Cao et al. *Zool. Res.* 2026, 47(1): 72−87
https://doi.org/10.24272/j.issn.2095-8137.2025.271
https://cstr.cn/32248.14.zr.2025.271

ZR | Zoological Research

Article Open Access

Highly divergent virome in wild rodents of Xinjiang, China: Implications for rodent origin of porcine reproductive and respiratory syndrome virus 1

Le Cao[1,#], Ying-Ying Ma[1,#], Guo-Wu Zhang[2,#], Jun Li[2], Wen-Jing Qi[2], Chuan-Chuan Wu[3], Meng-Xiao Tian[3], Yao Zhang[3], Zhuang-Zhi Zhang[4], Ming-Zhi Yan[2], Malike Aizezi[5], Yan-Peng Li[1,*], Chi-Yu Zhang[1,*], Wen-Bao Zhang[2,3,*]

[1] *Shanghai Public Health Clinical Center*, *Fudan University*, *Shanghai* 201508, *China*

[2] *State Key Laboratory of Pathogenesis*, *Prevention and Treatment of High Incidence Diseases in Central Asia*, *First Affiliated Hospital of Xinjiang Medical University*, *Urumqi*, *Xinjiang* 830054, *China*

[3] *College of Basic Medical Sciences*, *Xinjiang Medical University*, *Urumqi*, *Xinjiang* 830054, *China*

[4] *Veterinary Research Institute*, *Xinjiang Academy of Animal Sciences*, *Urumqi*, *Xinjiang* 830000, *China*

[5] *Xinjiang Uyghur Autonomous Region Center for Animal Disease Control and Prevention*, *Urumqi*, *Xinjiang* 830011, *China*

ABSTRACT

Rodents play a pivotal role in the maintenance and transmission of zoonotic viruses. The Yili River Valley, one of the most biodiverse regions in Xinjiang, functions as a critical biogeographic corridor linking China with Central and Western Asia, and historically with Europe via the ancient Silk Road. Despite its significance, the viral landscape of this region remains largely unexplored. To elucidate the virological landscape of this understudied region, meta-transcriptomic sequencing was conducted on wild rodent samples collected between 2020 and 2023, encompassing multiple host species and tissue types (liver, lung, spleen, and intestine). Analysis identified 18 vertebrate-associated viral families, including several of known zoonotic or evolutionary relevance, such as *Arteriviridae*, *Coronaviridae*, *Flaviviridae*, *Hantaviridae*, *Hepeviridae*, *Hepadnaviridae*, and *Picornaviridae*. Remarkably, over 80% of the detected viruses represented putative novel species, highlighting the vast and previously undocumented viral diversity harbored by rodents in this region. Viral community composition exhibited clear host- and tissue-specific patterns. Critically, novel rodent-derived arteriviruses (RtArteVs) in *Microtus obscurus* were identified, exhibiting approximately 86% nucleotide identity with porcine reproductive and respiratory syndrome virus 1 (PRRSV1). Additionally, phylogenetic and recombination analyses support the hypothesis that PRRSV1 emerged through ancestral recombination among divergent RtArteVs, implicating rodents as the likely origin of this economically significant swine pathogen. These findings expand current understanding of rodent viromes in an important biodiversity hotspot, revealing a substantial yet largely uncharted viral diversity. Furthermore, this study underscores the critical need for continued surveillance of viral groups with potential for viral spillover to humans and domestic animals, including *Arteriviridae*, *Flaviviridae*, *Hantaviridae*, *Hepeviridae*, and *Hepadnaviridae*.

Keywords: Virome; Rodent; Zoonotic transmission; Virus evolution

INTRODUCTION

Over recent decades, the global incidence of emerging and re-emerging infectious diseases (EIDs) has risen sharply, accompanied by an increasing frequency of cross-species viral transmission at the human-animal interface (Allen et al., 2017; Cleaveland et al., 2007). Notably, over 60% of human pathogens originate from zoonotic sources (Grange et al., 2021; Wiethoelter et al., 2015), with RNA viruses representing a disproportionate share due to their high mutation rates and inherent adaptability across host species. Historical and contemporary examples—ranging from rabies virus, Nipah virus, hantavirus, Ebola virus, H1N1 influenza virus, West Nile

Received: 15 August 2025; Accepted: 15 September 2025; Online: 16 September 2025

Foundation items: This work was supported by the Research on Genetic Evolution Characteristics and Key Techniques for Tracing of Highly Pathogenic Viruses, National Key Research and Development Program (2023YFC2605602), Xinjiang Science and Technology Department's "Tianshan Talents" Training Program-Science and Technology Innovation Team Project (2023TSYCTD0017), and National Natural Science Foundation of China (32370153)

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: alphaleeyp@hotmail.com; chiyu_zhang1999@163.com; wenbaozhang2013@163.com

virus, severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1), and Middle East respiratory syndrome coronavirus—demonstrate the recurrent emergence of high-impact pathogens from animal reservoirs. The recent global dissemination of SARS-CoV-2 and the ensuing COVID-19 pandemic further underscores the crucial need for proactive viral surveillance in wildlife (Holmes, 2024; Wu et al., 2020; Zhou et al., 2020), particularly in reservoir species inhabiting ecological hotspots (Harvey & Holmes, 2022; Parrish et al., 2008), to assess the likelihood of zoonotic transmission.

Rodents, representing the most species-rich order of mammals, comprise over 2 700 species across more than 33 families and occupy a wide range of natural and anthropogenically modified habitats worldwide (Blanga-Kanfi et al., 2009; Huchon et al., 2002; Li et al., 2023). Their ecological ubiquity, population density, and frequent contact with humans and domestic animals amplify the potential for interspecies pathogen exchange and the emergence of rodent-associated infectious diseases (Meerburg et al., 2009). Rodents have long been implicated in the transmission of human pathogens, including those responsible for plague and viral hemorrhagic fever (Meerburg et al., 2009; Yang et al., 2023). In the past two decades, high-throughput sequencing has led to the discovery of diverse rodent-associated viruses, including many with substantial zoonotic potential, such as members of *Hantaviridae*, *Flaviviridae*, *Arenaviridae*, *Reoviridae*, *Picornaviridae*, *Coronaviridae*, and *Togaviridae* (Kane et al., 2024; Li et al., 2023, 2025; Monteiro et al., 2025; Raghwani et al., 2023; Wu et al., 2018, 2021; Zhang et al., 2025). Rodent-borne pathogens can be transmitted to humans via direct routes such as biting, or through indirect mechanisms such as contamination of food and water by excreta. In addition, ectoparasitic arthropods, such as ticks, mites, and fleas, can serve as vectors for rodent-associated pathogens. Collectively, these ecological and epidemiological traits establish rodents as major natural reservoirs for an extensive array of viruses with the capacity to cross species barriers and initiate outbreaks in human populations (Keesing & Ostfeld, 2024).

China encompasses a wide range of ecological zones and supports a rodent fauna of approximately 200 species (Wang, 2003). Although rodent-associated viromes have been partially surveyed in several regions, knowledge of viral composition within rodents remains scarce, particularly in areas with distinctive geographic and ecological features. Viral diversity in wild rodents is influenced not only by host species but also by environmental variables such as elevation, vegetation coverage, and local biodiversity (Bergner et al., 2020; Chen et al., 2023; Rodríguez-Nevado et al., 2018; Tirera et al., 2021). Xinjiang, located in northwestern China, forms a major terrestrial corridor connecting China with Central and Western Asia and extending toward Europe. This expansive region, dominated by deserts, grasslands, and mountain systems, sustains an ecosystem that includes over 80 rodent species (Guzalnur et al., 2014; Zhang et al., 2023). Of note, previous studies have identified the etiological agents of several rodent- or vector-associated zoonoses in Xinjiang, including Crimean-Congo hemorrhagic fever caused by ticks (Guo et al., 2017; Sun et al., 2009; Zhang et al., 2004) and plague caused by *Yersinia pestis*. Despite these findings, the virome associated with wild rodents in this region remains poorly characterized relative to other mammalian reservoirs such as bats, leaving a critical gap in understanding of potential viral threats to human and animal health.

To elucidate the composition, evolutionary dynamics, and transmission patterns of rodent-associated viruses in this understudied region, rodents were systematically sampled from three ecologically distinct habitats in the Yili River Valley, an area recognized for its exceptionally high biodiversity within Xinjiang. Meta-transcriptomic sequencing was performed to comprehensively profile the virome across host species and tissue types, including liver, lung, spleen, and intestine. The resulting dataset delineates the virome landscape circulating within local rodent populations, revealing patterns of host specificity, tissue tropism, and evolutionary divergence across both established and previously unrecognized viral taxa. These findings advance current understanding of virus-host-environment interactions in a region of critical ecological and epidemiological relevance

MATERIALS AND METHODS

Sample collection

From July 2020 to October 2023, wild rodents were systematically captured across three ecologically distinct regions of the Yili River Valley in Xinjiang, including the northern Yili River Valley (G1), southern Yili River Valley (G2), and southern foothills of the Tianshan Mountains (G3). Sampling activities were conducted under a long-term ecological monitoring protocol approved by the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University (Approval No. 20140619-12). Sites were selected to represent diverse habitat types, including grasslands, forested zones, and areas proximal to water bodies, to maximize ecological and taxonomic coverage. In G1 and G2, four sites were selected per subregion, each spaced 20–50 km apart. Live traps were deployed in habitats with high rodent activity. Following capture, animals were maintained at 4°C during transport to nearby field laboratories. Dissections were subsequently performed under ether anesthesia to minimize suffering, and major organs—including liver, lung, spleen, and intestine—were collected. Tissues were immediately preserved in RNAlater (Thermo Fisher, USA) and stored at –80°C, and subsequently transferred to the central laboratory for further analysis. Species identification was initially based on external morphology by experienced field biologists and later confirmed by mapping the cytochrome c oxidase subunit I (*cox1*) and cytochrome b (*cytb*) genes to the NCBI database using BLAST.

RNA extraction, library construction, and sequencing

Tissue samples were pooled by rodent species, tissue type, and sampling location. For each pool, equivalent quantities of the same tissue type from 10–15 individuals of the same species and site were combined. A total of 640 individual tissue samples were processed into 69 pooled libraries (Supplementary Table S1). Each pool was homogenized in 1 mL of QIAzol lysis reagent (Qiagen, Germany) with the aid of stainless-steel beads using a Tissuelyzer system (Qiagen, Germany). Homogenates were cleared by centrifugation at 12 000 ×*g* for 10 min at 4°C. Total RNA was extracted from the supernatant using the RNeasy Plus Universal Mini Kit (Qiagen, Germany) following the manufacturer's instructions. RNA quality was determined using the Agilent 4200 Bioanalyzer (Agilent Technologies, USA) and quantified using a Qubit 4.0 fluorometer (Invitrogen, USA). Prior to processing, ribosomal RNA (rRNA) was removed from each sample pool

using the TIANSEQ Ribosomal RNA Remove Kit (TIANGEN, China). First-strand cDNA synthesis was performed using SuperScript IV Reverse Transcriptase (Invitrogen, USA) with random hexamer primers. Sequencing libraries were prepared using the Fast RNA-seq Lib Prep Kit v2 (ABclonal, China) and quantified on a Qubit 4.0 instrument (Invitrogen, USA). Libraries were sequenced on the NovaSeq platform (Illumina, USA) using paired-end mode with a read length of 150 bp and dual barcoding.

Data processing and bioinformatics analysis

For each sample, rRNA sequences were removed by mapping raw reads against the SILVA rRNA database (Release 138.1) using Bowtie2 (v.2.3.5.1) with the "–end-to-end" and "–fast" parameters. Adapter sequences and low-quality reads were removed using fastp (v.0.20.1). Reads with low sequence complexity or duplicates were discarded using Trimmomatic (v.1.2) with parameters: -lc_entropy=0.5 -lc_dust=0.5. The remaining high-quality, non-rRNA reads were assembled *de novo* into contigs using MEGAHIT (v.1.2.8) with default settings and a minimal contig length of 300 bp. Assembled contigs were then queried against the NCBI non-redundant protein database (nr, downloaded February 2024) using DIAMOND (v.2.0.15) with an e-value threshold of 1×10^{-5} to optimize sensitivity while reducing false positives. Viral contigs were further validated by comparison with the NCBI nt database (downloaded on December 2023) using BLASTN (v.2.11.0). To quantify viral abundance, clean non-rRNA reads were mapped to the assembled viral contigs using Bowtie2 (v.2.3.5.1). Resulting alignment files were sorted and indexed with SAMtools (v.0.1.18), and read counts per contig were calculated using BBMap (v.39.0.1, https://sourceforge.net/projects/bbmap/). Contigs supported by fewer than 10 reads were excluded to minimize false positives. Viral abundance was normalized and expressed as reads per million (RPM).

Vertebrate-associated viral families were identified through manual curation based on literature evidence of host specificity. Host range assignments were cross-referenced with data from the International Committee on Taxonomy of Viruses (ICTV, https://ictv.global) and Virus-Host Database (http://www.genome.jp/virushostdb/). Viral species classification was primarily guided by ICTV-specific demarcation criteria. For viral families lacking defined species thresholds, a provisional threshold of 80% amino acid identity to known viral species was applied.

Phylogenetic and recombination analysis

Putative open reading frames (ORFs) were predicted using ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/), and only viral sequences encoding proteins of at least 400 amino acids were retained for subsequent phylogenetic analysis. Reference protein sequences representing each relevant viral family were downloaded from ICTV. Viral sequences identified in this study were aligned to corresponding family-level reference proteins using MAFFT (v.7.425). Maximum-likelihood phylogenetic trees were constructed using IQ-TREE (v.2.1.4) based on the best-fit substitution model automatically selected by the software. Each tree was inferred from amino acid alignments with 1 000 bootstrap replicates. Trees were visualized using FigTree v.1.4.4 (https://tree.bio.ed.ac.uk/software/figtree/). Recombination analyses were conducted using the Recombination Detection Program v.4 (RDP4). Only recombination events supported by at least four independent algorithms, coupled with statistically significant *P*-values, were retained. To further resolve the evolutionary relationship between rodent-derived arteriviruses (RtArteVs) and porcine reproductive and respiratory syndrome virus (PRRSV), a Bayesian molecular clock analysis was performed on the *Helicase* gene using BEAST v.1.10.4. The analysis employed an uncorrelated lognormal relaxed clock, a GTR+F+G4 nucleotide substitution model, and a constant size tree prior. Two independent Markov chain Monte Carlo (MCMC) runs of 300 million generations were conducted, sampling every 30 000 generations. Parameter convergence was assessed using Tracer v.1.7.1, and a minimum effective sample size (ESS) of 200 was required. After discarding the first 10% of each run as burn-in, a maximum clade credibility (MCC) tree was generated and visualized in FigTree v.1.4.4.

Detection of novel viruses by specific RT-qPCR assays

Five representative viruses, including hantavirus, arterivirus, hepacivirus, hepevirus, and hepatovirus, were selected for confirmation and prevalence assessment across rodent hosts. Total RNA was extracted from individual tissue samples using the RNeasy Plus Universal Mini Kit (Qiagen, Germany). Virus-specific primers and probes were designed based on assembled contigs or full-length genomes. One-step reverse transcription quantitative PCR (RT-qPCR) assays were performed using the PrimerScript™ RT-PCR Kit (Perfect Real Time; Takara, Japan). Each 25 μL reaction contained 10 μL of one-step RT-PCR buffer (2×), 0.2 μL of TaKaRa Ex Taq HS (5 U/μL), 0.4 μL of PrimerScript RT Enzyme Mix II, 0.8 μL of each forward/reverse primer (10 nmol/L), 0.4 μL of probe (10 nmol/L), 3 μL of RNA, and nuclease-free water. Amplification was conducted under the following conditions: 42°C for 5 min, 95°C for 10 s, 40 cycles at 95°C for 5 s, and 63°C for 34 s.

For the detection of filovirus-like sequences, specific primers were designed based on assembled viral contigs: filo_F, 5'-TTTGTCCAACCCCTGCGTAA-3', filo_R, 5'-GCTGAGGATTAGGCAGGACC-3'. RNA was first transcribed into cDNA using SuperScript® III Reverse Transcriptase (Thermo Fisher Scientific, USA), followed by conventional PCR amplification with DreamTaq PCR master mix (Thermo Fisher Scientific, USA) using 3 μL of cDNA as template. Cycling conditions were: 95°C for 5 min, 35 cycles at 95°C for 10 s, 54°C for 30 s, and 72°C for 25 s.

Alpha and beta diversity, transmission network, and statistical analyses

Alpha diversity metrics, including Shannon diversity and Chao1 richness, were calculated to assess viral diversity and species richness within samples. Beta diversity was evaluated by principal coordinates analysis (PCoA) based on Bray-Curtis distances. Inter-group comparisons of continuous variables were performed using the Kruskal-Wallis test with Dunn's *post hoc* correction. PCoA-based group differences were determined using permutational multivariate analysis of variance (PERMANOVA). Host-virus association networks were constructed using the *igraph* package in R and visualized using Gephi (v.0.10.1). In these networks, nodes represent viral clusters defined by >90% nucleotide identity, while edges connecting the host node with a virus node indicate the presence of that virus in the corresponding host. Viral sequence clustering was performed using CD-HIT (v.4.8.1). All statistical analyses were conducted using R v4.2.0 and GraphPad Prism v.8.

RESULTS

Wild rodent sampling and tissue collection

To characterize the rodent-associated virome in Yili, Xinjiang, wild rodents were collected from three different geographical subregions (G1, G2, G3) across the Yili River Valley, a biodiversity-rich area in northwestern China (Figure 1A). Between July 2020 and October 2023, a total of 173 rodents were captured across six counties: Xinyuan, Nileke, Zhaosu, Jinghe, Yining, and Tekesi (Supplementary Table S1). Species identification was initially performed in the field based on morphological features and subsequently confirmed by mitochondrial *cox1* and *cytb* gene sequencing. The sampling encompassed four rodent species from two families, including *M. obscurus* (*n*=105), *Lagurus lagurus* (*n*=22), and *Ellobius talpinus* (*n*=15) within the *Cricetidae* family, and *Urocitellus undulatus* (*n*=30) within the *Sciuridae* family (Figure 1B). For each individual, intestine (including luminal contents), liver, lung, and spleen tissues were dissected under sterile conditions. A total of 640 tissue samples were collected and subsequently pooled into 69 composite samples based on species, tissue type, and geographic origin (Supplementary Table S2). Total RNA was extracted from each pooled sample and subjected to meta-transcriptomic sequencing for virome analysis.

Overview of rodent viromes

A total of 4 679.8 million paired-end reads were generated across all 69 sample pools, yielding an average of 66.9 million reads per pool (range: 30.9–126.4 million). Following quality filtering, 3 836.9 million clean reads were retained (range: 20.6–105.3 million; median: 55.6 million) for downstream assembly and taxonomic classification. From these, 3 936 879 viral reads were identified and assembled into 7 580 viral contigs. These contigs were taxonomically assigned to 59 viral families, including eukaryotic viruses (*n*=39), prokaryotic viruses (*n*=17), and unclassified viruses (*n*=3). To focus on potentially active and transmissible viruses, sequences belonging to *Retroviridae*—primarily representing endogenous retroelements integrated into host genomes—were excluded from further analysis. Eukaryotic viruses (excluding *Retroviridae*) accounted for the majority of viral reads (77.3%), while prokaryotic and unclassified viruses comprised 16.2% and 6.5%, respectively (Figure 2A; Supplementary Figure S1).

Within the eukaryotic virome, 84.3% of viral reads were attributed to vertebrate-associated viruses, while the remaining fraction originated from viruses known to infect invertebrates, plants, or fungi (Supplementary Figure S2). Based on known host associations, 18 vertebrate-related viral families were selected for focused analysis, including 15 RNA and three DNA viral families (Figure 2A; Supplementary Figure S1). *Arteriviridae* and *Hantaviridae* were the most abundant, representing 16.6% and 16.5% of total eukaryotic viral reads, respectively, followed by *Flaviviridae* (15.9%), *Picobirnaviridae* (13.1%), *Astroviridae* (7.8%), *Coronaviridae* (7.5%), *Parvoviridae* (7.3%), *Sedoreoviridae* (6.5%), *Circoviridae* (5.1%), and *Picornaviridae* (2.1%), as well as *Adenoviridae*, *Hepeviridae*, *Hepadnaviridae*, *Caliciviridae*, *Orthoherpesviridae*, *Anelloviridae*, *Filoviridae*, and *Paramyxoviridae* (Supplementary Table S3).

Viral species classification was primarily guided by ICTV demarcation criteria. For viral taxa lacking formal ICTV

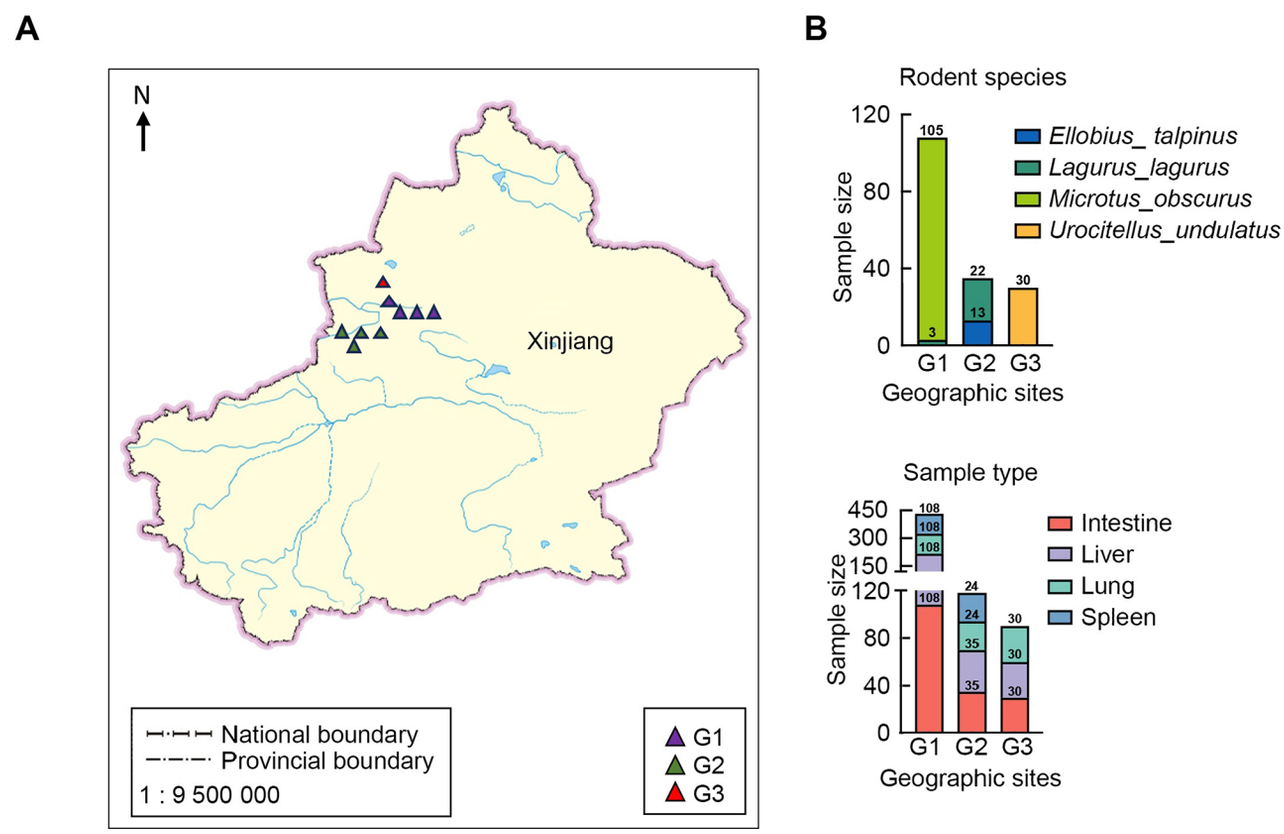

Figure 1 Distribution of sampling locations, host species, and tissue types

A: Map of rodent sampling sites across the Yili River Valley, Xinjiang, China. B: Bar charts showing number of individuals captured at each site, categorized by host species and organ type.

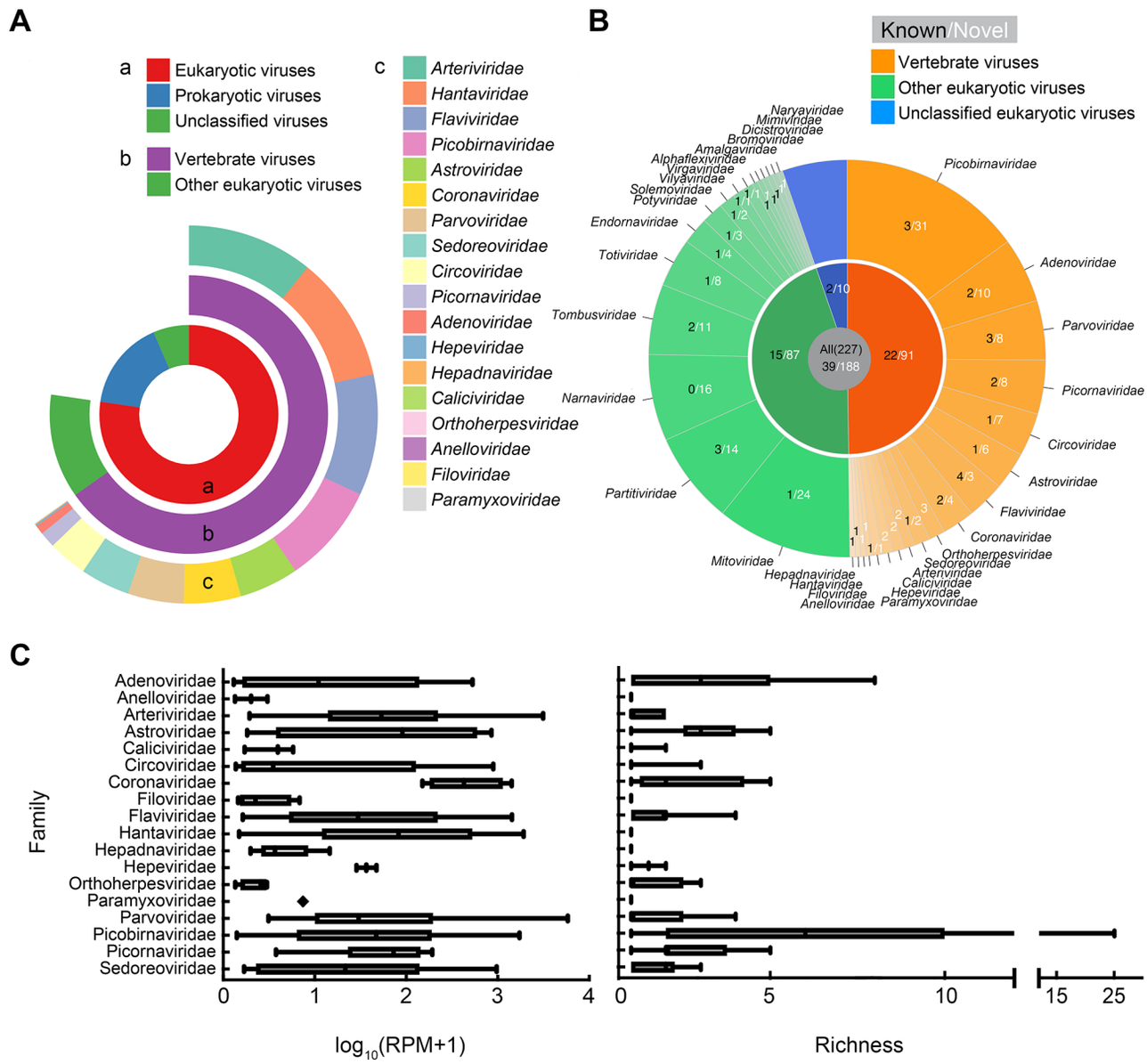

Figure 2 Viral composition of wild rodents from the Yili River Valley

A: Summary of viral categories detected across all samples, including (a) proportions of eukaryotic (excluding *Retroviridae*), prokaryotic, and unclassified viruses; (b) relative abundance of vertebrate-related versus other eukaryotic viruses; and (c) detailed composition of vertebrate-associated viral families. B: Taxonomic classification of all eukaryotic viruses identified. Newly discovered viruses and previously known viruses are labeled in white and black font, respectively. C: Viral load and species richness for each vertebrate-associated viral family across all samples.

thresholds, a provisional criterion of 80% amino acid identity was applied to designate putative novel species (Chen et al., 2023) (Figure 2B; Supplementary Table S4). In total, 227 viral species were identified, 188 (82.8%) of which were predicted to represent previously unrecognized species. Among these, 91 novel species spanning 18 viral families were associated with known vertebrate hosts. These results highlight an exceptionally high level of viral diversity within wild rodent populations. Subsequent analyses focused on the ecological characteristics of vertebrate-associated viruses, especially those with zoonotic potential. These viral families displayed marked differences in viral load and richness, suggesting distinct transmission dynamics and evolutionary histories across host taxa (Figure 2C). *Coronaviridae* exhibited the highest median viral load, followed by *Astroviridae*, *Hantaviridae*, *Picobirnaviridae*, and *Arteriviridae*. Although several families showed relatively high abundance, only *Astroviridae* demonstrated consistently elevated species richness (~3), reflecting considerable genetic heterogeneity. In contrast, most other families exhibited low richness despite high abundance. Notably, *Picobirnaviridae* displayed the highest richness among all viral families, albeit with considerable variability among samples.

Host distribution of vertebrate-associated viruses

Rodent host species harbored distinct virome compositions, reflecting divergent viral community structures. To assess inter-host variation, β-diversity and intra-community parameters, including viral load, species richness, and Shannon diversity, were analyzed. PCoA revealed distinct species-specific clustering patterns, with *M. obscurus* forming discrete clusters that were significantly separated from other rodent hosts (PERMANOVA, P=0.001), indicating strong divergence in virome structure at the viral family level

(Figure 3A; Supplementary Figure S3). Furthermore, *M. obscurus* exhibited the highest overall viral load (RPM), species richness, and α-diversity among all species analyzed (Figure 3B, C; Supplementary Figure S4), contributing 42.8% of all vertebrate-associated viral reads. These findings suggest that the *M. obscurus* population in this region serves as a major reservoir for multiple viruses. Seventeen of the 18 vertebrate-associated viral families (94.4%) were detected in *M. obscurus* (Figure 3D). Among these, *Flaviviridae*, *Arteriviridae*, and *Hantaviridae* were the most prevalent, detected in 25 (78.1%), 19 (59.4%), and 18 (56.3%) of *M. obscurus* libraries, respectively, followed by *Picobirnaviridae*, *Astroviridae*, and *Sedoreoviridae*. In *E. talpinus*, *Circoviridae* was the dominant family, present in 100% of libraries. In *L. lagurus* (gray steppe lemming), *Picobirnaviridae* exhibited the highest prevalence, detected in seven of 20 libraries (35%). *Hepadnaviridae* was the most prevalent viral family in *U. undulatus* (long-tailed ground squirrel), detected in seven of nine libraries (77.7%). Patterns of viral load mirrored those of prevalence. Notably, *M. obscurus* exhibited the highest viral load for *Flaviviridae*, *Arteriviridae*, and *Hantaviridae*, followed by *Picobirnaviridae* and *Astroviridae*. In contrast, *Circoviridae* and *Picobirnaviridae* demonstrated the highest viral load in *E. talpinus* and *L. lagurus*, respectively, while *Parvoviridae* exhibited the highest viral load in *U. undulatus*.

Host specificity was evident in nine viral families (Figure 3E; Supplementary Figure S5). Of these, eight families (*Arteriviridae*, *Hantaviridae*, *Astroviridae*, *Coronaviridae*, *Hepeviridae*, *Orthoherpesviridae*, *Anelloviridae*, and *Paramyxoviridae*) were exclusively detected in *M. obscurus*, while *Hepadnaviridae* was restricted to *U. undulatus*. Conversely, *Parvoviridae* and *Picobirnaviridae* were present across all host species. The remaining seven families were detected in two or three species, suggesting varying degrees of host range and transmission potential across rodent populations.

Organ tropism of vertebrate-associated viruses

To investigate tissue-specific patterns of viral distribution, virome composition and abundance were analyzed across different organs. PCoA revealed distinct organ-level clustering, with intestinal viromes forming separate clusters from those in other tissues (PERMANOVA, P=0.001; Figure 4A). The intestine exhibited higher overall diversity than the liver, lung, and spleen, although the difference was not statistically significant (Figure 4B; Supplementary Figure S6). Additionally, the intestine displayed significantly greater species richness and a higher Shannon diversity index compared to the other three organs, which showed relatively similar viral diversity levels (Figure 4C).

Among the 18 vertebrate-associated viral families, 15 (77.0%) were detected in intestinal samples (Figure 4D, E; Supplementary Figure S6), including *Adenoviridae*, *Caliciviridae*, and *Coronaviridae*, which were exclusively found in this organ. *Picobirnaviridae*, *Parvoviridae*, and *Adenoviridae* were the most prevalent in the intestine, detected in 18 (94.7%), 12 (63.2%), and 10 (55.6%) libraries, respectively. In the liver, *Flaviviridae* was the most frequently detected (9/19 libraries, 47.4%), while *Anelloviridae* was exclusively found in the liver, albeit with low prevalence (2/19, 10.5%). In the lung, *Hantaviridae* and *Flaviviridae* exhibited the highest prevalence, although at a relatively moderate rate (7/17, 41.2%). In the spleen, *Flaviviridae* also emerged as the most dominant, detected in seven of 14 libraries (50%) and found exclusively in this organ.

In terms of viral load, *Picobirnaviridae*, *Astroviridae*, *Coronaviridae*, *Parvoviridae*, and *Sedoreoviridae* exhibited the highest abundance in intestinal samples. *Flaviviridae* was the most abundant viral family in the liver, whereas *Hantaviridae* and *Arteriviridae* predominated in the lung. In the spleen, *Arteriviridae* was most abundant, followed by *Hantaviridae* and *Flaviviridae*.

Viral families exhibited either organ-restricted distribution (monotropism) or multi-organ distribution with preferential enrichment (pantropism). Five monotropic families showed strict organ specificity: *Adenoviridae*, *Caliciviridae*, and *Coronaviridae* were exclusively detected in intestinal tissue, while *Anelloviridae* and *Paramyxoviridae* were restricted to the liver and lung, respectively. Notably, no viral families were uniquely identified in the spleen. These patterns were consistent with known biological traits and transmission routes. For instance, *Adenoviridae* and *Caliciviridae* are predominantly enteric and support fecal-oral transmission, while *Paramyxoviridae* includes respiratory pathogens, consistent with its detection in lung tissue. Although *Anelloviridae* has been reported in multiple organs (Kikuchi et al., 2000; Okamoto et al., 2001), its presence here was limited to the liver, suggesting hepatic tropism in rodents.

Although pantropic families were detected in multiple organs, they exhibited clear organ preferences. *Astroviridae*, *Parvoviridae*, *Picobirnaviridae*, *Picornaviridae*, and *Sedoreoviridae* primarily targeted the intestine (RPM>2 000 in intestine vs. <500 in other tissues). *Flaviviridae* and *Circoviridae* showed elevated viral loads in the liver, consistent with their hepatotropic nature (Li et al., 2019). *Arteriviridae* and *Hantaviridae* demonstrated the highest viral loads in the lung (RPM>1 000), indicating pronounced pulmonary tropism. In contrast, the spleen harbored relatively low viral loads, suggesting it acts primarily as a secondary site reflecting systemic dissemination rather than primary replication. Collectively, these findings underscore the critical role of organ tropism in shaping virome composition and distribution within rodent hosts.

Diversification and evolutionary analyses of rodent-associated viruses

To resolve evolutionary relationships among viruses identified in this study, viral contigs from each sample were subjected to ORF prediction prior to phylogenetic reconstruction (Figure 5). Phylogenetic analyses revealed that rodent populations in the Yili River Valley harbor a diverse array of mammalian viruses with substantial genetic diversity, including lineages of considerable evolutionary significance.

Arteriviridae: Viruses in the family *Arteriviridae* are positive-sense single-stranded RNA viruses with genomes of approximately 12–16 kb and are characterized by strict host specificity and direct transmission in the absence of arthropod vectors. Two previously unrecognized phylogenetic subclades within *Betaarterivirus* were identified in *M. obscurus*, substantially expanding known arterivirus diversity in rodents. Intra-clade amino acid identity across the RNA-dependent RNA polymerase (RdRp) region was highly conserved (98%–100%), while inter-clade identity was slightly lower (97.4%–98.9%). Four near-complete RtArteVs were recovered and showed the closest relationship to porcine reproductive and respiratory syndrome virus 1 (PRRSV1) (GenBank

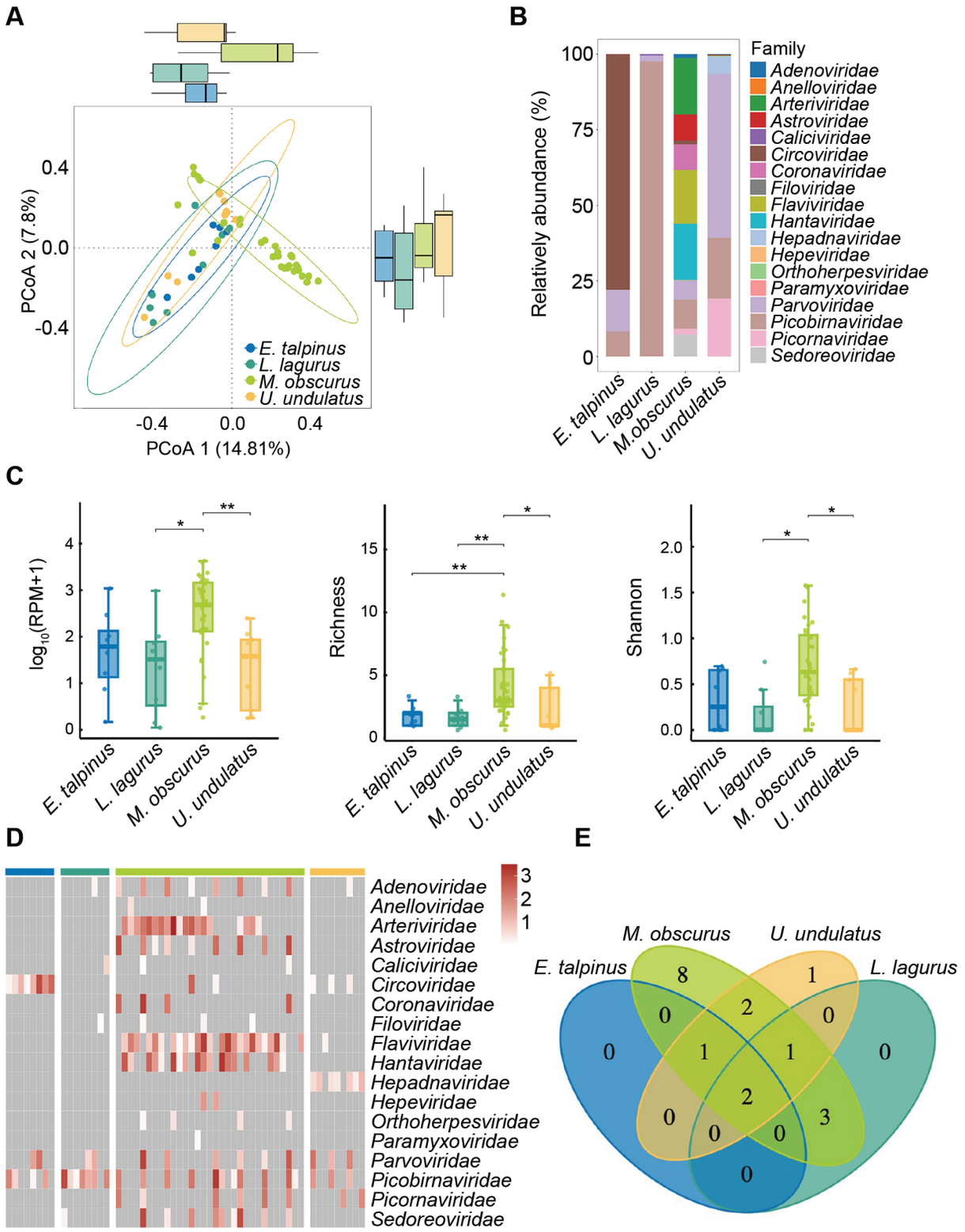

Figure 3 Comparative analysis of vertebrate-associated viruses across host species

A: Principal coordinate analysis (PCoA) illustrating differences in virome composition among host species. Edge bar charts show proportional contributions of each viral group along primary coordinates. B: Overview of relative abundance of vertebrate-associated viruses, categorized by viral family and host species. C: Comparison of viral loads and diversity across different host species, with viral loads expressed as $\log_{10}$(RPM+1) (reads per million of total reads) and diversity evaluated using species richness and Shannon index. Statistical analysis was conducted using Kruskal-Wallis test with Dunn's correction. *: P<0.05; **: P<0.01. D: Summary of viral spectrum at the family level across host species. Each column represents a pooled sample; each row corresponds to a viral family. Heatmap is grouped by host taxon. Viral abundance is expressed as $\log_{10}$(RPM+1). E: Venn diagram showing viral distribution patterns across four host species (*Microtus obscurus*, *Lagurus lagurus*, *Ellobius talpinus*, and *Urocitellus undulatus*).

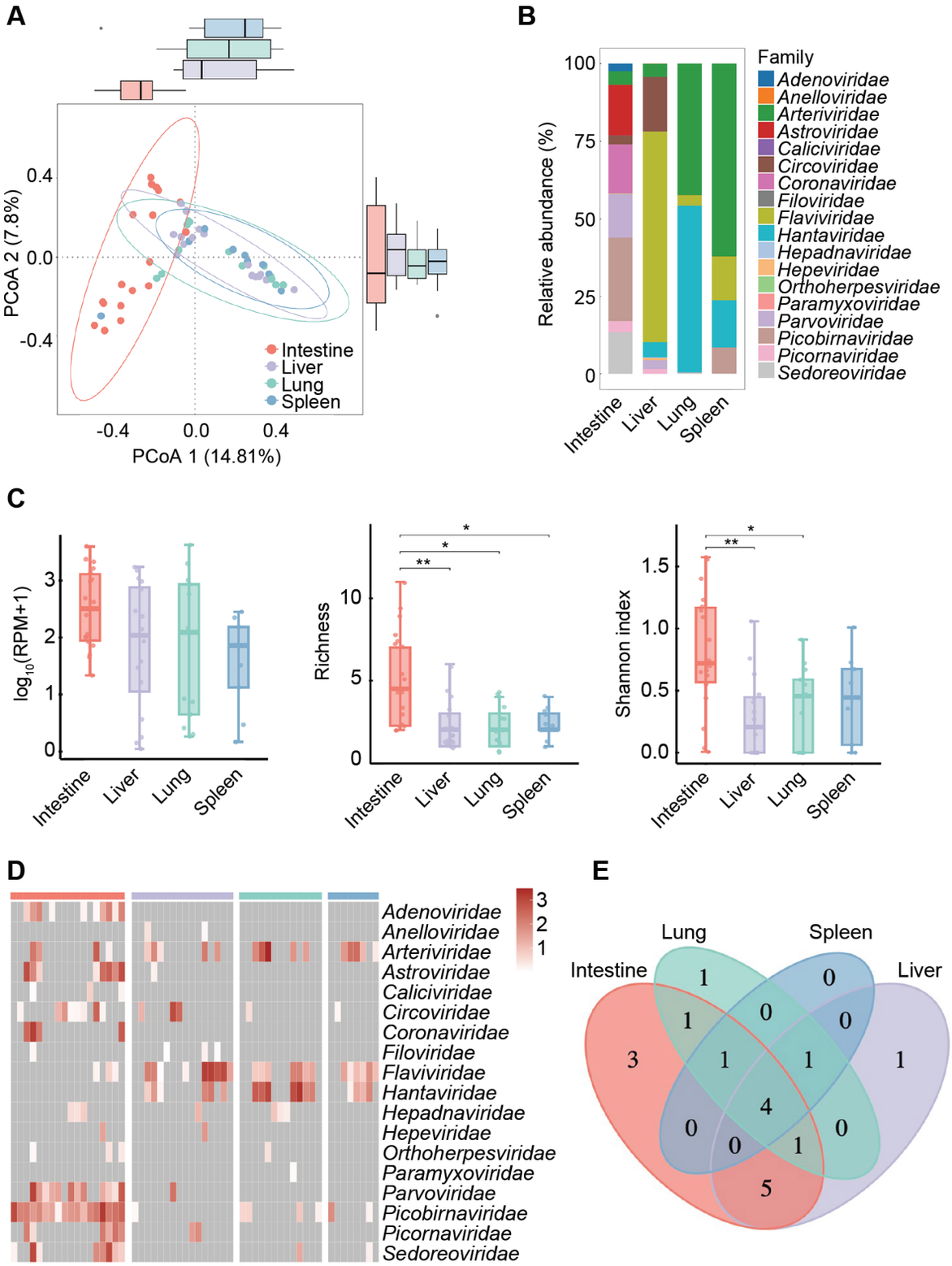

Figure 4 Comparative analysis of vertebrate-associated viruses across different organs

A: Principal coordinate analysis (PCoA) illustrating differences in virome composition among organs. Edge bar charts show proportional contributions of each viral group along primary coordinates. B: Overview of relative abundance of vertebrate-associated viruses, categorized by viral family and organs. C: Comparison of viral loads and diversity across different organs, with viral loads expressed as $\log_{10}$(RPM+1) and diversity evaluated using species richness and Shannon index. Statistical analysis was conducted using Kruskal-Wallis test with Dunn's correction. *: P<0.05; **: P<0.01. D: Summary of viral spectrum at the family level across organs. Each column represents a pooled sample; each row corresponds to a viral family. Heatmap is grouped by organ types. Viral abundance is expressed as $\log_{10}$(RPM+1). E: Venn diagram showing viral distribution patterns across different organs.

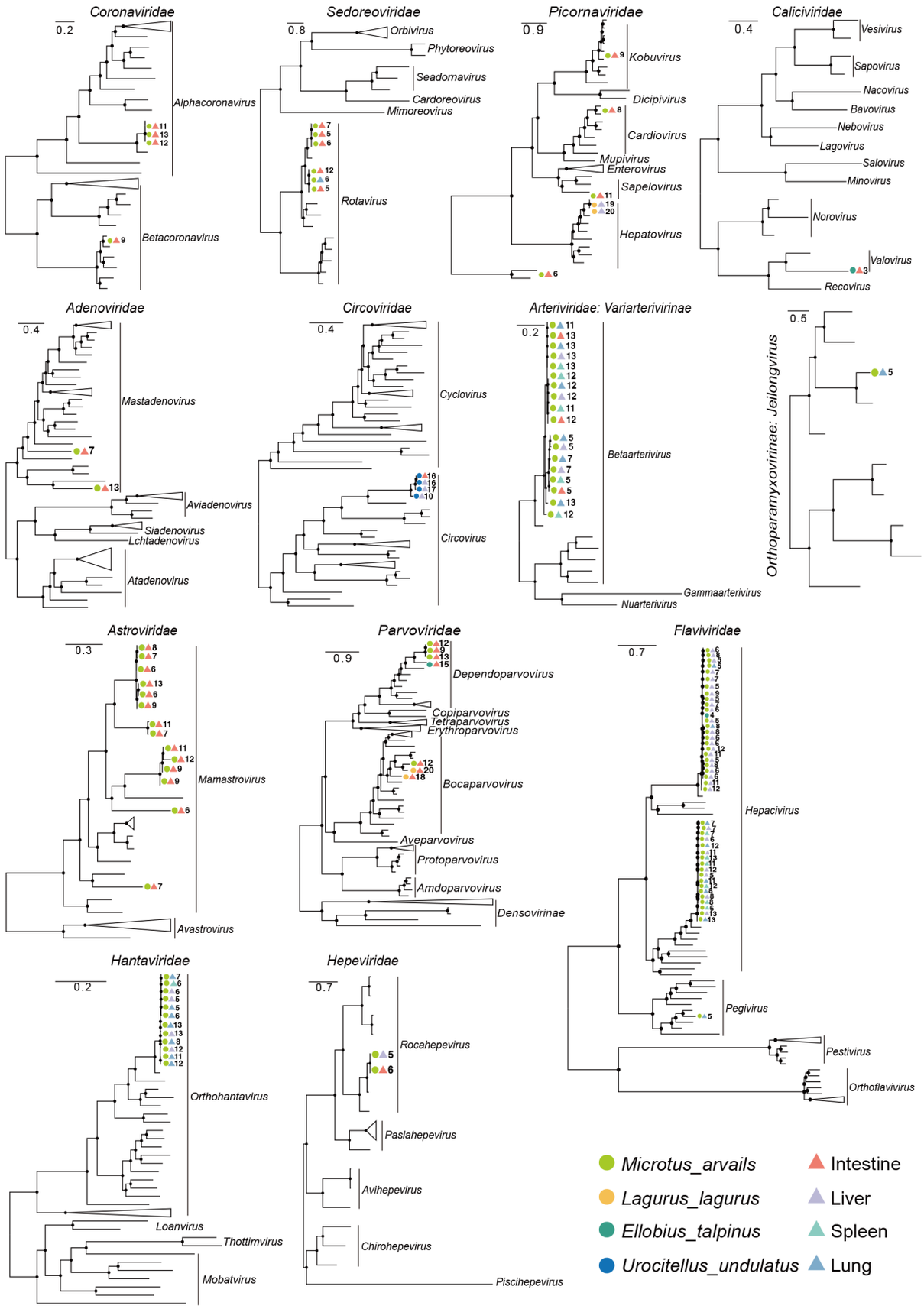

Figure 5 Phylogeny of vertebrate-associated viruses identified in this study

Maximum-likelihood phylogenetic trees were estimated based on amino acid sequences of putative open reading frames (ORFs) of viruses. Several clades corresponding to distinct genera from each tree were collapsed for clarity. Host species and tissue types are annotated with circular and triangular symbols, respectively, followed by the serial number of each sampling pool. All trees are midpoint rooted for clarity, with scale bars representing number of substitutions per site.

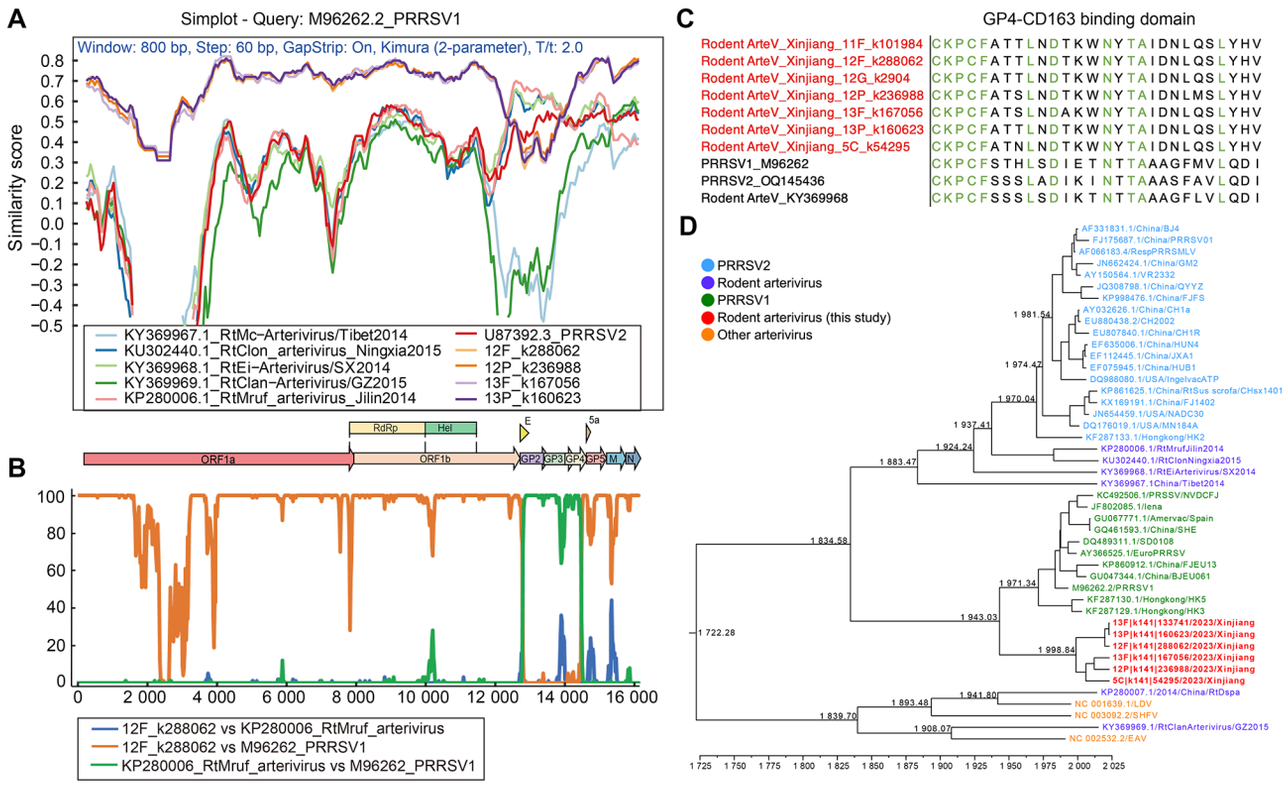

Figure 6 Genomic and evolutionary analyses of rodent arteriviruses (RtArteVs) identified in this study

A: Genome-wide similarity profiles among arteriviruses, with PRRSV1 used as the reference sequence. B: Recombination analysis across near-complete genomes of three representative arterivirus strains. C: Comparative analysis of amino acid variation within the CD163 receptor-binding region. Conserved residues are marked in green. D: Time-calibrated molecular phylogeny inferred from the arterivirus *helicase* gene. Maximum clade credibility (MCC) tree shows divergence times among PRRSV1/2, novel RtArteVs (bold), and historic arteriviruses.

accession No. M96262), sharing 65.9%–66.7% nucleotide identity across the whole genome (Figure 6A). At the RdRp region, these viruses shared 91.8%–92.8% amino acid identity with PRRSV1. In comparison, previously reported RtArteVs shared approximately 77% amino acid identity with PRRSV2 (Zhao et al., 2023). The relatively high genetic similarity between these newly identified RtArteVs and PRRSV1 supports an ancestral relationship. Consistent with this interpretation, genetic divergence between the RtArteVs and PRRSV1 ranged from 0.127 to 0.137, lower than the ICTV-defined interspecies threshold (0.174), indicating that the RtArteVs identified in *M. obscurus* belong to the same viral species as PRRSV1.

Genome-wide similarity profiling confirmed that PRRSV1 exhibited the greatest sequence similarity to these new RtArteVs across most genomic regions, whereas segments encompassing GP2, GP3, and GP4 were more closely related to previously reported RtArteV (KP280006) detected in 2014, suggesting historical recombination events between these viruses (Figure 6B). Formal recombination analysis using RDP corroborated this inference and identified a recombinant region spanning the GP2/GP3/GP4 coding region, supporting an origin of PRRSV1 through recombination among ancestral RtArteVs. Because GP4 mediates interaction with the cellular CD163 receptor during viral entry, comparative analysis of amino acid residues within the receptor-binding domain revealed substantial divergence between the newly identified RtArteV clade and PRRSV1 and PRRSV2 (Figure 6C). Bayesian molecular clock analysis based on the *helicase* gene, together with maximum-likelihood inference using the RdRp gene, consistently placed the novel RtArteVs in close proximity to PRRSV1 (Figures 6D, 7A). The most recent common ancestor (MRCA) of these RtArteVs and PRRSV1 was dated to 1943.0. In contrast, the MRCA of the novel RtArteVs was estimated to be more recent than that of PRRSV1 (1998.8 vs. 1971.3). The MRCA of the newly identified RtArteVs and previously reported RtArteVs was estimated to be 1834.6.

Astroviridae: Astroviruses (AstVs) are small, non-enveloped, single-stranded positive-sense RNA viruses (6.4–7.9 kb) classified into two genera: *Mamastrovirus* and *Avastrovirus* (Zhang et al., 2022). Phylogenetic analysis of sequences from *M. obscurus* intestinal samples revealed three distinct evolutionary clades, each exhibiting intra-clade amino acid identity exceeding 90% (range: 93.5%–99.5%). These clades clustered closely with previously identified rodent AstV (RtAstV) lineages, including strains from *Rattus tanezumi* (PQ678015) and *Microtus agrestis* (MN626434), with >80% amino acid identity in the RdRp region, forming a rodent-specific subclade. One divergent astrovirus from pool #7 shared only 64.2% amino acid identity with a canine astrovirus (WDW25692), indicating it likely represents a novel species. These findings underscore the considerable and underexplored diversity of RtAstVs.

Coronaviridae: Four rodent coronavirus (RtCoV) sequences were assembled from *M. obscurus* intestinal samples. Three clustered within the *Alphacoronavirus* genus, forming a monophyletic clade with 100% pairwise amino acid identity. The RdRp region of these RtCoVs shared 92.9% amino acid identity with Lucheng Rn rat coronavirus (GeneBank: QOE77284), an alphacoronavirus isolated from *Rattus* in Luchang, China (2017). The fourth sequence (RtCoV_9-C)

belonged to the *Betacoronavirus* genus and shared 97.7% amino acid identity with rabbit coronavirus HKU14 (YP_009944253), suggesting possible cross-species transmission between rodents and lagomorphs. In addition, RtCoV_9-C contained a conserved furin cleavage motif (RRXRR) also found in its close relatives, including rabbit coronavirus HKU14, RtMtuf CoV, and human OC43 coronavirus, suggesting a similar evolutionary trajectory (Supplementary Figure S7).

Flaviviridae: Flaviviruses are enveloped, positive-sense single-stranded RNA viruses classified into four genera: *Hepacivirus*, *Flavivirus*, *Pestivirus*, and *Pegivirus*. A total of 43 sequences belonging to the *Hepacivirus* genus and one sequence from the *Pegivirus* genus were identified in *M. obscurus* samples. Phylogenetic analysis revealed that these rodent hepacivirus (RtHV) sequences grouped into two distinct evolutionary clades. The first clade, comprising 25 sequences, exhibited intra-clade amino acid similarity ranging from 94.9% to 98.5% and clustered with hepacivirus_RMU10-3382 (YP_009506360) from *Myodes glareolus*, sharing 88.3% amino acid identity. The second clade demonstrated higher sequence conservation (99.2%–100%) and was most closely related to RtMc-HCV/Xizang2014 (KY370094) from *Neodon clarkei*, with an amino acid identity of 80.7%. The identified rodent pegivirus (RtPegV) sequence shared 89.2% amino acid identity with a strain isolated from *Phaiomys leucurus* (MW897328).

Hantaviridae: Hantaviruses possess segmented, linear, negative-sense or ambisense single-stranded RNA genomes and represent the largest family of negative-strand RNA viruses. While hantaviruses infect a wide range of hosts, spanning vertebrates, invertebrates, and plants, only rodent-borne hantaviruses (RtHanVs) are of clinical significance as zoonotic pathogens in humans (Kuhn & Schmaljohn, 2023; Laenen et al., 2019). All identified hantavirus sequences in this study were exclusively derived from *M. obscurus* tissues and formed a monophyletic clade within the *Orthohantavirus* genus. These RtHanVs exhibited high genetic homogeneity, with intra-clade amino acid identities ranging from 99.3% to 100%. Phylogenetically, this clade clustered with the hantavirus strain TOV.XJ2019 (GenBank accession No. MN183136.1), previously identified in *M. obscurus* from Xinjiang, sharing over 95% amino acid identity in the L segment.

Hepeviridae: Hepeviruses comprise small, non-enveloped, positive single-stranded RNA viruses. The subfamily *Orthohepevirinae* primarily infects mammals and birds and includes four genera: *Paslahepevirus* (humans and various domestic and wild mammals), *Avihepevirus* (birds), *Rocahepevirus* (rodents, shrews, and carnivores), and *Chirohepevirus* (bats). Two rodent hepatitis E virus (RtHEV) genomes were recovered from the intestine and liver of *M. obscurus*, sharing 99.8% pairwise amino acid identity. Both sequences clustered within the *Rocahepevirus* genus and exhibited 92.8%–92.9% amino acid identity with a previously reported hepevirus from *Eothenomys melanogaster* (AVP32824).

Sedoreoviridae: Six viral sequences belonging to *Sedoreoviridae* were assembled from *M. obscurus* intestine and lung samples, all of which clustered within the *Rotavirus* genus. Phylogenetic analysis revealed that these rodent rotaviruses (RtRotaVs) formed two distinct clades. Intra-clade amino acid identity among structural proteins ranged from 99.2% to 99.6%, while inter-clade identity ranged from 67.1% to 71.8%. These RtRotaVs showed limited similarity to established rotavirus sequences, with the closest match being the newly identified rotavirus A from *Myocastor coypu* (ULF48136), sharing 72.1%–82.3% amino acid identity. *Rotavirus A* (RVA) is a well-known pathogen associated with gastroenteritis and growth retardation in mammals and birds, causing significant economic losses in livestock production (Kishimoto et al., 2023). The divergence of these sequences suggests the presence of two novel rotavirus species in *M. obscurus*, potentially adapted to rodent-specific ecological niches.

Picornaviridae: Picornaviruses (PicoVs) exhibited the highest genus-level diversity among all detected families, encompassing four distinct genera, *Kobuvirus* (*n*=1), *Cardiovirus* (*n*=1), *Sapelovirus* (*n*=1), and *Hepatovirus* (*n*=2), as well as one unclassified picorna-like virus (*n*=1). *Hepatovirus* sequences were exclusively detected in the liver of *L. lagurus*, consistent with the hepatotropic nature of this genus. In contrast, sequences from the other genera were identified in intestinal tissues of *M. obscurus*. The kobuvirus exhibited 94.2% amino acid identity to Aichivirus SZAL6-KoV (KJ934637), originally isolated from *Coracias garrulus*. Similarly, the rodent cardiovirus clustered with *Cardiovirus* RtMrut-PicoV (NC_076099), showing 73.7% identity. The sapelovirus sequence displayed phylogenetic proximity to *Sapelovirus*/XZS03 (MW826472) from rat feces, with 58.7% amino acid identity. The novel picorna-like virus shared 59.9% identity with Wufeng rodent picorna-like virus 1 (OQ716149). *Hepatovirus A*, commonly known as hepatitis A virus (HAV), predominantly infects primate enterocytes and hepatocytes and causes acute viral hepatitis in humans (De Oliveira Carneiro et al., 2018). Two rodent hepatitis A virus (RtHAV) sequences clustered with a previously established sequence found in *Marmota monax* (KT229611), sharing 86.0% and 86.7% amino acid identity, respectively.

***Other* viruses**: Other mammalian viruses in these rodents included those from *Adenoviridae* (AdV), *Circoviridae* (CV), *Parvoviridae* (ParV), *Caliciviridae* (CaliV), *Orthoparamyxoviridae* (ParaV), and *Filoviridae* (FiloV). Two novel adenoviruses were identified in *M. obscurus*, each exhibiting substantial genetic divergence from known species. RtAdV_7C-K97534/XJ2023 clustered with red squirrel adenovirus 1 (KY427939), sharing 62.3% amino acid identity, while RtAdV_13C-K193307/XJ2023 showed 50.2% identity with murine adenovirus 2 (NC_014899). Both viruses exceeded the ICTV species demarcation threshold (85%–90% amino acid identity), supporting their classification as novel species.

Four rodent circoviruses (RtCVs) were detected in *U. undulatus*, forming a monophyletic clade within the *Circovirus* genus with intra-clade amino acid identity ranging from 89.9% to 100%. The closest known relative was a rodent circovirus (YP_010084723) from *Neodon clarkei* in Xizang, China, sharing 81.3%–82.3% identity.

Seven rodent parvovirus (RtParV) sequences were identified in intestinal tissues from *M. obscurus*, *L. lagurus*, and *U. undulatus*, including four *Dependoparvovirus* and three *Bocaparvovirus* sequences. *Dependoparvovirus* sequences shared 73.5%–99.3% identity among themselves and 79.6%–94.5% identity with canine parvovirus (WDW25764). Among the *Bocaparvovirus* sequences, one exhibited 99.6% identity to a strain from goats in Hubei, China (AYN07226),

while two others clustered with murine bocavirus (YP_010086823), sharing 69.1%–70.1% identity.

Single viral sequences from *Caliciviridae* and *Orthoparamyxoviridae* were identified in *M. obscurus*. The rodent calivivirus (RtCaliV) exhibited 92.8% amino acid identity to Stoat valovirus 1 (WRV02107), whereas the rodent paramyxovirus (RtParaV) was highly divergent, sharing only 32.1% identity with Wufeng *Eothenomys melanogaster jeilongvirus* 1 (UOL48909). Many of these sequences showed <80% amino acid identity to known viruses, underscoring a high degree of viral novelty in wild rodent populations.

Filoviridae sequences were detected in four libraries—three from *M. obscurus* and one from *L. lagurus*. Although the assembled sequences were short (<800 bp), viral RNA was confirmed by PCR in three libraries (Supplementary Figure S8). Subsequent Sanger sequencing and BLASTx analysis indicated that these sequences were most related (~48% identity) to the VP35 protein of *Cuevavirus* within *Filoviridae*, suggesting the presence of a novel filovirus in these rodent hosts.

Prevalence of rodent viruses with zoonotic or veterinary relevance

Rodents serve as significant reservoirs for a diverse array of viruses with potential for cross-species transmission. To assess their prevalence in wild populations, targeted qPCR screening was performed for five evolutionarily significant vertebrate-related viruses, including RtHanV, RtArteV, RtHV, RtHEV, RtHAV (Figure 7B). RtHV exhibited the highest prevalence (35.9%), followed by RtHanV (24.0%), with the remaining three viruses showing prevalence rates ranging from 14.4% to 21.6%. These results indicate substantial circulation of viruses with known or suspected zoonotic potential in wild rodent populations.

Host specificity and potential interspecies transmission

Analysis of vertebrate-related viruses at the species level revealed limited cross-species transmission among rodent hosts (Figure 8). Eight viral species were detected in more than one host species: six were shared between two species, and two were transmitted across three species. Most of these shared viruses (7/8) involved *M. obscurus*, highlighting its central role in viral maintenance and potential transmission. In contrast, viruses with greater zoonotic concern, such as RtHanV, RtArteV, RtHV, RtHEV, and RtHAV, were each confined to a single host species (*M. obscurus* or *U. undulatus*), with no evidence of cross-species transmission at the species level. This pattern implies constrained host ranges for these viruses under current ecological conditions.

DISCUSSION

Rodents, comprising the most taxonomically diverse order of mammals, represent key natural reservoirs for a wide spectrum of zoonotic viruses worldwide (Han et al., 2015). To systematically characterize the virome landscape of wild rodents in the Yili River Valley—a globally recognized hotspot of biodiversity—meta-transcriptomic sequencing was performed on multi-organ samples from four dominant rodent species. This investigation yielded the first region-wide survey of viral diversity among wild rodent species in Xinjiang, revealing extensive viral richness, organ-specific tropism, and phylogenetic diversity. In total, 227 vertebrate-related viral species were identified, 188 (82.8%) of which were putative novel lineages spanning 18 vertebrate-associated viral families. The majority comprised RNA viruses known to circulate in rodents, including RtHanV, RtHV, RtPegV, RtHEV, RtArteV, RtAstV, RtRotaV, RtCoVs, RtPicoV, RtParaV, and RtCaliV, as well as several DNA viruses (RtAdV, RtCV, and RtParV) (Li et al., 2023; Wu et al., 2021), expanding known viral diversity within these families.

Among the surveyed hosts, *M. obscurus* harbored the greatest virome diversity and abundance, with 17 of the 18 vertebrate-associated viral families (94.7%) detected in its tissues. This species supported replication of multiple high-priority zoonotic candidates, including RtHanV, RtArteV, RtHV, and RtHEV, underscoring its role as a core viral reservoir in this ecosystem and the risk of cross-species transmission to other mammals and humans. Nearly half of the identified families were restricted to a single rodent species, with only a limited number of viral species shared across different hosts, suggesting strict host tropism and significant cross-species transmission barriers. However, the apparently low rate of cross-species transmission may also reflect limited analytical sensitivity, potentially leading to an underestimation of viruses occurring at low abundance.

Different viruses preferentially replicate in primary or, in some cases, distinct organs, reflecting both their evolutionary origins and preferred transmission pathways. In this study, intestinal tissues, including luminal contents, harbored the greatest diversity and abundance of viruses, likely due to frequent dietary exposure to exogenous viral agents from diverse environmental sources. Numerous viruses demonstrated broad tissue tropism, with several detected across all four sampled organs (e.g., RtArteV, RtHV, and RtCV), although with differential abundance. These patterns suggest that certain viruses may disseminate or spill over from their primary site of replication to other organs when systemic viral loads are elevated—for instance, RtHV was most enriched in liver tissues, while RtArteV predominated in the lung (Chen et al., 2023). In contrast, several viruses displayed strict organ specificity, such as RtAdV, RtCaliV, and RtCoV in the intestine and RtParaV in the lung, consistent with their known tropisms and transmission routes (Jagirdhar et al., 2023).

Most identified viral sequences exhibited limited similarity to previously characterized taxa, highlighting the extensive and underexplored viral diversity circulating in rodent populations in Xinjiang. Notably, a distinct phylogenetic clade of novel RtArteVs closely related to PRRSV1 was identified in wild rodents, with higher sequence identity to PRRSV1 than to PRRSV2 or any other known arterivirus. This finding represents the first report of PRRSV1-related viruses in wild rodents. PRRSV1 is predominantly found in Europe, whereas PRRSV2 is prevalent in Asia and North America (Nelsen et al., 1999; Shi et al., 2010). PRRSV infection in pigs leads to severe reproductive disorders, including abortion, stillbirth, and fetal mummification, as well as respiratory pathology such as pneumonia and dyspnea, resulting in substantial economic loss to the global swine industry (Qiu et al., 2025; Sun et al., 2023). The detection of PRRSV1-related viruses in Xinjiang raises the possibility that PRRSV1 may have originated from ancestral RtArteVs in this region, with subsequent transmission along historical trade networks linking Central Asia and Europe, such as the ancient Silk Road. These results support the hypothesis that wild rodents constitute a natural reservoir of arteriviruses, analogous to the role of bats in

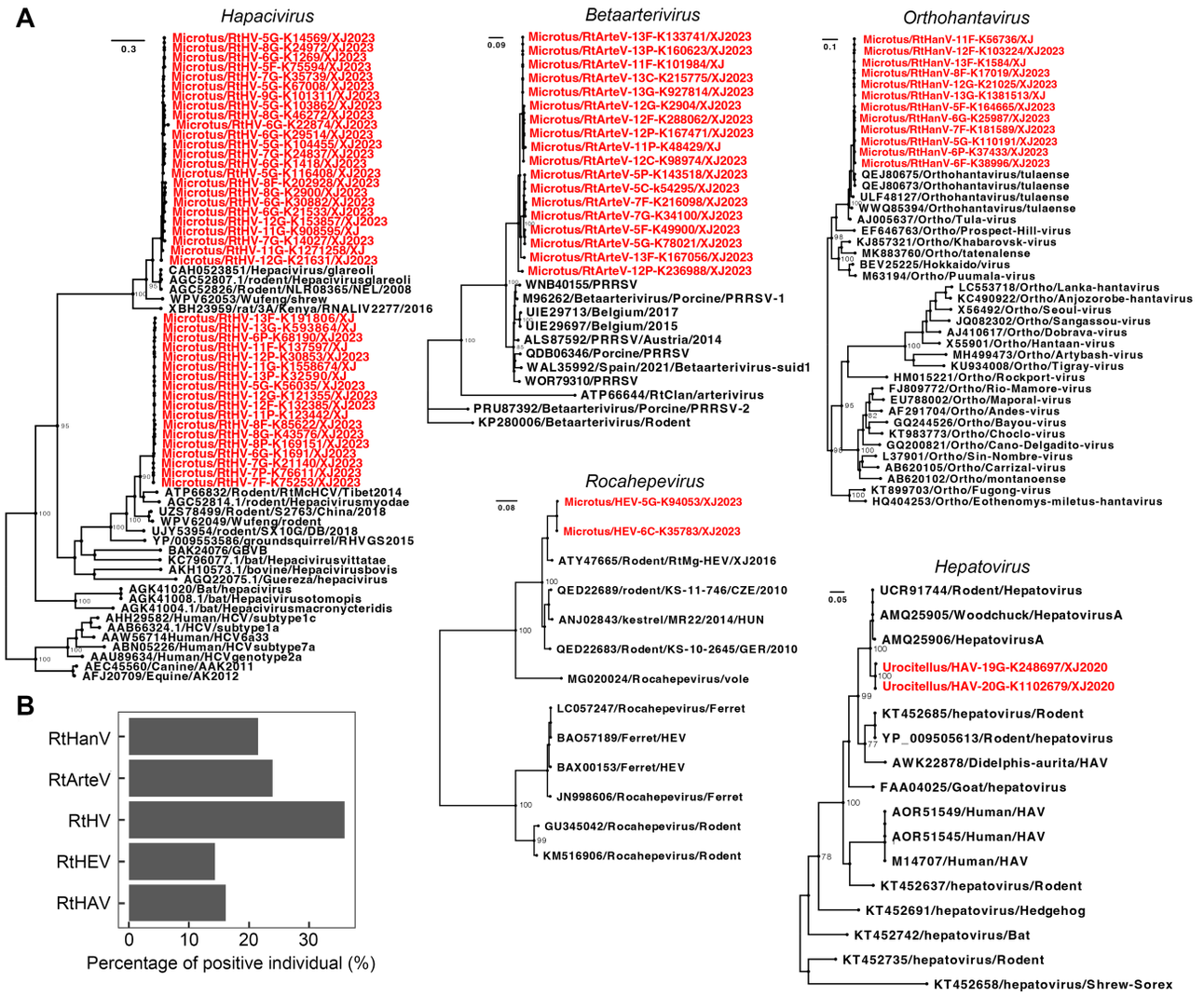

Figure 7 Evolutionary relationships and prevalence of viruses with zoonotic or veterinary relevance

A: Maximum-likelihood phylogenetic trees (1 000 bootstrap replicates) of rodent hantaviruses (RtHanVs), arteriviruses (RtArteVs), hepacivirus (RtHV), hepatitis E virus (RtHEV), and hepatitis A virus (RtHAV) sequences. Viruses identified in this study are marked in red; black-labeled sequences represent reference genomes and the closest related viruses from public databases. B: Prevalence of the five viral genera among different rodent species. Virus detection was performed using quantitative PCR (qPCR) on individual tissue samples.

coronavirus ecology. Furthermore, recombination events may have contributed to the emergence of PRRSV1-related viruses and their subsequent cross-species transmission and establishment in pig populations. Although the newly identified arteriviruses share limited sequence similarity with PRRSV1 and PRRSV2 at the receptor-binding domain, recombination within this region raises important questions regarding its potential role in enhancing viral infectivity, host adaptation, or interspecies transmission. Considering the significant economic burden imposed by arterivirus infections in swine, further investigation is critically needed to assess the prevalence of these novel RtArteVs in local animal populations and to evaluate their capacity for cross-species transmission to pigs.

Rodents represent key reservoirs for RtHanVs, many of which are known human pathogens (Manigold & Vial, 2014; Vial et al., 2023). The hantavirus sequences detected in this study showed high similarity to *Orthohantavirus tulaense* (TULV), previously reported in *Microtus* species, suggesting a high degree of host specificity within this genus. While most human HEV infections are caused by HEV-A, rat HEV-C has been confirmed as zoonotic and capable of causing persistent hepatitis in humans despite its considerable genetic divergence (Firth et al., 2014; Sridhar et al., 2021, 2018). Occupational exposure to swine and poultry is associated with elevated HEV infection risk (Wu et al., 2022), and the high prevalence of RtHanV and RtHEV in wild rodents raises concerns about potential spillover to humans, especially in agricultural settings where close contact with rodents is common. Both RtHEV and RtHAV are homologs of human HEV and HAV, respectively. Although direct zoonotic transmission of RtHEV or RtHAV has not been documented, these viruses possess deep evolutionary histories and may represent ancestral lineages of human-infecting hepatitis viruses (Drexler et al., 2015; Pybus & Thézé, 2016; Scheel et al., 2015; Smith & Simmonds, 2018). Their genetic diversity and evolutionary dynamics in wild rodents offer important insights into viral evolution, maintenance, and transmission in natural reservoirs, and may inform future assessments of zoonotic emergence potential (Mollentze et al., 2021).

Comparative analyses of rodent viromes across recent studies reveal significant geographical heterogeneity in viral

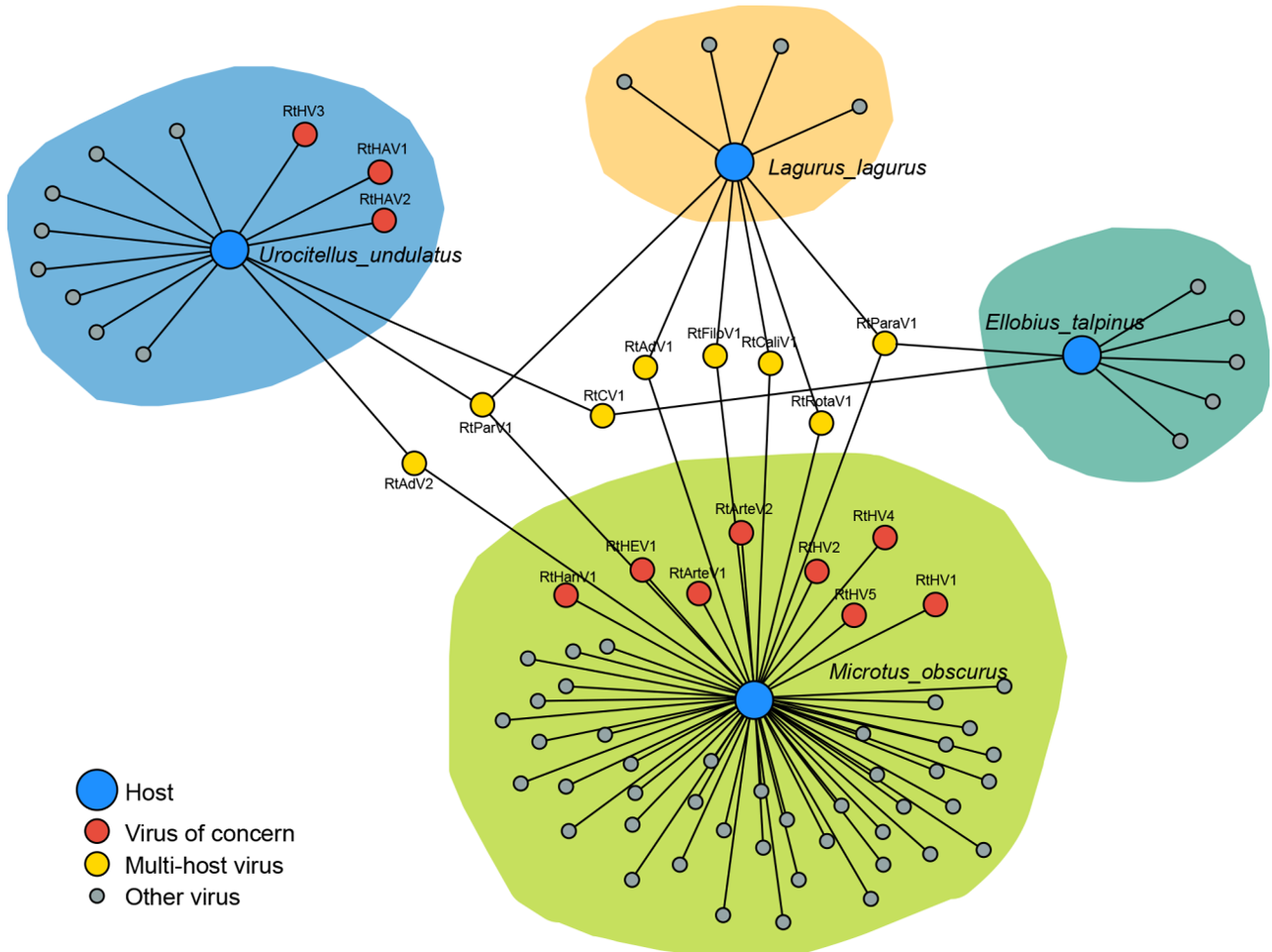

Figure 8 Transmission network of viral species across rodent hosts

Network shows host range and viral species shared among different rodent hosts. Blue nodes represent four different hosts; small nodes represent viral species. Viruses of concern, including rodent hantaviruses (RtHanVs), arteriviruses (RtArteVs), hepacivirus (RtHV), hepatitis E virus (RtHEV), and hepatitis A virus (RtHAV), are colored in red; viruses shared by different hosts are in yellow. Additional detected viruses include rodent *Adenoviridae* virus (RtAdV), *Parvoviridae* virus (RtParV), *Circoviridae* virus (RtCV), *Caliciviridae* virus (RtCaliV), rotavirus (RtRotaV), *Filoviridae* virus (RtFiloV), and *Orthoparamyxoviridae* virus (RtParaV).

family composition. While zoonotic *Arenaviridae* and *Phenuiviridae*, including pathogens such as Lassa fever virus and Rift Valley fever virus, have been frequently detected in rodents from Central China (Zhang et al., 2025) and Southeast Asia (Wu et al., 2021), neither was identified in the present study. Rodents from other regions of Xinjiang are predominantly associated with *Anelloviridae*, *Arenaviridae*, *Nairoviridae*, *Peribunyaviridae*, and *Parvoviridae*, but not *Arteriviridae*, *Flaviviridae*, *Hantaviridae*, and *Hepeviridae* (Tan et al., 2019). In tropical Hainan, rodents harbor *Arenaviridae* but not *Arteriviridae* or *Hantaviridae* (Li et al., 2023). Similarly, *Arteriviridae*, *Hantaviridae*, and *Hepeviridae* have not been detected (or only found at low abundance) in rodents from Yunnan and Guangdong (Kane et al., 2024; Zhuang et al., 2023). Rodents inhabiting the Xizang Plateau host *Arteriviridae*, *Flaviviridae*, and *Coronaviridae* but not *Arenaviridae*, *Phenuiviridae*, or *Hantaviridae* (He et al., 2022). Such variation likely reflects differences in host species composition, sample types, and ecological or geographical characteristics specific to each region. Additional contributing factors, including population density, habitat structure, and regional biodiversity gradients, also play a crucial role in shaping viral diversity. These findings underscore the need for expanded, systematic virome surveillance across diverse regions and taxa to improve understanding of viral ecology and potential spillover dynamics.

Several limitations should be acknowledged. First, despite the ecological significance of the Yili River Valley, the current sampling effort covered only a small proportion of the rodent population and habitat types, limiting generalizability. Second, although the study targeted ecologically dominant rodent species, this approach excluded many of the over 80 rodent species documented in Xinjiang, potentially biasing virome representation. Broader species coverage in future investigations will be essential to resolve the full spectrum of viral diversity. Third, the pooling of tissues from multiple individuals prior to sequencing likely reduced sensitivity for low-abundance viruses and precluded fine-scale analyses of viral evolution, transmission, and prevalence in individual animals. Lastly, viral diversity and ecology arise from complex interactions among host species, intermediate vectors or dietary exposures (e.g., fleas, ticks, and mites), and the surrounding environment. Comprehensive sampling that integrates these ecological components will not only improve characterization of virome diversity but also elucidate the ecological drivers and evolutionary processes that govern viral

emergence, maintenance, and transmission across interconnected niches.

CONCLUSIONS

This study demonstrated that wild rodents inhabiting the pasture ecosystems of the Yili River Valley in Xinjiang harbor an extensive diversity of viruses, with more than 80% representing previously uncharacterized species. The high prevalence of viruses with significant zoonotic potential, including hantavirus, hepacivirus, and the first identification of PRRSV1-related arteriviruses in wild rodents, underscores the importance of continuous monitoring and further investigations to assess the risk of cross-species transmission to humans and domestic animals. In addition, deeper understanding of the ecological and evolutionary drivers of viral emergence and transmission will be critical for anticipating and mitigating future outbreaks of infectious diseases.

DATA AVAILABILITY

Raw sequencing data were deposited in the Sequence Read Archive (SRA) (BioProjectID PRJNA1247200; BioSample No. SAMN47806273 to SAMN47806341, and SRA No. SRR33120695 to SRR33120763), Science Data Bank (Data DOI: 10.57760/sciencedb.j00139.00261), and Genome Sequence Archive (accession number PRJCA046616).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y.P.L., W.B.Z., and C.Y.Z. participated in the study design; G.W.Z., Y.Z., J.L., W.J.Q., C.C.W., and M.X.T. collected the samples; Y.Y.M., Y.P.L., Z.Z.Z., M.Z.Y., and M.A. performed the main experiments; L.C. and Y.Y.M. performed data analyses; L.C., Y.Y.M., and Y.P.L. interpreted the data; L.C. and Y.P.L. wrote the draft manuscript. C.Y.Z. and W.B.Z. contributed to manuscript revision. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We would like to thank Prof. Wen Zhang from Jiangsu University for his helpful suggestions regarding data interpretation and analysis.

REFERENCES

Allen T, Murray KA, Zambrana-Torrelio C, et al. 2017. Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications*, **8**(1): 1124.

Bergner LM, Orton RJ, Benavides JA, et al. 2020. Demographic and environmental drivers of metagenomic viral diversity in vampire bats. *Molecular Ecology*, **29**(1): 26–39.

Blanga-Kanfi S, Miranda H, Penn O, et al. 2009. Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evolutionary Biology*, **9**(1): 71.

Chen YM, Hu SJ, Lin XD, et al. 2023. Host traits shape virome composition and virus transmission in wild small mammals. *Cell*, **186**(21): 4662–4675. e12.

Cleaveland S, Haydon DT, Taylor L. 2007. Overviews of pathogen emergence: which pathogens emerge, when and why?. *Current Topics in Microbiology and Immunology*, **315**: 85–111.

De Oliveira Carneiro I, Sander AL, Silva N, et al. 2018. A novel marsupial hepatitis a virus corroborates complex evolutionary patterns shaping the Genus *Hepatovirus*. *Journal of Virology*, **92**(13): e00082–18.

Drexler JF, Corman VM, Lukashev AN, et al. 2015. Evolutionary origins of hepatitis A virus in small mammals. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(49): 15190–15195.

Firth C, Bhat M, Firth MA, et al. 2014. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal *Rattus norvegicus* in New York City. *mBio*, **5**(5): e01933–14.

Grange ZL, Goldstein T, Johnson CK, et al. 2021. Ranking the risk of animal-to-human spillover for newly discovered viruses. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(15): e2002324118.

Guo R, Shen S, Zhang YF, et al. 2017. A new strain of Crimean-Congo hemorrhagic fever virus isolated from Xinjiang, China. *Virologica Sinica*, **32**(1): 80–88.

Guzalnur Z, Mahmut H, Omar A, et al. 2014. Comparison of the rodent species composition between Kyrgyzstan and Xinjiang, China. *Journal of Arid Land Resources and Environment*, **28**: 94–98.

Han BA, Schmidt JP, Bowden SE, et al. 2015. Rodent reservoirs of future zoonotic diseases. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(22): 7039–7044.

Harvey E, Holmes EC. 2022. Diversity and evolution of the animal virome. *Nature Reviews Microbiology*, **20**(6): 321–334.

He XZ, Wang X, Fan GH, et al. 2022. Metagenomic analysis of viromes in tissues of wild Qinghai vole from the eastern Tibetan Plateau. *Scientific Reports*, **12**(1): 17239.

Holmes EC. 2024. The emergence and evolution of SARS-CoV-2. *Annual Review of Virology*, **11**(1): 21–42.

Huchon D, Madsen O, Sibbald MJJB, et al. 2002. Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. *Molecular Biology and Evolution*, **19**(7): 1053–1065.

Jagirdhar GSK, Pulakurthi YS, Chigurupati HD, et al. 2023. Gastrointestinal tract and viral pathogens. *World Journal of Virology*, **12**(3): 136–150.

Kane Y, Tendu A, Li RY, et al. 2024. Viral diversity in wild and urban rodents of Yunnan Province, China. *Emerging Microbes & Infections*, **13**(1): 2290842.

Keesing F, Ostfeld RS. 2024. Emerging patterns in rodent-borne zoonotic diseases. *Science*, **385**(6715): 1305–1310.

Kikuchi K, Miyakawa H, Abe K, et al. 2000. Indirect evidence of TTV replication in bone marrow cells, but not in hepatocytes, of a subacute hepatitis/aplastic anemia patient. *Journal of Medical Virology*, **61**(1): 165–170.

Kishimoto M, Kajihara M, Tabata K, et al. 2023. Isolation and characterization of distinct rotavirus A in bat and rodent hosts. *Journal of Virology*, **97**(1): e0145522.

Kuhn JH, Schmaljohn CS. 2023. A brief history of bunyaviral family *Hantaviridae*. *Diseases*, **11**(1): 38.

Laenen L, Vergote V, Calisher CH, et al. 2019. *Hantaviridae*: current classification and future perspectives. *Viruses*, **11**(9): 788.

Li FL, Wang B, Han PY, et al. 2025. Identification of novel rodent and shrew orthohepeviruses sheds light on hepatitis E virus evolution. *Zoological Research*, **46**(1): 103–121.

Li LL, Liu MM, Shen S, et al. 2019. Detection and characterization of a novel hepacivirus in long-tailed ground squirrels (*Spermophilus undulatus*) in China. *Archives of Virology*, **164**(9): 2401–2410.

Li YY, Tang CN, Zhang Y, et al. 2023. Diversity and independent evolutionary profiling of rodent-borne viruses in Hainan, a tropical island of China. *Virologica Sinica*, **38**(5): 651–662.

Manigold T, Vial P. 2014. Human hantavirus infections: epidemiology, clinical features, pathogenesis and immunology. *Swiss Medical Weekly*, **144**(1112): w13937.

Meerburg BG, Singleton GR, Kijlstra A. 2009. Rodent-borne diseases and

their risks for public health. *Critical Reviews in Microbiology*, **35**(3): 221–270.

Mollentze N, Babayan SA, Streicker DG. 2021. Identifying and prioritizing potential human-infecting viruses from their genome sequences. *PLoS Biology*, **19**(9): e3001390.

Monteiro AFM, Da Silva FS, Cruz ACR, et al. 2025. Viral diversity in wild rodents in the regions of Canaã de Carajás and Curionopólis, State of Pará, Brazil. *Frontiers in Microbiology*, **15**: 1502462.

Nelsen CJ, Murtaugh MP, Faaberg KS. 1999. Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. *Journal of Virology*, **73**(1): 270–280.

Okamoto H, Nishizawa T, Takahashi M, et al. 2001. Heterogeneous distribution of TT virus of distinct genotypes in multiple tissues from infected humans. *Virology*, **288**(2): 358–368.

Parrish CR, Holmes EC, Morens DM, et al. 2008. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews*, **72**(3): 457–470.

Pybus OG, Thézé J. 2016. Hepacivirus cross-species transmission and the origins of the hepatitis C virus. *Current Opinion in Virology*, **16**: 1–7.

Qiu YJ, Qiu M, Li SB, et al. 2025. Emergence, prevalence and evolution of porcine reproductive and respiratory syndrome virus 1 in China from 1994 to 2024. *Virology*, **605**: 110457.

Raghwani J, Faust CL, François S, et al. 2023. Seasonal dynamics of the wild rodent faecal virome. *Molecular Ecology*, **32**(17): 4763–4776.

Rodríguez-Nevado C, Lam TTY, Holmes EC, et al. 2018. The impact of host genetic diversity on virus evolution and emergence. *Ecology Letters*, **21**(2): 253–263.

Scheel TKH, Simmonds P, Kapoor A. 2015. Surveying the global virome: identification and characterization of HCV-related animal hepaciviruses. *Antiviral Research*, **115**: 83–93.

Shi M, Lam TTY, Hon CC, et al. 2010. Molecular epidemiology of PRRSV: a phylogenetic perspective. *Virus Research*, **154**(1-2): 7–17.

Smith DB, Simmonds P. 2018. Classification and genomic diversity of enterically transmitted hepatitis viruses. *Cold Spring Harbor Perspectives in Medicine*, **8**(9): a031880.

Sridhar S, Yip CCY, Wu SS, et al. 2018. Rat hepatitis E virus as cause of persistent hepatitis after liver transplant. *Emerging Infectious Diseases*, **24**(12): 2241–2250.

Sridhar S, Yip CCY, Wu SS, et al. 2021. Transmission of rat hepatitis E virus infection to humans in Hong Kong: a clinical and epidemiological analysis. *Hepatology*, **73**(1): 10–22.

Sun Q, Xu H, An TQ, et al. 2023. Recent progress in studies of porcine reproductive and respiratory syndrome virus 1 in China. *Viruses*, **15**(7): 1528.

Sun SR, Dai X, Aishan M, et al. 2009. Epidemiology and phylogenetic analysis of crimean-congo hemorrhagic fever viruses in xinjiang, China. *Journal of Clinical Microbiology*, **47**(8): 2536–2543.

Tan ZZ, Yu HJ, Xu L, et al. 2019. Virome profiling of rodents in Xinjiang Uygur Autonomous Region, China: isolation and characterization of a new strain of Wenzhou virus. *Virology*, **529**: 122–134.

Tirera S, De Thoisy B, Donato D, et al. 2021. The influence of habitat on viral diversity in neotropical rodent hosts. *Viruses*, **13**(9): 1690.

Vial PA, Ferrés M, Vial C, et al. 2023. Hantavirus in humans: a review of clinical aspects and management. *The Lancet Infectious Diseases*, **23**(9): e371–e382.

Wang YX. 2003. A Complete Checklist of Mammal Species and Subspecies in China: A Taxonomic and Geographic Reference. Beijing: China Forestry Publishing House. (in Chinese)

Wiethoelter AK, Beltrán-Alcrudo D, Kock R, et al. 2015. Global trends in infectious diseases at the wildlife-livestock interface. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(31): 9662–9667.

Wu F, Zhao S, Yu B, et al. 2020. A new coronavirus associated with human respiratory disease in China. *Nature*, **579**(7798): 265–269.

Wu JY, Lau EHY, Lu ML, et al. 2022. An occupational risk of hepatitis E virus infection in the workers along the meat supply chains in Guangzhou, China. *One Health*, **14**: 100376.

Wu ZQ, Du J, Lu L, et al. 2018. Detection of hantaviruses and arenaviruzses in three-toed jerboas from the Inner Mongolia Autonomous Region, China. *Emerging Microbes & Infections*, **7**(1): 1–3.

Wu ZQ, Han YL, Liu B, et al. 2021. Decoding the RNA viromes in rodent lungs provides new insight into the origin and evolutionary patterns of rodent-borne pathogens in Mainland Southeast Asia. *Microbiome*, **9**(1): 18.

Yang RF, Atkinson S, Chen ZQ, et al. 2023. *Yersinia pestis* and plague: some knowns and unknowns. *Zoonoses*, **3**(1): 5.

Zhang FY, Li Y, Jiang WM, et al. 2022. Surveillance and genetic diversity analysis of avian astrovirus in China. *PLoS One*, **17**(2): e0264308.

Zhang NL, Hu B, Zhang L, et al. 2025. Virome landscape of wild rodents and shrews in Central China. *Microbiome*, **13**(1): 63.

Zhang P, Zhang Y, Cao L, et al. 2023. A diverse virome is identified in parasitic flatworms of domestic animals in Xinjiang, China. *Microbiology Spectrum*, **11**(3): e0070223.

Zhang YZ, Xiao DL, Wang Y, et al. 2004. The epidemic characteristics and preventive measures of hemorrhagic fever with syndromes in China. *Chinese Journal of Epidemiology*, **25**(6): 466–469. (in Chinese)

Zhao ZY, Yu D, Ji CM, et al. 2023. Comparative analysis of newly identified rodent arteriviruses and porcine reproductive and respiratory syndrome virus to characterize their evolutionary relationships. *Frontiers in Veterinary Science*, **10**: 1174031.

Zhou P, Yang XL, Wang XG, et al. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579**(7798): 270–273.

Zhuang Z, Qian LL, Lu J, et al. 2023. Comparison of viral communities in the blood, feces and various tissues of wild brown rats (*Rattus norvegicus*). *Heliyon*, **9**(6): e17222.