



## Research Article

## Individual virome analysis reveals the general co-infection of mammal-associated viruses with SARS-related coronaviruses in bats

Hao-Rui Si<sup>a,c,1</sup>, Ke Wu<sup>b,1</sup>, Jia Su<sup>a,c,1</sup>, Tian-Yi Dong<sup>a,c</sup>, Yan Zhu<sup>a</sup>, Bei Li<sup>a</sup>, Ying Chen<sup>a</sup>, Yang Li<sup>a</sup>, Zheng-Li Shi<sup>a,b,\*</sup>, Peng Zhou<sup>b,d,\*</sup><sup>a</sup> Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 43000, China<sup>b</sup> Guangzhou National Laboratory, Guangzhou International Bio Island, Guangzhou 510005, China<sup>c</sup> University of Chinese Academy of Sciences, Beijing 100000, China<sup>d</sup> State Key Laboratory of Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical School, Guangzhou 510005, China

## ARTICLE INFO

## Keywords:

Bat  
Microbiome  
SARS-related coronavirus  
Co-infection  
Alphacoronavirus

## ABSTRACT

Bats are the natural reservoir hosts for SARS-related coronavirus (SARSr-CoV) and other highly pathogenic microorganisms. Therefore, it is conceivable that an individual bat may harbor multiple microbes. However, there is limited knowledge on the overall co-circulation of microorganisms in bats. Here, we conducted a 16-year monitoring of bat viruses in south and central China and identified 238 SARSr-CoV positive samples across nine bat species from ten provinces or administrative districts. Among these, 76 individual samples were selected for further metagenomics analysis. We found a complex microenvironment characterized by the general co-circulation of microbes from two different sources: mammal-associated viruses or environment-associated microbes. The later includes commensal bacteria, enterobacteria-related phages, and insect or fungal viruses of food origin. Results showed that 25% (19/76) of the samples contained at least one another mammal-associated virus, notably alphacoronaviruses (13/76) such as AlphaCoV/YN2012, HKU2-related CoV and AlphaCoV/Rf-HuB2013, along with viruses from other families. Notably, we observed three viruses co-circulating within a single bat, comprising two coronavirus species and one picornavirus. Our analysis also revealed the potential presence of pathogenic bacteria or fungi in bats. Furthermore, we obtained 25 viral genomes from the 76 bat SARSr-CoV positive samples, some of which formed new evolutionary lineages. Collectively, our study reveals the complex microenvironment of bat microbiome, facilitating deeper investigations into their pathogenic potential and the likelihood of cross-species transmission.

## 1. Introduction

The SARS-related coronavirus (SARSr-CoV) represents a group of enveloped, single-stranded positive-sense RNA viruses categorized within the *Betacoronavirus* genus of the *Coronaviridae* family, which caused two global pandemics in the past two decades: the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002 and the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in 2019 (Cui et al., 2019; Zhou et al., 2020). Bats are recognized as natural reservoirs for various viruses, including the presumed ancestor viruses for SARS-CoV and SARS-CoV-2 (Ge et al., 2013; Temmam et al., 2022). Thus, surveillance and risk assessment of bat-derived SARS-related

coronaviruses is imperative for future pandemic preparedness (Keusch et al., 2022).

Bats not only carry coronaviruses but also harbor various highly pathogenic viruses, such as Hendra, Nipah, Ebola, and Marburg virus, which could pose risks to people during fieldwork (Hu et al., 2015; Olival et al., 2017; Kohl et al., 2021). The presence of co-circulating microorganisms may exacerbate these risks by affecting viral pathogenicity, evolutionary trajectories, and the likelihood of cross-species transmission (Abu-Raddad et al., 2006; Gupta et al., 2008; Loving et al., 2010). Therefore, comprehending the range and distribution of co-circulating microorganisms in the SARSr-CoV positive samples is crucial for understanding viral pathogenicity. In addition, it provides valuable insights

\* Corresponding authors.

E-mail addresses: [zhou\\_peng@gzlab.ac.cn](mailto:zhou_peng@gzlab.ac.cn) (P. Zhou), [zshi@wh.iov.cn](mailto:zshi@wh.iov.cn) (Z.-L. Shi).<sup>1</sup> Hao-Rui Si, Ke Wu, and Jia Su contributed equally to this work.

into future surveillance priorities and cross-species transmission mechanisms (Zhu et al., 2023).

To identify potential co-circulating microorganisms, we collected 238 bat SARSr-CoV positive samples and sequenced 140 samples using next-generation sequencing (NGS). Subsequent analyses were performed on 76 samples to detect co-circulating microorganisms, including those potentially pathogenic microbes, by aligning reads to reference sequences in microbial databases (Huang et al., 2020; Gao et al., 2023; Han et al., 2023; Xie et al., 2024). Our findings reveal the complex micro-environment within a bat SARSr-CoV sample, uncovering 25 new viral genomes alongside SARSr-CoVs. In summary, our study provided an insight into the potential co-circulating microorganisms in bat samples positive for SARSr-CoV, which contributes to our understanding of the bat-borne microbiome.

## 2. Material and methods

### 2.1. Sample collection

Between 2004 and 2019, we obtained 238 bat SARSr-CoV positive samples, including 171 from anal swab, 62 from fecal samples, and 5 from intestines (Supplementary Table S1). Samples were collected, transported and stored according to standard procedures according to the animal ethics permit (WIVA05201705) issued by the Wuhan Institute of Virology. The samples were stored in  $-80^{\circ}\text{C}$  freezers until further processing.

### 2.2. SARSr-CoV detection

The samples were centrifuged at  $20,000 \times g$  for 10 min and 200  $\mu\text{L}$  of the supernatant was subjected to RNA extraction using the QIAamp 96 Virus QIAcube HT Kit (Qiagen) according to the manufacturer's instructions. Samples were detected using pan-CoV detection primers targeting a 440-nt fragment in the *RdRp* gene by one-step hemi-nested reverse transcription PCR (RT-PCR), as described previously (Hu et al., 2017). Amplified PCR products were examined by gel electrophoresis. Targeted products were purified using the gel extraction kit (Omega) and sequenced by an ABI 3730XL DNA Analyzer (Sangon Biotech). The newly obtained viral sequences were aligned with known CoV sequences from GenBank. Of 238 SARSr-CoV positive samples, 140 were further processed for NGS.

### 2.3. NGS analysis of co-circulating microorganisms

An unpurified RNA library (without removing host RNA) was constructed using MGIEasy RNA Library Prep Set (96 RXN) (MGI). Briefly, 10  $\mu\text{L}$  of total RNA was added to first strand synthesis buffer and random primers before incubating at  $87^{\circ}\text{C}$  for 6 min to generate RNA fragments larger than 300 nucleotides (nt). Following first and second strand cDNA synthesis, double-stranded cDNA was purified using Agencourt AMPure XP beads (Beckman Coulter Genomics). To obtain a library size larger than 300 nt, the library was amplified by PCR using the following conditions: initial denaturation at  $95^{\circ}\text{C}$  for 3 min; 17 cycles of denaturation for 30 s at  $95^{\circ}\text{C}$  followed by annealing for 30 s at  $56^{\circ}\text{C}$  and extension for 60 s at  $72^{\circ}\text{C}$  and a final extension step for 5 min at  $72^{\circ}\text{C}$  and a stop reaction at  $4^{\circ}\text{C}$  to hold. Libraries were purified using Agencourt AMPure XP beads (Beckman Coulter Genomics), eluted in 20  $\mu\text{L}$  nuclease-free water, visualised on a 1.5% agarose gel and quantified using a Bio-analyzer High Sensitivity DNA Assay (Agilent). Then paired-end (150-bp) sequencing of the RNA library was performed on the MGI DNBSEQ-T7 platform (v2.0). Cutadapt (v1.18) (Kechin et al., 2017) was used to remove low-quality and contaminative reads with a minimum length of 30 bp and other parameters at default settings.

MEGAHIT (v1.2.9) (Li et al., 2015) was used to assemble filtered reads into contigs with default parameters. Target contigs were selected

using Blastn (v2.12.0+) (Camacho et al., 2009) searching against a local virus database. About 25 new SARSr-CoV viral genomic sequences were obtained and checked using Geneious (v11.0.3) with moderate sensitivity and other parameters at default settings. Of the 140 high-throughput sequenced samples, 76 samples were selected for further analysis due to the acquisition of SARSr-CoV genomes after de novo assembly. The filtered reads were then mapped to the reference database (downloaded from NCBI, <https://www.ncbi.nlm.nih.gov/>, for virus; downloaded from silva, a high quality ribosomal RNA databases, <https://www.arb-silva.de/search/show/ssu/cart>, for fungi and bacteria) using BWA (v 0.7.17-r1188) (Li and Durbin, 2010) with BWA-MEM algorithm. Samtools (v 1.9) (Danecek et al., 2021) was used to index and sort sam files and an in-house script was used for statistical analysis of the mapping results. A uniform empirical relaxed criterion (coverage  $\geq 10\%$ , mapped read number  $\geq 200$  and average-depth  $\geq 1$ ) was used to filter out potential co-circulating microorganisms which were then displayed in R (v4.2.1) using ggplot2. Another more rigorous criterion (coverage  $\geq 60\%$ , mapped read number  $\geq 200$  and average-depth  $\geq 1$ ) was used to screen mammal-associated microbes, also assisted by the de novo assembly results. And the screened reads were also compared with the pathogens in the list of notifiable human infectious diseases from the National Health Commission of the People's Republic of China and known mammalian-associated viromes of bats that have been reported.

### 2.4. The spatial and temporal distribution of bat SARSr-CoV positives

The geographic distribution of 238 SARSr-CoV-positive samples was marked on map by ArcGis (v10.8). The corresponding host and temporal distribution were labeled in R (v 4.2.1) using packages: pheatmap. The statistical histograms of host, geographic, and sampling time were presented by GraphPad (v8.0.2).

### 2.5. The phylogeny analysis of full-length viral genomes

Reference sequences were downloaded from the Virus Metadata Resource (VMR) sheet of the International Committee on Taxonomy of Viruses (ICTV) (<https://ictv.global/vmr>). All 25 genomic sequences and corresponding reference sequences were aligned by mafft (v7.505) (Katoh and Standley, 2013) with default parameters. Then pairwise genomic identity was calculated in Bioedit (v7.1.3.0). IQtree (v2.2.2.7) (Minh et al., 2020) was used to construct phylogenies with the most suitable model detected automatically and bootstrap value set to 1000 after truncating the ends of aligned sequences. Visualization of the phylogenetic tree was applied in R (v 4.2.1) using ggtree packages (v3.10) (Yu et al., 2018).

### 2.6. The annotation of viral genomes

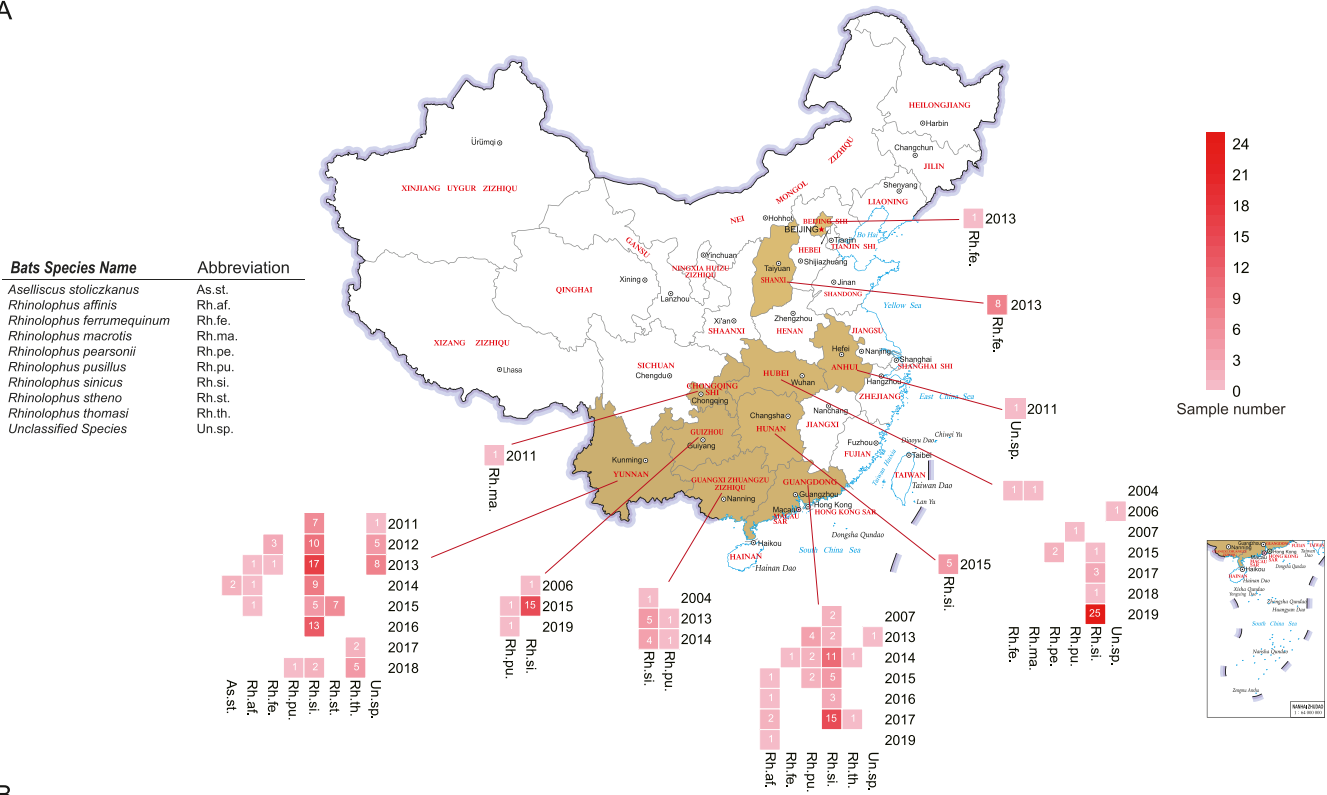
All 25 viral genomes were uploaded to ORFinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) for ORF search with parameters set by minimum ORF length of 75 nt, standard genetic code and start codon only using "ATG". After the ORF searching, blastp was used to identify the specific genes and also assisted by manual judgment. Visualization of the genomic structures was applied in R (v 4.2.1) using ggtree packages (v3.10) (Yu et al., 2018).

## 3. Results

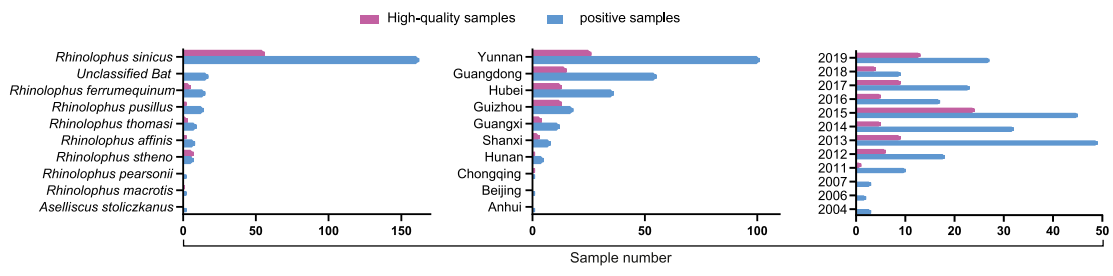
### 3.1. An extensive long-term monitoring of bat viruses in China

From 2004 to 2019, we conducted long-term monitoring of bat viruses in south and central China, which resulted in a total of 238 positive samples for SARSr-CoVs across nine different bat species from ten provinces or administrative districts (Fig. 1, Supplementary Table S1). Although multiple hosts and provinces were involved in the sampling

A



B



**Fig. 1.** The spatial and temporal distribution of bat SARSr-CoVs positive samples China. **A** The geographical distribution of viral positive samples and their host. The sampling time, location and host species information for an individual sample were shown in red rectangle in a map of China. Numbers were shown for more than one samples with similar information. See more details in [Supplementary Table S1](#). **B** Histograms of host, geographic, and sampling time information, the high-quality samples used for subsequent NGS analysis were in purple.

process, a sampling bias is apparent. The main hosts were *Rhinolophus sinicus* (n = 162, 68%), *Rhinolophus ferrumequinum* (n = 15, 6.30%), *Rhinolophus pusillus* (n = 14, 5.88%), while the major provinces were Yunnan (n = 101, 42.44%), Guangdong (n = 55, 23.11%), Hubei (n = 36, 15.13%).

### 3.2. Characterization of bat viromes in SARSr-CoV positive samples

To explore the potential co-circulating microbe, we selected 76 high-quality positive samples for SARSr-CoV (the genomic sequences of which were obtained and analyzed in previous study), primarily from south China, for virome sequencing. It is noteworthy that although the composition is uneven, the spatial and temporal distribution of these 76 samples (involving seven host species, eight provinces or administrative districts, and spanning nine years) is relatively diverse and representative (Fig. 1B, [Supplementary Table S1](#)).

The meta-transcriptomic sequencing generated over 7.5 billion reads in raw data post quality control, with an average of 99,989,899 reads per sample. De novo assembly yielded a total of 29,919,070 contigs. The detection of viruses involved mapping, where a predefined filter criterion was established (coverage  $\geq 10\%$ , mapped read number  $\geq 200$ , and average depth  $\geq 1$ ). Reads that passed the filter were then compared with

the pathogens in the list of notifiable human infectious diseases from the National Health Commission of the People's Republic of China and known mammalian-associated virome associated with bats. Taxonomic classification was then assigned based on the relevant references.

### 3.3. The co-circulating viruses in bat SARSr-CoV positive samples

Our analysis identified 19 phyla, 38 classes, 61 orders, 206 families and 1488 genera for viruses (Fig. 2, [Supplementary Table S2](#)). The primary virus taxa with a sample carrying probability exceeding 66.67% were summarized (Fig. 2).

Overall, these viruses can be classified into three categories, reflecting their sources: (1) Mammal-associated viruses: Apart from coronaviruses, the most prevalent viruses are bat herpesviruses, detected in 51 samples (67.1% positive), and primarily related to human betaherpesvirus 7. Other significant co-circulating viruses include orthobunyavirus, picornavirus, circovirus, adenovirus and polyomavirus, which we will analyze in details below. (2) Viruses associated with bat diet: A large variety of viruses related to plant or insect are prevalent in bats, which were probably obtained from diet. Viruses in *Bunyavirales* can be detected in 69 samples, with the most common families being *Phenuiviridae* and *Tospoviridae*. Although some viruses in *Phenuiviridae* can infect humans



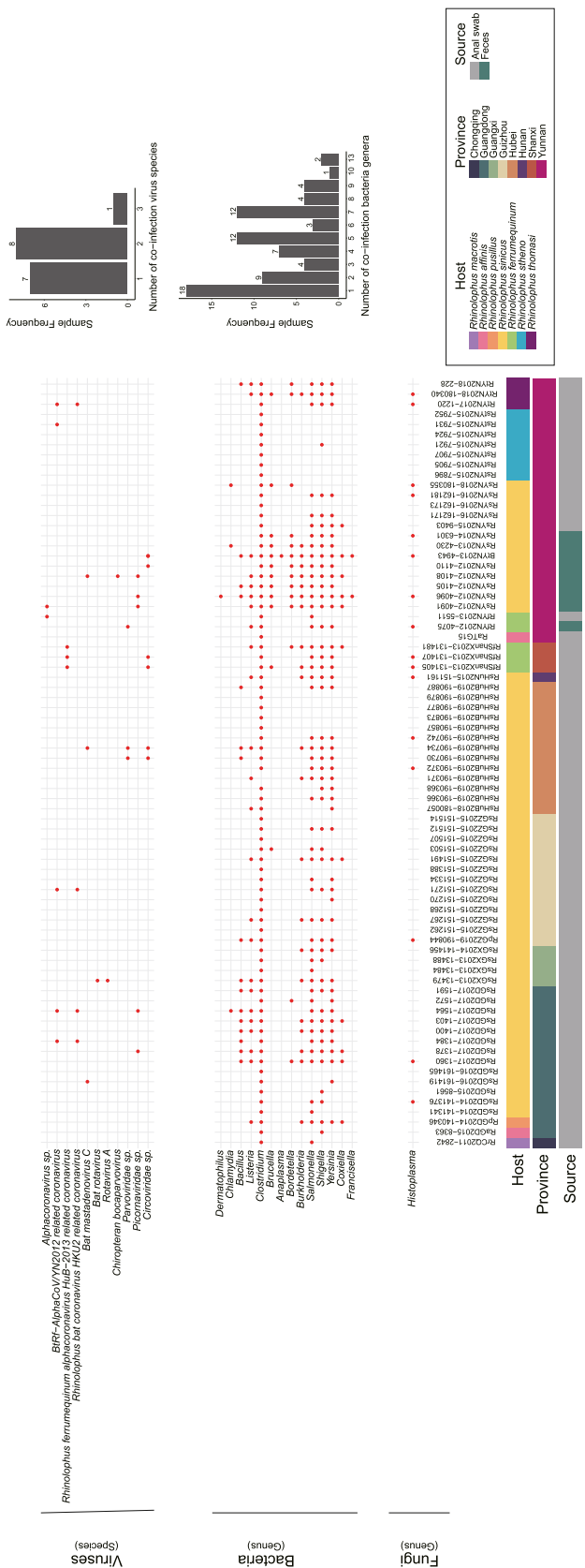


Fig. 3. The co-circulating mammal-associated microbes. The co-circulating mammal-associated microbes for each of the 76 samples are shown in dot, with the corresponding statistical histograms of co-circulating microbes displayed on the right. The host, geographic and sample type information are marked below.

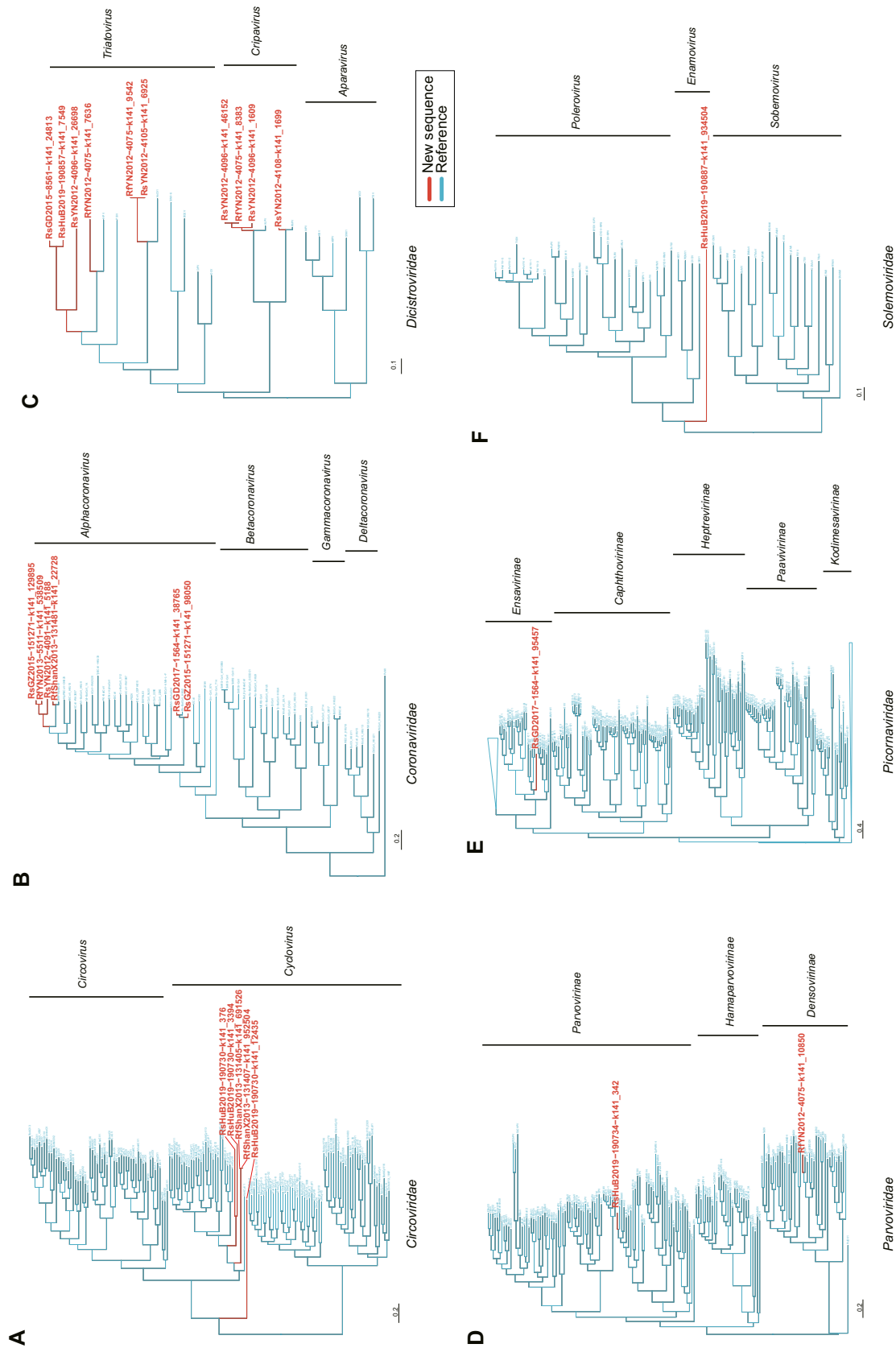
and cause disease (Yu et al., 2011), the primary virus species we detected belong to *Tenuivirus* genus, which only infect plants, sharing a host tropism with viruses in the *Tospoviridae* family (Falk and Tsai, 1998). Similarly, although *Poxviridae* includes pathogenic human viruses (smallpox and etc), the two primary virus species we detected in nearly 76 samples are all insect-associated poxviruses from *Betaentomopoxvirus* or *Deltaentomopoxvirus* genus. Other viruses in *Nudiviridae* associated with a diverse range of insects and crustaceans were also found. (3) Viruses associated with bat microbiome: Viruses related to enterobacteria, parasites or fungi are also prevalent. Amoeba-associated *Mimiviridae*, the largest virus ever found (with a DNA genome length more than 1 million bp) (Kutikhin et al., 2014), was detected in 53 bat samples. A common flagellate-infecting species (*Cafeteria roenbergensis virus*) in *Mimiviridae* is the sole virus species in our analysis. Moreover, tailed bacteriophage *Caudoviricetes* accounts for about 30% of all recognized virus species; the majority of viruses detected in our study belong to the unclassified *Caudoviricetes* order, suggesting that they are bacteriophages specific to bats. Furthermore, the fungus-associated *Chrysoviridae* (Sahin and Akata, 2018), an unclassified family includes mycoplasma phage phiMFV1 and saccharomyces cerevisiae killer virus M1, are all likely from the bat microbiome. There is also a large variety of retrovirus fragments classified under the *Revtraviricetes*, *Preplasmiviricota* and *Polydnaviriformidae*, which could originate from either retrovirus infection or endogenous viral elements (Fig. 2, Supplementary Table S2).

### 3.4. The co-circulating mammal-associated microbes

We then focused on mammal-associated microbes. To further ensure the accuracy at the species level, we raised the positive threshold to: coverage  $\geq 60\%$ , mapped read number  $\geq 200$  and average-depth  $\geq 1$  (Fig. 3, Supplementary Tables S3 and S4).

Overall, co-infection of SARSr-CoVs with other mammal-associated viruses is rather common in these positive samples. About 25% of samples (19/76) harbor at least one another mammalian virus. We also found that multiple alphacoronaviruses may co-circulate with SARSr-CoV. In addition to co-infection with bat HKU2-related coronavirus, we also observed co-infection with BtRf-AlphaCoV/YN2012 related coronavirus, and BtRf-AlphaCoV/Rf HuB-2013 related coronavirus. In total, 13 samples (17%) were co-infected with at least one alphacoronavirus. Moreover, bat picornavirus, bat adenovirus or adeno-associated virus, bat rotavirus, bat parvovirus, and bat circovirus may also co-circulate with SARSr-CoV. Finally, we found co-infection with three other viruses in one sample: AlphaCoV/YN2012, HKU2-related CoV and bat picornavirus, reflecting the complex microenvironment of bat virome.

We also identified several populations of bacteria or fungi that may be pathogenic to mammals. *Firmicutes* bacteria usually possess gram-positive cell wall structures and can produce endospores, rendering them resistant to desiccation and capable of surviving extreme conditions (Fujisaka et al., 2023). Among *Firmicutes*, bacteria species associated with *Listeria monocytogenes*, the causative agent of the highly virulent foodborne infection listeriosis, were detected in 24 samples (18 anal swabs and 6 fecal pellets) collected from seven provinces and four bat species (Ramaswamy et al., 2007). In three samples (BtYN2013-4943, RfShanX2013-131481 and RsYN2012-4096), the mapping coverage of the corresponding 16S ribosomal RNA reference gene are even more than 99%, with mapped read numbers exceeding 3000 and average depth exceeding 300 at the same time. Furthermore, species related to *Clostridium botulinum*, which can produce the neurotoxin botulinum leading to botulism, a severe flaccid paralytic disease in humans and animals, was detected in 75 samples (66 anal swabs and 9 fecal pellets) collected from eight regions and seven bat species with mapping coverage  $\geq 95\%$ , mapped read number  $\geq 3000$  and average depth  $\geq 150$  in 12% samples (n = 9) (Peck, 2009). *Proteobacteria*, a major phylum of Gram-negative bacteria encompassing pathogenic genera such as *Escherichia*, *Salmonella*, *Vibrio*, *Yersinia*, *Legionella*, as well as bacteria responsible for nitrogen fixation, were found in all samples (Rizzatti et al., 2017). Among *Proteobacteria*, various genera associated with



**Fig. 4.** Phylogenetic analysis of 25 new viral genomic sequences. The phylogenies of each new virus genomes were aligned with the representative virus in each of the viral family: **A** *Circoviridae*, **B** *Coronaviridae*, **C** *Dicistroviridae*, **D** *Panoviridae*, **E** *Picornaviridae*, **F** *Solemoviridae*. Sequences contributed by this study were marked in red and the reference sequences were in blue.

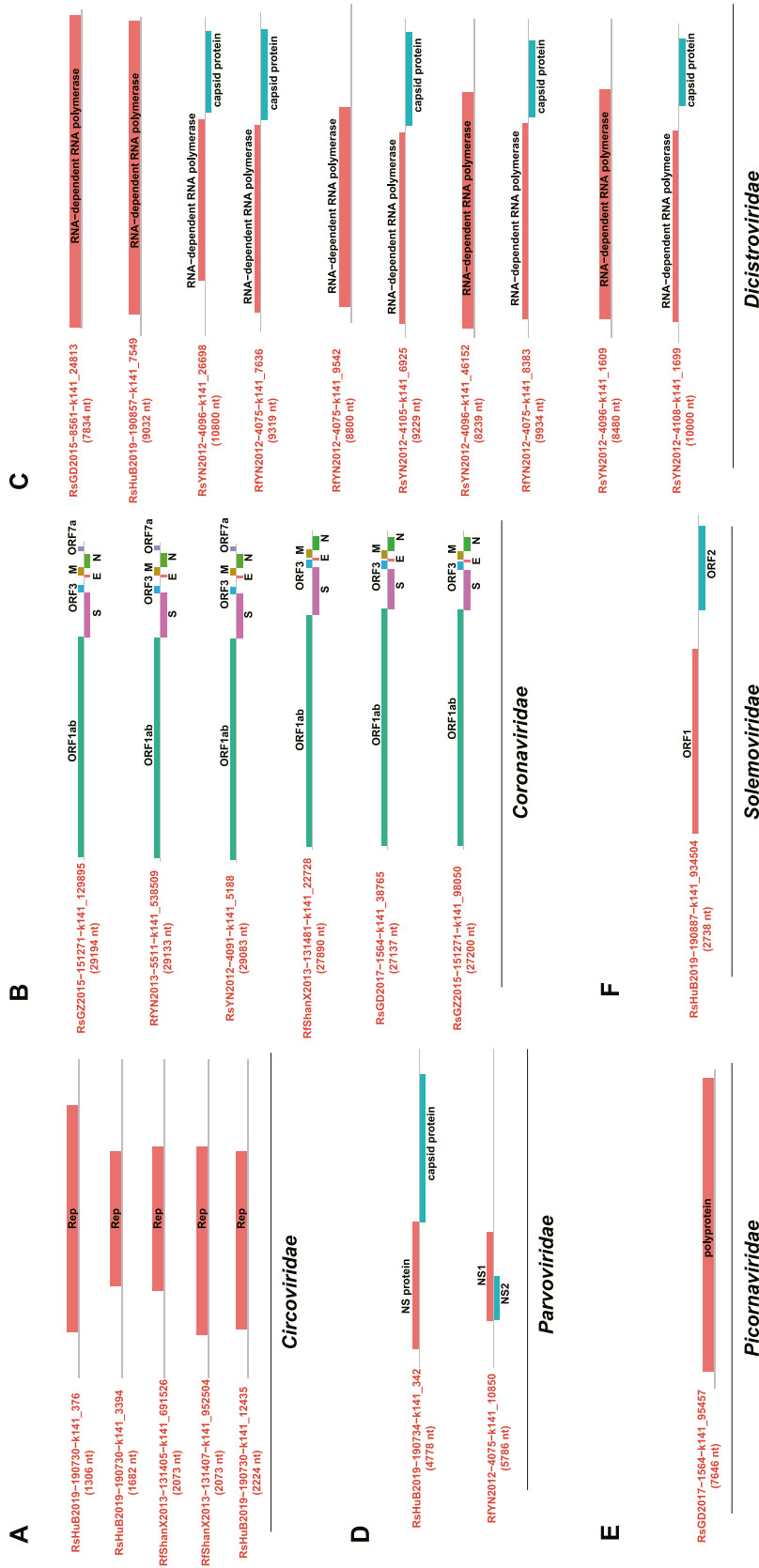


Fig. 5. Genome structures of 25 new viral sequences. The distributions of annotated ORFs on genomes were displayed in blocks with different colors for each of the viral family: **A** *Circoviridae*, **B** *Coronaviridae*, **C** *Dicistroviridae*, **D** *Parvoviridae*, **E** *Solemoviridae*, **F** *Dicistroviridae*. Genomes are represented by a thin gray line with all ORFs indicated.

pathogens, such as *Shigella* (causative agent of human shigellosis) and *Salmonella* (agent of Salmonellosis), were detected in 48 and 50 samples respectively, showing diverse host and geographical distribution (Gut et al., 2018; Sahin and Akata, 2018). *Bacteroidota*, detected in 72 samples, is composed of three large classes of Gram-negative bacteria widely distributed in the environment. They can act as opportunistic pathogens and symbiotic species highly adapted to the gastrointestinal tract (Rajilic-Stojanovic and de Vos, 2014). *Cyanobacteria*, detected in 53 samples, represent a phylum of autotrophic gram-negative bacteria capable of obtaining biological energy through photosynthesis. Certain *Cyanobacteria* have the potential to produce neurotoxins that are harmful to humans and animals, causing water blooms that contaminate water quality (Weirich and Miller, 2014). Additionally, we detected *Histoplasma capsulatum*, a fungus, in 16 samples (12 anal swabs and 4 fecal pellets) collected from six provinces and four bat species. This fungus is known to cause pulmonary and disseminated histoplasmosis in humans, livestock, and rodents, typically via the respiratory tract (Fig. 3, Supplementary Tables S3 and S4) (Kauffman, 2007).

### 3.5. Discovery of new viral lineages or variants

Finally, we tried to identify new viruses within the microbes co-infected with SARSr-CoVs and successfully obtained 25 full-length genomes (Supplementary Table S5). These genomes can be categorized into six distinct viral families: five from *Circoviridae*, six from *Coronaviridae*, one from *Picornaviridae*, two from *Parvoviridae* for mammal-associated viruses; ten from *Dicistroviridae* for insect viruses; and one from *Solemoviridae* for plant viruses based on the phylogenetic topology with the corresponding reference sequences designated by ICTV (Fig. 4, Supplementary Tables S6 and S7).

Regarding the mammal-associated viruses, we obtained five genomic sequences from the *Circoviridae* family, forming a new lineage under the *Cyclovirus* genus, showing pairwise identity to references ranging 8.1%–27% (Fig. 4A, Supplementary Table S8). Each of the five circoviruses possesses a replication-associated gene (*Rep*) of different length (Fig. 5A). The six coronavirus genomes are all belong to the *Alphacoronavirus* genus, exhibiting a typical genomic structure comprising four structural proteins (S, E, M, N) and other ORFs, although ORF7a is absent in three genomes (Figs. 4B and 5B). Among these, RsGD2017-1564-k141\_38765 and RsGZ2015-151271-k141\_98050 are HKU2-related *Alphacoronaviruses*, with 95% and 92% genome identity to Rh-BatCoV\_HKU2 respectively. The other three *Alphacoronaviruses* are closely related to BtRf-AlphaCoV, with genome identities ranging from 67% to 94% (Fig. 4B, Supplementary Table S8). The two parvovirus genomes we obtained belong to *Parvovirinae* and *Densovirinae*, with 61.2% and 64.5% identity to their closest references BtAAV and PamDV, respectively. The genome structures of these two parvoviruses appear to be different, exhibiting separate or encompassing forms of the two ORFs (Figs. 4D and 5D). The only picornavirus we obtained falls under the genus *Ensavirinae*, with pairwise identity to references ranging from 15% to 43% (Fig. 4E, Supplementary Table S8). Similarly, we discovered several potential new lineages of insect-associated viruses within *Dicistroviridae*, closely related to *Triatovirus* and *Cripavirus*, each displaying distinct genome structures and pairwise identity ranging from 33% to 96% (Figs. 4C and 5C, Supplementary Table S8). In the plant-associated *Solemoviridae*, we also obtained a novel genomic sequence (RsHuB2019-190887-k141\_934504) related to the genus *Enamovirus*, with a maximum pairwise identity of less than 17% (Fig. 4F, Supplementary Table S8).

## 4. Discussion

As the natural reservoir host for many pathogenic viruses, bats are believed to potentially carry multiple viruses, increasing viral pathogenicity and the risk of spillover (Federici et al., 2022; He et al., 2022; Wang et al., 2023; Su et al., 2023). However, this area remains understudied.

Here, we focused on bat SARSr-CoV positive samples and conducted metagenomic sequencing to characterize the microenvironment of 76 individual bats. Our data revealed a remarkably high proportion of co-infections with other mammal-associated viruses (at least one additional virus was found in 25% of the samples), along with general co-circulation of viruses or bacteria possibly originating from the environment. We also identified alphacoronaviruses in the samples (13/76), including BtRf-AlphaCoV/YN2012 related coronavirus and BtRf-AlphaCoV/Rf HuB-2013 related coronavirus, which had not been previously reported to co-circulate with SARSr-CoVs. Furthermore, our analysis led to the discovery of 25 new viral genomes, some of which represent novel evolutionary lineages.

Previous bat metagenomics sequencing was typically based on pooled samples, which missed key co-infection information in individual bats (Chen et al., 2023). Through traditional viral surveillance, only bat HKU2-related CoVs have been found co-circulating with SARSr-CoVs (Lau et al., 2007). Our work significantly filled these knowledge gaps in three aspects: (1) We identified common co-infection of other viruses, particularly alphacoronaviruses. Co-infection with different coronaviruses could increase the likelihood of recombination, potentially leading to the emergence of new viral variants. Additionally, coronaviruses may acquire genes from other viral species during co-infection, enhancing viral fitness for host jumping (Huang et al., 2016). (2) We elucidated the most probable sources of the bat virome, including mammal-associated viruses and environmental-associated viruses. Mammal-associated viruses likely circulated within bat populations long-term, while environmental-associated viruses reflected the bat's diet, bacteria, fungi, and even parasites commonly carried by a bat. (3) We primarily identified bacteria and fungi populations that may be pathogenic to humans. The bat microbiome remains understudied. Our data indicates that bats generally harbor bacteria associated with human disease (for example *Yersinia*) and fungi species linked to human respiratory diseases (*Histoplasma*), although further study is necessary to fully assess their potential for causing disease.

## 5. Conclusion

In conclusion, we analyzed the potential co-circulating microorganisms in 76 SARSr-CoV positive samples. Our findings offer new insights into the complex microenvironment of the bat microbiome, facilitating further in-depth analyses of their potential pathogenicity and the likelihood of cross-species transmission.

### Data availability

All of the 25 viral genomes have been uploaded to GenBase (Supplementary Table S5; <https://ngdc.cnpc.ac.cn/genbase/?lang=en>) and Science Data Bank (<https://doi.org/10.57760/sciencedb.08257>).

### Ethics statement

The animal experiments were approved by the Animal Ethics Committee of the Wuhan Institute of Virology with the animal ethics permit (WIVA05201705).

### Author contributions

Hao-Rui Si: conceptualization, project administration, visualization, writing-review & editing, data curation, formal analysis, methodology, software, writing-original draft. Ke Wu: project administration, software, writing-review & editing, formal analysis, methodology, visualization. Jia Su: data curation, formal analysis, methodology, software. Tian-Yi Dong: formal analysis, methodology, validation. Yan Zhu, Bei Li, Ying Chen, Yang Li: resources. Zheng-Li Shi, Peng Zhou: conceptualization, project administration, funding acquisition, writing-review & editing, supervision.



## Conflict of interest

Prof. Zheng-Li Shi is the Editor-in-Chief for *Virologica Sinica* and Prof. Peng Zhou is an editorial board member for *Virologica Sinica*. They were not involved in the editorial review or the decision to publish this article. The authors declare that they have no competing interests.

## Acknowledgements

We appreciate the help from bat sampling work by Prof. Zhang Li-Biao at Guangdong Academy of Sciences. This work was jointly supported by the National Key R&D Program of China (2021YFC2300901 to P.Z.), China Natural Science Foundation for outstanding scholars (82325032 to P.Z.) and the Self-Supporting Program of Guangzhou Laboratory (SRPG22-001 to P.Z.; GZNL2023A01001 to ZLS).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.virs.2024.06.008>.

## References

- Abu-Raddad, L.J., Patnaik, P., Kublin, J.G., 2006. Dual infection with hiv and malaria fuels the spread of both diseases in sub-Saharan Africa. *Science* 314, 1603–1606.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. Blast+: architecture and applications. *BMC Bioinformatics* 10, 421.
- Chen, Y.M., Hu, S.J., Lin, X.D., Tian, J.H., Lv, J.X., Wang, M.R., Luo, X.Q., Pei, Y.Y., Hu, R.X., Song, Z.G., Holmes, E.C., Zhang, Y.Z., 2023. Host traits shape virome composition and virus transmission in wild small mammals. *Cell* 186, 4662–4675 e4612.
- Cui, J., Li, F., Shi, Z.L., 2019. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 17, 181–192.
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of samtools and bcftools. *Gigascience* 10, giab008.
- Falk, B.W., Tsai, J.H., 1998. Biology and molecular biology of viruses in the genus tenuiviruses. *Annu. Rev. Phytopathol.* 36, 139–163.
- Federici, L., Masulli, M., De Laurenzi, V., Allocati, N., 2022. An overview of bats microbiota and its implication in transmissible diseases. *Front. Microbiol.* 13, 1012189.
- Fujisaka, S., Watanabe, Y., Tobe, K., 2023. The gut microbiome: a core regulator of metabolism. *J. Endocrinol.* 256, e220111.
- Gao, Z., Yu, L., Cao, L., Yang, M., Li, Y., Lan, Y., Tang, R., Huang, Y., Luan, G., Liu, Y., Yu, H., Jian, L., Zha, Y., Fan, Z., Bai, Y., Luo, M., He, M., Deng, S., 2023. Analysis of coexisting pathogens in nasopharyngeal swabs from covid-19. *Front. Cell. Infect. Microbiol.* 13, 1140548.
- Ge, X.Y., Li, J.L., Yang, X.L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., Zhang, Y.J., Luo, C.M., Tan, B., Wang, N., Zhu, Y., Cramer, G., Zhang, S.Y., Wang, L.F., Daszak, P., Shi, Z.L., 2013. Isolation and characterization of a bat sars-like coronavirus that uses the ace2 receptor. *Nature* 503, 535–538.
- Gupta, R.K., George, R., Nguyen-Van-Tam, J.S., 2008. Bacterial pneumonia and pandemic influenza planning. *Emerg. Infect. Dis.* 14, 1187–1192.
- Gut, A.M., Vasiljevic, T., Yeager, T., Donkor, O.N., 2018. Salmonella infection - prevention and treatment by antibiotics and probiotic yeasts: a review. *Microbiology (Reading)* 164, 1327–1344.
- Han, Y., Xu, P., Wang, Y., Zhao, W., Zhang, J., Zhang, S., Wang, J., Jin, Q., Wu, Z., 2023. Panoramic analysis of coronaviruses carried by representative bat species in Southern China to better understand the coronavirus sphere. *Nat. Commun.* 14, 5537.
- 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., Ge, X., Wang, L.F., Shi, Z., 2015. Bat origin of human coronaviruses. *Virology* 12, 221.
- 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.
- Huang, C., Liu, W.J., Xu, W., Jin, T., Zhao, Y., Song, J., Shi, Y., Ji, W., Jia, H., Zhou, Y., Wen, H., Zhao, H., Liu, H., Li, H., Wang, Q., Wu, Y., Wang, L., Liu, D., Liu, G., Yu, H., Holmes, E.C., Lu, L., Gao, G.F., 2016. A bat-derived putative cross-family recombinant coronavirus with a reovirus gene. *PLoS Pathog.* 12, e1005883.
- Huang, J., Jiang, E., Yang, D., Wei, J., Zhao, M., Feng, J., Cao, J., 2020. Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions. *Infect. Drug Resist.* 13, 567–576.
- Katoh, K., Standley, D.M., 2013. Mafft multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kauffman, C.A., 2007. Histoplasmosis: a clinical and laboratory update. *Clin. Microbiol. Rev.* 20, 115–132.
- Kechin, A., Boyarskikh, U., Kel, A., Filipenko, M., 2017. Cutprimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput. Biol.* 24, 1138–1143.
- Keusch, G.T., Amuasi, J.H., Anderson, D.E., Daszak, P., Eckerle, I., Field, H., Koopmans, M., Lam, S.K., Das Neves, C.G., Peiris, M., Perlman, S., Wacharapluesadee, S., Yadana, S., Saif, L., 2022. Pandemic origins and a one health approach to preparedness and prevention: solutions based on sars-cov-2 and other rna viruses. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2202871119.
- Kohl, C., Brinkmann, A., Radonic, A., Dabrowski, P.W., Muhldorfer, K., Nitsche, A., Wibbelt, G., Kurth, A., 2021. The virome of German bats: comparing virus discovery approaches. *Sci. Rep.* 11, 7430.
- Kutikhin, A.G., Yuzhalin, A.E., Brusina, E.B., 2014. Mimiviridae, marseilleviridae, and virophages as emerging human pathogens causing healthcare-associated infections. *GMS Hyg. Infect. Control* 9, Doc16.
- 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, D., Liu, C.M., Luo, R., Sadakane, K., Lam, T.W., 2015. Megahit: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph. *Bioinformatics* 31, 1674–1676.
- Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics* 26, 589–595.
- Loving, C.L., Brockmeier, S.L., Vincent, A.L., Palmer, M.V., Sacco, R.E., Nicholson, T.L., 2010. Influenza virus coinfection with bordetella bronchiseptica enhances bacterial colonization and host responses exacerbating pulmonary lesions. *Microb. Pathog.* 49, 237–245.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. Iq-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534.
- Olival, K.J., Hosseini, P.R., Zambrana-Torrel, C., Ross, N., Bogich, T.L., Daszak, P., 2017. Host and viral traits predict zoonotic spillover from mammals. *Nature* 546, 646–650.
- Peck, M.W., 2009. Biology and genomic analysis of clostridium botulinum. *Adv. Microb. Physiol.* 55, 183–265, 320.
- Rajilic-Stojanovic, M., de Vos, W.M., 2014. The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* 38, 996–1047.
- Ramaswamy, V., Cresence, V.M., Rejitha, J.S., Lekshmi, M.U., Dharsana, K.S., Prasad, S.P., Vijila, H.M., 2007. Listeria-review of epidemiology and pathogenesis. *J. Microbiol. Immunol. Infect.* 40, 4–13.
- Rizzatti, G., Lopetuso, L.R., Gibiino, G., Binda, C., Gasbarrini, A., 2017. Proteobacteria: a common factor in human diseases. *Biomed Res. Int.* 2017, 9351507.
- Sahin, E., Akata, I., 2018. Viruses infecting macrofungi. *Virusdisseas* 29, 1–18.
- Su, H., Wang, Y., Han, Y., Jin, Q., Yang, F., Wu, Z., 2023. Discovery and characterization of novel paramyxoviruses from bat samples in China. *Virology* 38, 198–207.
- Temmam, S., Vongphaytho, K., Baquero, E., Munier, S., Bonomi, M., Regnault, B., Douangbongpha, B., Karami, Y., Chretien, D., Sanamkay, D., Xayaphet, V., Paphaphan, P., Lacoste, V., Somlor, S., Lakeomany, K., Phommavanh, N., Perot, P., Dehan, O., Amara, F., Donati, F., Bigot, T., Nilges, M., Rey, F.A., van der Werf, S., Brey, P.T., Eloit, M., 2022. Bat coronaviruses related to sars-cov-2 and infectious for human cells. *Nature* 604, 330–336.
- Wang, J., Pan, Y.F., Yang, L.F., Yang, W.H., Lv, K., Luo, C.M., Wang, J., Kuang, G.P., Wu, W.C., Gou, Q.Y., Xin, G.Y., Li, B., Luo, H.L., Chen, S., Shu, Y.L., Guo, D., Gao, Z.H., Liang, G., Li, J., Chen, Y.Q., Holmes, E.C., Feng, Y., Shi, M., 2023. Individual bat virome analysis reveals co-infection and spillover among bats and virus zoonotic potential. *Nat. Commun.* 14, 4079.
- Weirich, C.A., Miller, T.R., 2014. Freshwater harmful algal blooms: toxins and children's health. *Curr. Probl. Pediatr. Adolesc. Health Care* 44, 2–24.
- Xie, L., Luo, G., Yang, Z., Wu, W.C., Chen, J., Ren, Y., Zeng, Z., Ye, G., Pan, Y., Zhao, W.J., Chen, Y.Q., Hou, W., Sun, Y., Guo, D., Yang, Z., Li, J., Holmes, E.C., Li, Y., Chen, L., Shi, M., 2024. The clinical outcome of covid-19 is strongly associated with microbiome dynamics in the upper respiratory tract. *J. Infect.* 88, 106118.
- Yu, G., Lam, T.T., Zhu, H., Guan, Y., 2018. Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Mol. Biol. Evol.* 35, 3041–3043.
- Yu, X.J., Liang, M.F., Zhang, S.Y., Liu, Y., Li, J.D., Sun, Y.L., Zhang, L., Zhang, Q.F., Popov, V.L., Li, C., Qu, J., Li, Q., Zhang, Y.P., Hai, R., Wu, W., Wang, Q., Zhan, F.X., Wang, X.J., Kan, B., Wang, S.W., Wan, K.L., Jing, H.Q., Lu, J.X., Yin, W.W., Zhou, H., Guan, X.H., Liu, J.F., Bi, Z.Q., Liu, G.H., Ren, J., Wang, H., Zhao, Z., Song, J.D., He, J.R., Wan, T., Zhang, J.S., Fu, X.P., Sun, L.N., Dong, X.P., Feng, Z.J., Yang, W.Z., Hong, T., Zhang, Y., Walker, D.H., Wang, Y., Li, D.X., 2011. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N. Engl. J. Med.* 364, 1523–1532.
- 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. *Virology* 38, 344–350.