Title: Comprehensive Infectome Analysis Reveals Diverse Infectious Agents with Zoonotic Potential in Wildlife

Authors: Xin Hou[1,2,3]*†, Jing Wang[1]†, Da-Xi Wang[4]†, Ye Chen[5]*†, Xin-Xin Li[6]†, Wei-Chen Wu[1], Shi-Jia Le[1,7], Shi-Qiang Mei[1], Han-Lin Liu[1], Yun Feng[8], Li-Fen Yang[8], Pei-Bo Shi[4,16,17], Zi-Rui Ren[4,16], Lin Yang[17,18], Hai-Jian He[9], Wan-Jie Song[6], Jie Chen[6], Meng-Meng Zhu[6], Zhi-Wen Jiang[6], Xing-Mei Li[10], Deyin Guo[11], Edward C. Holmes[12,13], Zi-Qing Deng[4]*, Mang Shi[1,14,15]*, Shuo Su[6]*

Affiliations:

[1]National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, State Key Laboratory for Biocontrol, School of Medicine, Shenzhen Campus of Sun Yat-sen University, Sun Yat-sen University, Shenzhen, China.
[2]Evolutionary Genomics of RNA Viruses, Institut Pasteur, Paris, France.
[3]National Key Laboratory of Chinese Medicine Modernization, Innovation Center of Yangtze River Delta, Zhejiang University, Jiaxing, China.
[4]BGI Research, Beijing, China.
[5]Key Laboratory of Fujian-Taiwan Animal Pathogen Biology, College of Animal Sciences, Fujian Agriculture and Forestry University, Fuzhou, China.
[6]Jiangsu Engineering Laboratory of Animal Immunology, Institute of Immunology, College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, China.
[7]Fujian Provincial Center for Disease Control and Prevention, Fuzhou, China.
[8]Department of Viral and Rickettsial Disease Control, Yunnan Provincial Key Laboratory for Zoonosis Control and Prevention, Yunnan Institute of Endemic Disease Control and Prevention, Dali, China.
[9]Agricultural College, Jinhua Polytechnic, Jinhua, China.
[10]Nanjing Hongshan Forest Zoo, Nanjing, China.
[11]Guangzhou National Laboratory, Guangzhou International Bio-Island, Guangzhou, China.
[12]School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia.
[13]Laboratory of Data Discovery for Health Limited, Hong Kong SAR, China.
[14]Shenzhen Key Laboratory for Systems Medicine in Inflammatory Diseases, Shenzhen Campus of Sun Yat-sen University, Sun Yat-sen University, Shenzhen, China.
[15]Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China.
[16]Shenzhen Key Laboratory of Unknown Pathogen Identification, BGl Research, Shenzhen, China.
[17]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China.
[18]MGI, BGI-Shenzhen, Shenzhen, China.

†These authors contributed equally to this work

*Correspondence to: fafuchenye@163.com; xin.hou@pasteur.fr; dengziqing@genomics.cn; shim23@mail2.sysu.edu.cn; shuosu@njau.edu.cn

Abstract

Understanding wildlife-pathogen interactions is crucial for mitigating zoonotic risk. Through meta-transcriptomic sequencing we profiled the infectomes of 1,922 samples from 67 mammalian species across China, uncovering a remarkable diversity of viral, bacterial, fungal, and parasitic pathogens. Of the 195 pathogens identified, 62 were novel, including a bi-segmented coronavirus in diseased lesser pandas, which we propose represents a new genus – Zetacoronavirus. The orders Carnivora and Rodentia exhibited the highest pathogen diversity and were implicated in numerous host-jumping events. Comparative analysis of diseased versus healthy animals revealed a trend of higher pathogen loads in the former, with possible differences in tissue tropisms. In total, 48 zoonotic and 17 epizootic pathogens were identified, with frequent cross-species transmission, emphasizing the potential for emerging public health threats. This study highlights the urgent need for wildlife pathogen surveillance to inform proactive disease management strategies.

Keywords: wild animal; infectomes; metatranscriptomics; pathogens; segmented coronaviruses

Main Text:

Introduction

Mammals host a diverse array of viruses, bacteria, eukaryotic parasites and fungi that can cause disease in humans and other animals (*1-7*). Cross-species transmission enables these pathogens to occasionally jump to humans, sometimes resulting in localized outbreaks or even large-scale epidemics and pandemics. To trace the origins of particular viral pathogens, and to help prevent future emergence events, considerable attention has been devoted to discovering viruses in wildlife using metagenomic approaches (*8-14*). While most of the viruses discovered primarily affect a single wildlife host species, others have crossed species boundaries, or even emerged in different mammalian orders, suggesting that they might be host generalists with the potential to emerge in multiple hosts, perhaps including humans (*15, 16*).

Although metagenomics has brought a new perspective on the diversity of the virosphere, the understandable emphasis on viral infections has meant that less is known about the roles played by bacteria and eukaryotic pathogens in disease emergence. However, reports of bacterial and fungal infections associated with human mortality, including those carrying antimicrobial resistance genes, highlight the need for a broader perspective on disease emergence (*17*). Meta-transcriptomic (i.e., total RNA) sequencing, which can simultaneously reveal viruses, bacteria, and eukaryotic pathogens, provides an efficient and increasingly cost-effective way to obtain a comprehensive understanding of the total "infectomes" of wildlife species (*18, 19*). Herein, we conducted an extensive individual-level infectome analysis in a diverse array of mammals from China, including common and rare species, domestic and wild animals, as well as diseased and healthy individuals. For many species, including primates (e.g., macaques and ring-tailed lemurs), artiodactyls (e.g., sika deer and muntjacs), and perissodactyls (e.g., horses and donkeys), such analyses are rare. Our particular aim was to systematically characterize the pathogen composition in a wide range of mammalian species, explore their interactions and cross-species transmission, and provide crucial insights for the early warning and control of human infectious diseases.

Results

The mammalian total infectome

Between 2018 to 2024 we collected 1,922 samples from 67 species across 11 mammalian orders in 20 Chinese provinces, comprising 1,283 diseased and 639 healthy samples (Fig. 1A, fig. S1 and table S1). This sampling included rare animal species such as tufted capuchins (*Sapajus apella*), snub-nosed monkeys (*Rhinopithecus roxellana*) Chinese water deer (*Hydropotes inermis*), six-banded armadillos (*Euphractus sexcinctus*), and caracals (*Caracal caracal*), which have not previously systematically examined for potential pathogens. Various organs relevant to disease outbreak, including liver, spleen, lung, kidney, intestine, and lymph nodes from diseased animals were collected. While faecal and respiratory swabs were obtained from healthy animals for comparison. Subsequently, we employed the meta-transcriptomic sequencing to profile the total infectome which led to the characterization of a whole spectrum of virus, bacteria, fungi, and parasitic pathogens within these samples. These include: (i) established pathogens, (ii) opportunistic pathogens, and (iii) previously undescribed microorganisms with pathogenic potential in mammals. Phylogenetic assessments and abundance estimations (RPM > 10) confirmed the presence of 105 RNA viral species, 42 DNA viral species, 34 bacterial pathogen species, ten fungal pathogen species, and four parasitic species (Fig. 1B, fig S2 and table S2). Among them, 62 novel pathogen species were discovered, comprising 58 species of viruses and four divergent members of the fungal genus *Pneumocystis*.

Our standardized analysis, adjusted for variation in sample sizes, revealed that the mammalian order Carnivora harboured the most diverse and abundant pathogens, followed closely by the Rodentia (Fig. 1C, fig. S2 and S3). While high viral diversity and abundance in Rodentia can likely be attributed to their often high population densities, the elevated viral diversity and abundance in Carnivora more likely reflects their unusually high population density in fur farms as opposed to natural environments. Among the pathogens identified, we found zoonotic viruses such as avian influenza viruses H5N1 and H5N6, *Alphacoronavirus 1*, *Norwalk virus*, *Rotavirus A*, *Hepatitis E virus* (HEV), and several viruses known to cause acute or severe diseases to mammals including *Morbillivirus canis*, *Rabbit hemorrhagic disease virus (RHDV)*, *Pangolin pestiviruses*, *Feline paramyxovirus*, *Respirovirus bovis and Respirovirus muris*, *Sapporo virus*, *Amdoparvovirus carnivoran1* and *Amdoparvovirus carnivoran3* (Fig. 1D, fig. S4

and S5A). In addition to the well-known mammalian-infecting viruses mentioned above, several previously uncharacterized viral agents at high abundance (≥10 RPM) were identified in diseased animals and warrant further attention. Also of note was the identification of two novel sapoviruses in mink and raccoon dogs, as well as a novel strain of *Hepatitis E virus* – specifically *Mink hepatitis E virus* – in mink populations from Shandong and Hebei provinces.

In addition to viral pathogens, we identified a diverse array of bacterial, fungal, and parasitic pathogens, many of which are well-documented for their association with human diseases and zoonotic transmission. For instance, *Clostridioides difficile* is a leading cause of healthcare-associated infections (*20*), while *Coxiella burnetii*, the causative agent of Q fever, poses significant zoonotic risks (*21*). Similarly, *Pasteurella canis* which is commonly linked to animal bites (*22*), a relative of *Aeromonas caviae* usually associated with gastrointestinal infections (*23*), and the presence of parasites such as *Giardia intestinalis* and *Toxoplasma gondii*, underscore the broad spectrum of pathogens transmitted from animal reservoirs to humans. Additionally, we detected several opportunistic pathogens, including *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Campylobacter jejuni*, *Eggerthella lenta*, and *Aspergillus fumigatus*, that are generally associated with mild or opportunistic infections in immunocompromised individuals (*24*). Notably, we also identified four divergent members of the fungal genus *Pneumocystis* which is associated with pneumonia in humans and mammalian hosts (*7*). These fungi were primarily identified in species within the orders Carnivora and Rodentia (Fig. 1 and fig. S5B). Together, these findings highlight the intricate interplay between wildlife, domestic animals, and human health, emphasizing the critical need for ongoing surveillance and proactive measures to monitor and mitigate zoonotic transmission risks.

A novel segmented coronavirus in diseased lesser pandas

Of particular note, we discovered a novel, segmented coronavirus in lesser (red) pandas (*Ailurus fulgens*) that exhibited signs of disease. Phylogenetic analysis using the RNA-dependent RNA polymerase (RdRP) gene placed this virus, provisionally denoted Zetacoronavirus 1, within the family *Coronaviridae*, but distinct from other known coronaviruses genera, including the putative *Epsiloncoronavirus* found in intestines and lungs of *Meles* (i.e. European badgers) *(25)* (Fig. 2A). Indeed, genomic analysis showed that this virus had less than 50% similarity to known

coronaviruses (Fig. S6). Importantly, this virus was primarily found in the lung and spleen tissues of lesser pandas in Jiangsu and Fujian provinces, with a high viral load in the lungs (RPM 7,396.26, Fig. 2B), suggesting a link to respiratory illness and high mortality. Despite efforts, we did not obtain the unsegmented genome typical of other coronaviruses. Rather, the virus appears to be segmented, consisting of at least two segments: one with ORF1ab and RdRP (1.47 kb, PQ685031) and another with the spike and nucleocapsid proteins (7.66 kb, PQ685032) (Fig. 2C). Rapid amplification of cDNA ends (RACE) confirmed this genome structure and revealed a highly conserved 700 bp sequence at the 5' terminus of both segments, with at least 11 SNPs between strains (Fig. 2D, Fig. S7). This structure resembles bi-segmented Nido-like viruses found in aquatic vertebrates, which exhibit an identical trailing sequence of 129 nucleotides at their respective 3’ ends (Fig. 2A and 2C) (*26*). Both contigs exhibited poly(A) tails at their 3’ termini, confirmed through 3’ RACE, further supporting the presence of a bi-segmented genome (Fig. 2D, Fig. S7). This conserved 700 bp sequence at the 5' terminus (5UTR-F-1: GGCAAGCAGTGGTATCAACGCA; 5UTR-F-2: ACCAACAGGCTTGACCTCTTGA) and poly(A) tails at the 3’ terminus were utilized in PCR attempts to amplify additional segments. However, no products were obtained, providing further evidence for the bi-segmented nature of the genome. We also utilized long-read amplicon sequencing data to validate the contiguity within both segments, especially the regions between the S and N genes, of Zetacoronavirus 1 (Fig. S7C).

Evolutionary analysis of RdRP and spike proteins revealed substantial genetic divergence of Zetacoronavirus 1 from known genera within the *Coronaviridae* (pairwise distances: RdRP > 0.781, spike > 1.025; Fig. 2E, Fig. S6B). We therefore proposed that this virus should be classified as a putative new genus, tentatively named Zetacoronavirus, based on its significant divergence and distinct genomic features. The virus segments, although small, contain key components for replication and structural (i.e., spike and nucleoprotein) functions. Detailed genome annotation revealed 13 non-structural proteins (NSPs) and an unconventional ribosomal frameshifting signal (TTTAAAT, instead of the canonical TTTAAAC) within segment 1 (PQ685031), encoding ORF1ab. Interestingly, three non-structural proteins – NSP1, NSP2, and NSP15 – were absent. Segment 2 (PQ685032), encoding the spike and nucleocapsid proteins, notably lacked the envelope and membrane structural proteins. These distinct features define a compact, bi-segmented genome architecture (Fig. 2), highlighting the hidden diversity of

mammalian coronaviruses and their largely unexplored genomic formations. These observations merit additional virus isolation and molecular validation.

Comparisons of infectomes in diseased and healthy wild animals

To assess disease associations of the pathogens identified here, we compared infectomes from diseased and healthy wild animals. The diseased group shown a higher incidence of RNA viruses, bacteria, fungi, and parasites, while DNA viruses were more common in the healthy group, suggesting opportunistic or less-pathogenic roles (Fig. 3A). However, differences in pathogen distribution were partly influenced by sample type: organ samples predominated in the diseased group, while swab samples were more common in the healthy group (Figs. 3B and 3C). Several pathogens, including *Influenza A virus*, *Morbillivirus canis*, *Alphacoronavirus 1*, were more abundant in diseased animals, indicating a potential link to illness (Fig. 3D). Conversely, pathogens like *Campylobacter jejuni* and *Klebsiella pneumoniae* were equally prevalent in both groups, suggesting opportunistic or subclinical roles (Fig. 3D).

We also examined tissue tropisms, which revealed greater pathogen richness in the gut and lungs compared to other organs such as the liver, spleen and kidney (Figs. S8A and S8B). Some pathogens, like AMPV carnivoran1 and RHDV detected in mink and rabbits, respectively, were present in multiple organs, suggesting systemic infections (Fig. 3E and fig. S8C). Other pathogens showed the expected patterns of organ specificity; for example, gastrointestinal pathogens were mainly found in the gut, while respiratory pathogens, including newly identified *Pneumocystis* species, were found only in lung tissues (Fig. 3E and fig. S8D). Overall, while many pathogens were distributed across multiple organs, most exhibited pronounced organ specificity. The novel Zetacoronavirus 1 appears to initially affect the lungs, although with evidence of spread to the spleen, raising concerns about its impact on multiple organs and systems.

Although healthy controls were not included, we identified key pathogens in several cases that may be associated with the observed diseases. For example, a ring-tailed lemur (*Lemur catta*) from Zhejiang province, which died in a zoo, exhibited necrotic liver lesions, hemorrhagic infarctions in a blackened and hardened spleen, severe pulmonary haemorrhage, and enlarged kidneys upon necropsy. A high load of *Toxoplasma gondii* was detected in various tissues, with

RPM values ranging from 82 to 5,245 suggesting it was the likely causative agent. In five mink (*Neovison vison*) from Liaoning Province, symptoms of emaciation and yellow-green mucoid diarrhea were observed, and fecal samples revealed multiple pathogens linked to gastrointestinal diseases, including *Canine norovirus 2*, *Mink enteric sapovirus*, *Mink coronavirus 1*, *Amdoparvovirus carnivoran 1*, and *Campylobacter jejuni*. Additionally, continuous monitoring from 2021 to 2022 of a single Malayan pangolin (*Manis javanica*) revealed increasing viral loads of *Pangolin pestivirus 4*, *Pangolin shanbavirus*, and *Pangolin copiparvovirus 1* in a deceased individual. These viral infections likely contributed to the animal's death, potentially linked to observed symptoms such as pericardial effusion, yellow discoloration, bladder distension, and ascites. Moreover, a tick-borne virus, Qingdao tick uukuvirus *(27)*, was detected in hedgehogs with lung tissue calcification.

Characterization of zoonotic and epizootic pathogens

We analysed the data generated here combined with all publicly available data to examine the frequency of host jumps among zoonotic and epizootic pathogens. These results revealed that species from the mammalian orders Carnivora, Rodentia, and Lagomorpha commonly host RNA viruses, DNA viruses, bacteria, fungi, and parasites (Fig. 4A). We then categorized the total set of 195 pathogens into three types: zoonotic (i.e., transmitting between animals and humans), epizootic (spreading across different animal orders), and order-specific (infecting species within the same order) (Fig. 4B and Table S3). In this manner we identified 48 zoonotic, 17 epizootic and 130 order-specific pathogens, with Carnivora and Rodentia serving as both sources and recipients of zoonotic spillovers (Fig. 4C and Fig. S9). Pathogens at high prevalence and that exhibited evidence of cross-species transmission were detected in multiple animal species across different orders (Fig. 4D). Examples include the highly pathogenic avian influenza virus (HPAIV) H5N1 and H5N6 in raccoon dogs and mink, respectively, *Qingdao tick uukuvirus* in hedgehogs, and *Morbillivirus canis* (*Canine distemper virus*, CDV) in raccoon dogs, civets, mink, and other carnivores, highlighting their potential for cross-species transmission. Phylogenetic analysis of the HA gene revealed that the H5N1 influenza virus identified in raccoon dogs from Shandong province was closely related to strains detected in chickens from Tottori and Chiba, all circulating in 2022. Similarly, the H5N6 influenza virus detected in mink from Shandong

clustered with a human-associated strain identified in 2021 (Fig. S10). These findings raise concern, particularly as H5N1 has caused the deaths of thousands of marine mammals in South America since 2022 (*28, 29*). We also identified known zoonotic pathogens such as *Sapporo virus*, *Campylobacter spp*., *Clostridium perfringens*, *Fusobacterium necrophorum*, *Toxoplasma gondii*, and *Coxiella burnetii* across multiple mammalian hosts, particularly in Zhejiang and Jiangsu provinces in southern China where nearly 20 cross-species transmissions were recorded here (Fig. 4E and 4F).

In addition to zoonotic pathogens, the data generated here highlight the importance of monitoring pathogens that are currently confined to individual animal populations. For example, despite their limited host range, pathogens like *Amdoparvovirus carnivoran1* (*Aleutian mink disease parvovirus*, AMDV), *Amdoparvovirus carnivoran3* (raccoon dog and arctic fox amdoparvovirus, RDAV), and pangolin pestiviruses could pose significant threats to the principal host as they are commonly associated with overt disease (Fig. S11).

Discussion

Wild animals are often linked to the emergence of major human diseases, making it crucial to identify existing or potential pathogens in these species. Such efforts can also help trace the origins of specific epidemics and assess the risk of future outbreaks, particularly the identification of host "generalist" pathogens that seemingly possess an enhanced ability to emerge in new species, including those from different mammalian orders (*4, 15, 16, 30*). The diverse array of pathogens uncovered in this study highlights the importance of adopting a comprehensive infectome approach to better understand and mitigate the risks of zoonotic disease transmission (*18, 19*). By profiling the spectrum of viral, bacterial, fungal, and parasitic pathogens we have demonstrated the intricate and interconnected nature of wildlife-pathogen dynamics. Traditional pathogen-specific or virus-focused studies risk overlooking important contributors to disease ecology, including bacterial, fungal, or parasitic agents capable of zoonotic or epizootic transmission. For example, the identification of opportunistic pathogens such as *Pseudomonas aeruginosa* and *Aspergillus fumigatus* underscores their potential roles in wildlife health and spillover events. As such, these findings emphasize the need for integrative

surveillance that encompasses all pathogen types, enabling more accurate risk assessments and informed, proactive interventions.

The complex interplay between wildlife pathogens and human health underscores the need for a collaborative, interdisciplinary approach to disease surveillance. By including diverse mammalian species, particularly under-sampled orders like Carnivora and Rodentia, this study highlights the critical importance of understanding pathogen reservoirs in both wild and managed animal populations. For instance, the high pathogen loads observed in fur-farmed carnivores suggest that anthropogenic pressures may significantly amplify spillover risks (*31*). Additionally, the discovery of novel pathogens such as Zetacoronavirus 1 in lesser pandas reinforces the role of wildlife as reservoirs for emerging infectious diseases. These findings call for enhanced collaboration among ecologists, veterinarians, epidemiologists, and public health experts to develop comprehensive frameworks for pathogen monitoring. Such frameworks should integrate genomic, ecological, and epidemiological data to better predict and mitigate future outbreaks.

Limitations

A key limitation of this study stems from differences in sample types collected from diseased and healthy groups, which may have influenced the observed variations in pathogen diversity and abundance. In particular, diseased samples primarily consisted of internal organ tissues, such as the liver, spleen, lungs, and intestines—sites that are more likely to harbor pathogens directly associated with clinical symptoms. In contrast, healthy samples were largely derived from fecal or respiratory swabs, which may capture a different subset of the infectome. These methodological discrepancies complicate the interpretation of whether the observed differences in pathogen richness and load between the groups are truly disease-associated or artifacts of sample type variability. Additionally, this study is geographically limited to China, highlighting the need to expand the framework to encompass a wider range of species and geographic regions for more comprehensive global health preparedness. Furthermore, while we relied on genomic tools for pathogen characterization, the isolation and functional study of novel agents in cell and animal models will be crucial for gaining deeper insights into their potential impacts on wildlife and human populations.

References and Notes

1. E. C. Holmes *et al.*, The origins of SARS-CoV-2: A critical review. *Cell* **184**, 4848-4856 (2021).
2. D. M. Skowronski *et al.*, Severe acute respiratory syndrome (SARS): a year in review. *Annu Rev Med* **56**, 357-381 (2005).
3. Y. M. Chen *et al.*, Host traits shape virome composition and virus transmission in wild small mammals. *Cell* **186**, 4662-4675 e4612 (2023).
4. W. T. He *et al.*, Virome characterization of game animals in China reveals a spectrum of emerging pathogens. *Cell* **185**, 1117-1129 e1118 (2022).
5. I. Ghasemzadeh, S. H. Namazi, Review of bacterial and viral zoonotic infections transmitted by dogs. *J Med Life* **8**, (2015).
6. P. Parola *et al.*, Update on tick-borne rickettsioses around the world: a geographic approach. *Clin Microbiol Rev* **26**, 657-702 (2013).
7. J. R. Kohler, B. Hube, R. Puccia, A. Casadevall, J. R. Perfect, Fungi that Infect Humans. *Microbiol Spectr* **5**, (2017).
8. K. G. Andersen *et al.*, Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus. *Cell* **162**, 738-750 (2015).
9. D. J. Park *et al.*, Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. *Cell* **161**, 1516-1526 (2015).
10. H. C. Metsky *et al.*, Zika virus evolution and spread in the Americas. *Nature* **546**, 411-415 (2017).
11. N. D. Grubaugh *et al.*, Travel Surveillance and Genomics Uncover a Hidden Zika Outbreak during the Waning Epidemic. *Cell* **178**, 1057-1071 e1011 (2019).
12. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
13. J. A. Lednicky *et al.*, Emergence of porcine delta-coronavirus pathogenic infections among children in Haiti through independent zoonoses and convergent evolution. *medRxiv*, (2021).
14. W. Shi *et al.*, Trafficked Malayan pangolins contain viral pathogens of humans. *Nat Microbiol* **7**, 1259-1269 (2022).
15. X. Cui *et al.*, Virus diversity, wildlife-domestic animal circulation and potential zoonotic viruses of small mammals, pangolins and zoo animals. *Nat Commun* **14**, 2488 (2023).
16. C. C. S. Tan, L. van Dorp, F. Balloux, The evolutionary drivers and correlates of viral host jumps. *Nat Ecol Evol* **8**, 960-971 (2024).
17. J. Larsen *et al.*, Emergence of methicillin resistance predates the clinical use of antibiotics. *Nature* **602**, 135-141 (2022).
18. J. Wang *et al.*, Total RNA sequencing of Phlebotomus chinensis sandflies in China revealed viral, bacterial, and eukaryotic microbes potentially pathogenic to humans. *Emerg Microbes Infect* **11**, 2080-2092 (2022).
19. X. Huang *et al.*, A total infectome approach to understand the etiology of infectious disease in pigs. *Microbiome* **10**, (2022).
20. J. E. Buddle, R. P. Fagan, Pathogenicity and virulence of Clostridioides difficile. *Virulence* **14**, 2150452 (2023).
21. J. Njeru, K. Henning, M. W. Pletz, R. Heller, H. Neubauer, Q fever is an old and neglected zoonotic disease in Kenya: a systematic review. *BMC Public Health* **16**, 297 (2016).

22. Z. T. Kern, O. M. Swartley, P. Neupane, N. Balakrishnan, E. B. Breitschwerdt, Pasteurella canis infective endocarditis in a dog. *Vet Microbiol* **229**, 14-19 (2019).
23. A. Fernandez-Bravo, M. J. Figueras, An Update on the Genus Aeromonas: Taxonomy, Epidemiology, and Pathogenicity. *Microorganisms* **8**, (2020).
24. N. Riccardi, G. A. Rotulo, E. Castagnola, Definition of Opportunistic Infections in Immunocompromised Children on the Basis of Etiologies and Clinical Features: A Summary for Practical Purposes. *Curr Pediatr Rev* **15**, 197-206 (2019).
25. G. Zamperin *et al.*, Discovery of a coronavirus in the Eurasian badger (Meles meles) belonging to a putative new genus. *Infect Genet Evol* **109**, 105406 (2023).
26. R. C. Edgar *et al.*, Petabase-scale sequence alignment catalyses viral discovery. *Nature* **602**, 142-147 (2022).
27. G. Hu *et al.*, Diversity Analysis of Tick-Borne Viruses from Hedgehogs and Hares in Qingdao, China. *Microbiology Spectrum* **11**, e05340-05322 (2023).
28. L. C. Caserta *et al.*, Spillover of highly pathogenic avian influenza H5N1 virus to dairy cattle. *Nature* **634**, 669-676 (2024).
29. T. Peacock *et al.*, The global H5N1 influenza panzootic in mammals. *Nature*, (2024).
30. N. D. Wolfe, C. P. Dunavan, J. Diamond, Origins of major human infectious diseases. *Nature* **447**, 279-283 (2007).
31. J. Zhao *et al.*, Farmed fur animals harbour viruses with zoonotic spillover potential. *Nature* **634**, 228-233 (2024).

Acknowledgments: This work was supported by the National Key Research and Development Program of China (2024YFC2607502), the National Natural Science Foundation of China (82341118, 32270160), the Shenzhen Science and Technology Program (KQTD20200820145822023), the Natural Science Foundation of Guangdong Province (2022A1515011854), the Guangdong Province “Pearl River Talent Plan” Innovation and Entrepreneurship Team Project (2019ZT08Y464), the Hong Kong Innovation and Technology Fund (ITF) (MRP/071/20X), and the Health and Medical Research Fund (COVID190206). E.C.H. is funded by a National Health and Medical Research Council (Australia) Investigator grant (GNT2017197) and by AIR@InnoHK administered by the Innovation and Technology Commission, Hong Kong Special Administrative Region, China. We gratefully acknowledge colleagues at BGI Research and the China National Genebank (CNGB) for library preparation, sequencing and data management.

Author contributions:

Conceptualization: X.H., E.C.H., M.S., S.S.;

Methodology: X.H., J.W., D.-X.W., E.C.H., Z.Q.D., M.S., S.S.;

Investigation: X.H., J.W., D-X.W., W-C.W., S-J.L., S-Q.M., H-L.L., Y.F., L-F.Y., P-B.S., Z-R.R., L.Y., Z-Q.D.;

Sampling: F-F.C., X-X.L., H-J.H., W-J.S., J.C., M-M.Z., Z-W.J., X-M.L., S.S.;

Visualization: X.H.;

Funding acquisition: D.Y., E.C.H., M.S., S.S.;

Writing – original draft: X.H., E.C.H., M.S., S.S.;

Writing – review & editing: All authors;

Supervision: X.H., Z.Q.D., M.S., S.S.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: The genome sequences of bi-segmented coronavirus identified in lesser pandas are available at the NCBI GenBank database under accession numbers PQ685031-PQ685032.

Supplementary Materials

Materials and Methods

Figs. S1 to S11

Tables S1 to S3

References (*32*–40)

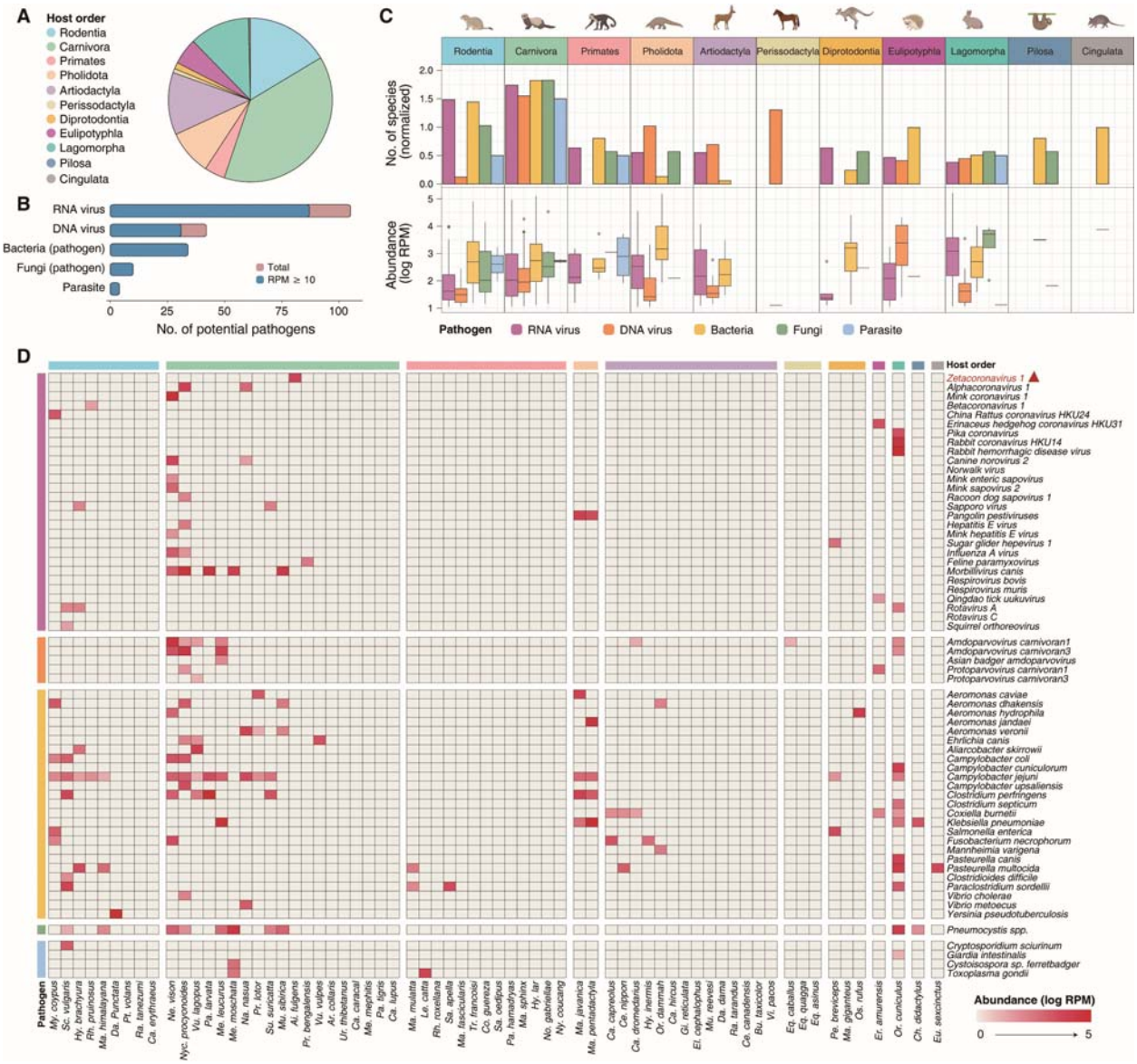

Fig. 1. Total infectome characterization of mammalian wildlife. (**A**) Number of samples collected at the level of mammalian order. (**B**) Number of potential viral, bacterial, fungal and parasitic pathogens identified in this study. The abundance levels of 29 pathogens fell below our classification threshold (i.e., 10 RPM) for positivity and were excluded from further analysis. (**C**) Pathogen richness and abundance at the animal order level. Pathogen richness was normalized based on sample size for comparison. (**D**) Distribution and abundance of representative pathogens with potential link to disease. The relative abundance of pathogens in each library was calculated and normalized by RPM. To remove potential contamination, only pathogens with an

abundance RPM ≥10 are shown. Pathogen groups and mammalian orders are shown by the colors on the heatmap.

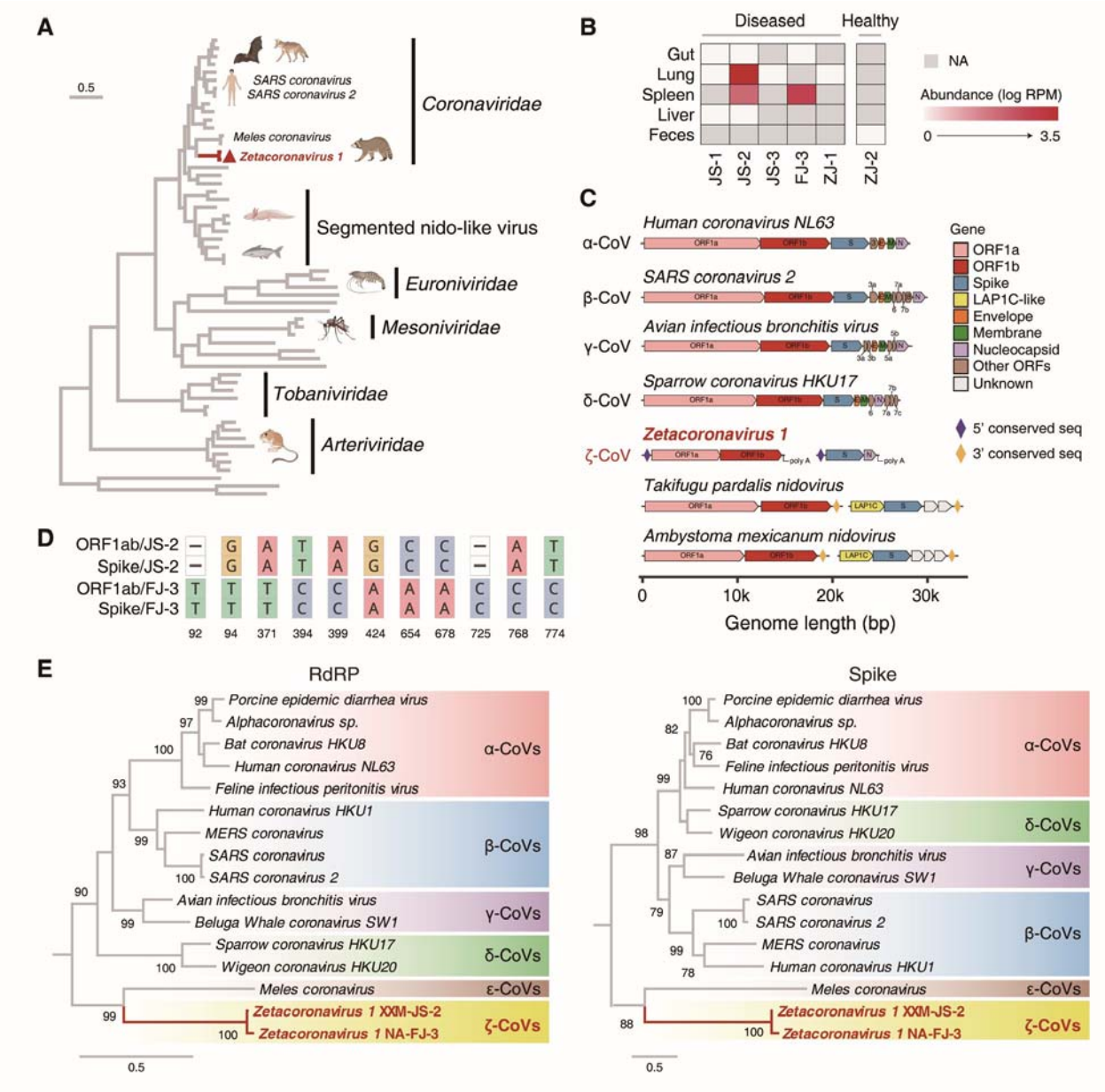

Fig. 2. Novel pathogenic and segmented coronavirus in diseased lesser pandas. (**A**) The phylogenetic position of the novel coronavirus – tentative species Zetacoronavirus 1 – identified in this study is depicted in the phylogenetic tree of the *Nidovirales* and marked by red bold font. (**B**) The prevalence and abundance of Zetacoronavirus 1 in lesser pandas. "NA" denotes the sample types are not available for sample collection in individuals. (**C**) The genomic structures of common coronaviruses, as well as Zetacoronavirus 1 and segmented Nido-like viruses. Zetacoronavirus 1 contains a conserved sequences of around 700bp in 5' end of each segment,

while the segmented Nido-like viruses contain an identical 129-nt trailing sequence in 3’ end of each segment. (**D**) Single nucleotide polymorphisms (SNPs) of the conserved 700 bp sequences in the two strains of Zetacoronavirus 1. (**E**) Phylogenetic tree of *Coronaviridae* using a maximum likelihood method based on RdRP and spike protein sequences. All trees were midpoint-rooted for clarity only, and the scale bar indicates 0.5 amino acid substitutions per site. Zetacoronavirus 1 is marked in red bold font.

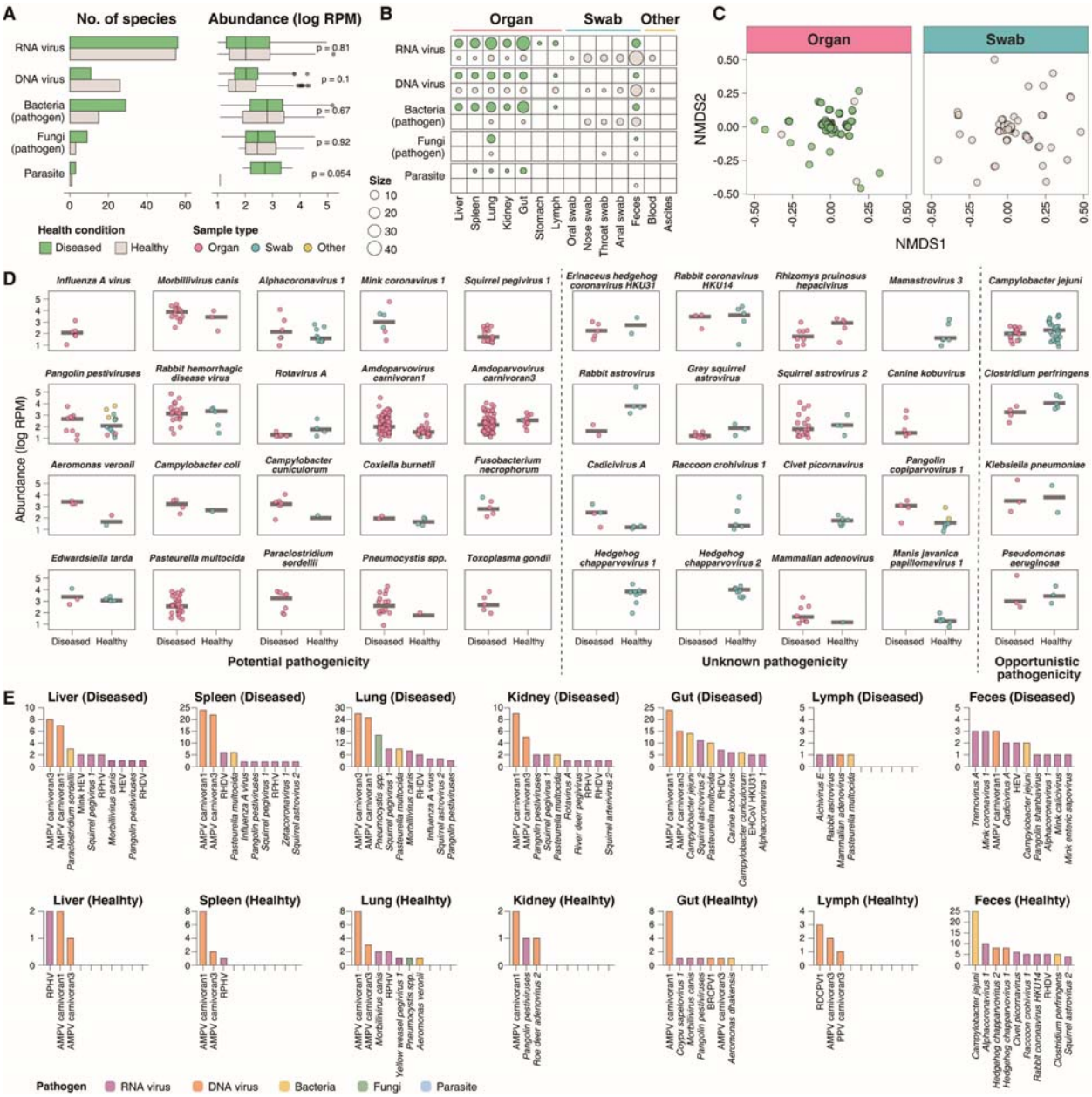

Fig. 3. Comparisons of infectomes in diseased and healthy wild animals. (**A**) Richness and abundance of each pathogen group are shown between diseased and healthy groups. The text in each pair box indicts the significance of abundance comparisons between diseased and healthy animals. (**B**) Number of viral, bacterial, fungal and parasitic pathogens identified at the sample type level. The circle size reflects the number of pathogens detected. (**C**) Non-metric multidimensional scaling (NMDS) analysis of a Bray-Curtis dissimilarity matrix calculated from organ and swab samples showing differences in pathogen distribution patterns between diseased and healthy animals. (**D**) Abundance comparisons between healthy and unhealthy animals for

pathogens detected at least five libraries (i.e., prevalence ≥5 libraries). (**E**) Prevalence of pathogens in diseased and healthy tissues, and feces.

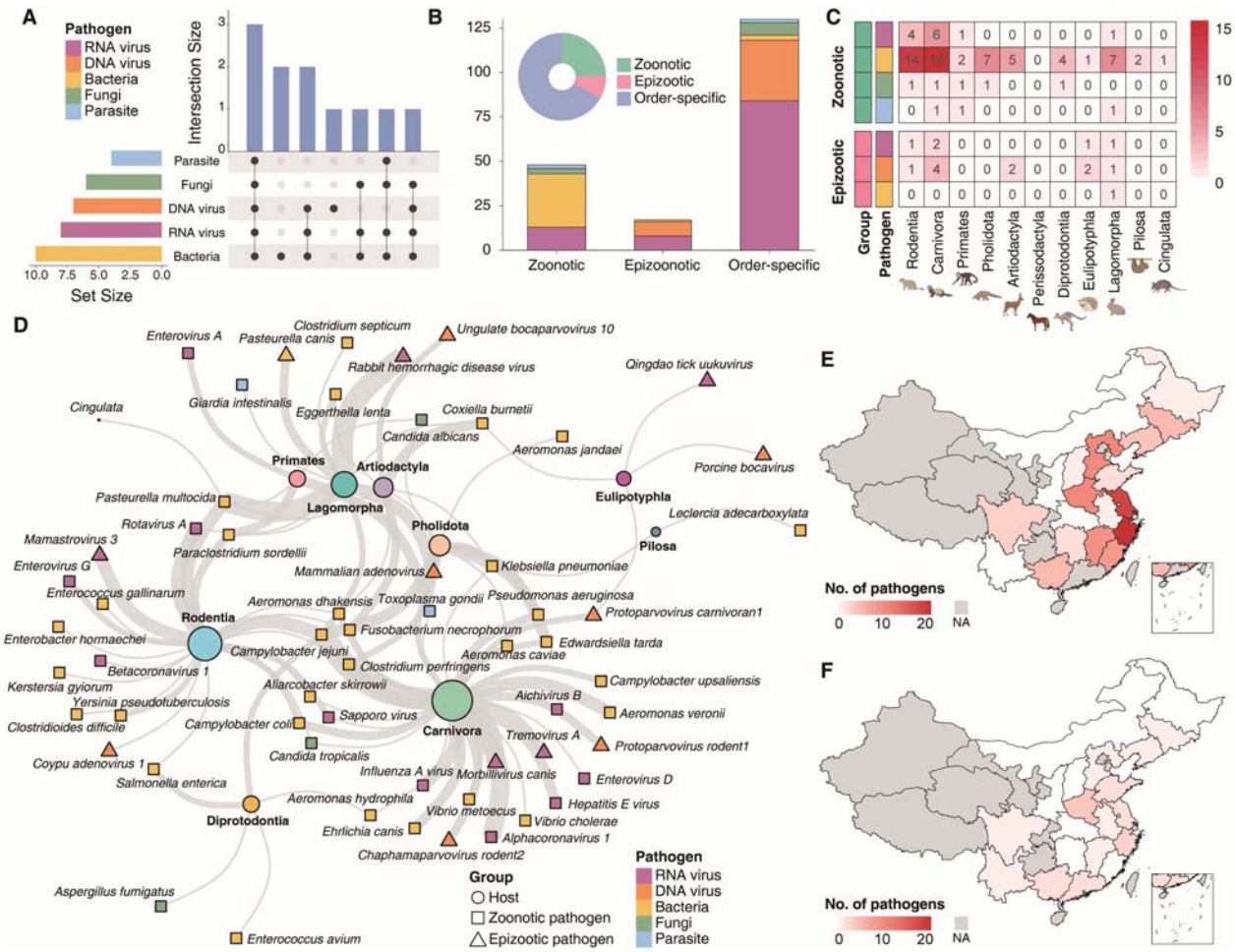

Fig. 4. Epidemiological patterns of zoonotic and epizootic pathogens. (**A**) The upset diagram illustrates the incidence of pathogen infection in mammalian orders. The left bar plot shows the number of infected animal orders of each pathogen group. The top right bar plot and crosspoint diagram shows the co-infection of pathogen groups in different animal orders. (**B**) Number of (i) zoonotic (i.e., capable of transmission between animals and humans), (ii) epizootic (i.e., capable of transmission between animals from different orders), and (iii) order-specific (i.e., exclusively infecting mammalian species within the same order) pathogens identified in this study. (**C**) The number of zoonotic and epizootic pathogens at the level of mammalian order. (**D**) Associations between potential zoonotic and epizootic pathogens and mammalian hosts. The size of the colored circles representing animal hosts indicates the abundance of potential zoonotic and epizootic pathogens identified within each mammalian order, while the thickness of the line reflects the prevalence of each pathogen detected in that particular mammalian order. The

number of zoonotic (**E**) and epizootic (**F**) pathogens at provincial level. “NA” shaped in grey denotes provinces for which the sample is not available.