that pangolin HKU4-related coronaviruses in *T. robustula* contain a furin protease cleavage site. Further studies and characterisations of bat CoVs should be conducted to provide additional insights into their host range and the evolutionary history of bat populations in China and Southeast Asia.

Data availability

The data presented in this study are available in the present article and supplementary materials. Viral *RdRp* sequences were deposited in GenBank under the accession numbers: MW600658–MW600715. Complete viral genomic sequences were deposited in GenBank under the accession numbers ON745165 and ON745166.

Ethics statement

The bats were captured with permission from the Guangdong Entomological Institute Administrative Panel on Laboratory Animal Care (No. GDEI-AE-2006001). The Administration of Laboratory Animals (Decree No. 2, State Science and Technology Commission, People's Republic of China) approved capture and sample collection.

Author contributions

Min Guo: conceptualisation, data curation, investigation, validation, writing-original draft preparation, writing-review, and editing. Kai Zhao: conceptualisation, formal analysis, methodology, validation, visualisation, writing-original draft preparation, writing-review and editing. Xingwen Peng: data curation, investigation, resources. Xiangyang He: investigation, resources. Jin Deng: investigation, resources. Bo Wang: writing-original draft. Xinglou Yang: conceptualisation, writing-original draft preparation, writing-review and editing, supervision, project administration, funding acquisition. Libiao Zhang: conceptualisation, writing-original draft preparation, writing-review and editing, supervision, project administration, funding acquisition.

Conflict of interest

The authors declare no conflicts of interest. The funders had no role the study design; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

Acknowledgements

This work was jointly funded by the Special Foundation for the National Science and Technology Basic Research Program of China (2021FY100303), the Guangdong Provincial Science and Technology Program (2021B1212050021, 2021B1212110003), the GDAS Special Project of Science and Technology Development (2021GDASYL-20210103052), and the Young Top-notch Talent Cultivation Program of Hubei Province and the Youth Innovation Promotion Association of the Chinese Academy of Sciences (CAS) (2019328). We thank Jie Liang for her valuable advice in this study, and Linmiao Li for her critical reading of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virs.2023.11.003.

References

Carrington, C.V., Foster, J.E., Zhu, H.C., Zhang, J.X., Smith, G.J., Thompson, N., Auguste, A.J., Ramkissoon, V., Adesiyun, A.A., Guan, Y., 2008. Detection and phylogenetic analysis of group 1 coronaviruses in South American bats. Emerg. Infect. Dis. 14, 1890–1893.

Chen, J., Yang, X., Si, H., Gong, Q., Que, T., Li, J., Li, Y., Wu, C., Zhang, W., Chen, Y., Luo, Y., Zhu, Y., Li, B., Luo, D., Hu, B., Lin, H., Jiang, R., Jiang, T., Li, Q., Liu, M., Xie, S., Su, J., Zheng, X., Li, A., Yao, Y., Yang, Y., Chen, P., Wu, A., He, M., Lin, X., Tong, Y., Hu, Y., Shi, Z.L., Zhou, P., 2023. A bat MERS-like coronavirus circulates in pangolins and utilizes human DPP4 and host proteases for cell entry. Cell 186, 850–863.e816.

Cui, J., Li, F., Shi, Z.L., 2019. Origin and evolution of pathogenic coronaviruses. Nat. Rev. Microbiol. 17, 181–192.

Drexler, J.F., Gloza-Rausch, F., Glende, J., Corman, V.M., Muth, D., Goettsche, M., Seebens, A., Niedrig, M., Pfefferle, S., Yordanov, S., Zhelyazkov, L., Hermanns, U., Vallo, P., Lukashev, A., Müller, M.A., Deng, H., Herrler, G., Drosten, C., 2010. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. J. Virol. 84, 11336–11349.

Fan, Y., Zhao, K., Shi, Z.L., Zhou, P., 2019. Bat coronaviruses in China. Viruses 11, 210.

He, W.T., Hou, X., Zhao, J., Sun, J., He, H., Si, W., Wang, J., Jiang, Z., Yan, Z., Xing, G., Lu, M., Suchard, M.A., Ji, X., Gong, W., He, B., Li, J., Lemey, P., Guo, D., Tu, C., Holmes, E.C., Shi, M., Su, S., 2022. Virome characterization of game animals in China reveals a spectrum of emerging pathogens. Cell 185, 1117–1129.e1118.

Hu, B., Zeng, L.P., Yang, X.L., Ge, X.Y., Zhang, W., Li, B., Xie, J.Z., Shen, X.R., Zhang, Y.Z., Wang, N., Luo, D.S., Zheng, X.S., Wang, M.N., Daszak, P., Wang, L.F., Cui, J., Shi, Z.L., 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog. 13, e1006698.

Lam, T.T., Jia, N., Zhang, Y.W., Shum, M.H., Jiang, J.F., Zhu, H.C., Tong, Y.G., Shi, Y.X., Ni, X.B., Liao, Y.S., Li, W.J., Jiang, B.G., Wei, W., Yuan, T.T., Zheng, K., Cui, X.M., Li, J., Pei, G.Q., Qiang, X., Cheung, W.Y., Li, L.F., Sun, F.F., Qin, S., Huang, J.C., Leung, G.M., Holmes, E.C., Hu, Y.L., Guan, Y., Cao, W.C., 2020. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature 583, 282–285.

Langeland, N., Moore, L.J., 1990. Reduction of HSV-1 binding to BHK cells after treatment with phosphatidylinositol-specific phospholipase C. FEBS Lett. 277, 253–256.

Latinne, A., Hu, B., Olival, K.J., Zhu, G., Zhang, L., Li, H., Chmura, A.A., Field, H.E., Zambrana-Torrelio, C., Epstein, J.H., Li, B., Zhang, W., Wang, L.F., Shi, Z.L., Daszak, P., 2020. Origin and cross-species transmission of bat coronaviruses in China. Nat. Commun. 11, 4235.

Lau, S.K., Woo, P.C., Li, K.S., Huang, Y., Wang, M., Lam, C.S., Xu, H., Guo, R., Chan, K.H., Zheng, B.J., Yuen, K.Y., 2007. 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.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B.T., Zhang, S., Wang, L.F., 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science 310, 676–679.

Sanchez, C.A., Li, H., Phelps, K.L., Zambrana-Torrelio, C., Wang, L.F., Zhou, P., Shi, Z.L., Olival, K.J., Daszak, P., 2022. A strategy to assess spillover risk of bat SARS-related coronaviruses in Southeast Asia. Nat. Commun. 13, 4380.

Shi, W., Shi, M., Que, T.C., Cui, X.M., Ye, R.Z., Xia, L.Y., Hou, X., Zheng, J.J., Jia, N., Xie, X., Wu, W.C., He, M.H., Wang, H.F., Wei, Y.J., Wu, A.Q., Zhang, S.F., Pan, Y.S., Chen, P.Y., Wang, Q., Li, S.S., Zhong, Y.L., Li, Y.J., Tan, L.H., Zhao, L., Jiang, J.F., Hu, Y.L., Cao, W.C., 2022. Trafficked Malayan pangolins contain viral pathogens of humans. Nat Microbiol 7, 1259–1269.

Tang, X.C., Zhang, J.X., Zhang, S.Y., Wang, P., Fan, X.H., Li, L.F., Li, G., Dong, B.Q., Liu, W., Cheung, C.L., Xu, K.M., Song, W.J., Vijaykrishna, D., Poon, L.L., Peiris, J.S., Smith, G.J., Chen, H., Guan, Y., 2006. Prevalence and genetic diversity of coronaviruses in bats from China. J. Virol. 80, 7481–7490.

Wacharapluesadee, S., Duengkae, P., Rodpan, A., Kaewpom, T., Maneeorn, P., Kanchanasaka, B., Yingsakmongkon, S., Sittidetboripat, N., Chareesaen, C., Khlangsap, N., Pidthong, A., Leadprathom, K., Ghai, S., Epstein, J.H., Daszak, P., Olival, K.J., Blair, P.J., Callahan, M.V., Hemachudha, T., 2015. Diversity of coronavirus in bats from Eastern Thailand. Virol. J. 12, 57.

Wang, L., Fu, S., Cao, Y., Zhang, H., Feng, Y., Yang, W., Nie, K., Ma, X., Liang, G., 2017. Discovery and genetic analysis of novel coronaviruses in least horseshoe bats in southwestern China. Emerg Microbes Infect 6, e14.

Wu, Z., Yang, L., Ren, X., Zhang, J., Yang, F., Zhang, S., Jin, Q., 2016. ORF8-related genetic evidence for Chinese horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. J. Infect. Dis. 213, 579–583.

Xiao, K., Zhai, J., Feng, Y., Zhou, N., Zhang, X., Zou, J.J., Li, N., Guo, Y., Li, X., Shen, X., Zhang, Z., Shu, F., Huang, W., Li, Y., Zhang, Z., Chen, R.A., Wu, Y.J., Peng, S.M., Huang, M., Xie, W.J., Cai, Q.H., Hou, F.H., Chen, W., Xiao, L., Shen, Y., 2020. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. Nature 583, 286–289.

Yang, J., Skaro, M., Chen, J., Zhan, D., Lyu, L., Gay, S., Kandeil, A., Ali, M.A., Kayali, G., Stoianova, K., Ji, P., Alabady, M., Bahl, J., Liu, L., Arnold, J., 2023. The species coalescent indicates possible bat and pangolin origins of the COVID-19 pandemic. Sci. Rep. 13, 5571.

Yang, T., 2016. A revised bat (Mammalia: Chiroptera) checklist for Guizhou, China. Biodivers. Sci. 24, 957–962.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., Fouchier, R.A., 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med. 367, 1814–1820.

Zhou, P., Fan, H., Lan, T., Yang, X.L., Shi, W.F., Zhang, W., Zhu, Y., Zhang, Y.W., Xie, Q.M., Mani, S., Zheng, X.S., Li, B., Li, J.M., Guo, H., Pei, G.Q., An, X.P., Chen, J.W., Zhou, L., Mai, K.J., Wu, Z.X., Li, D., Anderson, D.E., Zhang, L.B., Li, S.Y., Mi, Z.Q., He, T.T., Cong, F., Guo, P.J., Huang, R., Luo, Y., Liu, X.L., Chen, J., Huang, Y., Sun, Q., Zhang, X.L., Wang, Y.Y., Xing, S.Z., Chen, Y.S., Sun, Y., Li, J., Daszak, P., Wang, L.F., Shi, Z.L., Tong, Y.G., Ma, J.Y., 2018. Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. Nature 556, 255–258.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., Chen, H.D., Chen, J., Luo, Y., Guo, H., Jiang, R.D., Liu, M.Q., Chen, Y., Shen, X.R., Wang, X., Zheng, X.S., Zhao, K., Chen, Q.J., Deng, F., Liu, L.L., Yan, B., Zhan, F.X., Wang, Y.Y., Xiao, G.F., Shi, Z.L., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273.

Zhu, W, Huang, Y, Gong, J, Dong, L, Yu, X, Chen, H, Li, D, Zhou, L, Yang, J, Lu, S, 2023. A novel bat coronavirus with a polybasic furin-like cleavage site. Virol Sin 38, 344–350.