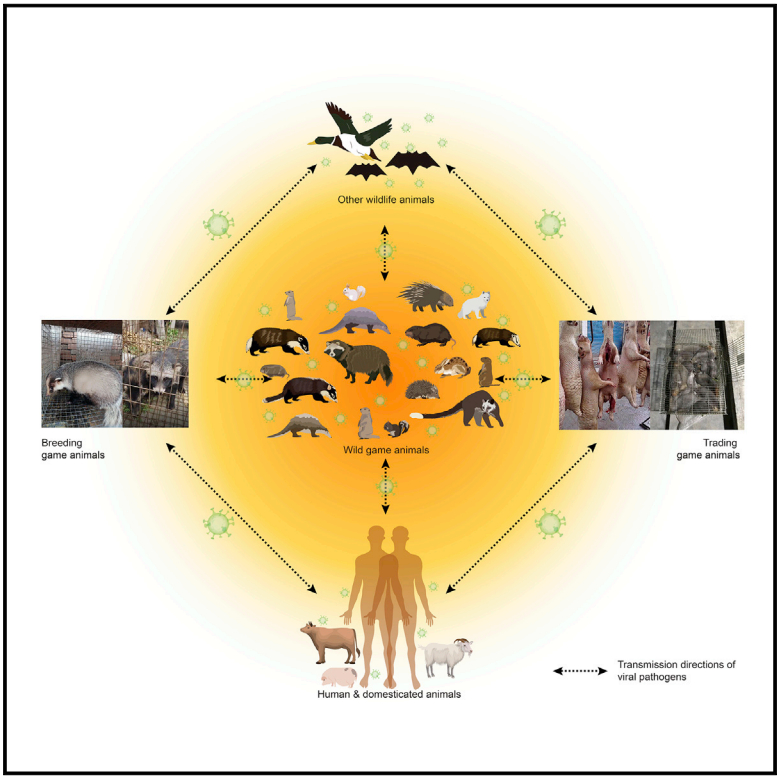
Cell

Virome characterization of game animals in China reveals a spectrum of emerging pathogens

Graphical abstract



Highlights

- 1,941 game animals from five mammalian orders were surveyed for viruses
- 102 mammalian-infecting viruses were discovered, 21 posing a potential risk to humans
- Civets carried a relatively higher number of potentially "high-risk" viruses
- Human-infecting viruses were also identified in game animals

Authors

Wan-Ting He, Xin Hou, Jin Zhao, ..., Edward C. Holmes, Mang Shi, Shuo Su

Correspondence

shuosu@njau.edu.cn (S.S.), edward.holmes@sydney.edu.au (E.C.H.), shim23@mail.sysu.edu.cn (M.S.)

In brief

An analysis of 18 different wildlife species traded or consumed as food across China identifies nearly 100 mammalian-infecting viruses, of which a subset has the potential to jump between species and cause respiratory diseases.

He et al., 2022, Cell *185*, 1117–1129
March 31, 2022
https://doi.org/10.1016/j.cell.2022.02.014

CellPress

Cell

CellPress

Article

Virome characterization of game animals in China reveals a spectrum of emerging pathogens

Wan-Ting He,[1,11] Xin Hou,[2,11] Jin Zhao,[1,11] Jiumeng Sun,[1] Haijian He,[7] Wei Si,[4] Jing Wang,[2] Zhiwen Jiang,[1] Ziqing Yan,[1] Gang Xing,[4] Meng Lu,[1] Marc A. Suchard,[5] Xiang Ji,[6] Wenjie Gong,[3] Biao He,[3] Jun Li,[8] Philippe Lemey,[9] Deyin Guo,[2] Changchun Tu,[3] Edward C. Holmes,[10,12,*] Mang Shi,[2,12,*] and Shuo Su[1,12,13,*]

[1]Jiangsu Engineering Laboratory of Animal Immunology, Institute of Immunology, College of Veterinary Medicine, Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, Nanjing 210095, China

[2]The Centre for Infection and Immunity Studies, School of Medicine, Shenzhen Campus of Sun Yat-sen University, Sun Yat-sen University, Shenzhen 518107, China

[3]Key Laboratory of Jilin Province for Zoonosis Prevention and Control, Institute of Military Veterinary, Academy of Military Medical Sciences, Changchun, Jilin 130062, China

[4]MOA Key Laboratory of Animal Virology, Zhejiang University, Hangzhou 310058, China

[5]Department of Biostatistics, Fielding School of Public Health, and Departments of Biomathematics and Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA

[6]Department of Mathematics, School of Science & Engineering, Tulane University, New Orleans, LA 70118, USA

[7]Agricultural College, Jinhua Polytechnic, Jinhua 320017, China

[8]Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong 999077, China

[9]Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory for Clinical and Epidemiological Virology, KU Leuven, Leuven 3000, Belgium

[10]Sydney Institute for Infectious Diseases, School of Life and Environmental Sciences and School of Medical Sciences, the University of Sydney, Sydney, NSW 2006, Australia

[11]These authors contributed equally

[12]Senior author

[13]Lead contact

*Correspondence: shuosu@njau.edu.cn (S.S.), edward.holmes@sydney.edu.au (E.C.H.), shim23@mail.sysu.edu.cn (M.S.)

https://doi.org/10.1016/j.cell.2022.02.014

SUMMARY

Game animals are wildlife species traded and consumed as food and are potential reservoirs for SARS-CoV and SARS-CoV-2. We performed a meta-transcriptomic analysis of 1,941 game animals, representing 18 species and five mammalian orders, sampled across China. From this, we identified 102 mammalian-infecting viruses, with 65 described for the first time. Twenty-one viruses were considered as potentially high risk to humans and domestic animals. Civets (*Paguma larvata*) carried the highest number of potentially high-risk viruses. We inferred the transmission of bat-associated coronavirus from bats to civets, as well as cross-species jumps of coronaviruses from bats to hedgehogs, from birds to porcupines, and from dogs to raccoon dogs. Of note, we identified avian *Influenza A virus* H9N2 in civets and Asian badgers, with the latter displaying respiratory symptoms, as well as cases of likely human-to-wildlife virus transmission. These data highlight the importance of game animals as potential drivers of disease emergence.

INTRODUCTION

Mammalian game animals are wild or semi-wild animals that are commonly traded and consumed as exotic food in China and other Asian countries (Huong et al., 2020; Philavong et al., 2020; Shivaprakash et al., 2021). They include rodents (such as porcupines, bamboo rats, and marmots), carnivores (such as civets, raccoon dogs, badgers, and foxes), pangolins, hedgehogs, and rabbits. These animals are normally either caught and raised locally or imported illegally from neighboring countries before being transferred to live animal (or "wet") markets for trading (Aditya et al., 2021; Damania, 2005; Holmes et al., 2021; Huong et al., 2020; Shivaprakash et al., 2021; Xiao et al., 2021). In recent decades, there has been a major expansion in commercial wildlife farming operations and their species diversity. The Huanan Seafood Wholesale Market in Wuhan, to which many of the early COVID-19 cases were linked (Sun et al., 2020a; Worobey, 2021), is a notable example of a live animal market. Poor hygiene conditions and close contact between animals and humans, as well as a wide mix of species within live animal markets and the restaurants they serve, make them an ideal breeding ground for emerging infectious diseases.

Unsurprisingly, consuming, capturing, processing, and/or trading game animals has been linked to several infectious

disease outbreaks with grave public health consequences. Early cases of both SARS-CoV and SARS-CoV-2 were identified in animal handlers at animal markets in Guangdong (Skowronski et al., 2005) and Hubei provinces (Sun et al., 2020a), respectively, and close relatives of SARS-CoV and SARS-CoV-2 have been identified in civets (Guan et al., 2003), raccoon dogs (Guan et al., 2003), and pangolins (Lam et al., 2020; Zhang et al., 2020). These are the most popular exotic game animals and subject to frequent trading and human consumption (Chen et al., 2008), although their role as direct hosts for the transmission of these viruses to humans is often uncertain. In an analogous manner, multiple lineages of human immunodeficiency virus (HIV), the causative agent of acquired immune deficiency syndrome (AIDS), likely originated from the hunting or handling of primate carcasses in central and central-west Africa (Aghokeng et al., 2010; Locatelli and Peeters, 2012).

As game animals are frequently associated with important human diseases, it is of obvious importance to identify existing or potential pathogens within these species so that they can be used to trace the origins of specific epidemics and provide a risk assessment of the most likely sources of future outbreaks. Since the first SARS-CoV outbreak in 2002/2003, virus discovery studies have been performed in a variety of game animal species including civets, pangolins, marmots, and badgers. Initially, these studies utilized virus isolation, consensus PCR, and Sanger sequencing (Banks et al., 2002; Dai et al., 2018; Guan et al., 2003; Kent et al., 2018; Techangamsuwan et al., 2015; Wang et al., 2019, 2020), whereas in recent years more attention has been directed toward metagenomic high-throughput sequencing (Ao et al., 2017; Liu et al., 2019; Luo et al., 2018; Bodewes et al., 2014). In addition to SARS-like viruses, these studies have identified several infectious agents that are of direct importance for human infection, such as *Rotavirus A* from civets and raccoon dogs (Abe et al., 2010; Shao et al., 2008) and *Orthohepevirus A* in wild boars (Arnaboldi et al., 2021). Both of these viruses are known to infect humans as well as a wide range of mammalian hosts (Dóró et al., 2015; Nimgaonkar et al., 2018; Pavio et al., 2015). Despite this, there have been few systematic investigations of the virome in game animals, especially in China where their consumption is commonplace. For example, one study identified a virus commonly associated with pneumonia in rodents—*Murine respirovirus*—and in pangolins, suggesting the epizootic potential of viruses in game animals (Liu et al., 2019).

Since the emergence of SARS-CoV in 2002, China has been the focus of widespread viral surveillance in wildlife animals, and a number of SARS-CoV-2-related viruses have been discovered in bats (Zhou et al., 2020, 2021). However, with the exception of pangolins, there has been little investigation of game animals, even though they have close contact with humans and domestic animals and, hence, provide a link to other wildlife species. To help fill this gap, we performed a systematic meta-transcriptomic (i.e., total RNA sequencing) investigation of the viromes of 18 species of game animals representing five mammalian orders collected across China. Many of the species were investigated for the first time within a metagenomic framework. Our goal was to reveal the diversity and abundance of vertebrate-associated viruses in these game animals and assess which species have the greatest potential to carry viruses that could eventually emerge in human populations.

RESULTS

Game animals studied

Between 2017 and 2021, we performed a large-scale survey of viral pathogens in game animals commonly consumed as exotic food in China (Figure 1A). The majority of our animal samples (99.07% [1,923/1,941]) were collected after February 2020 and concurrent with the SARS-CoV-2 pandemic. Only a subset of the pangolin samples (n = 11) and Malayan porcupine samples (n = 7) was collected between 2017 and 2019.

The complete sample collection comprised 1,941 animals from 18 species representing five mammalian orders: Rodentia, Pholidota, Carnivora, Eulipotyphla, and Lagomorpha (Figure 1B). Many of these species were shown on the game animal price list at the Huanan Seafood Wholesale Market in Wuhan, including civets (*P. larvata*) that have been implicated in the emergence of SARS-CoV (Wu et al., 2005). Although most of these animals were maintained in artificial breeding sites that supply animal markets and zoos (Figure 1D, upper panel), some were obtained from their natural habitats. In addition, some animals had obvious signs of infectious diseases and some even died during collection (Figure 1D, middle panel), presenting with symptoms such as paralysis in porcupines, anorexia and convulsion in pangolins, flu-like symptoms (nasal excretions) in Himalayan marmots, and runny nose in raccoon dogs (Figure 1C). Other animals exhibited no overt signs of diseases. Respiratory and fecal samples were collected from these animals across 20 provinces in China (Figure 1D, bottom panel), subsequently organized into 239 pools according to species, location, health, and living conditions (Table S1) for meta-transcriptomic sequencing. This process yielded 3,150.84 billion nucleotide bases of sequence reads for virus discovery and characterization.

Virome characterization

Despite the very large number of viruses discovered, we focused on those associated with vertebrates, comprising: (1) vertebrate-specific viruses that exhibited relatively close phylogenetic relationships to virus families or genera already known to infect vertebrates and (2) vector-borne viruses previously associated with both vertebrates and arthropods. Other viruses likely associated with non-vertebrate hosts derived from animal diet, co-infecting parasites, or endosymbionts were not considered further.

A total of 102 vertebrate-associated viral species from 13 viral families were identified based on meta-transcriptomic sequencing and confirmed by RT-PCR and Sanger sequencing (Figure 2A; Table S2). This comprised 16 species of DNA virus belonging to the genera *Dependoparvovirus*, *Bocaparvovirus*, *Protoparvovirus*, *Amdoparvovirus*, and *Chapparvovirus* from the family *Parvoviridae* and 86 RNA virus species belonging to the genera/families *Picornaviridae*, *Astroviridae*, *Paramyxoviridae*, *Orthomyxoviridae*, *Orthopneumoviruis*, *Flaviviridae*, *Reoviridae*, *Coronaviridae*, *Caliciviridae*, *Tobaniviridae*, *Hepeviridae*, and *Birnaviridae* (Figure 2B). Among these, viruses of the *Picornaviridae*, *Astroviridae*, *and Parvovirdae* were the most commonly detected and showed relatively high abundance and prevalence in many animal species, whereas

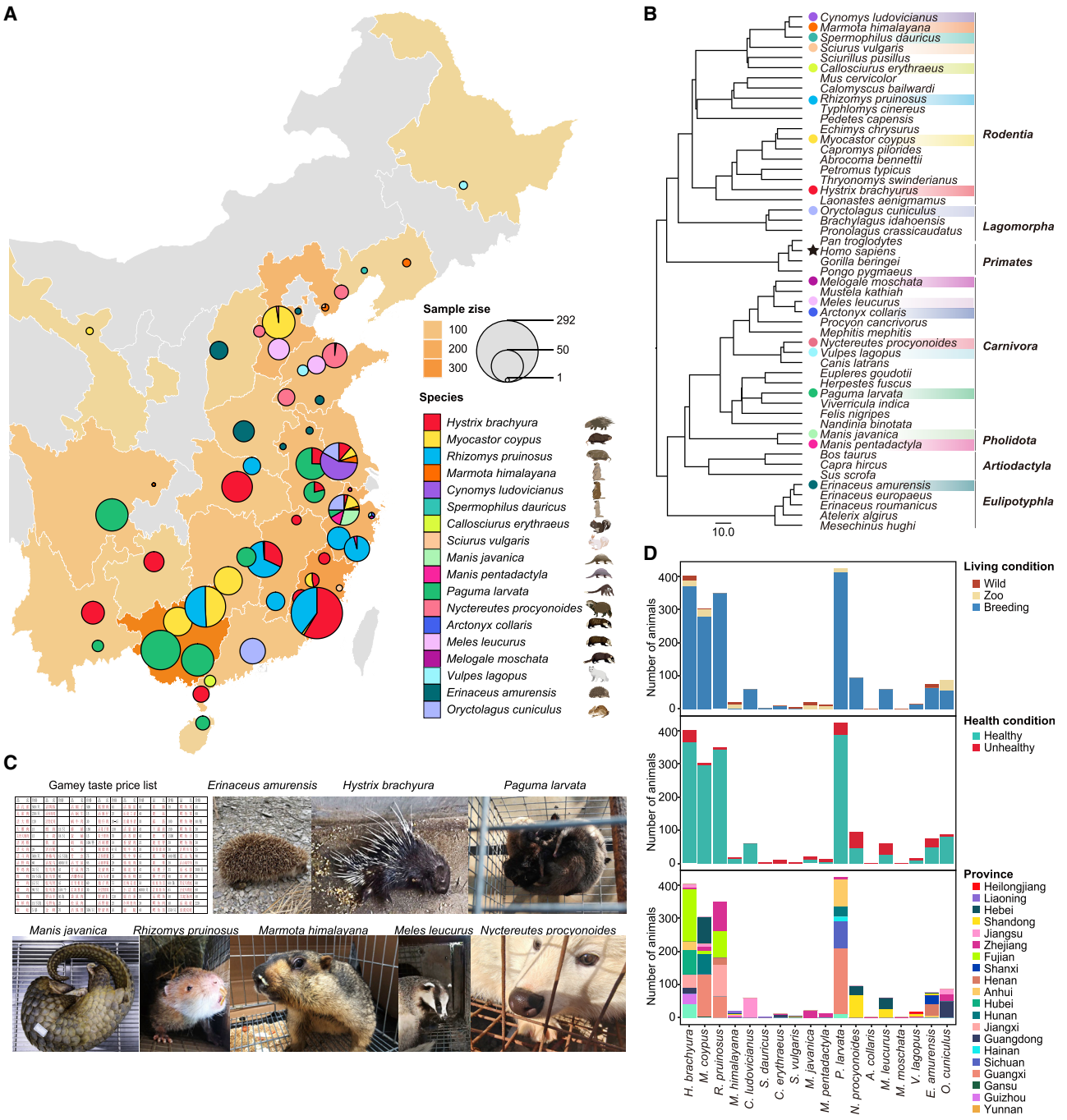


Figure 1. Game animals analyzed in this study

(A) Circles on the map show the geographical locations in China where the game animals were surveyed (n = 3,111 samples from 1,941 individuals) between June 2017 and December 2021. Circle colors indicate the animal species according to the legend, and the Chinese provinces sampled are shown in shades of orange.

(B) Phylogenetic relationships of the game animal hosts surveyed here and related representative mammalian species. Circles denote the species included in this study, and the star highlights the phylogenetic position of humans.

(C) A game animal price list from an animal market in China, and pictures of representative some animals surveyed here.

(D) Distribution of game animal samples by living condition (top panel), health condition (middle panel), and sampling province (bottom panel). See also Table S1.

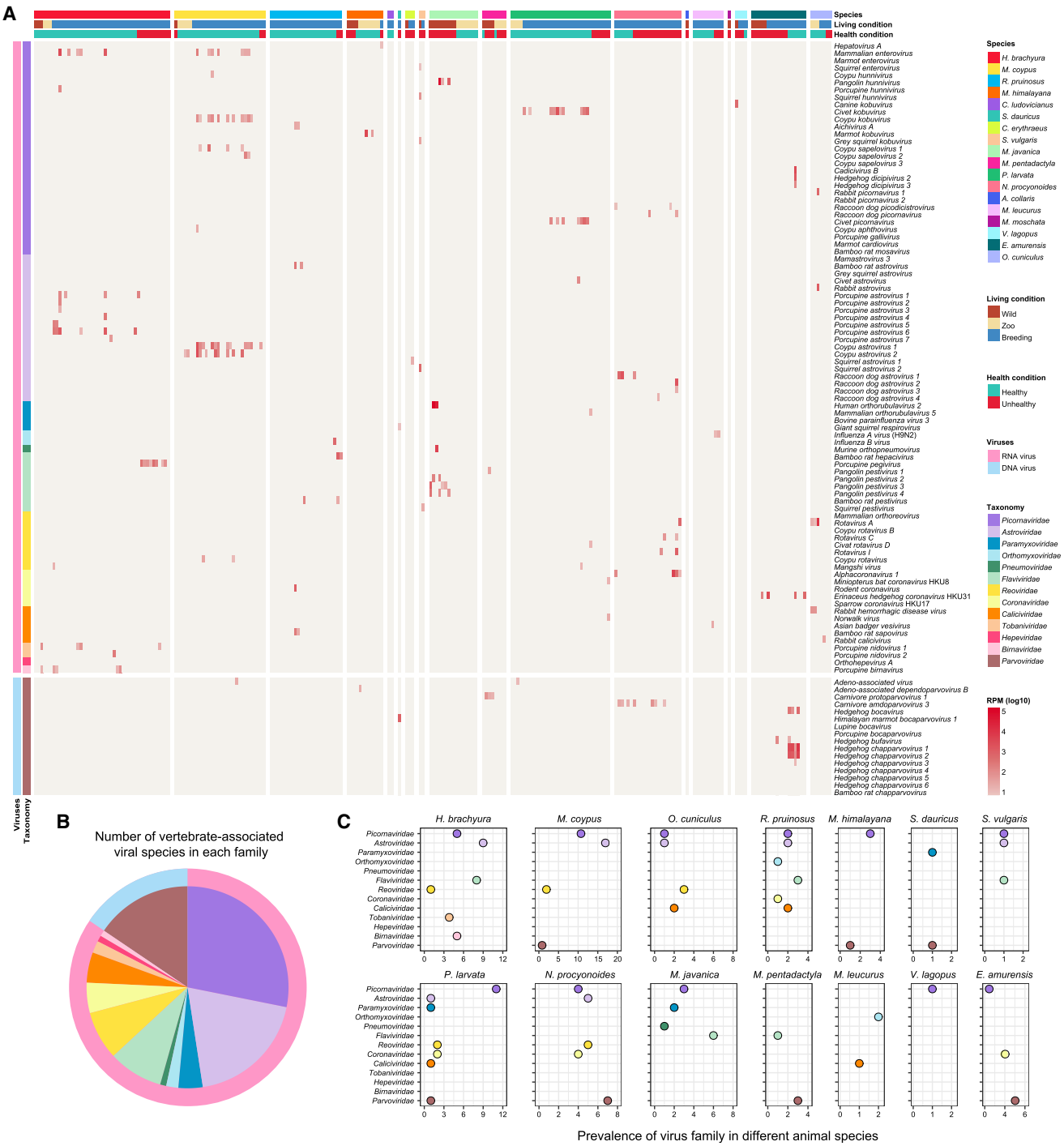


Figure 2. Overview of the vertebrate-associated virome in game animals

(A) Distribution and abundance of vertebrate-associated viruses in game animal species. The relative abundance of viruses in each library was calculated and normalized by the number of mapped reads per million total reads (RPM). To remove potential contamination, only viruses with an abundance RPM > 10 are shown. The living and health condition of the animals in question were recorded at the time of sampling and reflected in the corresponding colors. Viral species from 13 families are shown, with each family indicated by the colors on the heatmap.

(B) Number of vertebrate-associated viral species in each family.

(C) Prevalence (i.e., number of positive libraries) of each viral family in the animal species surveyed. Species with no vertebrate-associated viruses are not shown. See also Table S2.

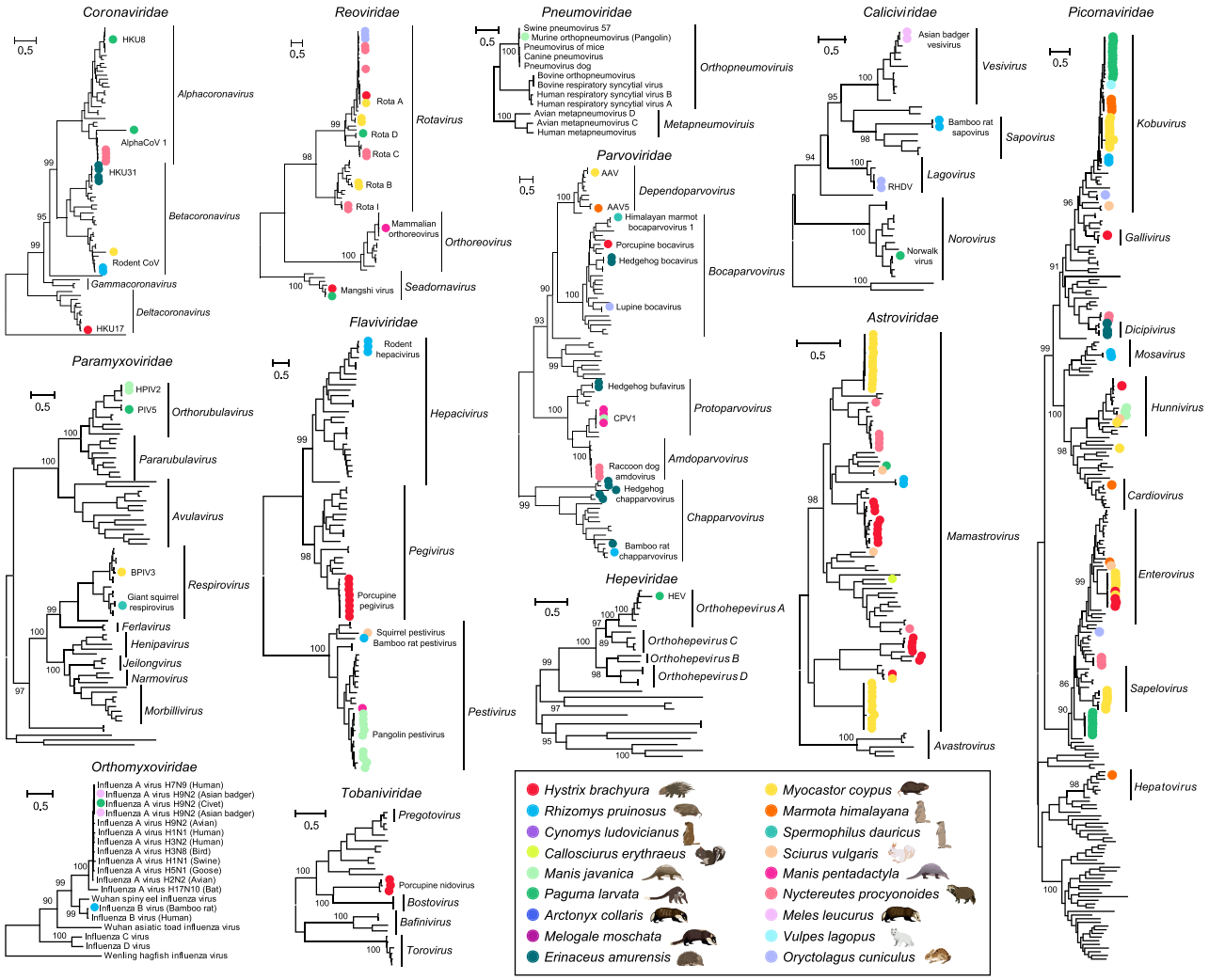


Figure 3. Inter-specific phylogenetic relationships of 12 major vertebrate-associated virus families

Each phylogenetic tree was estimated using a maximum likelihood method based on conserved viral proteins (RNA viruses = RdRp domain; DNA viruses = capsid protein). All trees were midpoint-rooted for clarity only, and the scale bar indicates 0.5 amino acid substitutions per site. Bootstrap values are shown for major nodes. Within each phylogenetic tree, the viruses newly identified here are marked by with solid circles and colored by animal host as shown in the legend at the bottom of the figure.

other viral families were more sporadically detected (Figure 2C). For example, coronaviruses were identified in bamboo rats, civets, raccoon dogs, and hedgehogs; influenza viruses were identified in civets, Asian badgers, and bamboo rats; caliciviruses were identified in bamboo rats, rabbits, civets, and Asian badgers; and the genera *Orthorubulavirus* (*Paramyxoviridae*) and *Pestivirus* (*Flaviviridae*) were mainly identified in pangolins (Figure 2C). No viruses closely related to either SARS-CoV or SARS-CoV-2 (or other sarbecoviruses) were detected in any of the animals examined.

Evolutionary history of vertebrate-associated viruses

Phylogenetic analyses of the viruses identified here revealed that many had close evolutionary relationships (>80% nucleotide sequence identity) to viruses known to cause diseases in other wildlife species, domestic animals, or even humans, thereby greatly expanding their host range (Figures 3, 4, and S1). In particular, viruses associated with human infection, namely *Influenza A virus* (subtype H9N2), *Influenza B virus*, *Norwalk virus*, *Orthohepevirus A*, *Human parainfluenza virus 2* (HPIV2), *Rotavirus A and C*, and *Mammalian orthoreovirus*, were identified in some game animals for the first time.

Evolutionary analyses of the newly identified H9N2 viruses based on the HA and NA segments revealed that the sequences obtained here comprised two clusters, both belonging to the BJ94-like lineage that is the major epizootic strain in China; this lineage is of poultry origin and has been recently associated with a number of human cases (Bi et al., 2020) (Figures 4 and S2). Additional analyses based on surface and internal gene sets revealed that the viruses identified in this study possessed the 155T, 183N, and 226L amino substitutions in the HA1 receptor-binding

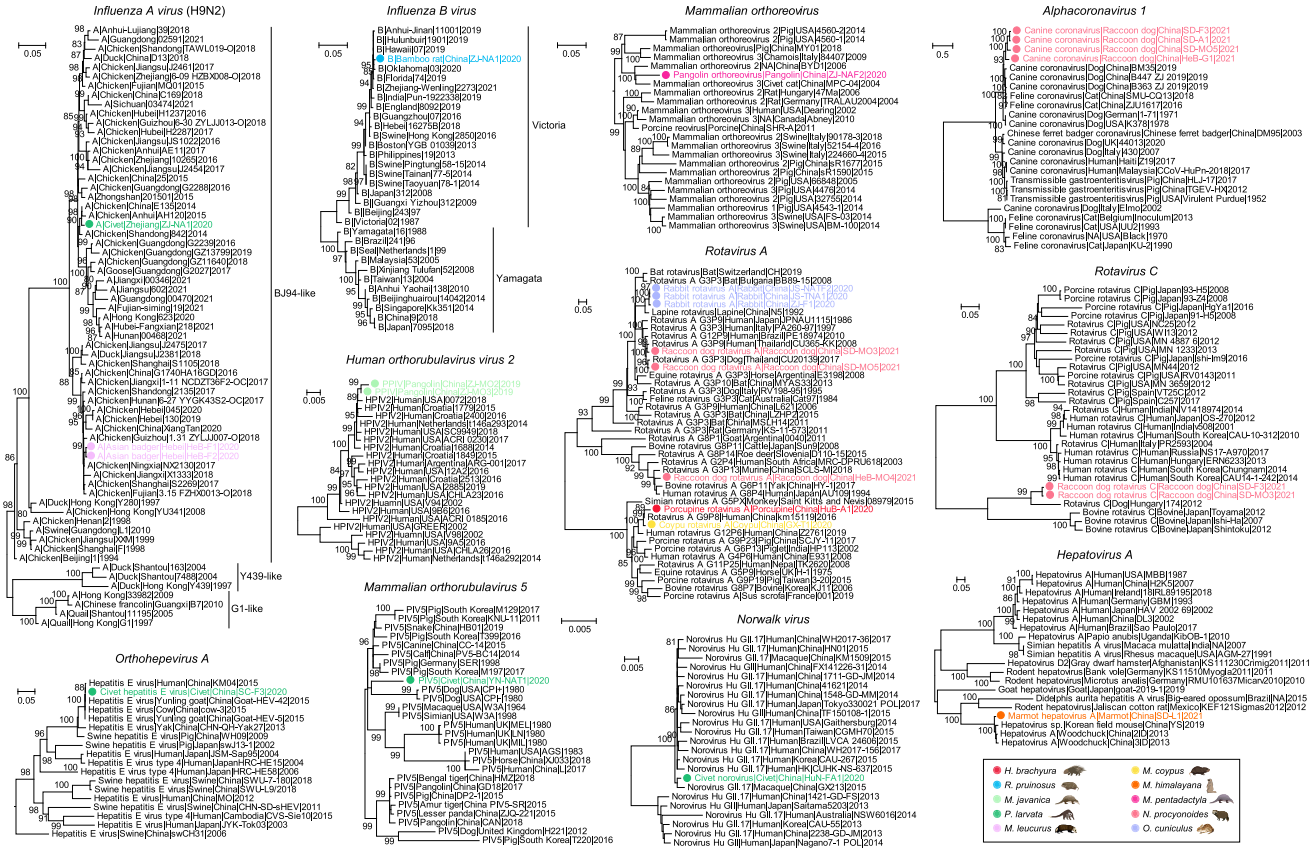


Figure 4. Intra-specific phylogenetic diversity of pathogenic human-infecting viruses identified in game animals

Each phylogenetic tree was estimated using a maximum likelihood method based on conserved viral proteins (influenza viruses = HA gene; AlphaCoV 1 = S1 gene; other RNA viruses = RdRp domain). The trees were midpoint-rooted for clarity only, and the scale bar represents the number of nucleotide substitutions per site. For clarity, only support values >80% are shown. Within each phylogeny, the names in black represented published viral genomes, whereas the viruses newly identified here are marked by with solid circles and colored by animal hosts as shown in the key. Virus names indicate viral type (subtype), host species, sampling location, strain name, and year, from left to right.

domain (RBD) and the I292V substitution (in the civet virus only) in the PB2 protein that have been associated with human infection or adaptation to human cell lines (Gao et al., 2019; Li et al., 2014; Sun et al., 2020b). Similarly, amino acid substitutions, such as 368V in PB1, and 356R, 100A, and 409N in the PA gene, have been associated with the infection of mammalian hosts (Chen et al., 2006; Herfst et al., 2012; Sun et al., 2020b; Xu et al., 2016) (Table S3).

In contrast, HPIV2, *Norwalk virus* and *Influenza B virus*, previously thought to be human-specific, were detected in pangolins, civets, and bamboo rats, respectively, and at moderate or high abundance (29.64–293,198.54 reads per million total reads [RPM]). Furthermore, we identified four canine coronaviruses (CCoV, species alphacoronavirus 1, or AlphaCoV1) in raccoon dogs suffering from diarrhea. Notably, these viruses shared 93.65%–94.27% genome identity with a recombinant CCoV strain—CCoV-HuPn-2018—recently reported to cause diseases in humans (Lednicky et al., 2021a; Vlasova et al., 2022) (Figure 4). The four new raccoon dog CCoV genomes also exhibited a complex recombination history, involving a divergent S2 protein (Figure S3). Specifically, although phylogenetic analysis of a breakpoint-free segment in ORF1AB suggested a relatively close, paraphyletic relationship with the genotype 2 subtype a CCoV from Chinese dogs sampled in 2019 (Figure S3C), the raccoon dog CCoVs form a distinct monophyletic lineage in a breakpoint-free S segment (i.e., S2). These novel raccoon dog CCoV genomes, therefore, represent a previously unsampled lineage of canine coronaviruses.

Our results similarly suggested cross-species transmission among animal viruses (Figure S1). Bat-associated coronavirus HKU8 was identified in a civet (98.66% amino acid sequence identity in the RdRp protein, Figure S4); avian-associated coronavirus HKU17 was identified in porcupines (in both oral and fecal swabs, 94.27% identity); dog-associated canine coronavirus was identified in raccoon dogs (97.05% identity; see above); mouse-associated *Murine orthopneumovirus* was identified in pangolins (98.56% identity); bovine-associated *Bovine respirovirus 3* was identified in a rodent species—coypu (98.91% identity); and multiple *Carnivore protoparvovirus 1* lineages were identified in both Malayan and Chinese pangolins (96.60%–99.89% identity). Interestingly, our data also suggested that *Mangshi virus* (genus *Seadornavirus* and *Reoviridae*), a mosquito-borne virus of unknown vertebrate host but with the capacity to replicate in vertebrate

cell lines (i.e., a vector-borne virus) (Wang et al., 2015), was present in both porcupine and civet samples. The remaining viruses identified, such as *Rabbit hemorrhagic disease virus* (RHDV), *Aichivirus A*, *Rodent coronavirus* (*Embecovirus*), and *Carnivore amdoparvovirus 3*, were more reflective of their known host range.

In addition to the existing viral species, we identified 65 previously undescribed putative viral species of vertebrates, the majority of which belonged to the *Picornaviridae* (n = 21), *Astroviridae* (n = 17), *Flavivirdae* (n = 8), and *Parvoviridae* (n = 10), with others belonging to the *Reoviridae* (*Coypu rotavirus*, *Coypu rotavirus B*, and *Civet rotavirus D*), *Tobaniviridae* (*Porcupine nidovirus*), and *Caliciviridae* (*Asian badger vesivirus* and *Bamboo rat sapovirus*) (Figure 3). Interestingly, some of the newly discovered viruses or virus groups, including *Mammalian enterovirus*, civet and coypu *kobuviruses*, porcupine and coypu astroviruses, *Porcupine pegivirus*, *Pangolin pestivirus*, *Bamboo rat sapovirus*, and *Hedgehog chapparvovirus*, had both high prevalence and abundance in their respective host species (Figure 2) and sometimes formed species-specific clusters (Figure 3). However, their disease manifestation and potential threat to humans is unclear.

The virome of diseased game animals

We also characterized the virome of sick and/or deceased animals, often using mixed tissue samples. For individual animals with a clear record of clinical symptoms, the majority (9/21) were infected with at least one virus species, whereas others (7/21) carried two or more viruses (Figure 5A). *Influenza A virus* H9N2 was detected in both pools samples from Asian badgers presenting with respiratory symptoms; different combinations of *Pangolin hunnivirus*, *Human orthorubulavirus 2*, *Murine orthopneumovirus*, and a diverse group of pangolin pestiviruses were detected in individual samples of Malayan pangolins suffering from gastroenteritis, pneumonia, and multiple organ hemorrhage; *Miniopterus bat coronavirus* HKU8 and *Norwalk virus* were simultaneously detected in a pooled fecal sample obtained from a civet population experiencing diarrheal symptoms; different combinations of enteric viruses, including astroviruses, rotaviruses, and AlphaCoV 1, were detected in raccoon dog populations with diarrhea symptoms; and a novel pegivirus (*Porcupine pegivirus*) was detected in individual samples and eight Malayan porcupines pools, of which four presented with depressive-like behavior, paralysis, and trichomadesis symptoms. Conversely, no viral pathogens were detected in some deceased animals that had no recorded clinical symptoms and, hence, an unknown cause of death, including masked civets (n = 5), hedgehogs (n = 4), raccoon dogs (n = 3), and coypus (n = 2) (Figure 5A). In addition to those in deceased animals, a variety of pathogens were detected in seemingly healthy animals (Figure 5B). For example, *Influenza A virus* H9N2 was identified in a healthy civet, although the same virus may be the cause of the influenza-like symptoms in the Asian badgers studied here. Similarly, *Orthohepevirus A*, *Mangshi virus*, and *Mammalian orthorubulavirus 5* were detected in healthy civets, and a common human pathogen, *Influenza B virus*, was identified in a seemingly healthy bamboo rat population (Figure 5B).

Characterization of viruses at high emergence potential

Among the 102 vertebrate-associated viruses identified here, we characterized the epidemiological patterns of 21 that we considered to pose a greater risk for infection of humans or other animal species. This risk assessment was simply based on perceived zoonotic potential (i.e., likely ability to infect humans) and/or tendency to jump species barriers and infect other animal hosts (Figure 6A). Most of these putative high-risk viruses exhibited a strong geographic structure, clustering according to Chinese province. For example, *Miniopterus bat coronavirus* HKU8 was found in Hunan province, *Bovine respirovirus 3* was found in Hebei province, and *Influenza B virus* was found in Zhejiang province, whereas *Mammalian orthorubulavirus 5* was found in Yunnan province (Figure 6B). In contrast, several viruses had a wider geographical range. For example, *Influenza A virus* H9N2 was identified in Hebei province (northern China) and Zhejiang province (eastern China), *Alphacoronavirus 1* was identified in Shandong (eastern China), Hebei (central China), and Guangxi provinces (southern China), whereas *Rotavirus A* was identified in Shandong (eastern China), Jiangxi (eastern China), Hubei (central China), and Guangxi provinces (southern China), and others (Figure 6B). Several viruses exhibited high prevalence as well as signatures of cross-species transmission, appearing in multiple libraries within the same or different species. For example, *Rotavirus A* was detected in four animal species (Figure 6C), whereas *Mangshi virus* was discovered in both porcupines and civets, although human infection has not yet been identified.

Finally, we compared the likelihood of different game animals carrying pathogens of greater potential zoonotic risk and the presence of these pathogens among different species (Figure 6D). Among the animals studied, civets had the highest number of potentially high-risk viruses (n = 7), followed by porcupines (n = 4), coypus (n = 3), bamboo rats (n = 3), raccoon dogs (n = 3), and Malayan pangolins (n = 3), although sample sizes vary. Hence, there is a risk of cross-species transmission to other groups of game animals, and it is notable that Malayan and Chinese pangolins were not surveyed for viruses until very recently.

DISCUSSION

We performed a large-scale survey of game animals that are commonly consumed for food or medical use in China, among which many species were examined for the first time, and some have been banned by the Chinese Government for trading or artificial breeding since the onset of the COVID-19 pandemic. This analysis identified a wide diversity of both previously described and novel virus species, some of which may pose a direct threat to human health. Notably, we did not identify any SARS-CoV-2-like or SARS-CoV-like sequences in our samples, including in Malayan pangolins from which SARS-CoV-2-like viruses have previously been reported (Lam et al., 2020; Liu et al., 2019). However, although some frequently traded animals were included in this study, such as civets, raccoon dogs, and pangolins, the sample size and diversity of game animals analyzed were relatively small considering the total number consumed in China and Southeast Asia, and we did not analyze internationally smuggled animals. The ongoing surveillance of game animals in wildlife trading hotspots worldwide clearly remains of utmost importance in understanding the origin of SARS-CoV-2.

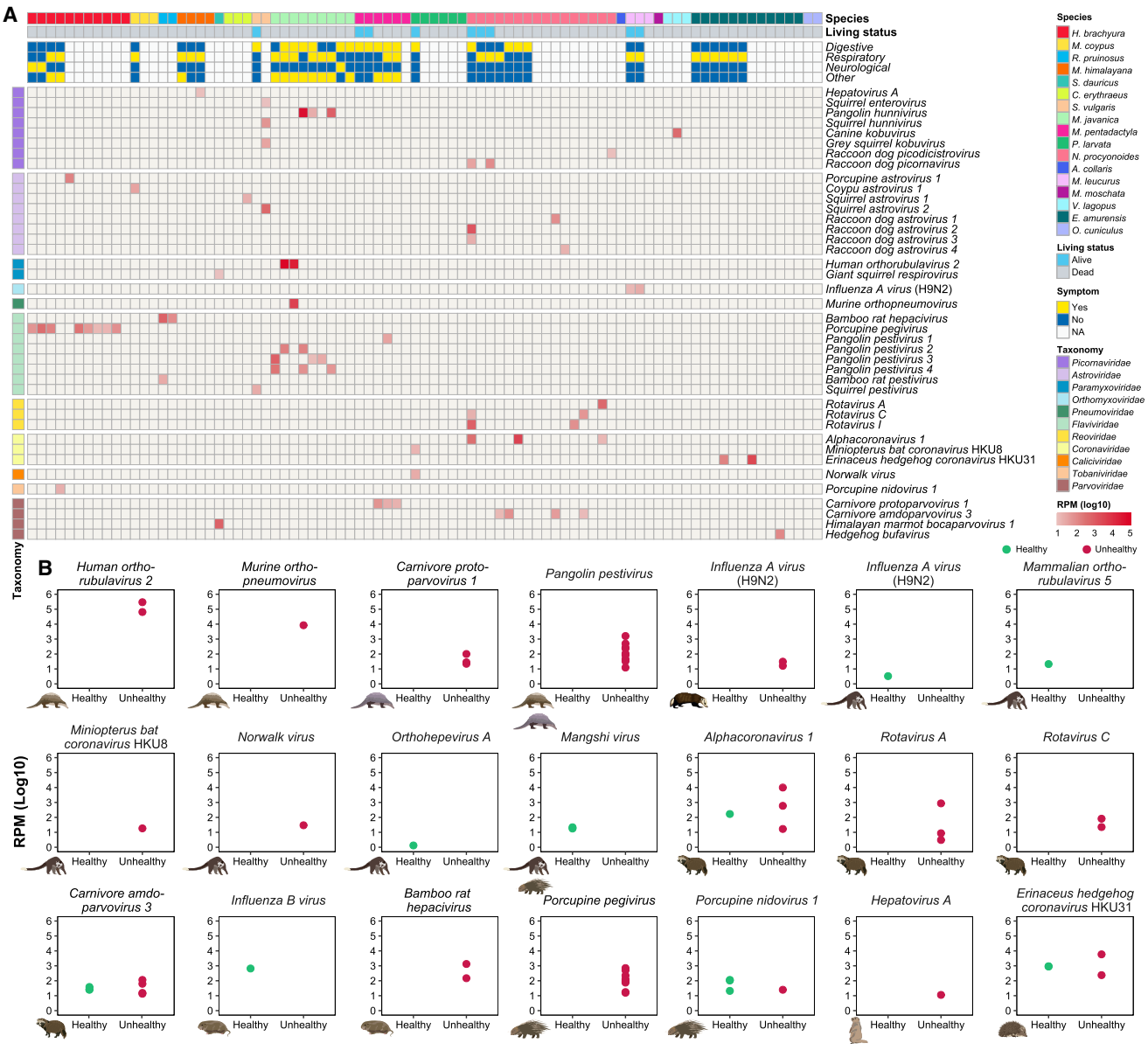


Figure 5. Virome characterization in diseased game animals

(A) Distribution and abundance of vertebrate-associated viruses in diseased game animals. The relative abundance of each virus in each library was calculated and normalized by the number of mapped reads per million total reads (RPM). To reduce the possibility of contamination, only RPM > 10 are shown. Disease symptoms were recorded when sampling and classified into digestive (intestine and stomach related), respiratory (lung related and influenza-like), neurological (paralysis, convulsion, and trauma) and other (spleen- and liver-related, injury, weakness, loss of appetite, etc.) symptoms. Living status (i.e., dead or alive) was recorded when sampling and depicted by different colors.

(B) Abundance comparisons between healthy and unhealthy game animals (see symbols) for 21 pathogenic viruses.

Although no SARS-like viruses were identified, we did detect a potential cross-species transmission event involving a bat-associated coronavirus. Specifically, we identified a bat-associated virus—Miniopterus BatCoV-HKU8—in a civet that represents a different mammalian order (Carnivora). Genomic comparisons revealed a high sequence identity across most of the virus genome (>93.84% amino acid identity) with the exception of the N-terminal proportion of the S1 subunit of the spike protein (36.52%–81.74% amino acid identity) that contained the RBD. In addition, potential bat-to-hedgehog (*Erinaceus amurensis hedgehog coronavirus* HKU31, Ea-HedCoV HKU31) spillovers or host switches were identified. Specifically, viruses identified from the hedgehog clustered within the MERS-like virus group (Figure 3) previously identified in Amur hedgehogs and European hedgehogs (Figure S1) (Corman et al., 2014; Lau et al., 2019), suggesting that hedgehogs are likely important reservoirs for

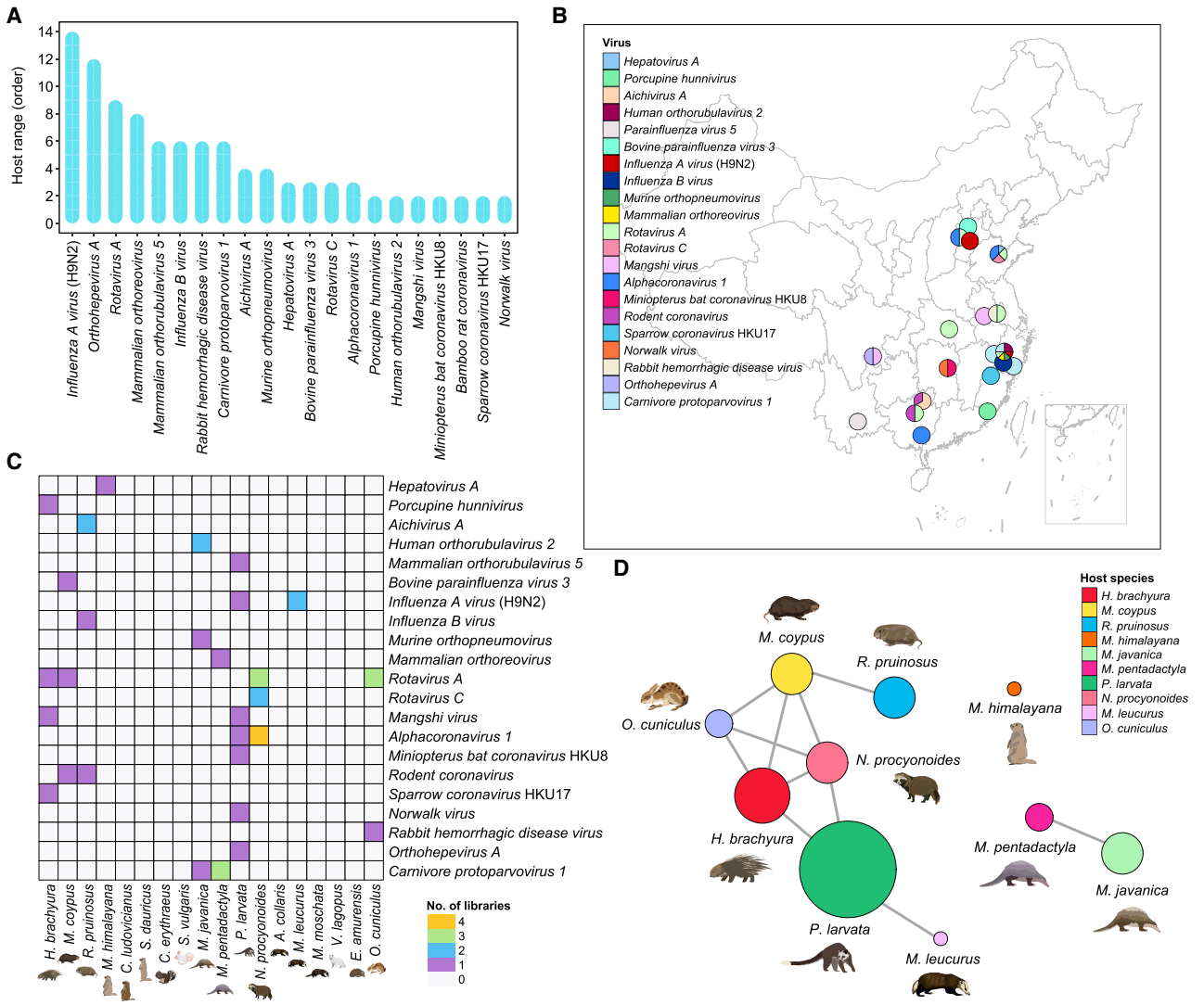


Figure 6. Epidemiological patterns of "high-risk" viruses with the potential to infect humans or other mammals

(A) Host range of potential zoonotic viruses as indicated by the level of the number of mammalian orders infected. For known viruses, a statistical analysis of host range was conducted based on host species information retrieved from NCBI and expanded with the hosts newly identified here. For unknown viruses, a statistical analysis of host range was conducted based only on the host species information obtained in this study.

(B) Geographical locations of potential zoonotic viruses identified in this study. Any viruses shared between more than two hosts were included and regarded as potentially zoonotic viruses.

(C) Distribution of potential zoonotic viruses in each host species.

(D) Associations between potential zoonotic viruses and game species hosts. The size of the colored circles indicates the number of potential zoonotic viral species carried by each host species, whereas the thickness of the line indicates the number of potential zoonotic viral species shared by each host species.

MERS-like viruses. Collectively, we suggest that cross-species transmission from bats to game animals may represent a possible pathway for the virus to move from bats to humans. Indeed, bat-associated alpha- and beta-coronaviruses are the ancestors for those viruses that cause disease outbreaks in humans, including SARS-CoV (Rota et al., 2003; Ksiazek et al., 2003), SARS-CoV-2 (Zhou et al., 2020), MERS-CoV (Assiri et al., 2013; Zaki et al., 2012), HCoV-229E (Fau and Procknow, 1966; Ntumvi et al., 2021), and HCoV-NL63 (van der Hoek et al., 2006), although they commonly pass through the so-called "intermediate" hosts such as civets and raccoon dogs for SARS-CoVs (Guan et al., 2003), camels for MERS-CoVs (Dudas et al., 2018), and alpacas for HCoV-229E (Crossley et al., 2012).

In addition to bat-associated viruses, coronaviruses associated with other mammalian species have also impacted human populations, including canine coronaviruses associated with infection of patients from Malaysia (Vlasova et al., 2022) and Haiti (Lednicky et al., 2021a) who suffered from febrile or lower respiratory tract symptoms, as well as porcine deltacoronaviruses in three Haitian children with acute undifferentiated febrile illness (Lednicky et al.,

2021b). Strikingly, related viruses were also identified in game animals (i.e., civets, raccoon dogs, and porcupines), exhibiting 94% and 85% nucleotide identity with the respective human CCoV strain and porcine deltacoronaviruses, respectively. This further supports the notion that these viruses can acquire expansive host ranges over relatively short evolutionary timescales. Our study, therefore, highlights the fact that coronaviruses are subject to relatively frequent host jumps among mammalian species and, therefore, may pose a direct threat to wildlife animal handlers or during food consumption.

Another virus that might pose an immediate threat to human health is avian influenza virus H9N2 that has increased in prevalence in live poultry markets workers in China (Li et al., 2017). H9N2 has gradually replaced H5N6 and H7N9 as the most prevalent AIV subtype in both chickens and ducks (Bi et al., 2020) and has caused numerous human infections in China. Indeed, in comparison with other currently circulating avian influenza viruses, H9N2 has clear potential to infect humans because almost all subtype H9 AIVs possess human-type receptor-binding ability (Bi et al., 2016; Liu et al., 2014). The viruses discovered in this study belonged to the BJ94-like lineage that harbors a number of genomic features associated with replication in mammalian and human cells, although they seemingly have limited pathogenicity in humans (Gao et al., 2019; Song et al., 2014). It is, therefore, of considerable significance that we detected H9N2 in two game animal species—Asian badgers and civets from northern and eastern China, respectively (Figures 6B and 6C). Furthermore, the Asian badgers displayed obvious respiratory symptoms, suggesting a potential respiratory transmission route that increases the risk of disease transmission to humans.

Our study also revealed that game animals are important hosts for viruses that are related to diseases in human and/or domestic animals. Not only did we sample common zoonotic pathogens such as *Rotavirus A*, *Rotavirus C*, *Alphacoronavirus 1, Orthohepevirus A*, *Mammalian orthoreovirus*, and *Mammalian orthorubulavirus 5* that are known to be harbored by a wide range of mammalian hosts, but we also identified viruses previously thought to be specific to certain host groups. For example, we present the first evidence of human pathogens such as *Influenza B virus*, HPIV2, and *Norwalk virus* in bamboo rats, Malayan pangolins, and civets, indicating human-to-wildlife transmission. Of note, *Influenza B viruses* have previously been identified in other mammalian species including seals and pigs (Osterhaus et al., 2000; Ran et al., 2015). Similarly, Miniopterus BatCoV-HKU8, *Bovine respirovirus 3*, *Carnivore protoparvovirus 1*, and *Influenza A virus* H9N2, which are predominantly found in bats, cattle, carnivores, and avian hosts, respectively, were now detected in a wide range of game animals from different mammalian orders than their perceived reservoir host groups. It is currently unclear how these viruses are maintained in these game animal species. Nevertheless, their capability to carry these viruses, even for a short period of time, may contribute to the virus transmission chains or the emergence of new variants. Indeed, the potential for game animals to act as a separate transmission chain might in part explain the sometimes sudden and unexpected emergence of new viral variants in humans or domestic animals, as observed in several norovirus outbreaks (Villabruna et al., 2019). Our results also show that these viruses were present in seemingly healthy animals and that there is ongoing transmission among different species of game animals (Figure 5). This again highlights the risk of close contact among game animals.

More broadly, our results provide important insights into those game animals and their viruses that might lead to the next pandemic or epizootic. Indeed, the viruses discovered in this study can be divided into two categories: those with restricted host range and those with the capability to infect animals from different mammalian orders and that, therefore, seem capable of overcoming host genetic barriers. The latter category merits most attention as they underpin the great majority of human epidemics and pandemics.

Limitations of the study

Despite its large scale, this study has several limitations. For some of the species considered such as pangolins, badgers, and foxes, the available sample size was limited, largely due to small population sizes and the endangered status of these animals. This prevented us from performing a more systematic comparison of viral diversity and abundance. In addition, meta-transcriptomics does not provide direct evidence of host associations. Although only mammalian-infecting viruses were targeted, it is possible that viruses identified from fecal samples might have originated from the consumption of other vertebrate animals, such that future confirmation of infection is required. Finally, some of the analyses utilized pooled samples, reducing the sensitivity of virus discovery to the extent that low abundance of viruses might go undetected.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Sample collection
  - RNA extraction, library preparation, and sequencing
  - Virus discovery and confirmation
  - Estimation of virus prevalence and abundance
  - Evolutionary analysis
  - Analysis of recombination in raccoon dog coronaviruses
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2022.02.014.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of Outstanding Youth Fund in China (NSFC grant no. 31922081), the

Fundamental Research Funds for the Central Universities (grant No. Y0201900459), and the Bioinformatics Center (BIC) of Nanjing Agricultural University. The genomic epidemiology of novel raccoon dog coronavirus was supported by the National Key Research and Development Program of China (grant No. 2021YFD1801101). M.S. was supported by the National Key Research and Development Program of China, the Shenzhen Science and Technology Program (KQTD20200820145822023), the Guangdong Province "Pearl River Talent Plan" Innovation and Entrepreneurship Team Project (2019ZT08Y464), and the Health and Medical Research Fund (COVID190206). E.C.H. was supported by an Australian Research Council Australia Laureate Fellowship (FL170100022). P.L. acknowledges support from the Research Foundation—Flanders ("Fonds voor Wetenschappelijk Onderzoek—Vlaanderen," G051322N, G0D5117N, and G0B9317N). P.L. and M.A.S. acknowledge funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement no. 725422—ReservoirDOCS), from the Wellcome Trust through project 206298/Z/17/Z (Artic Network) and from the NIH grant R01 AI153044.

We give special thanks to Alibaba Cloud Computing Co. Ltd. for providing the computational resources for rapid data processing. We thank Professor Jiyong Zhou, Zhejiang University, for his guidance and help.

AUTHOR CONTRIBUTIONS

S.S., M.S., and E.C.H. designed and supervised the research. W.-T.H., J.Z., J.S., H.H., W.S., J.W., Z.J., Z.Y., G.X., M.L., and C.T. collected and processed the samples. J.Z., J.S., Z.J., M.L., and Z.Y. performed the Sanger sequencing and molecular detection. X.H. performed the genome assembly and analysis of abundance. X.H., W.-T.H., J.Z., and M.L., performed the sequence annotation. X.H., W.-T.H., P.L., M.S., and S.S. performed the genomic and evolutionary analysis and interpretation. X.J., M.A.S., W.G., B.H., J.L., and D.G. assisted in the data interpretation. S.S., M.S., E.C.H., X.H., and W.-T.H. wrote the paper. All authors reviewed and edited the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 2, 2021
Revised: January 10, 2022
Accepted: February 10, 2022
Published: February 16, 2022

REFERENCES

Abe, M., Yamasaki, A., Ito, N., Mizoguchi, T., Asano, M., Okano, T., and Sugiyama, M. (2010). Molecular characterization of rotaviruses in a Japanese raccoon dog (Nyctereutes procyonoides) and a masked palm civet (Paguma larvata) in Japan. Vet. Microbiol. *146*, 253–259. https://doi.org/10.1016/j.vetmic.2010.05.019.

Aditya, V., Goswami, R., Mendis, A., and Roopa, R. (2021). Scale of the issue: mapping the impact of the COVID-19 lockdown on pangolin trade across India. Biol. Conserv. *257*, 109136. https://doi.org/10.1016/j.biocon.2021.109136.

Aghokeng, A.F., Ayouba, A., Mpoudi-Ngole, E., Loul, S., Liegeois, F., Delaporte, E., and Peeters, M. (2010). Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. Infect. Genet. Evol. *10*, 386–396. https://doi.org/10.1016/j.meegid.2009.04.014.

Ao, Y., Li, X., Li, L., Xie, X., Jin, D., Yu, J., Lu, S., and Duan, Z. (2017). Two novel Bocaparvovirus species identified in wild Himalayan marmots. Sci. China Life Sci. *60*, 1348–1356. https://doi.org/10.1007/s11427-017-9231-4.

Arnaboldi, S., Righi, F., Carta, V., Bonardi, S., Pavoni, E., Bianchi, A., Losio, M.N., and Filipello, V. (2021). Hepatitis E virus (HEV) spread and genetic diversity in game animals in Northern Italy. Food Environ. Virol. *13*, 146–153. https://doi.org/10.1007/s12560-021-09467-z.

Assiri, A., McGeer, A., Perl, T.M., Price, C.S., Al Rabeeah, A.A., Cummings, D.A., Alabdullatif, Z.N., Assad, M., Almulhim, A., Makhdoom, H., et al. (2013). Hospital outbreak of Middle East respiratory syndrome coronavirus. N. Engl. J. Med. *369*, 407–416. https://doi.org/10.1056/NEJMoa1306742.

Banks, M., King, D.P., Daniells, C., Stagg, D.A., and Gavier-Widen, D. (2002). Partial characterization of a novel gammaherpesvirus isolated from a European badger (Meles meles). J. Gen. Virol. *83*, 1325–1330. https://doi.org/10.1099/0022-1317-83-6-1325.

Bi, Y., Chen, Q., Wang, Q., Chen, J., Jin, T., Wong, G., Quan, C., Liu, J., Wu, J., Yin, R., et al. (2016). Genesis, evolution and prevalence of H5N6 avian influenza viruses in China. Cell Host Microbe *20*, 810–821. https://doi.org/10.1016/j.chom.2016.10.022.

Bi, Y., Li, J., Li, S., Fu, G., Jin, T., Zhang, C., Yang, Y., Ma, Z., Tian, W., Li, J., et al. (2020). Dominant subtype switch in avian influenza viruses during 2016–2019 in China. Nat. Commun. *11*, 5909. https://doi.org/10.1038/s41467-020-19671-3.

Bodewes, R., Ruiz-Gonzalez, A., Schapendonk, C.M., van den Brand, J.M., Osterhaus, A.D., and Smits, S.L. (2014). Viral metagenomic analysis of feces of wild small carnivores. Virol. J. *11*, 89. https://doi.org/10.1186/1743-422X-11-89.

Boni, M.F., Posada, D., and Feldman, M.W. (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics *176*, 1035–1047. https://doi.org/10.1534/genetics.106.068874.

Brister, J.R., Ako-Adjei, D., Bao, Y., and Blinkova, O. (2015). NCBI viral genomes resource. Nucleic Acids Res. *43*, D571–D577. https://doi.org/10.1093/nar/gku1207.

Bruen, T.C., Philippe, H., and Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination. Genetics *172*, 2665–2681. https://doi.org/10.1534/genetics.105.048975.

Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods *18*, 366–368. https://doi.org/10.1038/s41592-021-01101-x.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Chen, G.-W., Chang, S.-C., Mok, C.-k., Lo, Y.-L., Kung, Y.-N., Huang, J.-H., Shih, Y.-H., Wang, J.-Y., Chiang, C., Chen, C.-J., and Shih, S.-R. (2006). Genomic signatures of human versus avian influenza A viruses. Emerg. Infect. Dis. *12*, 1353–1360. https://doi.org/10.3201/eid1209.060276.

Chen, J.P., Andersen, D.H., Veron, G., Randi, E., and Zhang, S.Y. (2008). Isolation and characterization of polymorphic microsatellite markers for the masked palm civet (Paguma larvata). Biochem. Genet. *46*, 392–397. https://doi.org/10.1007/s10528-008-9157-7.

Clewley, J.P. (1995). Macintosh sequence analysis software. DNAStar's Lasergene. Mol. Biotechnol. *3*, 221–224. https://doi.org/10.1007/BF02789332.

Corman, V.M., Kallies, R., Philipps, H., Göpner, G., Müller, M.A., Eckerle, I., Brünink, S., Drosten, C., and Drexler, J.F. (2014). Characterization of a novel Betacoronavirus related to middle East respiratory syndrome coronavirus in European hedgehogs. J. Virol. *88*, 717–724. https://doi.org/10.1128/JVI.01600-13.

Crossley, B.M., Mock, R.E., Callison, S.A., and Hietala, S.K. (2012). Identification and characterization of a novel alpaca respiratory coronavirus most closely related to the human coronavirus 229E. Viruses *4*, 3689–3700. https://doi.org/10.3390/v4123689.

Dai, X., Shang, G., Lu, S., Yang, J., and Xu, J. (2018). A new subtype of eastern tick-borne encephalitis virus discovered in Qinghai-Tibet Plateau, China. Emerg. Microbes Infect. *7*, 74. https://doi.org/10.1038/s41426-018-0081-6.

Damania, B.R. (2005). An economic assessment of wildlife farming and conservation. Conserv. Biol. *19*, 1222–1233. https://doi.org/10.1111/j.1523-1739.2005.00149.x.

Dóró, R., Farkas, S.L., Martella, V., and Bányai, K. (2015). Zoonotic transmission of rotavirus: surveillance and control. Expert Rev. Anti Infect. Ther. *13*, 1337–1350. https://doi.org/10.1586/14787210.2015.1089171.

Dudas, G., Carvalho, L.M., Rambaut, A., and Bedford, T. (2018). MERS-CoV spillover at the camel-human interface. Elife *7*, e31257. https://doi.org/10.7554/eLife.31257.

Fau, H.D., and Procknow, J.J. (1966). A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med. *121*, 190–193. https://doi.org/10.3181/00379727-121-30734.

Gao, W., Zu, Z., Liu, J., Song, J., Wang, X., Wang, C., Liu, L., Tong, Q., Wang, M., Sun, H., et al. (2019). Prevailing I292V PB2 mutation in avian influenza H9N2 virus increases viral polymerase function and attenuates IFN-β induction in human cells. J. Gen. Virol. *100*, 1273–1281. https://doi.org/10.1099/jgv.0.001294.

Gibbs, M.J., Armstrong, J.S., and Gibbs, A.J. (2000). Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics *16*, 573–582. https://doi.org/10.1093/bioinformatics/16.7.573.

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science *302*, 276–278. https://doi.org/10.1126/science.1087139.

Guillén-Servent, A., and Francis, C.M. (2006). A new species of bat of the Hipposideros bicolor group (Chiroptera: Hipposideridae) from Central Laos, with evidence of convergent evolution with Sundaic taxa. Acta Chiropterol. *8*, 39–61, 10.3161/1733-5329(2006)8[39:ANSOBO]2.0.CO;2.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. *52*, 696–704. https://doi.org/10.1080/10635150390235520.

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. *22*, 160–174. https://doi.org/10.1007/BF02101694.

Herfst, S., Schrauwen, E.J.A., Linster, M., Chutinimitkul, S., de Wit, E., Munster, V.J., Sorrell, E.M., Bestebroer, T.M., Burke, D.F., Smith, D.J., et al. (2012). Airborne transmission of influenza A/H5N1 virus between ferrets. Science *336*, 1534–1541. https://doi.org/10.1126/science.1213362.

Holmes, E.C., Goldstein, S.A., Rasmussen, A.L., Robertson, D.L., Crits-Christoph, A., Wertheim, J.O., Anthony, S.J., Barclay, W.S., Boni, M.F., Doherty, P.C., et al. (2021). The origins of SARS-CoV-2: a critical review. Cell *184*, 4848–4856. https://doi.org/10.1016/j.cell.2021.08.017.

Huong, N.Q., Nga, N.T.T., Long, N.V., Luu, B.D., Latinne, A., Pruvot, M., Phuong, N.T., Quang, L.T.V., Hung, V.V., Lan, N.T., et al. (2020). Coronavirus testing indicates transmission risk increases along wildlife supply chains for human consumption in Viet Nam, 2013–2014. PLoS One *15*, e0237129. https://doi.org/10.1371/journal.pone.0237129.

Huson, D.H. (1998). SplitsTree: a program for analyzing and visualizing evolutionary data. Bioinformatics *14*, 68–73. https://doi.org/10.1093/bioinformatics/14.1.68.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Kent, A., Ehlers, B., Mendum, T., Newman, C., Macdonald, D.W., Chambers, M., and Buesching, C.D. (2018). Genital tract screening finds widespread infection with mustelid gammaherpesvirus 1 in the European badger (Meles meles). J. Wildl. Dis. *54*, 133–137. https://doi.org/10.7589/2016-12-274.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., et al. (2003). A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. *348*, 1953–1966. https://doi.org/10.1056/NEJMoa030781.

Lam, T.T.-Y., Jia, N., Zhang, Y.-W., Shum, M.H.-H., Jiang, J.-F., Zhu, H.-C., Tong, Y.-G., Shi, Y.-X., Ni, X.-B., Liao, Y.-S., et al. (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature *583*, 282–285. https://doi.org/10.1038/s41586-020-2169-0.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

Lau, S.K.P., Luk, H.K.H., Wong, A.C.P., Fan, R.Y.Y., Lam, C.S.F., Li, K.S.M., Ahmed, S.S., Chow, F.W.N., Cai, J.-P., Zhu, X., et al. (2019). Identification of a novel Betacoronavirus (Merbecovirus) in Amur hedgehogs from China. Viruses *11*, 980. https://doi.org/10.3390/v11110980.

Lednicky, J.A., Tagliamonte, M.S., White, S.K., Blohm, G.M., Alam, M.M., Iovine, N.M., Salemi, M., Mavian, C., and Morris, J.G. (2021a). Isolation of a novel recombinant canine coronavirus from a visitor to Haiti: further evidence of transmission of coronaviruses of zoonotic origin to humans. Clin. Infect. Dis. *28*, ciab924. https://doi.org/10.1093/cid/ciab924.

Lednicky, J.A., Tagliamonte, M.S., White, S.K., Elbadry, M.A., Alam, M.M., Stephenson, C.J., Bonny, T.S., Loeb, J.C., Telisma, T., Chavannes, S., et al. (2021b). Independent infections of porcine Deltacoronavirus among Haitian children. Nature *600*, 133–137. https://doi.org/10.1038/s41586-021-04111-z.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., and 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. https://doi.org/10.1093/bioinformatics/btv033.

Li, X., Shi, J., Guo, J., Deng, G., Zhang, Q., Wang, J., He, X., Wang, K., Chen, J., Li, Y., et al. (2014). Genetics, receptor binding property, and transmissibility in mammals of naturally isolated H9N2 avian influenza viruses. PLoS Pathog. *10*, e1004508. https://doi.org/10.1371/journal.ppat.1004508.

Li, X., Tian, B., Jianfang, Z., Yongkun, C., Xiaodan, L., Wenfei, Z., Yan, L., Jing, T., Junfeng, G., Tao, C., et al. (2017). A comprehensive retrospective study of the seroprevalence of H9N2 avian influenza viruses in occupationally exposed populations in China. PLoS One *12*, e0178328. https://doi.org/10.1371/journal.pone.0178328.

Liu, D., Shi, W., and Gao, G.F. (2014). Poultry carrying H9N2 act as incubators for novel human avian influenza viruses. Lancet *383*, 869. https://doi.org/10.1016/S0140-6736(14)60386-X.

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. https://doi.org/10.3390/v11110979.

Locatelli, S., and Peeters, M. (2012). Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. AIDS *26*, 659–673. https://doi.org/10.1097/QAD.0b013e328350fb68.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., and Ray, S.C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. *73*, 152–160. https://doi.org/10.1128/JVI.73.1.152-160.1999.

Luo, X.-L., Lu, S., Jin, D., Yang, J., Wu, S.-S., and Xu, J. (2018). Marmota himalayana in the Qinghai-Tibetan Plateau as a special host for bi-segmented and unsegmented picobirnaviruses. Emerg. Microbes Infect. *7*, 20. https://doi.org/10.1038/s41426-018-0020-6.

Martin, D., and Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. Bioinformatics *16*, 562–563. https://doi.org/10.1093/bioinformatics/16.6.562.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. *1*, vev003. https://doi.org/10.1093/ve/vev003.

Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. *30*, 1188–1195. https://doi.org/10.1093/molbev/mst024.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient

methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Nimgaonkar, I., Ding, Q., Schwartz, R.E., and Ploss, A. (2018). Hepatitis E virus: advances and challenges. Nat. Rev. Gastroenterol. Hepatol. *15*, 96–110. https://doi.org/10.1038/nrgastro.2017.150.

Ntumvi, N.F., Ndze, V.N., Gillis, A., Doux Diffo, J.L., Tamoufe, U., Takuo, J.-M., Mouiche, M.M.M., Nwobegahay, J., LeBreton, M., Rimoin, A.W., et al. (2021). Wildlife in Cameroon harbor diverse coronaviruses including many isolates closely related to human coronavirus 229E. Preprint at bioRxiv. https://doi.org/10.1101/2021.09.03.458874.

Osterhaus, A.D., Rimmelzwaan, G.F., Martina, B.E., Bestebroer, T.M., and Fouchier, R.A. (2000). Influenza B virus in seals. Science *288*, 1051–1053. https://doi.org/10.1126/science.288.5468.1051.

Pavio, N., Meng, X.J., and Doceul, V. (2015). Zoonotic origin of hepatitis. E. Curr. Opin. Virol. *10*, 34–41. https://doi.org/10.1016/j.coviro.2014.12.006.

Philavong, C., Pruvot, M., Reinharz, D., Mayxay, M., Khammavong, K., Milavong, P., Rattanavong, S., Horwood, P.F., Dussart, P., Douangngeun, B., et al. (2020). Perception of health risks in Lao market vendors. Zoonoses Public Health *67*, 796–804. https://doi.org/10.1111/zph.12759.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596. https://doi.org/10.1093/nar/gks1219.

Ran, Z., Shen, H., Lang, Y., Kolb, E.A., Turan, N., Zhu, L., Ma, J., Bawa, B., Liu, Q., Liu, H., et al. (2015). Domestic pigs are susceptible to infection with influenza B viruses. J. Virol. *89*, 4818–4826. https://doi.org/10.1128/JVI.00059-15.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Peñaranda, S., Bankamp, B., Maher, K., Chen, M.-H., et al. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science *300*, 1394–1399. https://doi.org/10.1126/science.1085952.

Salminen, M., Carr, J.K., Burke, D.S., and McCutchan, F.E. (1995). Identification of breakpoints in intergenotype recombinants of HIV type 1 by bootscanning. AIDS Res. Hum. Retrovir. *11*, 1423–1425. https://doi.org/10.1089/aid.1995.11.1423.

Sawyer, S. (1989). Statistical tests for detecting gene conversion. Mol. Biol. Evol. *6*, 526–538. https://doi.org/10.1093/oxfordjournals.molbev.a040567.

Sayers, E.W., Beck, J., Bolton, E.E., Bourexis, D., Brister, J.R., Canese, K., Comeau, D.C., Funk, K., Kim, S., Klimke, W., et al. (2021). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. *49*, D10–D17. https://doi.org/10.1093/nar/gkaa892.

Shao, Y.-H., Han, Z.-X., Chen, L.-F., Kong, X.-G., and Liu, S.-W. (2008). Isolation and identification of A reovirus from masked civet cats (Paguma larvata). Bing Du Xue Bao *24*, 376–382.

Shivaprakash, K.N., Sen, S., Paul, S., Kiesecker, J.M., and Bawa, K.S. (2021). Mammals, wildlife trade, and the next global pandemic. Curr. Biol. *31*, 3671–3677.e3. https://doi.org/10.1016/j.cub.2021.06.006.

Skowronski, D.M., Astell, C., Brunham, R.C., Low, D.E., Petric, M., Roper, R.L., Talbot, P.J., Tam, T., and Babiuk, L. (2005). Severe acute respiratory syndrome (SARS): a year in review. Annu. Rev. Med. *56*, 357–381. https://doi.org/10.1146/annurev.med.56.091103.134135.

Smith, J.M. (1992). Analyzing the mosaic structure of genes. J. Mol. Evol. *34*, 126–129. https://doi.org/10.1007/BF00182389.

Song, W., Wang, P., Mok, B.W.-Y., Lau, S.-Y., Huang, X., Wu, W.-L., Zheng, M., Wen, X., Yang, S., Chen, Y., et al. (2014). The K526R substitution in viral protein PB2 enhances the effects of E627K on influenza virus replication. Nat. Commun. *5*, 5509. https://doi.org/10.1038/ncomms6509.

Sun, J., He, W.-T., Wang, L., Lai, A., Ji, X., Zhai, X., Li, G., Suchard, M.A., Tian, J., Zhou, J., et al. (2020a). COVID-19: epidemiology, evolution, and cross-disciplinary perspectives. Trends Mol. Med. *26*, 483–495. https://doi.org/10.1016/j.molmed.2020.02.008.

Sun, X., Belser, J.A., and Maines, T.R. (2020b). Adaptation of H9N2 influenza viruses to mammalian hosts: a review of molecular markers. Viruses *12*, 541. https://doi.org/10.3390/v12050541.

Tavare, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences *17*, 57–86.

Techangamsuwan, S., Banlunara, W., Radtanakatikanon, A., Sommanustweechai, A., Siriaroonrat, B., Lombardini, E.D., and Rungsipipat, A. (2015). Pathologic and molecular virologic characterization of a canine distemper outbreak in farmed civets. Vet. Pathol. *52*, 724–731. https://doi.org/10.1177/0300985814551580.

van der Hoek, L., Pyrc, K., and Berkhout, B. (2006). Human coronavirus NL63, a new respiratory virus. FEMS Microbiol. Rev. *30*, 760–773. https://doi.org/10.1111/j.1574-6976.2006.00032.x.

Villabruna, N., Koopmans, M.P.G., and de Graaf, M. (2019). Animals as reservoir for human Norovirus. Viruses *11*, 478. https://doi.org/10.3390/v11050478.

Vlasova, A.N., Diaz, A., Damtie, D., Xiu, L., Toh, T.-H., Lee, J.S.-Y., Saif, L.J., and Gray, G.C. (2022). Novel canine coronavirus isolated from a hospitalized pneumonia patient, East Malaysia. Clin. Infect. Dis. *74*, 446–454. https://doi.org/10.1093/cid/ciab456.

Wang, J., Li, H., He, Y., Zhou, Y., Meng, J., Zhu, W., Chen, H., Liao, D., and Man, Y. (2015). Isolation and genetic characterization of Mangshi virus: a newly discovered Seadornavirus of the Reoviridae family found in Yunnan Province, China. PLoS One *10*, e0143601. https://doi.org/10.1371/journal.pone.0143601.

Wang, X., Chen, W., Xiang, R., Li, L., Chen, J., Zhong, R., Xiang, H., and Chen, J. (2019). Complete genome sequence of parainfluenza virus 5 (PIV5) from a Sunda pangolin (Manis javanica) in China. J. Wildl. Dis. *55*, 947–950.

Wang, S.-L., Tu, Y.-C., Lee, M.-S., Wu, L.-H., Chen, T.-Y., Wu, C.-H., Tsao, E.H.-S., Chin, S.-C., and Li, W.-T. (2020). Fatal canine parvovirus-2 (CPV-2) infection in a rescued free-ranging Taiwanese pangolin (Manis pentadactyla pentadactyla). Transbound. Emerg. Dis. *67*, 1074–1081. https://doi.org/10.1111/tbed.13469.

Worobey, M. (2021). Dissecting the early COVID-19 cases in Wuhan. Science *374*, 1202–1204. https://doi.org/10.1126/science.abm4454.

Wu, D., Tu, C., Xin, C., Xuan, H., Meng, Q., Liu, Y., Yu, Y., Guan, Y., Jiang, Y., Yin, X., et al. (2005). Civets are equally susceptible to experimental infection by two different severe acute respiratory syndrome coronavirus isolates. J. Virol. *79*, 2620–2625. https://doi.org/10.1128/JVI.79.4.2620-2625.2005.

Xiao, X., Newman, C., Buesching, C.D., Macdonald, D.W., and Zhou, Z.M. (2021). Animal sales from Wuhan wet markets immediately prior to the COVID-19 pandemic. Sci. Rep. *11*, 11898. https://doi.org/10.1038/s41598-021-91470-2.

Xu, G., Zhang, X., Gao, W., Wang, C., Wang, J., Sun, H., Sun, Y., Guo, L., Zhang, R., Chang, K.-C., et al. (2016). Prevailing PA mutation K356R in avian influenza H9N2 virus increases mammalian replication and pathogenicity. J. Virol. *90*, 8105–8114. https://doi.org/10.1128/JVI.00883-16.

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. *39*, 306–314. https://doi.org/10.1007/BF00160154.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., and Fouchier, R.A. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med. *367*, 1814–1820. https://doi.org/10.1056/NEJMoa1211721.

Zhang, T., Wu, Q., and Zhang, Z. (2020). Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. Curr. Biol. *30*, 1346–1351.e2. https://doi.org/10.1016/j.cub.2020.03.022.

Zhou, H., Ji, J., Chen, X., Bi, Y., Li, J., Wang, Q., Hu, T., Song, H., Zhao, R., Chen, Y., et al. (2021). Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. Cell *184*, 4380–4391.e14. https://doi.org/10.1016/j.cell.2021.06.008.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature *579*, 270–273. https://doi.org/10.1038/s41586-020-2012-7.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, recombinant proteins, and critical commercial assays		
TRIzol	Thermofisher	Cat. #15596026
TruSeqTM Stranded Total RNA Sample Preparation Kit	Illumina	Cat. #20020597
DNase I	TAKARA	Cat. #2270B
Ribo-ZeroTM rRNA Removal Kits	Illumina	Cat. #20037135
Biological samples		
Samples are described in Table S1	This paper	N/A
Deposited data		
Raw sequencing reads after QC	This paper	NCBI-SRA BioProject: PRJNA793740 and PRJNA795267
Data for all viral genomes	This paper	NCBI-Genbank: OM132156-OM132190, OM450919-OM451215, OM451218-OM451219, OM480510-480545, OM864518-OM864519
Data for all phylogenetic tree	This paper	https://doi.org/10.6084/m9.figshare.18120236
Software and algorithms		
bbmap v38.62	https://sourceforge.net/projects/bbmap/	https://sourceforge.net/projects/bbmap/
MEGAHIT v1.2.8	Li et al., 2015	https://github.com/voutcn/megahit
Diamond v0.9.25.26	Buchfink et al., 2021	https://github.com/bbuchfink/diamond
SeqMan v7.0	Clewley, 1995	https://www.dnastar.com/software/
Geneious v2019.2.1	The Biomatters development team	https://www.geneious.com/
Bowtie2 v2.3.5.1	Langmead and Salzberg, 2012	http://bowtiebio.sourceforge.net/bowtie2
TransDecoder v5.5.0	http://transdecoder.github.io	http://transdecoder.github.io
MAFFT v7.475	Katoh and Standley, 2013	https://mafft.cbrc.jp/alignment/software/
TrimAL v1.2rev59	Capella-Gutiérrez et al., 2009	https://github.com/scapella/trimal
PhyML v3.1	Guindon and Gascuel, 2003	http://www.atgc-montpellier.fr/phyml/
Simplot v3.5.1	Lole et al., 1999	https://www.mybiosoftware.com/simplot-3-5-1-sequence-similarityplotting.html
SplitsTree v4.15.1	Huson, 1998	https://software-ab.informatik.uni-tuebingen.de/download/splitstree4/welcome.html
RDP4 v4.97	Martin et al., 2015	http://web.cbio.uct.ac.za/~darren/rdp.html
IQ-Tree v2.0.3	Minh et al., 2020	http://www.iqtree.org/
Other		
Sequencing systems	Illumina	Novaseq 6000
NCBI GenBank database	Sayers et al., 2021	https://www.ncbi.nlm.nih.gov/genbank/
RefSeq (viral genomes)	Brister et al., 2015	https://www.ncbi.nlm.nih.gov/refseq/
SILVA database	Quast et al., 2013	www.arb-silva.de

RESOURCE AVAILABILITY

Lead contact

Requests for resources, reagents, and further information should be forwarded to and will be fulfilled by the lead contact, Shuo Su (shuosu@njau.edu.cn).

Materials availability

This study did not generate new unique reagents. The materials used and generated in this study are available from the lead contact upon reasonable request with a completed Materials Transfer Agreement.

Data and code availability

The sequence reads after QC generated from the 239 libraries generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) database under the BioProject accession numbers PRJNA793740 and PRJNA795267. The genome sequences of all the viruses generated in this study have also been deposited in GenBank and assigned accession numbers OM132156-OM132190, OM450919-OM451215, OM451218-OM451219, OM480510-OM480545, and OM864518-OM864519. All sequence alignments (fasta format), and phylogenetic trees (newick format) are available at https://doi.org/10.6084/m9.figshare.18120236.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All the sources of bioinformatic data used in the analysis are provided in the key resources table.

METHOD DETAILS

Sample collection

Sampling of game animals was performed between June 2017 and December 2021 at locations representing the natural habitat of animals, artificial breeding sites and zoos across 20 provinces in China (Table S1). Accordingly, a total of 3111 samples from 1941 animal individuals and 18 mammalian species (Xiao et al., 2021) were obtained. These comprised: *Hystrix brachyura* (Malayan porcupine, *n*=402), *Myocastor coypus* (coypu, *n*=303), *Rhizomys pruinosus* (hoary bamboo rat, *n*=349), *Marmota himalayana* (Himalayan marmot, *n*=19), *Cynomys ludovicianus* (prairie dog, *n*=60), *Spermophilus dauricus* (daurian ground squirrel, *n*=3), *Callosciurus erythraeus* (Pallas's squirrel, *n*=11), *Sciurus vulgaris* (Eurasian red squirrel, *n*=4) from the order Rodentia; *Manis javanica* (Malayan pangolin, *n*=21), *Manis pentadactyla* (Chinese pangolin, *n*=12) from the order Pholidota; *Paguma larvata* (civet, *n*=423), *Nyctereutes procyonoides* (raccoon dog, *n*=95), *Arctonyx collaris* (hog badger, *n*=1), *Meles leucurus* (Asian badger, *n*=60), *Melogale moschata* (Chinese ferret-badger, *n*=1), *Vulpes lagopus* (Arctic fox, *n*=16) from the order Carnivora; *Erinaceus amurensis* (Amur hedgehog, *n*=75) from the order Eulipotyphla; and *Oryctolagus cuniculus* (rabbit, *n*=86) from the order Lagomorpha (Figures 1A and 1B). Pharyngeal, anal, nasal and fecal swabs were obtained from living animals, whereas tissue samples (including heart, liver, spleen, lung, kidney, intestine, stomach, and brain) were obtained from deceased animals. All animal specimens were kept in dry ice before transfer to a -80°C freezer for storage.

Host species information was initially identified by experienced field biologists upon capture based on morphological characteristics, then confirmed by Sanger sequencing of the cytochrome b (*CytB*) gene (Guillén-Servent and Francis, 2006). The procedures for sampling and sample processing were approved by the ethics committee of Institute of Military Veterinary Medicine (No. IACUC of AMMS-11-2020-012).

RNA extraction, library preparation, and sequencing

Samples were pooled according to the species, location, health condition, feeding (living) condition, for subsequent RNA extraction and library construction. Detailed information of each pool is shown in Table S1. Total RNA from each pool was extracted using TRIzol® Reagent according to the manufacturer's instructions (Invitrogen). Genomic DNA was subsequently removed using DNase I (Takara). Meta-transcriptome library preparation of each pool was carried out using the TruSeq™ Stranded Total RNA Sample Preparation Kit from Illumina (San Diego, CA) after removal of host ribosomal RNA using Illumina Ribo-Zero™ rRNA Removal Kits (San Diego, CA). Paired-end (150 bp) sequencing of each RNA library was performed using the Illumina Novaseq 6000 platform.

Virus discovery and confirmation

For each library, sequencing reads were first quality controlled using bbduk.sh (https://sourceforge.net/projects/bbmap/; parameters, maq=10 qtrim=r trimq=10 ftl=1 minlen=90). The remaining reads were assembled *de novo* using MEGAHIT (Li et al., 2015) (version 1.2.8) deploying default parameters. The assembled contigs were compared against the NCBI non-redundant protein database (nr) using Diamond blastx (version 0.9.25.26) (Buchfink et al., 2021). The *E*-value cut-off was set at 1*E*-3 to maintain high sensitivity at a low false-positive rate. Taxonomic lineage information was obtained for the top blast hit of each contig, and those classified under kingdom "Viruses" were identified as probable virus hits. Contigs with unassembled overlaps were merged using the SeqMan program implemented in the Lasergene software package version 7.1 (DNAstar) (Clewley, 1995). The final virus genomes were

verified by mapping reads to the corresponding contigs and inspecting the mapping results using Geneious (Kearse et al., 2012). The presence of viruses was further verified using RT-PCR and nested RT-PCR on the original RNA samples with primer sets designed based on the viral genome sequences. The PCR products were then sequenced and compared with the original template. For incomplete viral genomes, the internal gaps were filled by RT-PCR and Sanger sequencing, whereas the genome termini were determined using RACE analyses.

Among the virus contigs described here, those likely associated with vertebrates, specifically vertebrate-specific viruses and vector-borne vertebrate viruses, were preliminarily identified based on the taxonomic lineage information of the blastx results and confirmed by phylogenetic analyses. Those grouping within a vertebrate-specific or vector-borne virus family/genus/group were subsequently identified as vertebrate viruses. We employed a simplified criterion of <80% nucleotide identity at the whole viral genome level to assign new virus species which were then verified by detailed phylogenetic analysis (see below).

Estimation of virus prevalence and abundance

Each virus sequence was first assigned to specific virus species based on the level of sequence similarity to each other and to reference sequences. To estimate the relative abundance of each virus species in each library, quality-trimmed reads were first mapped to the SILVA database (www.arb-silva.de, version 132.1) using Bowtie2 (version 2.3.5.1) (Langmead and Salzberg, 2012) to remove reads associated with ribosomal RNA. Unmapped reads were subsequently mapped to confirmed viral genomes using the "end-to-end" setting, and the abundance of each virus species was estimated as the number of mapped reads per million total reads (RPM) in each library. To limit false-positives, we set the RPM greater than or equal to 10 as providing evidence for a positive virus hit.

Evolutionary analysis

To infer the evolutionary history of all vertebrate-associated viruses here, we first assigned these viral sequences into their respective (vertebrate associated) viral family or genus based on the Diamond blastx results. The amino acid sequences of the viral replicase proteins (i.e. RNA viruses = RdRp domain; DNA viruses = capsid protein) of these virus genomes were then aligned with those of related viruses downloaded from GenBank using MAFFT (Katoh and Standley, 2013) (version 7.475), employing the L-INS-I algorithm. Following sequence alignment, all ambiguous aligned regions were removed using TrimAL (Capella-Gutiérrez et al., 2009) (version 1.2). Phylogenetic trees were then estimated for the sequence alignment of each family using the maximum likelihood method available in PhyML (version 3.1) (Guindon and Gascuel, 2003), employing the LG model of amino acid substitution and a subtree pruning and the regrafting (SPR) branch swapping algorithm. A similar procedure was utilized for intra-specific virus phylogenetic trees, with the exception that greater sequence identity meant that genome-scale nucleotide sequences were used in sequence alignments and a general-time reversible (GTR) substitution model was employed for phylogenetic inference.

Analysis of recombination in raccoon dog coronaviruses

We tested for the presence of recombination using the PHI-test (Bruen et al., 2006). Neighbor-Net reconstructions were performed using SplitsTree version 4 (Huson, 1998) based on Hasegawa-Kishino-Yano (Hasegawa et al., 1985) pairwise distances with empirical base frequencies (excluding gap sites). Edge support was calculated using 1,000 bootstrap replicates. For visualization, we filtered out splits with bootstrap support values <75%. Recombination patterns for CCoV-2 and the four novel raccoon dog coronaviruses were further analyzed using RDP4 (Martin et al., 2015). We used conservative settings for recombination detection in order to identify relative sizeable breakpoint-free regions. Specifically, our RDP4 analysis combined the RDP method (with a window size of 100 bp) (Martin and Rybicki, 2000), the Geneconv method (Sawyer, 1989), the MaxChi method (considering windows of 100 variable positions) (Smith, 1992), the BootScan method (with a window size of 500 bp) (Salminen et al., 1995), the SiScan method (with a window size of 500 bp) (Gibbs et al., 2000) and the 3Seq method (Boni et al., 2007). We identified mosaic patterns in all genomes based on those recombination events detected by >3 methods. For the two largest breakpoint-free regions identified by the RDP4 analysis, we performed maximum likelihood phylogenetic inference using IQ-Tree (version 2.0.3) (Minh et al., 2020), utilizing the GTR substitution model (Tavare, 1986) with a gamma-distributed rate variation among sites (Yang, 1994). Bootstrap support values were calculated using the ultrafast bootstrap method (Minh et al., 2013) with 1,000 replicates.

QUANTIFICATION AND STATISTICAL ANALYSIS

All computational and statistical analyses were performed with the open-source software tools referenced in the STAR Methods along with the described procedures.

Supplemental figures

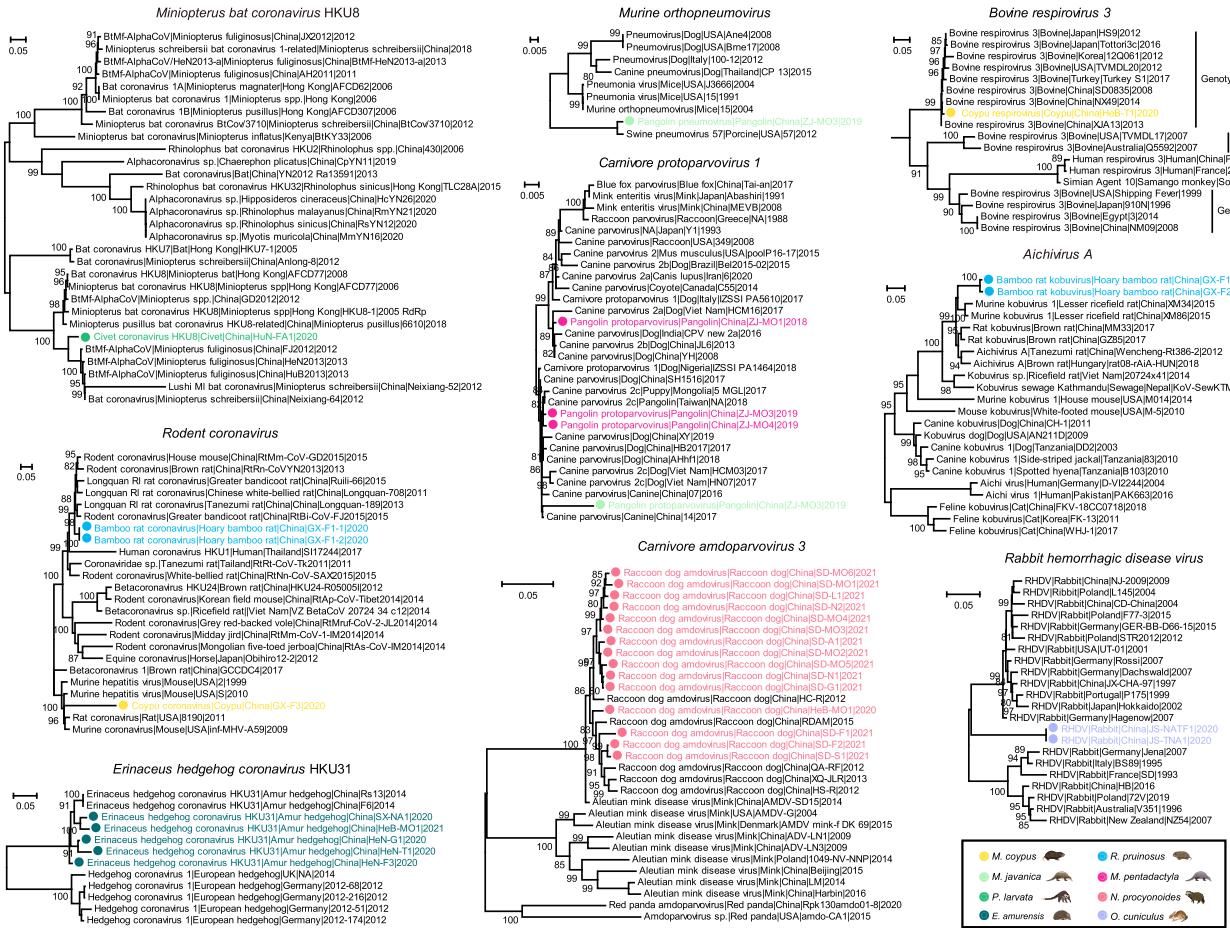
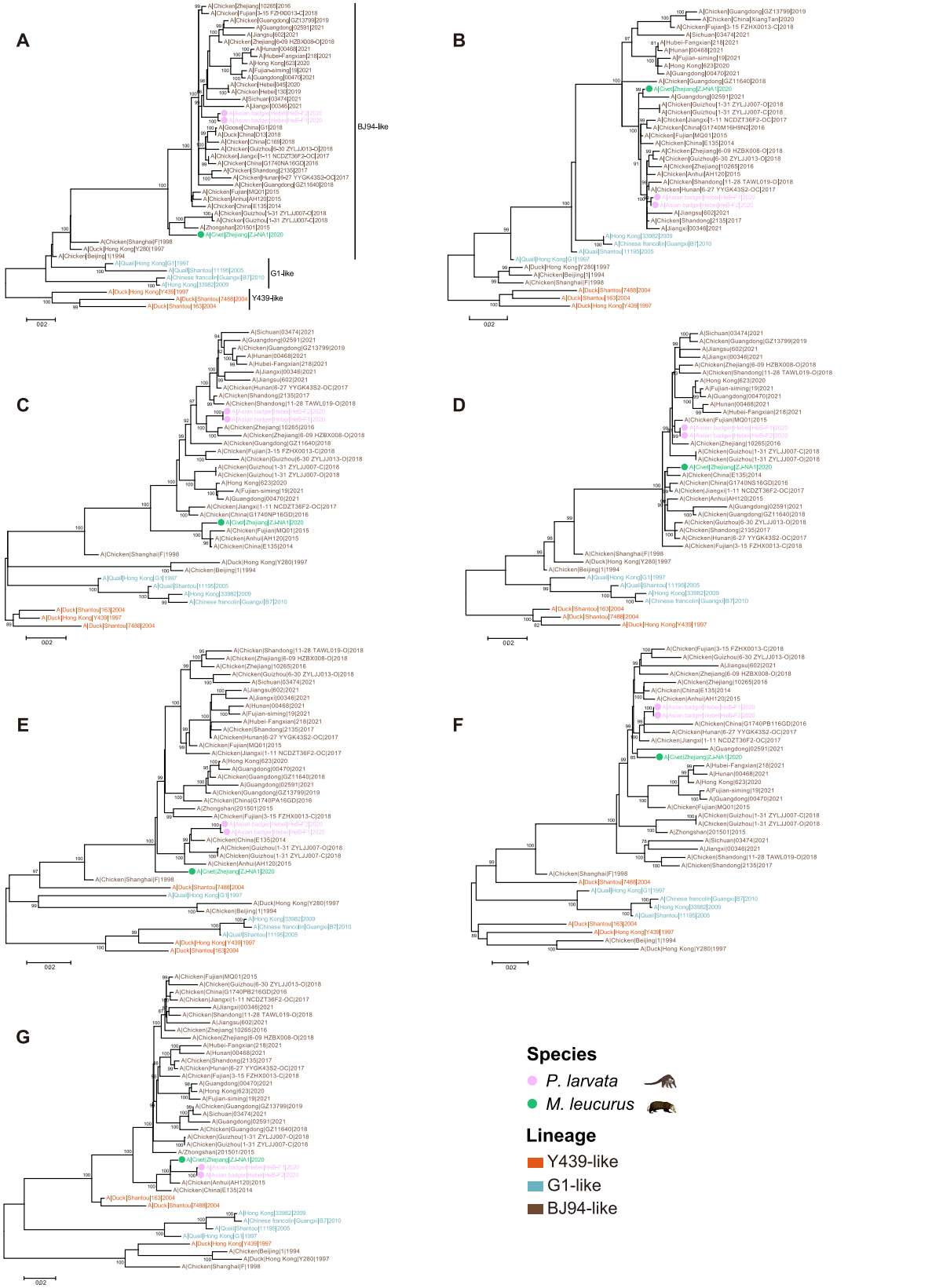


Figure S1. Intra-specific phylogenetic diversity of non-human-infecting pathogenic viruses identified in game animals, related to Figure 4

Each phylogenetic tree was estimated using a maximum likelihood method based on nucleotide sequence alignments of RdRp protein gene (RNA viruses) and capsid protein gene (DNA viruses). The trees were midpoint-rooted with the scale bar denoting the number of nucleotide substitutions per site. For clarity, only support values >80% are shown. Within each phylogeny, virus names in black represent published viral genomes, whereas those newly identified here are marked with solid circles and colored by animal host (see animal symbols in the legend bar). Virus names indicate viral type (subtype), host species, sampling location, strain name, and year, from left to right. For HKU8 and the rodent coronavirus, we specified the host as detailed as by the common name or Latin name.



(legend on next page)

Figure S2. H9N2 phylogenetic trees of the neuraminidase (NA), matrix protein (MP), nucleocapsid protein (NP), nonstructural protein (NS), polymerase PA (PA), polymerase PB1 (PB1), and polymerase PB2 (PB2) genes of H9N2 subtype AIVs, related to Figure 4

(A) The maximum-likelihood tree of the NA gene of H9N2 AIVs. The tree was inferred using PhyML employing the general-time reversible (GTR) nucleotide substitution model. Viruses were classified into three lineages—BJ/94-like, G1/97-like, and Y439/97-like—marked by different colors. The labels of the H9 isolates sequenced here were highlighted in light purple and green, whereas the light purple and green dots represent the hosts.

(B–G) The ML trees of the MP, NP, NS, PA, PB1, and PB2 genes of AIVs H9N2 estimated using the same procedure as described above. All the phylogenetic trees were rooted to the branch where A|Duck|Hong Kong|Y439|1997 was located. Support values >80% are shown.

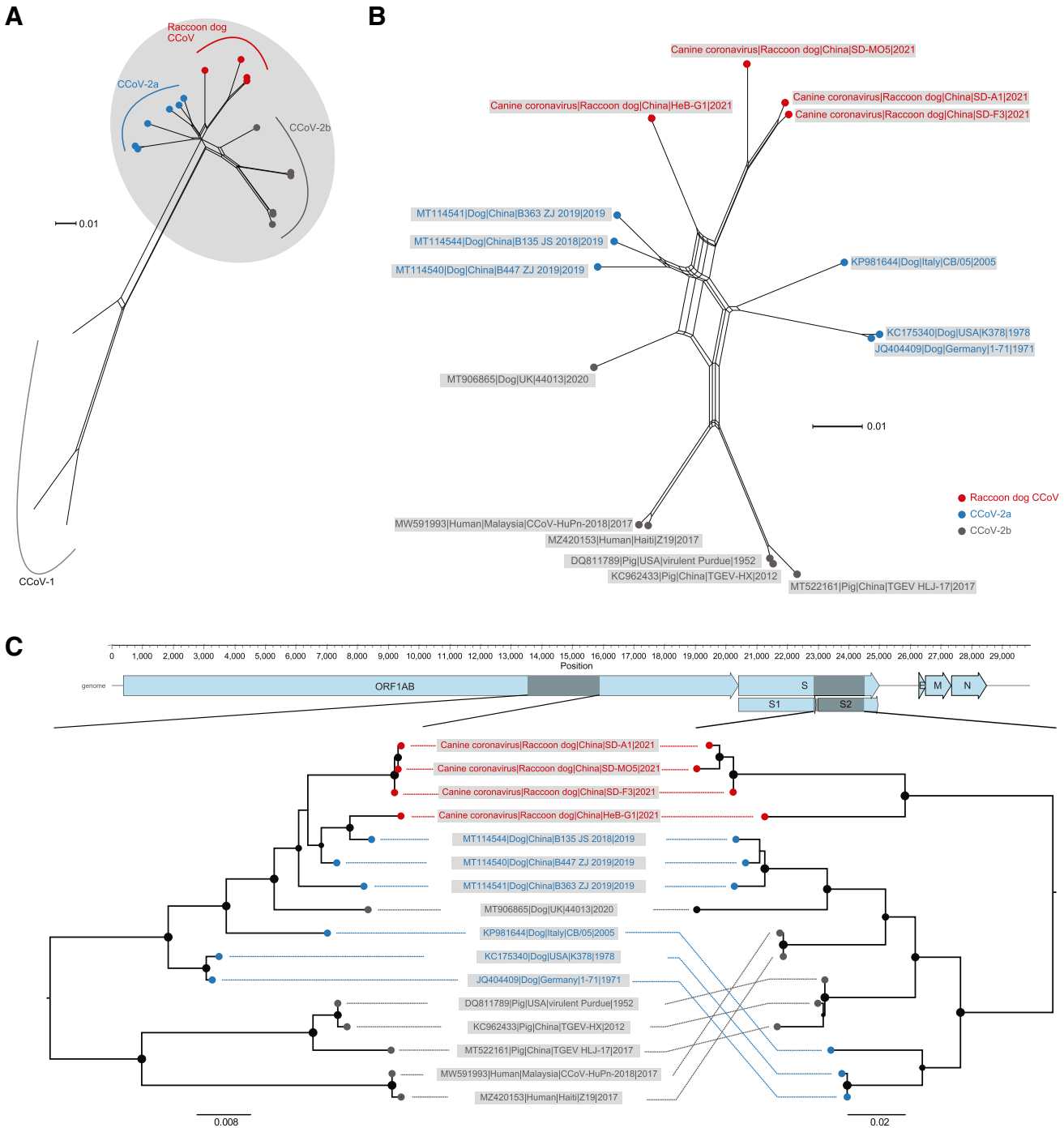


Figure S3. Recombination analysis of canine coronaviruses (CCoVs) including four novel raccoon dog CCoVs, related to Figure 4

(A) Neighbor-Net reconstruction based on complete genomes for canine coronavirus genotype 1 (CCoV-1) and genotype 2 (CCoV-2, subtype a (blue) and b (gray) viruses and four novel CCoVs from raccoon dogs (red).

(B) Neighbor-Net reconstruction based on the same data set excluding CCoV-1. In both Neighbor-Net reconstructions, edges are supported by bootstrap support values $\geq$ 75%.

(C) Phylogenetic tree of a breakpoint-free 2,322-bp segment in ORF1AB (left) and a 1,675-bp segment in S (~S2, right) from the alignment used in (B). The sizes of the internal node circles are proportional to bootstrap support values.

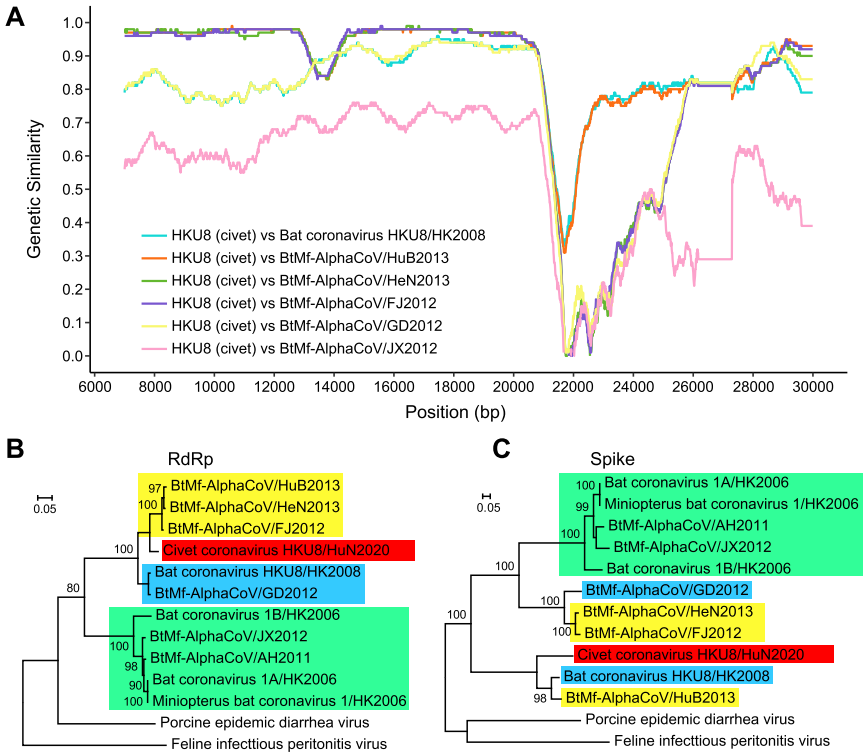


Figure S4. Recombination analysis of *Miniopterus bat coronavirus* HKU8 in the civet *P. larvata*, related to Figure 4

(A) Recombination analysis of *Civet coronavirus* HKU8 using SimPlot. Genome-scale similarity comparisons of *Civet coronavirus* HKU8 (query) against Bat coronavirus HKU8/HK2008 (blue), BtMf-AlphaCoV/HuB2013 (orange), BtMf-AlphaCoV/HeN2013 (green), BtMf-AlphaCoV/FJ2012 (purple), BtMf-AlphaCoV/GD2012 (yellow), and BtMf-AlphaCoV/JX2012 (pink).

(B and C) Maximum likelihood phylogenetic trees of the RdRp (B) and spike protein (C) of *Civet coronavirus* HKU8 and its closest relatives. The trees were midpoint-rooted and the scale bar represents the number of nucleotide substitutions per site. Major clades/lineages are shaded yellow, red (HKU8 identified in this study), blue, and green. The phylogenetic incongruence observed suggests that *Civet coronavirus* HKU8 has a recombination origin from bat coronaviruses.