Microbiome

RESEARCH

Open Access

Virome landscape of wild rodents and shrews in Central China

Nailou Zhang[1†], Bing Hu[2†], Li Zhang[3], Min Gan[1], Qingwen Ding[1], Kai Pan[2], Jinbo Wei[1], Wen Xu[1], Dan Chen[1], Shaolong Zheng[1], Kun Cai[2*] and Zhenhua Zheng[1*]

Abstract

Background Wild rodents and shrews serve as vital sentinel species for monitoring zoonotic viruses due to their close interaction with human environments and role as natural reservoirs for diverse viral pathogens. Although several studies have explored viral diversity and assessed pathogenic risks in wild rodents and shrews, the full extent of this diversity remains insufficiently understood.

Results We conducted high-throughput sequencing on 1113 small mammals collected from 97 townships across seven cities in Hubei Province during 2021, supplemented by publicly available data from 2014 and 2016–2017. This analysis revealed a diverse array of novel viruses spanning several viral families, including *Arenaviridae*, *Hepeviridae*, *Chuviridae*, *Paramyxoviridae*, *Arteriviridae*, *Nodaviridae*, *Rhabdoviridae*, *Dicistroviridae*, *Astroviridae*, and *Picornaviridae*. Phylogenetic analysis and genome structure characterization highlighted the discovery of these novel viruses, enhancing our understanding of viral diversity and evolution.

Key host species such as *Chodsigoa smithii*, *Anourosorex squamipes*, *Niviventer niviventer*, and *Apodemus agrarius* were identified as significant contributors to viral circulation, making them crucial targets for future surveillance. Additionally, the central Plain of Hubei Province was recognized as a critical geographic hub for viral transmission, underscoring its importance in monitoring and controlling viral spread.

Machine learning models were employed to assess the zoonotic potential of the identified viruses, revealing that families such as *Arenaviridae*, *Coronaviridae*, *Hantaviridae*, *Arteriviridae*, *Astroviridae*, *Hepeviridae*, *Lispiviridae*, *Nairoviridae*, *Nodaviridae*, *Paramyxoviridae*, *Rhabdoviridae*, *Picornaviridae*, and *Picobirnaviridae* possess a high likelihood of infecting humans. Notably, rodent-derived *Rotavirus A*, *HTNV*, and *SEOV* displayed almost complete amino acid identity with their human-derived counterparts, indicating a significant risk for human outbreaks.

Conclusion This study provides a comprehensive virome landscape for wild rodents and shrews in Central China, highlighting novel viruses and the critical roles of specific host species and regions in viral transmission. By identifying key species and hotspots for viral spread and assessing the zoonotic potential of the discovered viruses, this research enhances our understanding of virus ecology and the factors driving zoonotic disease emergence. The findings emphasize the need for targeted surveillance and proactive strategies to mitigate the risks of zoonotic spillovers, contributing to global public health preparedness.

†Nailou Zhang and Bing Hu contributed equally to this work.

*Correspondence:
Kun Cai
ckreal@163.com
Zhenhua Zheng
zhengzh@wh.iov.cn
Full list of author information is available at the end of the article

Keywords Rodents, Shrews, Virus landscape, Zoonotic potential, Viral transmission, Hantaviruses

Introduction

Wild animals, particularly those near human populations, often serve as reservoirs for pathogens capable of crossing species barriers, leading to small, limited outbreaks [1]. The ongoing expansion of industrialization and agriculture exacerbates this issue by increasingly encroaching on natural ecosystems, thereby facilitating greater human exposure to zoonotic pathogens [1–5]. To mitigate the risks posed by zoonotic diseases, it is essential to strengthen local surveillance systems and identify high-risk pathogens and early spillover events [6]. Rodents are known to carry at least 80 zoonotic pathogens, posing significant public health risks [7–10]. Notable zoonotic diseases transmitted by rodents include salmonellosis, leptospirosis, Lassa fever, and hantavirus-related illnesses [11]. Additionally, rodents act as amplifier hosts or carriers for arthropods in the transmission of vector-borne diseases such as plague and leishmaniasis, linking humans, wildlife, and arthropods [7]. Their asymptomatic carriage of pathogens enables long-term transmission, potentially triggering outbreaks through contaminated water, urine, or consumption as food [12].

Recent advancements in high-throughput sequencing technologies have significantly expanded our understanding of the dissemination of zoonotic viruses among rodent populations, including those in the families *Arenaviridae*, *Flaviviridae*, *Hantaviridae*, *Phenuiviridae*, *Nairoviridae*, *Paramyxoviridae*, *Hepeviridae*, *Coronaviridae*, *Picobirnaviridae*, *Picornaviridae*, *Nodaviridae*, *Reoviridae*, *Astroviridae*, *Rhabdoviridae*, *Sedoreoviridae*, *Caliciviridae*, *Arteriviridae*, *Dicistroviridae*, and *Spinareoviridae* [2–4, 13–33]. These studies have provided a broad overview of viral diversity, distribution, and evolution. However, the specific roles of different rodent subspecies or individuals in pathogen circulation remain poorly understood. Not all rodent species contribute equally to pathogen circulation; some species are critical due to their ecological roles, population density, or specific interactions with other species, while others play a less significant role [11, 34, 35]. Understanding these differences and identifying key reservoir species are crucial for accurate assessments of transmission dynamics and for the effective control of epidemics.

Since 2010, viromics has markedly upgraded our capacity to distinguish novel viral sequences. However, the sheer volume of viral data now available necessitates further refinement to assess the potential threats posed by these viruses [14, 20]. Evaluating the pathogenicity and zoonotic risk of newly identified viruses involves reconstructing intact virions and conducting rigorous testing in cell cultures or animal models—processes that are complex, labor-intensive, and not easily scalable. Advances in computational biology, particularly machine learning, offer promising solutions. Tools such as SpillOver and ZoonoticRank analyze viral sequence data to estimate cross-species transmission risk, considering factors such as codon usage, nucleotide composition, amino acid patterns, and dinucleotide frequencies [36, 37]. These approaches help prioritize viruses for targeted surveillance and public health interventions, thereby enhancing disease monitoring and resource allocation to prevent future pandemics.

Hubei Province, located in central China along the upper Yangtze River basin, features a subtropical monsoon climate. Historically, Hubei has experienced outbreaks of various infectious diseases, including COVID-19, hemorrhagic fever, hand-foot-and-mouth disease, hepatitis E, and measles [38–42]. Therefore, monitoring zoonotic diseases in rodents within this region is imperative. Leveraging a decade of high-throughput sequencing data, we investigate the geographic and species-specific distribution of viruses in rodents and shrews. By analyzing viral diversity and host-virus relationships, we gain insights into the factors influencing viral spread. Furthermore, we conduct detailed phylogenetic analyses of newly identified viruses, assessing their genetic and genome structure diversity and their potential to cause zoonotic diseases. Our findings provide insights into the molecular epidemiology of mammalian-associated viruses and will inform strategies for the surveillance and control of zoonotic pathogens in the region.

Methods

Sample collection

Field sampling was conducted from March to December 2021 across 97 townships in seven cities of Hubei Province, central China: Huangshi, Jingmen, Jingzhou, Shiyan, Xianning, Xiangyang, and Yichang. Sampling sites included diverse habitats such as forests, fields, caves, villages, and urban areas. Small mammals, primarily rodents (seven species) and shrews (one species) were captured using cages or rat clips. A total of 1113 specimens were collected: Huangshi ($n=64$), Jingmen ($n=164$), Jingzhou ($n=202$), Shiyan ($n=212$), Xianning ($n=171$), Xiangyang ($n=200$), and Yichang ($n=100$) (Fig. 1). Species identification was performed by expert

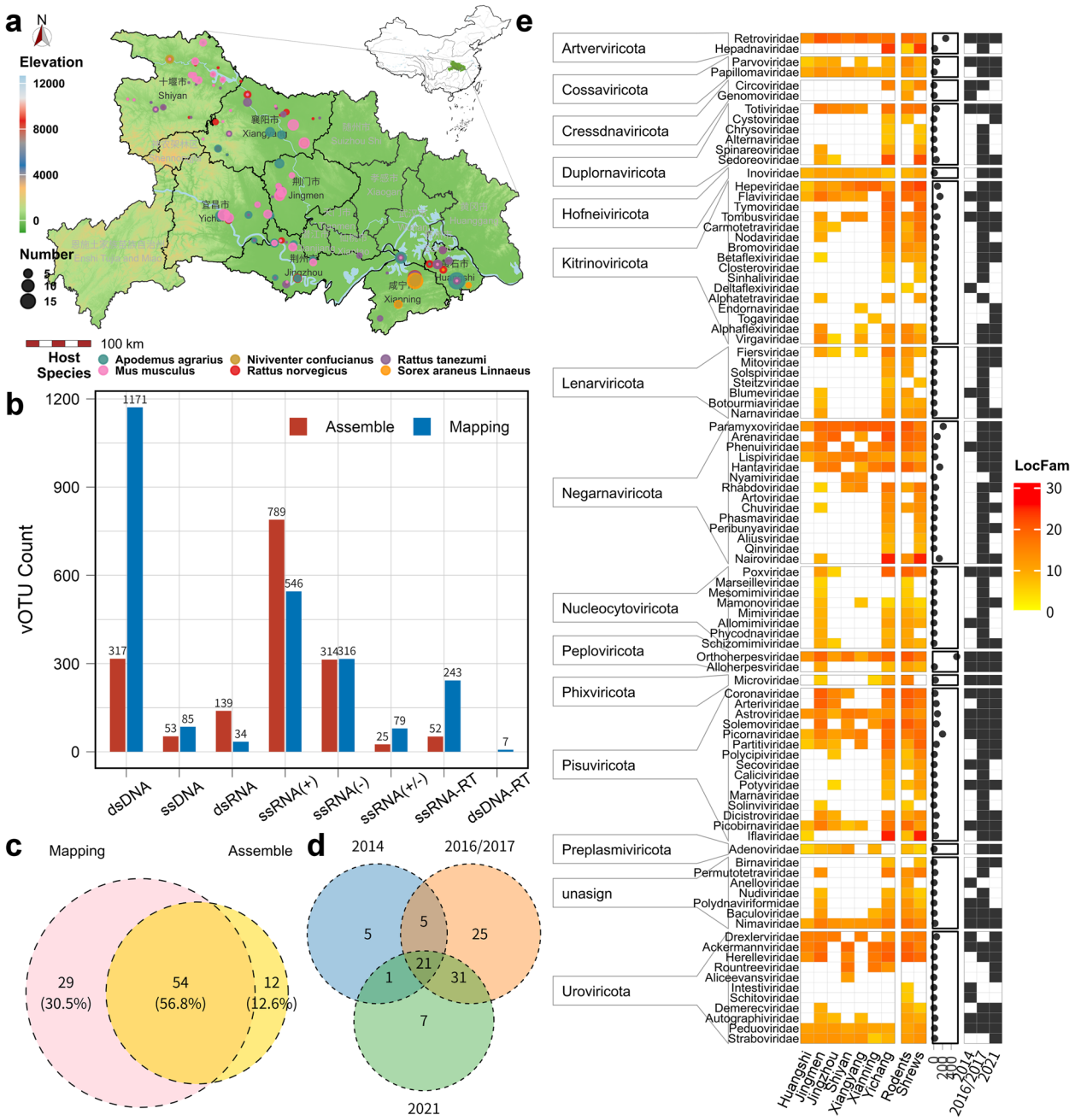

Fig. 1 Sampling locations, species distribution, and viral diversity in small wild mammals from Hubei Province, China. **a** Map of Hubei Province showing the sampling locations and species distribution of small wild mammals, including both rodents and shrews, collected in 2021. **b** Comparative analysis of viral genomic composition in wild mammals from Hubei Province, using both reads-mapping and de novo assembly methods. **c** Comparison of viral detection at the family level based on reads-mapping versus de novo assembly methods. **d** Temporal comparison of viral diversity at the family level across three different time periods in wild mammals. **e** Summary of the viral spectrum at the family level in wild mammals from Hubei Province. From left to right, the panel includes: (i) viral abundance (log2-transformed count values) across different phyla and families, and sampled regions; (ii) viral abundance (log2-transformed count values) in shrews and rodents; (iii) the number of viral species within each family; and (iv) detection status of viruses across three different years

biologists based on morphological characteristics. Tissue samples (heart, liver, spleen, lung, kidney) were preserved in RNAlater or viral transport medium, packed on dry ice, and transferred to the laboratory for storage at − 80 °C.

DNA/RNA extraction

RNAlater-preserved samples were homogenized using a DS1000 tissue disruptor (NZK Biotech, Wuhan, China) with 3 mm grinding beads. The resulting supernatant was pooled into subsets of ten, and DNA/RNA was extracted using the VNP-32P Automatic Nucleic Acid Extractor and Virus DNA/RNA Extraction Kit (Vazyme Biotech, Nanjing, China) following the manufacturer's instructions. For high-quality RNA required for genomic and cDNA amplification, the supernatant underwent further purification using the FastPure® Viral DNA/RNA Mini Kit (Vazyme). Specifically, 200 µl of supernatant was mixed with 500 µl Buffer VL, loaded onto FastPure® RNA Columns, and centrifuged at 12,000 rpm. The RNA was subsequently eluted with RNase-free water and stored at − 80 °C. For next-generation sequencing, RNA was extracted with the QIAamp® Viral RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. All procedures were carried out in a sterile, nuclease-free environment.

High-throughput sequencing

Tissues (liver, spleen, lung, kidney) from the same species or locations were pooled and homogenized. Due to the limited availability of rare species or organs, some samples were insufficient for reliable RNA extraction and library construction. To address this, 7 to 110 samples were combined, resulting in 30 libraries (Table S1). Ribosomal RNA was depleted prior to library preparation, and the libraries were paired-end sequenced on the MiSeq platform (150 or 100 bp) by BGI, China.

Collection and processing of publicly available data

Publicly available raw data files (*.sra) were downloaded from NCBI, including 103 files from project PRJNA953205 [13] and 3 files from project PRJNA375958 (Table S1) [16]. The samples in PRJNA953205, collected from 12 rodents and 6 shrew species in Hubei Province during 2016–2017, included pooled tissues from the kidney, liver, lung, spleen, intestine, and feces. Libraries were prepared using the KAPA RNA HyperPrep Kit with RiboErase (HMR) (KAPA Biosystems) for rRNA removal and sequenced on the Illumina Novaseq 6000 platform (paired-end, 150 bp reads) by Novogene. Species identification, initially performed based on morphological characteristics, was later confirmed through sequencing of a 600–700 nucleotide region near the 5′ end of the cytochrome c oxidase I (COI) gene [13]. The samples in PRJNA375958, collected in 2014 from 2 rodent species in Hubei Province, comprised pooled pharyngeal and anal swabs. rRNA was removed during library preparation, and sequencing was conducted on the Illumina HiSeq2500 platform (single-end, 100 bp reads) [16]. The raw SRA files were converted to fastq format using the fastq-dump utility from NCBI's SRA Toolkit (https://ncbi.github.io/sra-tools/).

Viral sequence identification

Raw sequencing reads were first quality-trimmed using Fastp [43] to remove adapters and low-quality sequences. Host reads were filtered out by mapping the trimmed reads against a custom rodent and shrew database, derived from the NCBI nt database, using Kraken2. The remaining reads were de novo assembled with Megahit, polished using Pilon, and further refined with CAP3. Assembled contigs were compared against the nr protein database using Diamond Blastx and the nt database using blastn, both with an *e*-value threshold of 1×10^{-5}. Sequences annotated as "Viruses" based on the top 15 blast hits, identified using TaxonKit, were considered potential viral sequences. Viral unique operational taxonomic units (vOTUs) were obtained with MMseqs2 using parameters: easy-linclust –min-seq-id 0.99 -e 0.01 -c 0.99 –cov-mode 1 –cluster-mode 2 –ker-per-seq-scale 0.4 –spaced-ker-mode 0. vOTUs exceeding 1000 bases in length were retained for further analysis (Table S2). Virus-host associations were referenced from the Virus-Host Database (http://www.genome.jp/virushostdb/, release 209), and any viral family containing members known to infect mammals was classified as mammal-associated.

Viral quantification

To quantify viral abundance, clean reads (excluding host reads) were mapped to vOTUs and a custom viral sequence database using Bowtie2. A custom viral database was created by clustering all viral sequences from the NCBI Virus database (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/) using MMseqs2 to obtain representative sequences. Initially, viral sequences were downloaded and subjected to quality control. All U bases were converted to T, and non-ATGC bases were replaced with N to represent ambiguous nucleotides. Sequences with more than 10% ambiguous bases were excluded. The remaining sequences were clustered at a 99% similarity threshold using MMseqs2, and the most representative sequence was selected based on overall similarity rather than length. This approach mitigates issues with segmented viruses, such as hantaviruses, where genome segments (L, M, and S) may be concatenated during

submission (e.g., ON361561, JN542542), which could result in incomplete viral data if the longest sequence is chosen. For SARS-CoV-2 and HIV, which comprise approximately 67% and 9% of sequences in the database, 1000 sequences from each were randomly selected to represent these viruses. The resulting SAM files were processed with BBmap's built-in pipeline to generate viral abundance tables. Custom scripts were used to compile coverage and abundance tables for each reference sequence (Table S3). Viral sequences with continuous coverage of fewer than 1000 bases, spanning more than 2 segments, 20 reads in abundance, and 300 bases, were merged into vOTU abundance tables for subsequent alpha and beta diversity analysis.

Diversity analysis

Alpha diversity, including observed richness, Shannon index, and Simpson index, was calculated for each library. Differences in virome composition across sample years, locations, and mammalian groups were tested using the Kruskal–Wallis rank sum test. Beta diversity was assessed through principal coordinate analysis (PCoA) based on the Bray–Curtis dissimilarity matrix using the Vegan package. Statistical analysis of viral diversity and abundance was conducted using t-tests or Wilcoxon tests by the ggpubr package, and visualized with ggplot2. Core and specific viruses between groups were explored using ggvenn and UpSetR.

Annotation of viral genomes

Viral ORFs were predicted using getorf, selecting ORFs longer than 150 nucleotides and reporting only the longest ORF if a shorter one was fully nested within it. Predicted ORFs were annotated by BLASTP against the UniProt protein database with an *e*-value threshold of 1×10^{-5}, and protein coordinates on the genome were obtained using tblastn. A final GenBank-like file was generated using custom scripts. The translated amino acid sequences were further analyzed with HMMER and InterProScan to identify conserved motifs. Genome structures were visualized using the R package gggeomes.

Viral phylogeny analysis

Phylogenetic analysis was performed for each viral family using the RNA-dependent RNA polymerase (RdRp) proteins of RNA viruses. RdRp proteins and associated host metadata were sourced from NCBI Virus, ICTV, and Virus-Host Database (https://www.genome.jp/virushostdb/). RdRp motifs were extracted by HMMER. For mammalian viruses within the same viral family, RdRp amino acid sequences were aligned using the E-INS-i algorithm in MAFFT, retaining ambiguously aligned regions. Maximum likelihood phylogenetic trees were constructed using FastTree, and amino acid identity percentages were calculated with CIAlign (https://github.com/KatyBrown/CIAlign). The trees were visualized using the ggtree package.

Inferring zoonotic potential

Zoonotic potential was assessed using machine learning models to predict the likelihood of viruses infecting humans, based on host range features encoded in their genomes. A threshold value of 0.293 was applied as reported in previous studies [36]. These models used genomic composition biases, such as codon, amino acid, and dinucleotide frequencies, and compared viral sequences with human RNA transcripts to calculate similarity.

Single nucleotide polymorphism (SNP) detection

Publicly available viral sequences derived from humans were aligned to vOTUs using the MAFFT method in Augur, which maintains reading frames through codon-aware alignments [44]. SNPs were identified with snp-sites, normalized with bcftools, and functionally annotated using SnpEff. Bcftools stats were employed to analyze synonymous and non-synonymous mutations, intergenic regions, gene-specific variations, and nucleotide transitions and transversions. Genetic similarity was visualized with Simplot plots generated using ggmsa. Linear B-cell epitopes were predicted using the IEDB tool (http://tools.iedb.org/main/bcell/), and their genomic coordinates on contigs were mapped with tblastn. These analyses provided insights into viral evolution and potential immunogenicity.

Results

Viromic profiling of wild small mammals over nearly a decade

To construct a comprehensive viromic profile of wild small mammals in Hubei, China, we sampled 1113 individuals from 5 rodent species and one shrew species across 97 townships in seven cities in 2021 (Fig. 1a). These samples were collected from diverse habitats, including forests, fields, villages, and urban areas. Samples were pooled based on geographic location, species, and tissue type, resulting in 30 grouped samples. High-throughput sequencing was then performed after the removal of ribosomal RNA. Additionally, we incorporated data from two publicly available high-throughput datasets: PRJNA953205 and PRJNA375958. The former includes samples collected between 2016 and 2017 from 12 rodent species and six shrew species in Jingmen and

Yichang [13], while the latter includes samples from two rodent species and one shrew species collected in Hubei in 2014 [16]. Both datasets employed ribosomal RNA-depleted sequencing methods.

Due to nucleotide degradation during the sampling, storage, and RNA extraction processes, not all viral sequences could be assembled de novo. Therefore, we employed two complementary strategies: de novo assembly and read mapping to a viral reference database. We considered alignment results valid if they met the following criteria: (1) a coverage length exceeding 1000 bp with a viral abundance higher than 20, or (2) a coverage length shorter than 1000 bp but with more than three coverage regions and a viral abundance higher than 20. For de novo assembly, we polished contigs using Pilon, extended sequences with Cap3, and merged sequences with MMSeq to remove redundancy, discarding sequences shorter than 1000 bp. This process identified 2267 viral contigs, 578 of which could not be assigned to any existing viral taxonomy and were excluded from further analysis.

At the genomic composition level, the assembly-based approach identified a broader spectrum of RNA viruses, detecting 55 out of 62 RNA virus families identified by both methods. However, some ssRNA(+) viruses (e.g., *Alphaflexiviridae*, *Endornaviridae*, *Deltaflexiviridae*, *Solinviviridae*, *Solspiviridae*, *Togaviridae*) and dsRNA viruses (e.g., *Chrysoviridae*) could not be identified through assembly alone. The read mapping strategy, meanwhile, was more effective in identifying DNA viruses (Fig. 1b and Table S1).

Overall, 54 viral families were detected by both strategies, with the read-mapping approach identifying more viral families (Fig. 1c). We found that 21 viral families were consistently detected in small mammals across the region, while others were detected only in specific years, suggesting the presence of underexplored viral diversity. This underscores the need for continuous long-term surveillance of small mammal populations to construct a comprehensive and detailed viral landscape.

By merging results from read mapping for sequences with max coverage lengths below 1000 bp with those from de novo assembly, we compiled a final viromic profile of the region's small mammals (Fig. 1e). This profile revealed that viruses carried by wild small mammals in the region span at least 14 phyla and 95 families. Among sequences with more than 100 viral operational taxonomic units (vOTUs), the *Orthoherpesviridae* family exhibited the highest diversity, with 528 sequences or fragments, followed by *Retroviridae* (274 sequences), *Paramyxoviridae* (217 sequences), *Picornaviridae* (201 sequences), *Flaviviridae* (142 sequences), *Hantaviridae* (136 sequences), and *Nairoviridae* (125 sequences) (Fig. 1e). The prevalence and abundance of viruses varied significantly across different families. Several families, including *Astroviridae*, *Hepeviridae*, *Inoviridae*, *Lispiviridae*, *Nimaviridae*, *Orthoherpesviridae*, *Papillomaviridae*, *Paramyxoviridae*, *Peduoviridae*, *Retroviridae*, and *Straboviridae*, were found to be persistently circulating across multiple regions and species, indicating their potential for long-term endemicity.

Viral transmissions among wild small mammals

Despite belonging to different orders, shrews and rodents share similar habitats and behaviors, facilitating viral transmission between the two groups. Our study identified viruses from 59 different viral families in both shrews and rodents, with no significant differences in overall viral diversity between the two groups (Fig. 2a, b). To further investigate the role of wild small mammals as key nodes in viral transmission within their ecological environments, we constructed a bipartite network Venn diagram illustrating virus sharing among species (Fig. 2c). While *Chodsigoa smithii* carried a higher diversity of viruses, *Anourosorex squamipes* occupied a more central position in the network due to its higher number of shared viruses with other host species, making it a crucial species in the viral transmission network. Similarly, among rodents, *Niviventer niviventer* harbored a higher diversity of viruses, yet *Apodemus agrarius* held a more central position due to its greater number of shared viruses with other hosts (Fig. 2c). Given the broader range of species included in the 2016 dataset, we conducted a beta diversity analysis on this subset to explore whether host species drive viral diversity. The results indicate that species differences among rodents do not significantly influence viral diversity (Fig. 2d). In contrast, species diversity among shrews emerged as the primary factor influencing the diversity of viruses they harbor (Fig. 2e).

Regional patterns of viral transmission

We next assessed the influence of geographical factors on viral dissemination among wild small mammals. Alpha diversity analysis revealed substantial variability in viral diversity across different regions. Notably, small mammals from Shiyan City exhibited the highest viral diversity, followed by those from Jingzhou (Fig. 3a). This variability may be linked to the distinct geographical characteristics of these regions, although differences in pooling strategies and sample sizes might also have contributed to the observed patterns. A Venn diagram illustrating the distribution of viral families across different regions highlighted the

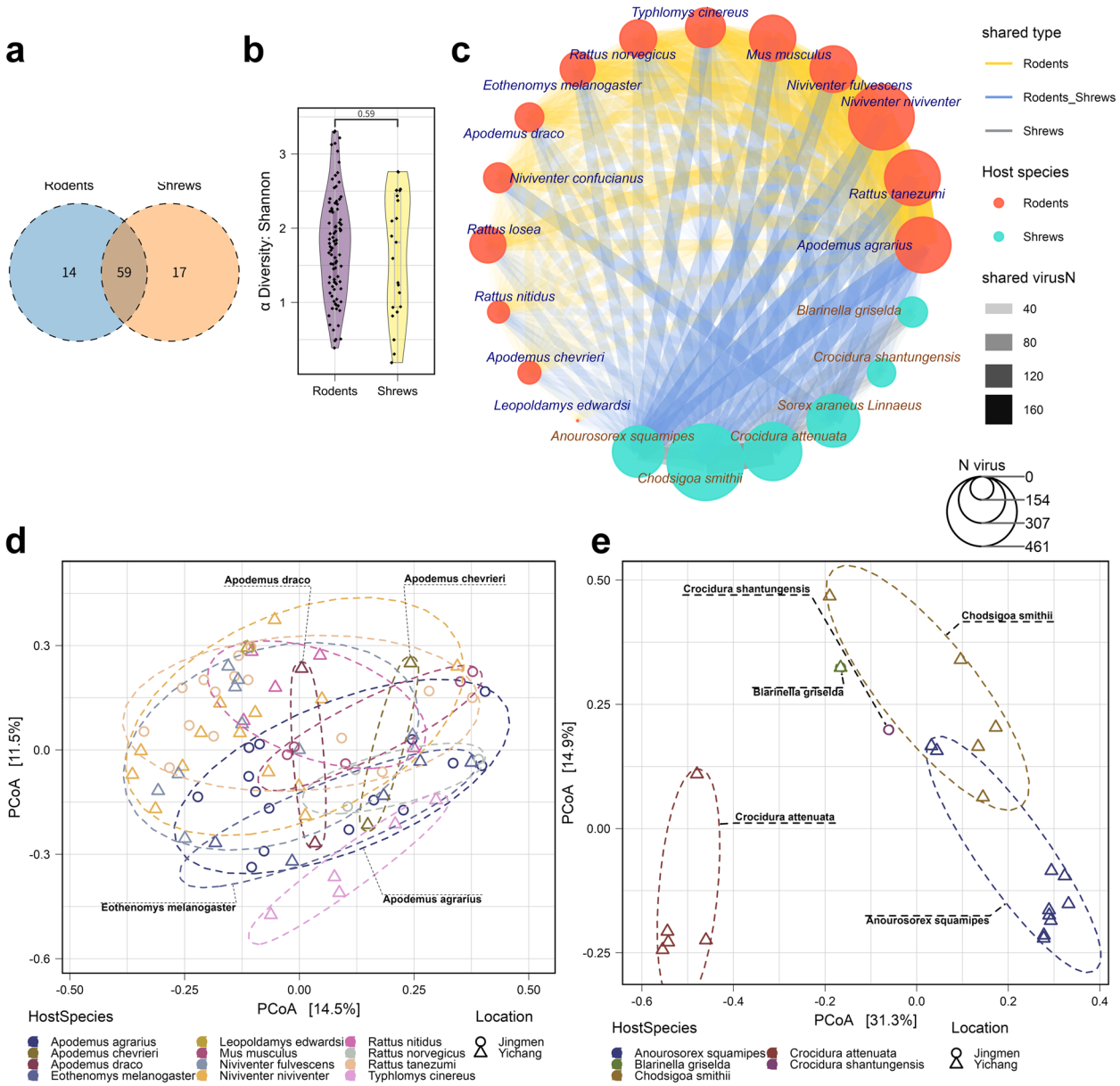

Fig. 2 Distribution and circulation of viruses in small wild mammals. a Comparison of viral prevalence at the family level between shrews and rodents. **b** Comparison of alpha diversity of viral species between shrews and rodents. **c** Binary network analysis illustrating shared viruses between different host species. **d** Beta diversity analysis of viruses at the species level in rodents sampled during 2016/2017. **e** Beta diversity analysis of viruses at the species level in shrews sampled during 2016/2017

presence of families such as *Astroviridae*, *Flaviviridae*, *Hepeviridae*, *Inoviridae*, *Lispiviridae*, *Nimaviridae*, *Orthoherpesviridae*, *Papillomaviridae*, *Paramyxoviridae*, *Peduoviridae*, *Retroviridae*, *Picobirnaviridae*, *Picornaviridae*, and *Straboviridae* in multiple areas (Fig. 3b). When constructing a binary virus-sharing network and mapping it onto a geographic map, we observed that adjacent regions, such as Shiyan and Xiangyang or Jingzhou and Xianning, exhibited lower levels of viral sharing compared to regions within the central Plain of Hubei Province, including Yichang, Jingmen, and Jingzhou (Fig. 3c). This suggests that the central Plain, particularly the areas of Yichang, Jingmen, and Jingzhou, may serve as hubs for viral dispersion or aggregation among small mammals. The ecological and environmental factors contributing to this potential role of the central Plain in viral dissemination require further investigation.

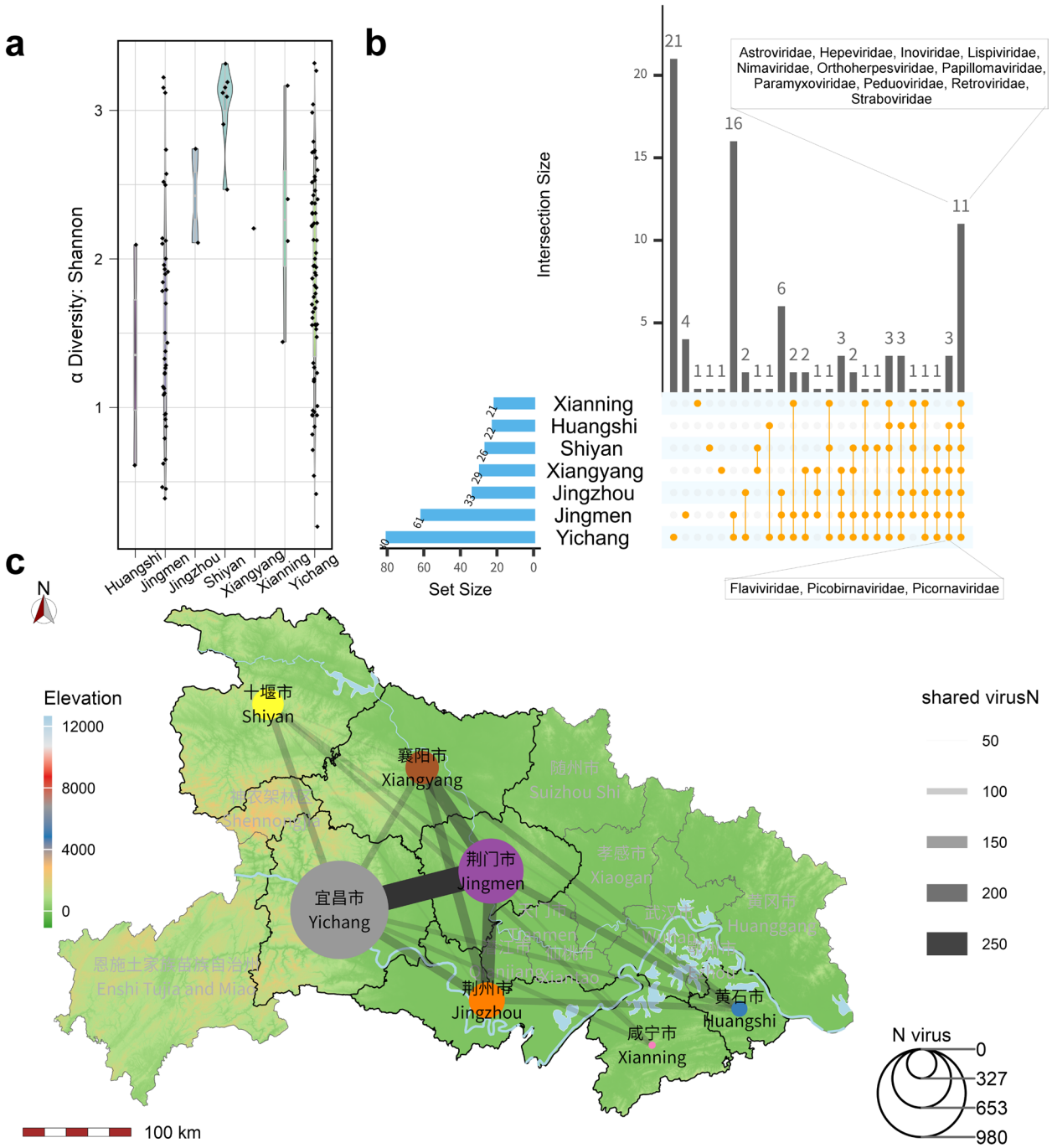

Fig. 3 Geographic distribution and circulation of viruses in Hubei Province. **a** Comparison of alpha diversity of viral species among different regions within Hubei Province. **b** Comparison of viral prevalence at the family level across different regions in small wild mammals. **c** Binary network illustrating shared viruses (at the species level) among different regions within Hubei Province

Evolutionary phylogeny of mammalian RNA virus RdRp

Given the extensive range of viruses harbored by wild small mammals, we assessed the potential spillover risks of mammalian-related and vector-borne RNA viruses to humans by focusing on the characterization and phylogenetic analysis of key RNA viruses from several families, including *Arenaviridae*, *Arteriviridae*, *Astroviridae*, *Caliciviridae*, *Chuviridae*, *Coronaviridae*,

Dicistroviridae, *Lispiviridae*, *Flaviviridae*, *Hantaviridae*, *Hepeviridae*, *Spinareoviridae*, *Picobirnaviridae*, *Sedoreoviridae*, *Nodaviridae*, *Nyamiviridae*, *Paramyxoviridae*, *Phenuiviridae*, *Rhabdoviridae*, *Nairoviridae*, *Picornaviridae*, and *Qinviridae*. We extracted the RNA-dependent RNA polymerase regions from these viruses to conduct detailed phylogenetic analyses. Additionally, we retrieved and analyzed corresponding RdRp amino acid sequences from all relevant viral families listed in the ICTV database (Fig. 4).

Arenaviridae

Arenaviridae is a family of ambisense RNA viruses characterized by a bisegmented genome, comprising the L and S segments [45]. This family includes five genera: *Antennavirus*, *Hartmanivirus*, *Innmovirus*, *Mammarenavirus*, and *Reptarenavirus*. *Mammarenaviruses*, known for infecting mammals, are classified into Old-World and New-World complexes and include several Biosafety Level 4 pathogens, such as Lassa fever virus [45]. In our study, we identified nine *Arenaviridae* contigs encompassing both L and S segments. Phylogenetic analysis of RdRp sequences classified these contigs within the Old-World arenaviruses, revealing that three rodent-derived contigs clustered with *Wenzhou virus* (91.2–99.7% amino acid identity) and one with *Lijiang virus* (65.2% amino acid identity), indicating significant evolutionary divergence within these sequences.

Arteriviridae

Arteriviridae is a family of enveloped, positive-sense RNA viruses currently comprising six subfamilies and thirteen genera [46]. We identified 52 *Arteriviridae* contigs in samples from rodents and shrews. Notably, we detected a novel arterivirus clade in rodents, with seven sequences forming a sister group to known arteriviruses, exhibiting only 17–20% RdRp amino acid identity with sequences recognized by the ICTV but 73.3–100% identity with the closest related viruses. Additionally, four contigs from Yichang rodents and shrews formed a sister clade positioned between *Muarterivirus* and *Epsilonarterivirus*, showing 42–56% RdRp amino acid sequence identity. Other contigs grouped with *Muarterivirus*, *Nuarterivirus*, and *Lambdaarterivirus*, displaying 51.7–74.7% RdRp amino acid identity.

Astroviridae

Astroviridae is a family of small, non-enveloped, positive-sense RNA viruses that cause diarrheal disease in mammals and birds [47]. In rodent and shrew samples, 39 *Astroviridae* contigs were identified. Five sequences showed close relation to astroviruses, with 10–20% identity in RdRp sequences from ICTV and 97.2–98.3% amino acid identity with the nearest viruses. Other contigs clustered with *Mamastrovirus*, displaying 58.5–98.3% amino acid identity with the closest viruses.

Caliciviridae

Caliciviridae is a family of non-enveloped, positive-sense RNA viruses comprising eleven genera: seven that infect mammals (*Lagovirus*, *Norovirus*, *Nebovirus*, *Recovirus*, *Sapovirus*, *Valovirus*, *Vesivirus*), two that infect birds (*Bavovirus*, *Nacovirus*), and two that infect fish (*Minovirus*, *Salovirus*) [48]. In shrew samples, we identified four *Caliciviridae* contigs. Phylogenetic analysis of a contig longer than 100 amino acids in the RdRp region revealed 98.6% identity with the closest known virus (OQ716201) but only 13.82–24.65% RdRp amino acid identity with sequences recognized by the ICTV, suggesting the potential existence of a new genus within the *Caliciviridae* family.

Chuviridae

Chuviridae is a family of negative-sense RNA viruses characterized by remarkable genomic diversity, encompassing non-segmented linear, non-segmented circular (e.g., *Morsusvirus*), bisegmented linear, and bisegmented circular forms [49]. We identified 22 *Chuviridae* contigs from rodent and shrew samples. Most of these contigs were distributed across various genera within arthropod-borne *Chuviridae*, with several sequences occupying phylogenetically distinct positions between established clades. This suggests a broader genetic diversity within *Chuviridae* than previously recognized. Notably, three sequences clustered with *Tacheng tick virus 4*—a circular virus—with RdRp amino acid sequence identities ranging from 40.13 to 47.82%.

Coronaviridae

Coronaviridae is a family of enveloped, positive-sense RNA viruses with a non-segmented genome, divided into four genera: *Alphacoronavirus*, *Betacoronavirus*,

(See figure on next page.)

Fig. 4 Phylogeny of mammal-associated RNA viruses. The phylogenetic tree shows RNA viruses associated with mammals. Blue and brown labels represent viruses detected exclusively in rodents and shrews, respectively, while red labels indicate viruses detected in both groups. Asterisks indicate novel viruses, and droplets indicate known viruses. Clades with viruses from the same genus are shaded in light grey, with corresponding genus names in blue. Host type icons are provided after genus names: buffalo (mammals), mosquito (arthropods), chicken (birds), fish (fish), frog (amphibians), leaf (plants), mushroom (fungi), gecko (reptiles), turtle (turtles), and curled line (nematodes)

Fig. 4 (See legend on previous page.)

Gammacoronavirus, and *Deltacoronavirus* [50]. Members of *Coronaviridae* infect a wide range of hosts, including mammals and birds, and are notable for their potential to cause significant diseases, such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) [50]. We identified 19 contigs associated with *Coronaviridae*, 18 of which belong to *Betacoronavirus*. These sequences exhibit nucleotide identities ranging from 89.7 to 100% with closely related sequences, including *Jingmen Apodemus agrarius betacoronavirus 1* and *Severe acute respiratory syndrome-related coronavirus*. However, due to the lengths of these sequences, which range from 1025 to 2738 nucleotides and fall short of encoding the RdRp region, phylogenetic analysis based on RdRp amino acid sequences could not be performed, limiting definitive classification. Additionally, we identified a full-length *Alphacoronavirus* sequence that shares 99.05% nucleotide identity with the closest known virus. Phylogenetic analysis places this sequence on the same branch as *Lucheng Rn rat coronavirus.*

Dicistroviridae

Dicistroviridae is a family of small, non-enveloped, positive-sense single-stranded RNA viruses distinguished by their unique dicistronic genome structure [51]. Primarily associated with insects, these viruses hold significant ecological and agricultural importance [51]. The family is currently classified into three genera—*Cripavirus*, *Triatovirus*, and *Aparavirus*—though formal genus demarcation criteria are yet to be fully established [51]. We identified 18 *Dicistroviridae* contigs from rodent and shrew samples, with seven containing RdRp sequences of sufficient length for phylogenetic analysis. These sequences clustered with both insect-borne and mammalian-associated viruses, showing 89.6% to 99.1% nucleotide identity with the closest known sequences. This discovery expands our understanding of *Dicistroviridae*'s host range and evolutionary dynamics, indicating a possible extension of this viral family into mammalian hosts.

Lispiviridae

Lispiviridae is a family of negative-sense, single-stranded RNA viruses with typically non-segmented genomes [52]. These non-enveloped viruses infect a diverse range of hosts, including mammals, birds, and arthropods [52]. In this study, we identified 23 contigs associated with *Lispiviridae*, 11 of which encoded RdRp sequences sufficient for phylogenetic analysis. Eight rodent-derived sequences clustered with mammalian-associated *Usmuvirus* species, showing 61.3 to 98.9% amino acid sequence identity with the closest known sequences. Two sequences clustered with avian-associated *Avesvirus*, displaying 42.8 to 46.6% RdRp amino acid identity, indicating significant evolutionary divergence. Additionally, a sequence from a shrew clustered with *Aranbvirus*, sharing 62.1% amino acid identity. These findings underscore the genetic diversity within the *Lispiviridae* family and suggest potential cross-species transmission, which could lead to the emergence of novel viral lineages.

Flaviviridae

Flaviviridae are enveloped, positive-sense single-stranded RNA viruses with non-segmented genomes, known for their impact on both human and animal health [53]. In this study, we identified 103 contigs belonging to *Flaviviridae*, primarily within the genera *Hepacivirus*, *Pegivirus*, and *Pestivirus*. Within the *Pestivirus* clade, sequences were segregated by host species, with amino acid identities ranging from 96.7 to 100% in shrew-derived sequences and 77.6 to 82.7% in rodent-derived sequences. The *Hepacivirus* clade exhibited broader diversity, with contigs spanning five distinct clades and amino acid similarities to the closest known sequences ranging from 31.2 to 100%. Notably, sequences from both shrews and rodents clustered with *giraffe pestivirus* from bats, sharing 82.4 to 90% amino acid similarity. This suggests a potential cross-species transmission of *Pestivirus G-related viruses* between bats and rodents, highlighting the ecological complexity and zoonotic potential of this viral family.

Hantaviridae

Hantaviridae are enveloped, negative-sense RNA viruses with a tri-segmented genome (L, M, and S segments) and are divided into New World and Old World hantaviruses based on their geographical distribution [54]. These viruses are clinically significant as they cause hantavirus pulmonary syndrome (HPS) and hemorrhagic fever with renal syndrome (HFRS) in humans [54]. Our study identified 22 contigs associated with *Hantaviridae*, covering all three genome segments. These sequences exhibited 97.4 to 100% amino acid identity with known hantaviruses. Phylogenetic analysis of RdRp sequences classified these contigs within the Old World hantaviruses, distributed across four distinct phylogenetic clusters, including *Seoul virus*. The identification of diverse hantavirus serotypes in local rodent and shrew populations highlights the wide circulation of these viruses, emphasizing the need for continuous surveillance to monitor hantavirus diversity and assess zoonotic transmission risks.

Hepeviridae

Hepeviridae are non-enveloped, positive-sense single-stranded RNA viruses with non-segmented genomes [55]. They are of significant public health concern due to their

association with hepatitis E, which can be transmitted zoonotically or from human to human [55]. In this study, we identified 21 Hepeviridae contigs from rodent and shrew samples, with the majority found in both host types. Phylogenetic analysis of RdRp amino acid sequences revealed that most contigs clustered within the *Rocahepevirus* genus, showing 88.5 to 98.19% amino acid identity with known sequences. Notably, two shrew-derived sequences exhibited lower RdRp amino acid identities, ranging from 20 to 29%. These results indicate the presence of diverse *Hepeviridae* variants in local rodent and shrew populations, emphasizing the need for further exploration of their genetic diversity.

Spinareoviridae

Spinareoviridae are non-enveloped, double-stranded RNA viruses with segmented genomes, encompassing multiple genera that infect a wide range of hosts, including insects and vertebrates [56]. In this study, we identified 10 *Spinareoviridae* contigs from shrew samples. Phylogenetic analysis revealed that one contig clustered with *Orthoreovirus*, showing 82 to 85.79% RdRp amino acid sequence identity. Another contig formed a sister clade with both *Aquareovirus* and *Orthoreovirus*, exhibiting 26.96 to 31.15% RdRp amino acid sequence identity, indicating a distant evolutionary relationship. The remaining contigs clustered with arthropod-borne *Cypovirus*, with 30.24 to 39.68% RdRp amino acid sequence identity. These findings highlight the genetic diversity of *Spinareoviridae* in shrews, suggesting the potential discovery of novel strains or species.

Picobirnaviridae

Picobirnaviridae are non-enveloped, double-stranded RNA viruses with segmented genomes, identified in a variety of hosts, including humans, and are potentially associated with gastrointestinal diseases, particularly in immunocompromised individuals [57]. In this study, we identified 22 *Picobirnaviridae* contigs. Phylogenetic analysis revealed that seven rodent-derived contigs clustered with human *Picobirnavirus* strains (AB186898 and AF246939), with RdRp amino acid sequence identities ranging from 22.84 to 63.58%. Additionally, contigs clustering with *BeiHai picobirna-like virus 7* (K884062) formed a distinct clade, sharing 23.26% RdRp amino acid sequence identity. Two other rodent-derived contigs clustered with *Picobirnavirus* strain *Eq. 3* (KR902507), exhibiting 37.92 to 42.40% RdRp amino acid sequence identity. These findings suggest the presence of novel or highly divergent strains of *Picobirnaviridae* in rodent and shrew populations.

Sedoreoviridae

Sedoreoviridae are non-enveloped, double-stranded RNA viruses with segmented genomes, typically comprising 10 to 12 segments [58]. This family includes the genus *Rotavirus*, which is significant due to its role in gastrointestinal diseases and complex evolutionary relationships [58]. In this study, we identified 41 contigs related to *Sedoreoviridae*. Phylogenetic analysis revealed that these contigs clustered within two major clades of *Rotavirus*. Specifically, the contigs displayed 57.01 to 59.43% RdRp amino acid sequence identity with *Rotavirus I* and 64.83 to 84.15% identity with *Rotavirus A*. The identification of these sequences highlights the genetic diversity of *Sedoreoviridae* in local rodent and shrew populations.

Nodaviridae

Nodaviridae are non-enveloped, positive-sense single-stranded RNA viruses with segmented genomes, typically consisting of two segments [59]. We identified 41 contigs belonging to *Nodaviridae*. Phylogenetic analysis of RdRp amino acid sequences revealed their distribution across at least three clades of *Alphanodavirus*. Specifically, one rodent-derived sequence clustered with *Boolarra virus* (AF329080), sharing 85.64% RdRp amino acid sequence identity. Eight sequences from shrews clustered with *Nodamura virus* (AF174533), displaying 46.91 to 53.71% RdRp amino acid sequence identity. Four additional shrew-derived sequences clustered with *Pariacoto virus* (AF171942), showing 27.67 to 38.54% RdRp amino acid sequence identity. Notably, three sequences formed a distinct clade with only 10.53 to 17.94% RdRp amino acid sequence identity to known viruses. These findings emphasize the genetic diversity of *Nodaviridae* in local rodent and shrew populations, suggesting the presence of novel or highly divergent strains.

Nyamiviridae

Nyamiviridae are non-enveloped, negative-sense single-stranded RNA viruses with a non-segmented genome, except for the genus *Tapwovirus*, which has a segmented genome [60]. These viruses primarily infect invertebrates but have also been detected in vertebrates [60]. In this study, we identified two *Nyamiviridae* contigs from rodent samples, encoding segment 1 and segment 2 of Tapwovirus. These sequences show high amino acid sequence identity with *Tapwovirus cesti,* at 98.6% and 98.8% for segments 1 and 2, respectively. Phylogenetic analysis reveals that these contigs cluster with *Wenzhou Tapwovirus* (KX884436.1), sharing 98.8% RdRp amino acid sequence identity.

Paramyxoviridae

Paramyxoviridae are enveloped, negative-sense single-stranded RNA viruses with non-segmented genomes [61]. This diverse family includes genera such as *Jeilongvirus*, *ParaJeilongvirus*, *Henipavirus*, and *Morbillivirus* [61].

Notably, *Henipavirus* includes highly pathogenic viruses like *Hendra virus* and *Nipah virus*, which cause severe respiratory and encephalitic diseases. *Morbilliviruses*, including *Measles virus*, *Rinderpest virus*, and *Canine distemper virus*, are known for systemic infections with significant morbidity and mortality [61]. In our study, we identified 117 contigs within *Paramyxoviridae* from rodents and shrews. The majority of rodent-derived contigs were distributed across eight distinct clades within the genera *Jeilongvirus* and *ParaJeilongvirus*. Specifically, ten sequences clustered with *Beilong virus* (DQ100461), showing 92.66 to 100% RdRp amino acid sequence identity, indicating close genetic relationships. Four contigs formed a novel clade closely related to *Beilong virus*, with 81.29 to 89.65% RdRp amino acid sequence identity. Five contigs clustered with *J virus* (AY90001), displaying 76.31% to 86.7% RdRp amino acid sequence identity. Five additional contigs grouped with *Tailam virus* (JN689227), exhibiting 80.76 to 85.35% RdRp amino acid sequence identity. Six contigs clustered with *Belerina virus* (MG573140), showing 70.62 to 78.83% RdRp amino acid sequence identity. Two contigs were within the *Narmovirus* clade, with 66.87 to 81.06% RdRp amino acid sequence identity. Three contigs formed a branch closely related to *Morbillivirus*, with 54.91 to 58.58% RdRp amino acid sequence identity. Shrew-derived contigs predominantly clustered within the *Parahenipavirus* clade, with RdRp amino acid sequence identities ranging from 61.95 to 73.67% compared to *Mojang virus*. These findings underscore the substantial genetic diversity of Paramyxoviridae in rodent and shrew populations. Our study reveals the presence of multiple novel or divergent strains, providing critical insights into the evolutionary dynamics and epidemiological significance of this virus family.

Phenuiviridae

Phenuiviridae are enveloped, negative-sense single-stranded RNA viruses with segmented genomes, typically comprising three segments [62]. This family includes genera such as *Bandavirus* and *Phlebovirus*, with highly pathogenic members like the *Severe Fever with Thrombocytopenia Syndrome Virus* (SFTSV), which is linked to severe human disease [62]. We identified 16 contigs belonging to *Phenuiviridae*. Phylogenetic analysis of RdRp amino acid sequences revealed these contigs distributed across four distinct clades. One contig clustered with *Tanzavirus* (MN062090), with an RdRp amino acid sequence identity of 33.96%. Another contig grouped with *Wenrivirus*, showing 40.52% RdRp amino acid sequence identity. Additionally, five contigs were classified within the *Mobuvirus* clade, known for its arthropod-borne viruses, with RdRp amino acid sequence identities ranging from 21.56 to 25.75%. These results highlight the substantial genetic diversity of *Phenuiviridae* in shrew populations and identify potentially novel or divergent strains within this viral family.

Rhabdoviridae

Rhabdoviridae comprises enveloped, negative-sense single-stranded RNA viruses with non-segmented genomes [63]. This diverse family includes genera such as *Lyssavirus*, *Vesiculovirus*, and *Novirhabdovirus*, which infect a broad range of hosts, including humans, animals, plants, and insects [63]. Notably, some members of this family, such as the *Rabies virus* (*Lyssavirus*), pose significant public health risks due to their zoonotic potential and high mortality rates [63]. We identified 39 contigs associated with *Rhabdoviridae* in shrew and rodent samples. Phylogenetic analysis of the RdRp amino acid sequences revealed that these contigs are distributed across five distinct genera or form sister clades, including *Ledantevirus*, *Alphanemrhavirus*, *Almendravirus*, *Betanemrhavirus*, and *Novirhabdovirus*. The contigs exhibited varying levels of RdRp amino acid sequence identity with their closest relatives: 97.29 to 98.25% with *Ledantevirus*, 62.03 to 62.20% with *Alphanemrhavirus*, 33.98 to 37.86% with *Almendravirus*, 28.39 to 36.53% with *Betanemrhavirus*, and 21.20 to 22.56% with *Novirhabdovirus*. Furthermore, a novel clade was identified between *Almendravirus* and *Betanemrhavirus*, with RdRp amino acid sequence identities ranging from 29.79 to 31.77% and 26.31 to 29.67%, respectively. These findings highlight the substantial genetic diversity and evolutionary complexity of Rhabdoviridae within rodent and shrew populations, underscoring the discovery of potentially novel or highly divergent strains.

Nairoviridae

Nairoviridae are enveloped, negative-sense single-stranded RNA viruses characterized by segmented genomes, typically comprising three segments [64]. This family includes the genus *Orthonairovirus*, which infects a diverse range of hosts, including humans, animals, and ticks. Notably, members of this family, such as the *Crimean-Congo hemorrhagic fever virus*, pose significant public health risks due to their high pathogenicity and zoonotic potential [64]. In this study, we identified 50 contigs related to *Nairoviridae*. Phylogenetic analysis of the RdRp amino acid sequences revealed that these contigs are distributed across three distinct clades within the *Orthonairovirus*. Two groups of sequences clustered with *Huangpi tick virus* (KM817667) and *Orthonairovirus thiaforaense* (KR537450), showing RdRp amino acid sequence identities of 60.07% and 44.78 to 60.00%, respectively. Additionally, a contig from Yichang shrews formed a novel clade, with RdRp amino acid sequence identities ranging from 25.24 to 59.22% relative to other *Orthonairovirus* sequences.

Picornaviridae

Picornaviridae are non-enveloped, positive-sense single-stranded RNA viruses with non-segmented genomes [65]. This diverse family encompasses numerous genera, including *Enterovirus*, *Hepatovirus*, and *Cardiovirus*, which infect a broad spectrum of hosts, ranging from humans and animals to plants [65]. Certain members, such as *Enterovirus* and *Hepatovirus*, are particularly noteworthy due to their pathogenicity and potential for widespread transmission, making them significant public health concerns [65]. We identified 95 contigs affiliated with *Picornaviridae*. Phylogenetic analysis based on RdRp amino acid sequences revealed that these contigs are distributed across ten genera: *Parabovirus*, *Rabovirus*, *Kunsagivirus*, *Hepatovirus*, *Ampivirus*, *Kobuvirus*, *Rosavirus*, *Cardiovirus*, *Mupivirus*, and *Hunnivirus*. The RdRp amino acid sequence identities with the closest related sequences within these genera ranged as follows: 72.28 to 91.77% for *Parabovirus*, 95.71% for *Rabovirus*, 77.16 to 94.10% for *Kunsagivirus*, 66.44% for *Hepatovirus*, 15.28 to 28.05% for *Ampivirus*, 63.18 to 66.34% for *Kobuvirus*, 97.00% for *Rosavirus*, 77.45 to 89.83% for *Cardiovirus*, 79.36 to 86.35% for *Mupivirus*, and 89.16 to 96.52% for *Hunnivirus*. Additionally, some contigs formed novel clades, indicating the potential discovery of new lineages withinPicornaviridae. These findings highlight the genetic diversity of *Picornaviridae* in rodent and shrew populations, underscoring the presence of potentially novel or highly divergent strains.

Qinviridae

Qinviridae is a recently described family of non-enveloped, negative-sense single-stranded RNA viruses characterized by segmented genomes, typically consisting of four segments [66]. Although not as extensively studied as other viral families, *Qinviridae* is notable for its unique evolutionary adaptations and its ability to infect a diverse range of invertebrate hosts, particularly insects. In this study, we identified two contigs belonging to *Qinviridae* from shrews. Phylogenetic analysis based on RdRp amino acid sequences revealed that these contigs clustered with the previously reported *Wuhan insect virus* (KX883998) from the same region, with RdRp amino acid sequence identities ranging from 21.84 to 45.04%. These findings suggest the presence of *Qinviridae* in mammalian hosts, particularly shrews, indicating a potentially broader host range than previously recognized.

Genomic and protein domain diversity in mammalian RNA viruses

Although the characterization of RdRp provided valuable insights, we recognized the need for a more detailed exploration of viral genomic structures and functional domain diversity. Therefore, we conducted a comprehensive visualization and analysis of the genomic structures of 139 viruses harboring near-full-length RdRp sequences (Fig. 5). Our analysis revealed that most viral sequences maintained the classical genomic structures typical of their respective viral families (e.g., *Sedoreoviridae*, *Spinareoviridae*, *Flaviviridae*, *Arenaviridae*, *Hantaviridae*, *Nairoviridae*, *Peribunyaviridae*, *Phenuiviridae*, *Lispiviridae*, *Nyamiviridae*, *Astroviridae*). However, some viruses exhibited distinct structural variations.

Hepeviridae

The canonical genome structure of rodent hepatitis E virus (Rocahepevirus) typically includes the arrangement 5′-ORF4-ORF1-ORF3-ORF2-3′, where ORF4 and ORF1 have significant overlaps at the 5′ end, and ORF3 and ORF2 overlap at the 5′ end [55]. In some strains, ORF3 and ORF1 also show minor overlaps at the 3′ end of ORF1 [55]. Our findings revealed that the rodent hepatitis E virus circulating in this region exhibits the classic viral structure but with two distinct variants: one with overlapping ORF3 and ORF1 regions and another without. Phylogenetic analysis based on RdRp sequences supported the presence of two clades, indicating the coexistence of at least two distinct genomic structures of the hepatitis E virus in the region. Additionally, we identified a novel clade of hepatitis E virus from shrews, which exhibited a completely different genome structure. In this variant, the RdRp-coding ORF1 completely overlaps with the capsid protein-coding ORF2, and two additional open reading frames (ORFs X1 and X2) were present. This highlights the structural diversity of hepatitis E viruses among small mammals in the region.

Nodaviridae

Nodamura virus is typically a segmented RNA virus with two genome segments: RNA-1, encoding RdRp and the RNAi suppressor protein B2, and RNA-2, encoding the coat protein precursor [59]. We identified *Nodamura virus* strains in shrews and rodents from Yichang with the classic RNA-1 genome structure. Additionally, we discovered two unsegmented Nodamura virus strains, where the ORFs encoding RdRp and coat protein were located on a single contig. These strains showed 100% amino acid identity with previously reported strains OQ715788 and OQ715813.

Chuviridae

Chuviridae genomes exhibit structural diversity, with both circular and linear forms [49]. We identified a circular genome in *Morsusvirus* (Yichang Shrews chu-like virus/2016/2017 9), with the structure G-N-X1-RdRp, and a linear genome in *Rochuvirus* (Yichang Shrews chu-like virus/2016/2017 3).

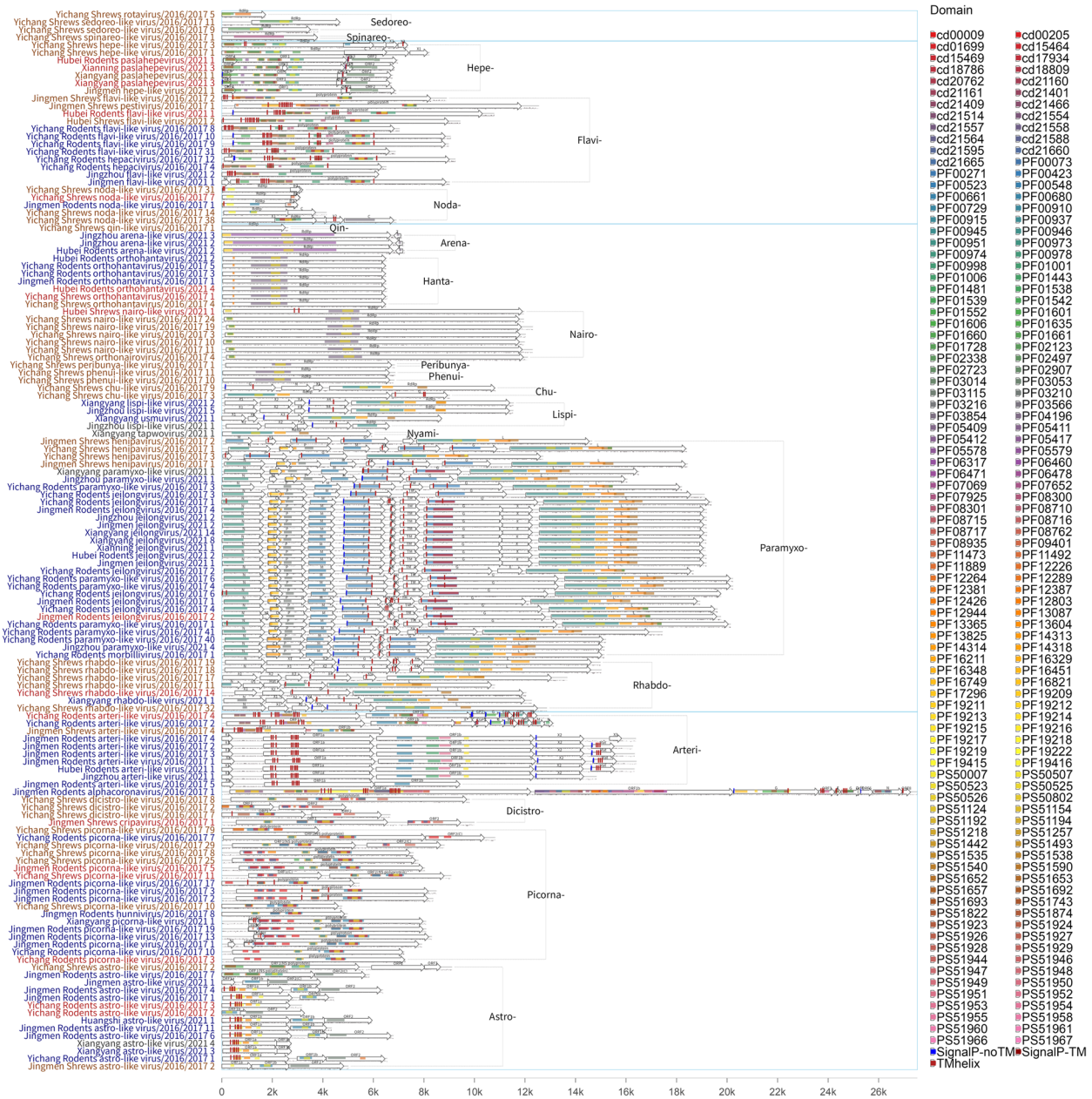

Fig. 5 Genomic architecture of near-complete mammal-associated viral genomes. Virus strains are listed on the left, with font colors corresponding to Fig. 4. White arrows indicate open reading frames (ORFs). Colored blocks represent different viral domains, with blue vertical rectangles indicating signal peptides and dark red vertical rectangles indicating predicted transmembrane regions. Different viral families are enclosed in light grey boxes, and viruses from different phyla are enclosed in light blue boxes

Paramyxoviridae

The canonical genome structure of viruses within Paramyxoviridae is 5′-N-P/C-M-F-G-L-3′ [61]. In this study, we identified several strains from shrews in Yichang and Jingzhou that belong to *Henipavirus*. Notably, these strains possess an additional open reading frame (ORF) of unknown function located between the M and F genes. Similarly, in a *Narmovirus* strain (Jinzhou paramyxo-like virus/2021 1), an ORF of unknown function was identified between the N and P/C genes. Additionally, two sequences closely related to *Belerina virus* (MN561699) were found to contain an ORF encoding a transmembrane protein (TM) between the F and G genes. In the *Jeilongvirus genus*, ORFs encoding both a small hydrophobic protein (SH) and TM were observed between the F and G genes; some strains also had an additional ORF

of unknown function between the G and L genes. Strains belonging to *Morbillivirus* were found to harbor an ORF encoding SH between the F and G genes. These findings reveal considerable genomic diversity within Paramyxoviridae, suggesting the evolution of novel viral lineages with potentially distinct biological functions.

Rhabdoviridae

The typical genome structure of viruses in *Rhabdoviridae* is 5′-N-P-M-G-L-3′ [63]. We identified a unique strain from shrews where only the L gene exhibited high conservation, containing multiple conserved protein domains. The other genes showed significant structural variation. While the N and G genes were identifiable through protein homology, the remaining genes were difficult to assign due to their divergence. This suggests a potential novel evolutionary path within *Rhabdoviridae*, emphasizing the importance of continued surveillance and characterization of these viruses.

Arteriviridae

The canonical genome structure of viruses in *Arteriviridae* is 5′-ORF1a-ORF1b-E-GP2-GP3-GP4-GP5-M–N-3′ [46]. In this study, we obtained three contigs displaying the classic arterivirus genome structure. However, due to potential sequencing or assembly errors, the X1 region in Yichang Rodents arteri-like virus/2016/2017 2 appears to be part of ORF1b. Additionally, we identified a new branch of arteriviruses in the phylogenetic tree with a completely different genome structure, characterized by an ORF of unknown function at the 5′ end and an additional ORF (X2) at the end of ORF1b. This may represent a novel genus within *Arteriviridae*, highlighting the ongoing evolution and diversification of arteriviruses.

Dicistroviridae

Viruses within *Dicistroviridae* typically have a genome structure of 5′-ORF1-ORF2-3′, where ORF1 encodes non-structural proteins and ORF2 encodes capsid proteins [51]. In this study, we identified a strain from shrews (Yichang Shrews dicistro-like virus/2016/2017 8) that encodes a single ORF. Functional domain analysis suggests that this ORF could encode both non-structural and capsid proteins, indicating a deviation from the typical dicistrovirus genome structure. This may represent a novel lineage within *Dicistroviridae*.

Picornaviridae

Most viruses in *Picornaviridae* possess a single ORF [65]. However, we identified a *Dicipivirus* (Yichang Shrews picorna-like virus/2016/2017 11) with two ORFs: ORF1 encoding capsid proteins and ORF2 encoding non-structural proteins. Additionally, Yichang Shrews picorna-like virus/2016/2017 29 and Yichang Rodents picorna-like virus/2016/2017 7 formed a new clade within *Picornaviridae*, each with two ORFs encoding non-structural proteins (ORF1) and capsid proteins (ORF2). These findings indicate significant genomic diversity within *Picornaviridae* and suggest the emergence of novel viral lineages.

These discoveries underscore the remarkable genomic and structural diversity of RNA viruses circulating among small wild mammals in the region. The identification of atypical genome structures in several virus families points to the presence of novel viral lineages, consistent with phylogenetic analyses based on RdRp sequences. This highlights the importance of comprehensive genomic exploration to understand the full spectrum of viral diversity and evolution.

Identifying viruses with high zoonotic potential

Prioritizing viruses based on zoonotic potential is critical for strategic resource allocation in surveillance and the assessment of viral spillover risks. We applied ZoonoticRank, a machine learning framework that synthesizes multiple virus-host interaction parameters—including genomic features, host range, and ecological context—to rank viruses according to their likelihood of cross-species transmission. This method not only optimizes monitoring efforts but also enhances our ability to anticipate and mitigate potential outbreaks, thereby bolstering public health preparedness.

Here, ZoonoticRank identified 19 viral sequences with "very high" zoonotic potential, 206 with "high" potential, and 410 with "medium" potential. Viruses classified with "very high" zoonotic potential predominantly belong to single-stranded RNA positive-sense (ssRNA(+)) families such as *Arteriviridae*, *Dicistroviridae*, *Flaviviridae*, *Nodaviridae*, and *Picornaviridae*, as well as single-stranded RNA negative-sense (ssRNA(-)) families like *Arenaviridae*, *Hantaviridae*, *Nairoviridae*, *Paramyxoviridae*, and *Rhabdoviridae* (Fig. 6a). Despite the sequences analyzed not being full-length genomes, the top 50 viruses with the highest predicted zoonotic potential included well-documented zoonotic pathogens such as *Mammarenaviruses*, *Betacoronaviruses*, *Hepaciviruses*, *Pestiviruses*, *Orthohantaviruses*, and *Orthonairoviruses* (Fig. 6b). The alignment of these predictions with known zoonotic agents underscores the robustness of machine learning-based approaches in assessing the risk of cross-species transmission.

Pathogenic RNA viruses in humans

Viruses with high sequence similarity to human-derived viruses may have a reduced capacity to breach species barriers and infect humans. To identify such viruses and assess the nucleotide changes necessary for potential rodent-to-human transmission, we employed BLASTn to align assembled viral

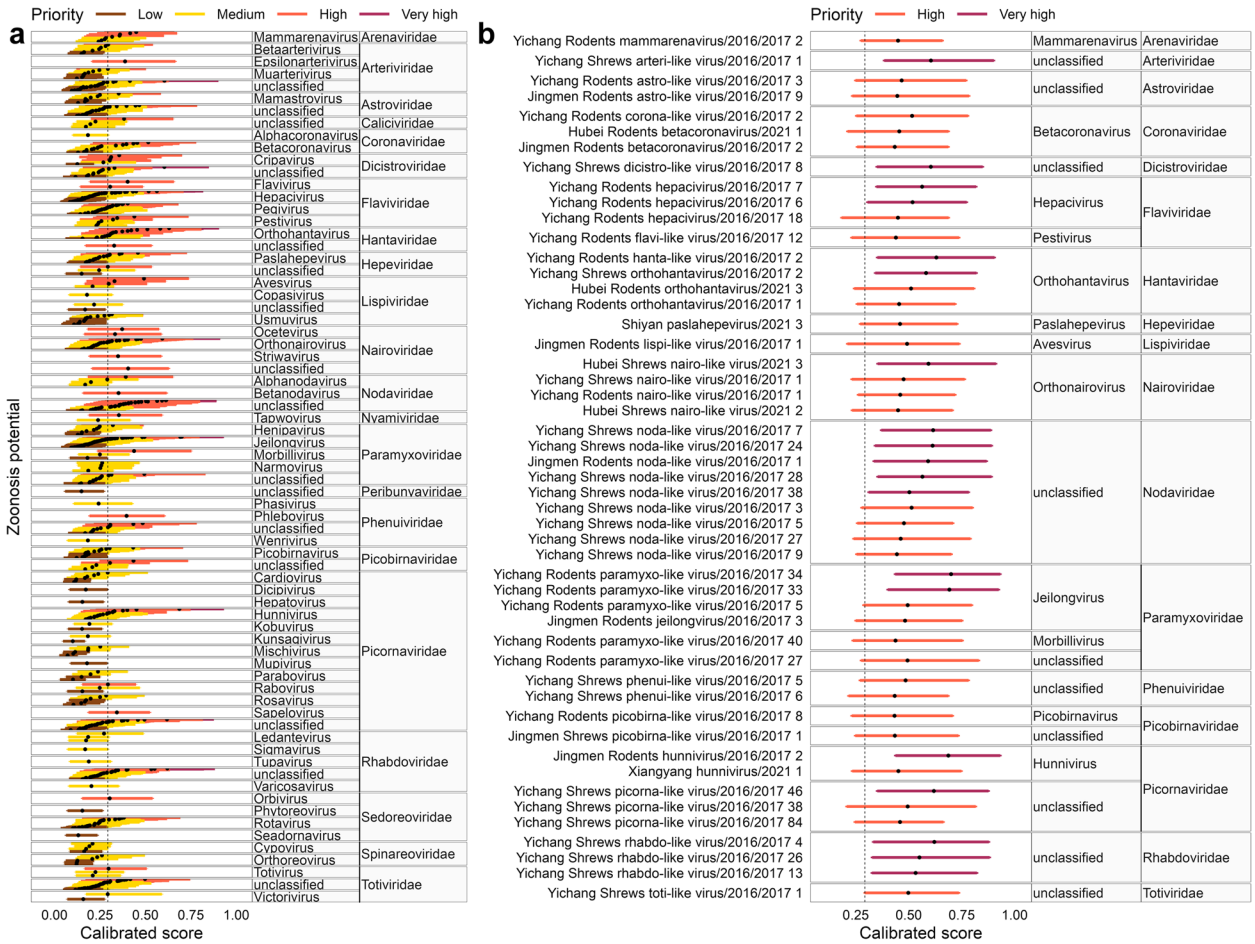

Fig. 6 Prediction of zoonotic potential. a Viral genera and families are depicted in distinct boxes. **b** The top 50 viruses with the highest predicted zoonotic potential are listed

contigs against an extensive viral database. We focused on sequences with over 90% similarity and more than 80% coverage to human-derived or human-isolated viral sequences. This approach identified two mammalian-associated viruses: *Orthohantavirus* and *Rotavirus*, each with seven sequences meeting these criteria, alongside one bacteriophage. The rodent-origin *Rotavirus* sequences exhibited 90.01–96.37% nucleotide identity, while *Orthohantavirus* sequences showed 91.58–98.47% nucleotide identity. *Rotavirus* is known for causing severe gastroenteritis in infants, characterized by watery diarrhea and vomiting. *Orthohantaviruses*, primarily rodent-borne, are significant zoonotic pathogens in humans, causing HFRS and HPS.

Given the pathogenic potential of *Orthohantaviruses*, our study focused on the genetic divergence between rodent-derived and human-derived viruses. *Orthohantaviruses* are enveloped, single-stranded RNA viruses with genomes comprising three segments: S (small), M (medium), and L (large). The S segment encodes the nucleocapsid protein (NP), the M segment encodes glycoproteins Gn and Gc (GP), and the L segment encodes the RdRp. These viral proteins are crucial for viral replication, immune evasion, and pathogenesis.

Our findings revealed that rodent-derived *Hantaan virus* (HTNV) sequences were highly similar to local human-derived strains, whereas rodent-derived *Seoul virus* (SEOV) were closely related to the Orthohantavirus seoulense strain Hu02-258/NGS from Korea. Two HTNV L segment sequences shared 98.32–99.95% amino acid identity with human-derived viral sequences, the M segment showed 98.73–100%, and the S segment showed 97.67–99.77%. Additionally, two SEOV S segment sequences exhibited 99.35–99.40% amino acid identity.

The high degree of amino acid sequence similarity suggests that rodent-derived *Orthohantaviruses* may require minimal nucleotide changes to breach species barriers and spill over into humans. To further explore convergent mutation patterns between rodent-derived and human-derived *Orthohantavirus* sequences, we aligned human-derived Orthohantavirus sequences to rodent-derived

Orthohantavirus. Single nucleotide polymorphisms (SNPs) were rare in the 5′ UTR regions, with most sequence variations occurring in coding regions (Fig. 7). The HTNV exhibited 10, 14, and 27 non-synonymous mutations in the L, M, and S segments, respectively, with 4, 2, and 14 mutations located within B cell linear epitope regions (Fig. 7a–c). SEOV had 14 non-synonymous mutations in the L segment, with 9 within B cell epitope regions (Fig. 7d). Mutations within antigenic epitopes accounted for 14–64% of all non-synonymous mutations.

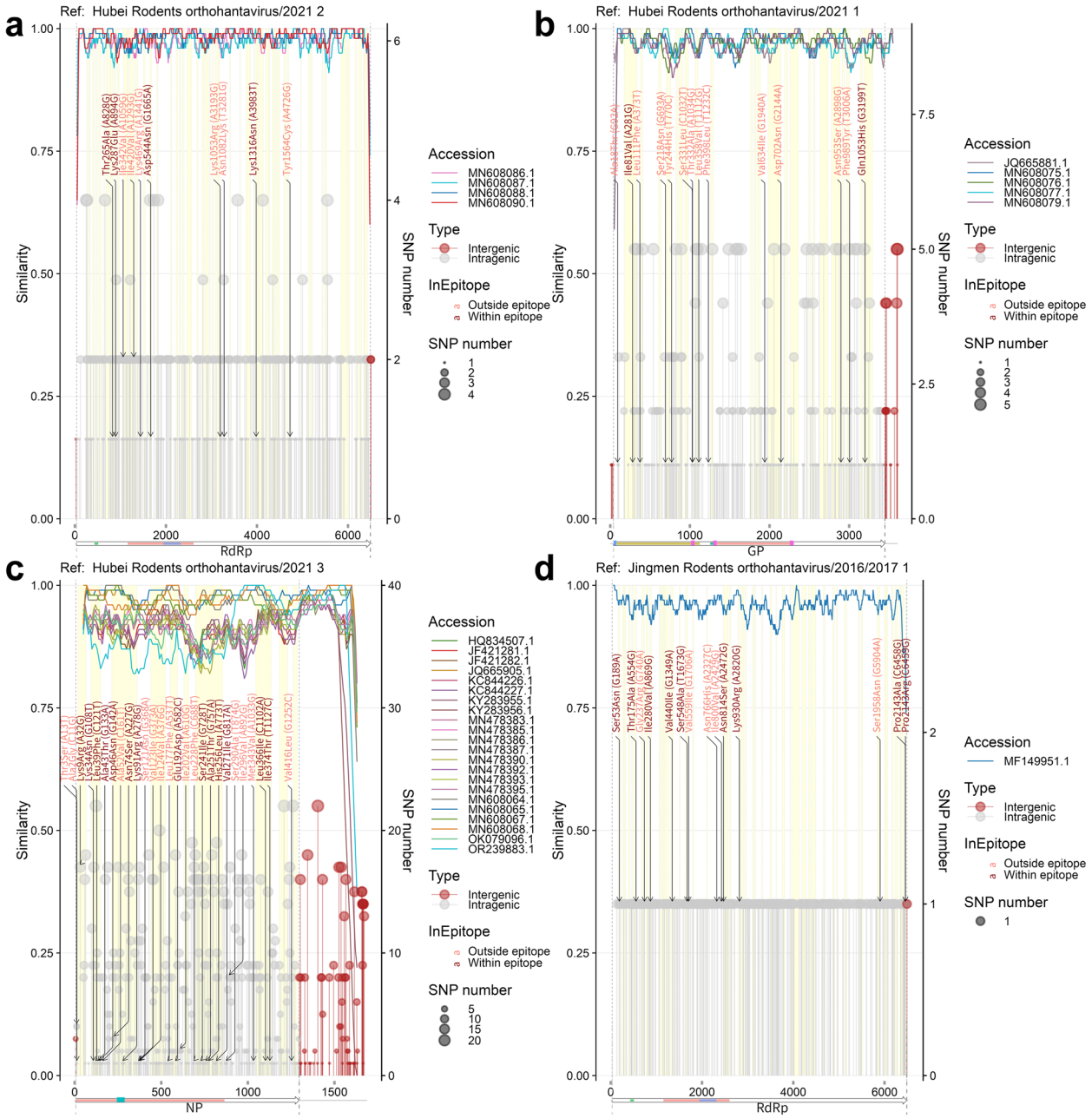

Fig. 7 Comparison of hantaviruses from rodents and humans. **a**–**d** Comparison of HTNV-L, HTNV-M, HTNV-S, and SEOV-L segments between human-derived and rodent-derived hantaviruses. Each panel includes: (i) Simplot analysis of genetic similarity using rodent-derived hantavirus as a reference; (ii) amino acid variation, with dark red mutations located in antigenic regions and light red mutations in non-antigenic regions; (iii) circles representing nucleotide variation, with circle size corresponding to the number of variations and colors indicating gene regions (grey for coding regions, red for 5′ and 3′ UTRs); and (iv) genome structure, with white arrows indicating ORFs and colored blocks representing different protein domains. The light yellow background highlights antigenic regions

Discussion

In recent decades, the emergence of infectious diseases originating from wildlife has become a significant global health concern, with small mammals playing a critical role in the transmission of zoonotic viruses. Our study, leveraging high-throughput sequencing, maps the viral landscape of wild small mammals over the past decade. Findings reveal that species such as *Anourosorex squamipes* and *Apodemus agrarius* are significant intermediaries in viral transmission, while *Rattus norvegicus* exhibits the highest viral diversity. The central Plain of Hubei Province emerges as a pivotal region for both viral spread towards surrounding cities and the enrichment of viral types in these areas. Phylogenetic and genomic analyses identified numerous novel mammalian-related viruses, including *Arenaviridae*, *Hantaviridae*, *Nairoviridae*, *Paramyxoviridae*, *Rhabdoviridae, Arteriviridae*, *Dicistroviridae*, *Flaviviridae*, *Nodaviridae*, and *Picornaviridae*. Notably, zoonotic risk assessments highlight the significant pathogenic potential of *Mammarenaviruses*, *Betacoronaviruses*, *Hepaciviruses*, *Pestiviruses*, *Orthohantaviruses*, and *Orthonairoviruses*. Of particular concern is the *orthohantavirus*, which displayed a 100% amino acid identity with human-derived strains. This high degree of genetic similarity, coupled with minimal nucleotide variation, suggests a substantial potential for immune evasion and cross-species transmission. These findings underline the urgent need for ongoing surveillance and the development of targeted public health interventions to mitigate the risks posed by these emerging zoonotic viruses.

Previous studies have revealed that host species characteristics are the primary determinants of viral diversity in mammals, with geographical factors playing a supporting role by influencing the environmental context in which these species live and interact [10, 16]. The biological and ecological traits inherent to certain species, such as immune system complexity, longevity, reproductive strategies, and social behaviors, contribute significantly to the diversity of viruses they carry. Geographic factors such as climate, habitat type, and local biodiversity can also modulate viral diversity by affecting host population dynamics and interspecies interactions [8]. Our study observed that the central Plain of Hubei Province acts as both a zone of viral spread towards surrounding cities and a zone of viral enrichment in these cities, aligning with the idea that certain landscapes, due to their ecological complexity and the convergence of multiple host species, can serve as hotspots for viral emergence [67]. Our study also revealed distinct patterns in the viral diversity of shrews compared to rodents. Shrews, unlike rodents, tend to carry viruses that exhibit a high degree of host species specificity. This finding is supported by recent research on six shrew species along the eastern coast of China, which revealed significant interspecies differences in viral diversity [68]. The species-specific viral diversity in shrews may be attributed to their unique ecological niches and lower rates of interspecies contact compared to rodents, which are more likely to encounter and transmit viruses to a variety of other species, including humans [69]. While our results support that hosts taxonomy is crucial in shaping the virome, particularly in shrews, they also emphasize the significant role of geographic factors in regions like the central Plain. This dual influence underscores the complexity of predicting viral diversity patterns and suggests that both host and environmental factors must be considered in tandem to fully understand the dynamics of viral emergence and transmission. Future research should adopt a holistic approach, integrating both host and environmental variables to better predict and mitigate the risks of viral emergence.

The heterogeneous nature of disease emergence across different host species is a critical factor in understanding viral transmission dynamics. Our binary network analysis of shared viruses among host species highlights the importance of identifying key host species and ranking their transmission potential. Our findings align with previous research indicating that certain rodent species, such as *Rattus norvegicus* and *Apodemus agrarius,* are particularly effective in harboring and transmitting a broad array of zoonotic viruses, including *hantaviruses* and *coronaviruses* [34, 69]. The ability to rank host species by their viral transmission potential has significant implications for public health strategies. By identifying and targeting the key species with the highest transmission potential, more focused and effective interventions can be implemented. For instance, rodent control programs could prioritize high-risk species like *Rattus norvegicus* in urban areas, where their interactions with humans are most frequent and consequential [70]. Beyond direct control measures, natural ecological interactions can also be leveraged to mitigate viral transmission risk. The presence of predators and competitors in the ecosystem can naturally limit the population densities of key host species, thereby reducing the overall transmission of viruses. For example, predators can reduce *hantavirus* prevalence by controlling the population of rodent hosts [71]. Similarly, competitive interactions among rodent species can alter the composition of the viral community by suppressing dominant species with higher transmission potentials [72]. Implementing strategies that support or mimic these natural ecological processes could serve as a sustainable approach to managing the risk of viral outbreaks.

Limitations

This study highlights the limitations of current pooling strategies, which can obscure important details regarding viral diversity and distribution among host species. Pooling samples from multiple individuals may dilute or mask the presence of low-abundance viruses, leading to an incomplete picture of the virome within specific host populations. Recent single-animal high-throughput sequencing has shown significant potential for overcoming these limitations. By sequencing individual animals, researchers can identify a broader spectrum of viruses, including those present at low levels, and establish more precise links between specific viruses, their host species, and environmental factors [13, 67, 73, 74].

In this study, machine learning was employed to rank the pathogenic potential of viruses. *Mammarenavirus* is classified with "high" zoonotic potential. However, functional studies are critical for assessing their pathogenicity, transmissibility, and zoonotic potential [67]. Recent evidence suggests that some rodent-borne *Mammarenavirus* can replicate in human cell lines, indicating a potential for human infection [75]. The absence of in vitro and in vivo characterization limits our ability to directly assess the replication, pathogenicity, and host range of the identified viruses. Serological studies have detected evidence of human infections with rodent-borne *Wenzhou virus*, suggesting possible human circulation [76]. Here, we did not conduct serological assays to determine the extent of human infections in our study region.

Conclusion

This study provides critical insights into the viral diversity harbored by wild small mammals, particularly rodents and shrews. The findings suggest that the central Plain of Hubei Province serves as a pivotal area for viral transmission among small mammals, with significant implications for understanding zoonotic risks and guiding future epidemiological investigations. Future research should focus on functional studies, targeted interventions, and a deeper understanding of the ecological and evolutionary drivers of viral emergence to prevent zoonotic disease outbreaks.

Abbreviations

COI	Cytochrome c oxidase I
HFRS	Hemorrhagic fever with renal syndrome
HPS	Hantavirus pulmonary syndrome
HTNV	Hantaan virus
ICTV	International Committee on Taxonomy of Viruses
MERS	Middle East respiratory syndrome
NP	Nucleocapsid protein
PCoA	Principal coordinate analysis
RdRp	RNA-dependent RNA polymerase
SARS	Severe acute respiratory syndrome
SEOV	Seoul virus
SNPs	Single nucleotide polymorphisms
vOTUs	Viral unique operational taxonomic units

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-025-02059-0.

Supplementary Material 1. Table S1. Sample metadata.

Supplementary Material 2. Table S2. Summary of the Viruses detected in this study.

Supplementary Material 3. Table S3. The expression of the Viruses detected in this study.

Acknowledgements

We would like to extend our heartfelt thanks to our research team for their exceptional collaboration and collective efforts throughout this project. Their dedication, diverse expertise, and willingness to share knowledge were crucial to the successful completion of this study. We are grateful to each team member for their commitment and for the invaluable contributions they made to this work.

Authors' contributions

N.L. and Z.H. designed and supervised the project. N.L. wrote the manuscript. N.L. performed the bioinformatic analysis. B.H., L.Z., M.G., K.P., and Q.W. performed the PCR and sequencing experiments. B.H., S.L., J.B., W.X., M.G., Q.W., and D.C. contributed to the animal collection and sampling treatments. The authors read and approved the final manuscript.

Funding

This research was supported by grants from the National Key Research and Development Program of China (2023YFC2605504), Major Project of Guangzhou National Laboratory (No. GZNL2023A01001), Hubei Provincial Natural Science Foundation Innovation Group (No. 2024AFA024), Hubei Provincial Natural Science Foundation General Project (No. 2024AFB1001), Hubei Provincial Public Health Outstanding Young Talents Project, and Chu Tian Excellence Program (Medical and Health). The funders had no role in the study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Data availability

The raw sequencing read data generated in 2021 have been deposited in the National Genomics Data Center (NGDC), Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation (CNCB) repository under BioProject accession number PRJCA030413 (available at https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA030413) and GSA accession number CRA019228 (https://ngdc.cncb.ac.cn/gsa/browse/CRA019228). Viral contigs assembled from this 2021 dataset are accessible in GenBase, a component of NGDC, under accession numbers C_AA087507.1 to C_AA087723.1 (see https://ngdc.cncb.ac.cn/genbase/browse/0004769). Additionally, this study utilized publicly available data from prior research conducted in 2014 and 2016–2017 in Hubei province, under BioProject accession numbers PRJNA375958 and PRJNA953205, respectively. Viral contigs assembled from these datasets are available on Zenodo (https://zenodo.org/records/13979871).
Additionally, this study utilized publicly available data from prior research conducted in 2014 and 2016–2017 in Hubei province, under BioProject accession numbers PRJNA375958 and PRJNA953205, respectively. Viral contigs assembled from these datasets are available on Zenodo (https://zenodo.org/records/13979871).

Declarations

Ethics approval and consent to participate

This study received approval from the Experimental Ethics Committee of the Hubei Provincial Center for Disease Control and Prevention (Project No. 2021–034–02).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details
[1]State Key Laboratory of Virology and Biosafety, Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan 430071, China. [2]Hubei Provincial Center for Disease Control and Prevention, Wuhan, Hubei 430079, China. [3]Xianning Municipal Center for Disease Control and Prevention, Xianning, Hubei 437000, China.

Received: 23 October 2024 Accepted: 7 February 2025
Published online: 03 March 2025

References

1. Gibb R, Redding DW, Chin KQ, Donnelly CA, Blackburn TM, Newbold T, et al. Zoonotic host diversity increases in human-dominated ecosystems. Nature. 2020;584:398–402.
2. Williams SH, Che X, Oleynik A, Garcia JA, Muller D, Zabka TS, et al. Discovery of two highly divergent negative-sense RNA viruses associated with the parasitic nematode, Capillaria hepatica, in wild Mus musculus from New York City. J Gen Virol. 2019;100:1350–62.
3. Williams SH, Che X, Garcia JA, Klena JD, Lee B, Muller D, et al. Viral diversity of house mice in New York City. mBio. 2018;9:e01354-17.
4. Kane Y, Tendu A, Li R, Chen Y, Mastriani E, Lan J, et al. Viral diversity in wild and urban rodents of yunnan province. Emerg Microbes Infect. 2024;13:2290842.
5. Blasdell KR, Morand S, Laurance SGW, Doggett SL, Hahs A, Trinh K, et al. Rats and the city: Implications of urbanization on zoonotic disease risk in Southeast Asia. Proc Natl Acad Sci. 2022;119:e2112341119.
6. Daszak P, Olival KJ, Li H. A strategy to prevent future epidemics similar to the 2019-nCoV outbreak. Biosafety and Health. 2020;2:6–8.
7. Meerburg BG, Singleton GR, Kijlstra A. Rodent-borne diseases and their risks for public health. Crit Rev Microbiol. 2009;35:221–70.
8. Mollentze N, Streicker DG. Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts. Proc Natl Acad Sci USA. 2020;117:9423–30.
9. Han BA, Kramer AM, Drake JM. Global patterns of zoonotic disease in mammals. Trends Parasitol. 2016;32:565–77.
10. Luis AD, Hayman DTS, O'Shea TJ, Cryan PM, Gilbert AT, Pulliam JRC, et al. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proceedings Biological sciences. 2013;280:20122753.
11. Akhtar N, Hayee S, Idnan M, Nawaz F, BiBi S. Rodents human zoonotic pathogens transmission: historical background and future prospects. In: Shah MM, editor. Rodents and their role in ecology, medicine and agriculture. London: IntechOpen; 2023.
12. Rabiee MH, Mahmoudi A, Siahsarvie R, Kryštufek B, Mostafavi E. Rodent-borne diseases and their public health importance in iran. PLoS Negl Trop Dis. 2018;12:e0006256.
13. Chen Y-M, Hu S-J, Lin X-D, Tian J-H, Lv J-X, Wang M-R, et al. Host traits shape virome composition and virus transmission in wild small mammals. Cell. 2023;186:4662–75.
14. Yi B, Deng Q, Guo C, Li X, Wu Q, Zha R, et al. Evaluating the zoonotic potential of RNA viromes of rodents provides new insight into rodent-borne zoonotic pathogens in Guangdong. China One Health. 2023;17:100631.
15. Wu Z, Han Y, Liu B, Li H, Zhu G, Latinne A, et al. 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. 2021;9:18.
16. Wu Z, Lu L, Du J, Yang L, Ren X, Liu B, et al. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. Microbiome. 2018;6:178.
17. Cui X, Fan K, Liang X, Gong W, Chen W, He B, et al. Virus diversity, wildlife-domestic animal circulation and potential zoonotic viruses of small mammals, pangolins and zoo animals. Nat Commun. 2023;14:2488.
18. Raghwani J, Faust CL, François S, Nguyen D, Marsh K, Raulo A, et al. Seasonal dynamics of the wild rodent faecal virome. Mol Ecol. 2022;32:4763–76.
19. Zhu W, Yang J, Lu S, Huang Y, Jin D, Pu J, et al. Novel pegiviruses infecting wild birds and rodents. Virologica Sinica. 2022;37:208–14.
20. Zhu W, Yang J, Lu S, Jin D, Pu J, Wu S, et al. RNA virus diversity in birds and small mammals from Qinghai-Tibet Plateau of China. Front Microbiol. 2022;13:780651.
21. Zhu W, Yang J, Lu S, Jin D, Wu S, Pu J, et al. Discovery and evolution of a divergent coronavirus in the plateau pika from China that extends the host range of alphacoronaviruses. Front Microbiol. 2021;12:755599.
22. Wu Z, Liu B, Du J, Zhang J, Lu L, Zhu G, et al. Discovery of diverse rodent and bat pestiviruses with distinct genomic and phylogenetic characteristics in several Chinese provinces. Front Microbiol. 2018;9:2562.
23. Wu Z, Du J, Lu L, Yang L, Dong J, Sun L, et al. Detection of Hantaviruses and Arenaviruzses in three-toed jerboas from the Inner Mongolia Autonomous Region. Emerg Microbes Infect. 2018;7:35.
24. Wang W, Lin X-D, Zhang H-L, Wang M-R, Guan X-Q, Holmes EC, et al. Extensive genetic diversity and host range of rodent-borne coronaviruses. Virus Evol. 2020;6:veaa078.
25. Wang B, Cai C-L, Li B, Zhang W, Zhu Y, Chen W-H, et al. Detection and characterization of three zoonotic viruses in wild rodents and shrews from Shenzhen city. China Virologica Sinica. 2017;32:290–7.
26. Tirera S, de Thoisy B, Donato D, Bouchier C, Lacoste V, Franc A, et al. The influence of habitat on viral diversity in neotropical rodent hosts. Viruses-Basel. 2021;13:1690.
27. Tan Z, Yu H, Xu L, Zhao Z, Zhang P, Qu Y, et al. Virome profiling of rodents in Xinjiang Uygur Autonomous Region, China: Isolation and characterization of a new strain of Wenzhou virus. Virology. 2019;529:122–34.
28. Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, Delwart EL. The fecal viral flora of wild rodents. PLoS Pathog. 2011;7:e1002218.
29. Peterson AC, Sharma H, Kumar A, Ghersi BM, Emrich SJ, Vandegrift KJ, et al. Rodent virus diversity and differentiation across post-Katrina New Orleans. Sustainability. 2021;13:8034.
30. Hansen TA, Mollerup S, Nguyen N, White NE, Coghlan M, Alquezar-Planas DE, et al. High diversity of picornaviruses in rats from different continents revealed by deep sequencing. Emerg Microbes Infect. 2016;5:1–8.
31. Wang D, Yang X, Ren Z, Hu B, Zhao H, Yang K, et al. Substantial viral diversity in bats and rodents from East Africa: insights into evolution, recombination, and cocirculation. Microbiome. 2024;12:72.
32. Du H, Zhang L, Zhang X, Yun F, Chang Y, Tuersun A, et al. Metagenome-Assembled Viral Genomes Analysis Reveals Diversity and Infectivity of the RNA Virome of Gerbillinae Species. Viruses-Basel. 2022;14:356.
33. Lwande OW, Mohamed N, Bucht G, Ahlm C, Olsson G, Evander M. Seewis hantavirus in common shrew (Sorex araneus) in Sweden. Virol J. 2020;17:198.
34. Clay CA, Lehmer EM, Jeor SS, Dearing MD. Sin nombre virus and rodent species diversity: a test of the dilution and amplification hypotheses. PLoS ONE. 2009;4:e6467.
35. Gervasi SS, Civitello DJ, Kilvitis HJ, Martin LB. The context of host competence: a role for plasticity in host–parasite dynamics. Trends Parasitol. 2015;31:419–25.
36. Mollentze N, Babayan SA, Streicker DG. Identifying and prioritizing potential human-infecting viruses from their genome sequences. PLoS Biol. 2021;19:e3001390.
37. Grange ZL, Goldstein T, Johnson CK, Anthony S, Gilardi K, Daszak P, et al. Ranking the risk of animal-to-human spillover for newly discovered viruses. Proc Natl Acad Sci. 2021;118:e2002324118.
38. Cai J, Huang S, Zhang P, Yang W. Epidemiological characteristics and the incidence trend prediction of measles in hubei province, 2005–2018. J Public Health Prev Med. 2020;31:42–5.
39. Wang Q, Yue M, Yao P, Zhu C, Ai L, Hu D, et al. Epidemic trend and molecular evolution of HV family in the main hantavirus epidemic areas from 2004 to 2016, in P.R. China. Front Cell Infect Microbiol. 2020;10:584814.
40. Li J, Ling J, Liu D, Liu J, Liu Y, Wei F, et al. Genetic characterization of a new subtype of Hantaan virus isolated from a hemorrhagic fever with renal syndrome (HFRS) epidemic area in Hubei Province. China Arch Virol. 2012;157:1981–7.
41. Cao K, Wu X, Yang M, Chen C, Zhang X, Jiang D, et al. Prevalence of hepatitis E virus in China from 1997 to 2022: a systematic review and meta-analysis. Front Public Health. 2023;11:1243408.
42. Wang W, Rosenberg MW, Chen H, Gong S, Yang M, Deng D. Epidemiological characteristics and spatiotemporal patterns of hand, foot, and mouth disease in hubei, China from 2009 to 2019. PLoS ONE. 2023;18:e0287539.
43. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.

44. Huddleston J, Hadfield J, Sibley TR, Lee J, Fay K, Ilcisin M, et al. Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. J Open S Softw. 2021;6:2906.
45. Radoshitzky SR, Buchmeier MJ, Charrel RN, Gonzalez J-PJ, Günther S, Hepojoki J, et al. ICTV virus taxonomy profile: Arenaviridae 2023. J Gen Virol. 2023;104:1891.
46. Brinton MA, Gulyaeva AA, Balasuriya UBR, Dunowska M, Faaberg KS, Goldberg T, et al. ICTV Virus Taxonomy Profile: Arteriviridae 2021. J Gen Virol. 2021;102:001632.
47. Astroviridae. https://ictv.global/report_9th/RNApos/Astroviridae. Accessed 12 Feb 2025
48. Vinjé J, Estes MK, Esteves P, Green KY, Katayama K, Knowles NJ, et al. ICTV virus taxonomy profile: Caliciviridae. J Gen Virol. 2019;100:1469–70.
49. Kuhn JH, Dheilly NM, Junglen S, Paraskevopoulou S, Shi M, Di Paola N. ICTV virus taxonomy profile: Jingchuvirales 2023: this article is part of the ICTV Virus Taxonomy Profiles collection. J Gen Virol. 2023;104:001924.
50. Woo PCY, De Groot RJ, Haagmans B, Lau SKP, Neuman BW, Perlman S, et al. ICTV virus taxonomy profile: coronaviridae 2023: this article is part of the ICTV virus taxonomy profiles collection. J Gen Virol. 2023;104:001843.
51. Valles SM, Chen Y, Firth AE, Guérin DMA, Hashimoto Y, Herrero S, et al. ICTV virus taxonomy profile: dicistroviridae. J Gen Virol. 2017;98:355–6.
52. Li J-M, Wang F, Ye G, Paraskevopoulou S. ICTV virus taxonomy profile: lispiviridae 2023. J Gen Virol. 2023;104:001869.
53. Simmonds P, Becher P, Bukh J, Gould EA, Meyers G, Monath T, et al. ICTV virus taxonomy profile: flaviviridae. J Gen Virol. 2017;98:2–3.
54. Bradfute SB, Calisher CH, Klempa B, Klingström J, Kuhn JH, Laenen L, et al. ICTV virus taxonomy profile: Hantaviridae 2024: this article is part of the ICTV virus taxonomy profiles collection. J Gen Virol. 2024;105:001975.
55. Purdy MA, Drexler JF, Meng X-J, Norder H, Okamoto H, Van Der Poel WHM, et al. ICTV virus taxonomy profile: Hepeviridae 2022: this article is part of the ICTV virus taxonomy profiles collection. J Gen Virol. 2022;103:001778.
56. Matthijnssens J, Attoui H, Bányai K, Brussaard CPD, Danthi P, del Vas M, et al. ICTV Virus Taxonomy Profile: Spinareoviridae 2022. J Gen Virol. 2022;103:1781.
57. Delmas B, Attoui H, Ghosh S, Malik YS, Mundt E, Vakharia VN, et al. ICTV virus taxonomy profile: Picobirnaviridae. J Gen Virol. 2019;100:133–4.
58. Matthijnssens J, Attoui H, Bányai K, Brussaard CPD, Danthi P, del Vas M, et al. ICTV virus taxonomy profile: sedoreoviridae 2022. J Gen Virol. 2022;103:1782.
59. Sahul Hameed AS, Ninawe AS, Nakai T, Chi SC, Johnson KL, ICTV Report Consortium. ICTV virus taxonomy profile: nodaviridae. J Gen Virol. 2019;100:3–4.
60. Dietzgen RG, Firth AE, Jiāng D, Junglen S, Kondo H, Kuhn JH, et al. ICTV virus taxonomy profile: Nyamiviridae 2021. J Gen Virol. 2021;102:1681.
61. Rima B, Balkema-Buschmann A, Dundon WG, Duprex P, Easton A, Fouchier R, et al. ICTV virus taxonomy profile: paramyxoviridae. J Gen Virol. 2019;100:1593–4.
62. Sasaya T, Palacios G, Briese T, Di Serio F, Groschup MH, Neriya Y, et al. ICTV virus taxonomy profile: Phenuiviridae 2023. J Gen Virol. 2023;104:1893.
63. Walker PJ, Freitas-Astúa J, Bejerman N, Blasdell KR, Breyta R, Dietzgen RG, et al. ICTV virus taxonomy profile: Rhabdoviridae 2022. J Gen Virol. 2022;103:1689.
64. Kuhn JH, Alkhovsky SV, Avšič-Županc T, Bergeron É, Burt F, Ergünay K, et al. ICTV Virus Taxonomy Profile: Nairoviridae 2024. J Gen Virol. 2024;105:1974.
65. Zell R, Delwart E, Gorbalenya AE, Hovi T, King AMQ, Knowles NJ, et al. ICTV virus taxonomy profile: picornaviridae. J Gen Virol. 2017;98:2421–2.
66. Wolf YI, Koonin EV, Krupovic M, Kuhn JH. ICTV virus taxonomy profile: Qinviridae 2023. J Gen Virol. 2023;104:1905.
67. Carroll D, Daszak P, Wolfe ND, Gao GF, Morel CM, Morzaria S, et al. The global virome project. Science. 2018;359:872–4.
68. Zhang J-T, Hu Z-Y, Tang F, Liu Y-T, Tan W-L, Ma X-F, et al. Decoding the RNA viromes in shrew lungs along the eastern coast of China. Npj Biofilms Microbiomes. 2024;10:1–13.
69. Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. Proc Natl Acad Sci. 2015;112:7039–44.
70. Himsworth CG, Parsons KL, Jardine C, Patrick DM. Rats, cities, people, and pathogens: a systematic review and narrative synthesis of literature regarding the ecology of rat-associated zoonoses in urban centers. Vector Borne Zoonotic Dis. 2013;13:349–59.
71. Guterres A, De Lemos ERS. Hantaviruses and a neglected environmental determinant. One Health. 2018;5:27–33.
72. Eleftheriou A, Kuenzi AJ, Luis AD. Heterospecific competitors and seasonality can affect host physiology and behavior: key factors in disease transmission. Ecosphere. 2021;12:e03494.
73. Liu Q, Cui F, Liu X, Fu Y, Fang W, Kang X, et al. Association of virome dynamics with mosquito species and environmental factors. Microbiome. 2023;11:101.
74. Wang J, Pan Y, Yang L, Yang W, Lv K, Luo C, et al. Individual bat virome analysis reveals co-infection and spillover among bats and virus zoonotic potential. Nat Commun. 2023;14:4079.
75. Luo X-L, Lu S, Qin C, Shi M, Lu X-B, Wang L, et al. Emergence of an ancient and pathogenic mammarenavirus. Emerg Microbes Infect. 2023;12:e2192816.
76. Guo L, Liu S, Song J, Han L, Zhang H, Wu C, et al. Seroprevalence of Wenzhou virus in China. Biosaf Health. 2020;2:152–6.

Publisher's Note