One Health 17 (2023) 100631

Contents lists available at ScienceDirect

One Health

journal homepage: www.elsevier.com/locate/onehlt

Evaluating the zoonotic potential of RNA viromes of rodents provides new insight into rodent-borne zoonotic pathogens in Guangdong, China

Boyang Yi [a,b], Qiang Deng [a,b], Cheng Guo [c], Xiaokang Li [a,b], Qin Wu [a,b], Renyun Zha [a,b], Xianhua Wang [d,*], Jiahai Lu [a,b,e,f,g,**]

[a] School of Public Health, Sun Yat-Sen University, Guangzhou 510080, China
[b] One Health Center of Excellence for Research & Training, Sun Yat-Sen University, Guangzhou 510080, China
[c] Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York 10032, USA
[d] School of Health Sciences, Guangzhou Xinhua University, Guangzhou 510520, China
[e] National Medical Products Administration Key Laboratory for Quality Monitoring and Evaluation of Vaccines and Biological Products, Guangzhou 510080, China
[f] Hainan Key Novel Thinktank "Hainan Medical University 'One Health' Research Center", Haikou 571199, China
[g] Key Laboratory of Tropical Diseases Control, Sun Yat-Sen University, Ministry of Education, Guangzhou 510080, China

ARTICLE INFO

Keywords:
Rodents
Viromes
Zoonotic potential
One health
Guangdong

ABSTRACT

Emerging and re-emerging infectious diseases have been on the rise, with a significant proportion being zoonotic. Rodents, as the natural reservoirs of numerous diverse zoonotic viruses, pose a substantial threat to human health. To investigate the diversity of known and unknown viruses harbored by rodents in Guangdong (southern province of China), we conducted a comprehensive analysis of viral genomes through metagenomic sequencing of organs from 194 rodents. Our analysis yielded 2163 viral contigs that were assigned to 25 families known to infect a wide range of hosts, including vertebrates, invertebrates, amoebas, and plants. The viral compositions vary considerably among different organs, but not in rodent species. We also assessed and prioritized zoonotic potential of those detected viruses. Ninety-two viral species that are either known to infect vertebrates and invertebrates or only vertebrates were identified, among which 21 are considered high-risk to humans. The high-risk viruses included members of the *Hantavirus*, *Picobirnaviruses*, *Astroviruses* and *Pestivirus*. The phylogenetic trees of four zoonotic viruses revealed features of novel viral genomes that seem to fit evolutionarily into a zone of viruses that potentially pose a risk of transmission to humans. Recognizing that zoonotic diseases are a One Health issue, we approached the problem of identifying the zoonotic risk from rodent-transmitted disease in the Guangdong province by performing next-generation sequencing to look for potentially zoonotic viruses in these animals.

1. Introduction

In recent years, various emerging and re-emerging infectious diseases have been increasing, of which >60% are zoonotic diseases. This trend presents a serious threat to global public health and social development, resulting in >1 billion illnesses and millions of deaths annually [1,2]. Multiple factors can influence the risk of transmission of zoonotic viruses, including animal population density, land-use issues that alter the extent of human-animal contact, climatic factors, nutrient availability, and other factors [3].

Rodents are among the most speciose mammals in the world, with 33 families and 2277 species, accounting for approximately 43% of all mammal species [4]. They have a global distribution, ranging from highly anthropized to pristine natural habitats, and closely interact with the human ecosystem, acting as a conduit for the spread of viruses between people and other wildlife [5,6]. Rodents can directly or indirectly transmit various pathogenic microorganisms, such as those in the families *Arenaviridae, Flaviviridae, Hantaviridae, Picobirnaviridae, Picornaviridae, Reoviridae, and Rhabdoviridae* [7]. Some of these, such as *Hantavirus* and *Arenavirus*, can spillover and produce severe human disease, while the rodent may be asymptomatic [8]. New zoonotic viruses, including *Coronaviruses* and *Picornaviruses*, are continuously being

* Corresponding author.
** Corresponding author at: School of Public Health, Sun Yat-Sen University, Guangzhou 510080, China.
E-mail addresses: wangxianhua@xhsysu.edu.cn (X. Wang), lujiahai@mail.sysu.edu.cn (J. Lu).

https://doi.org/10.1016/j.onehlt.2023.100631
Received 11 May 2023; Received in revised form 18 September 2023; Accepted 19 September 2023
Available online 20 September 2023

found [9,10]. Furthermore, there are at least 85 known zoonotic pathogens carried by rodents, which are the primary hosts for not only viral but also non-viral zoonotic diseases (*e.g.*, leptospirosis, salmonellosis, yersiniosis, pathogenic *E. coli*, campylobacterosis, giardiasis, and Lyme disease) [11]. Therefore, rodents, as important reservoirs of zoonotic pathogens, are extremely threatening to humans [12].

China, with its vast biodiversity, harbors around 200 rodent species from 12 families [13,14]. Guangdong Province, located in the south of China, has a tropical and subtropical monsoon climate that is conducive to the growth and reproduction of rodents. It also has the largest population of humans in China, with 126.84 million people in 2021. The extensive human traffic and the climate provide an ideal environment for the transmission of pathogens between rodents and other animals [15]. Historically, Guangdong has experienced a high incidence of outbreaks of various emerging infectious diseases (EIDs), including Severe Acute Respiratory Syndrome (SARS), Avian Influenza, rodent-borne scrub typhus, and *Hantavirus* Pulmonary Syndrome [5,16–18]. Given the vulnerability to zoonotic diseases, public health departments in Guangdong have implemented routine measures for detecting and monitoring known zoonotic pathogens (such as *Hantaviruses, Hepatitis viruses* and *Mammarenaviruses*) in rodents [19–21]. However, it is crucial to estimate the diversity and number of unknown viruses in rodents to monitor and control potential EIDs transmitted by rodents. Recent public health events all indicate that spillover can lead to great disasters globally if this barrier is breached [22,23].

Consequently, an extensive survey of viral communities in rodents, including viral diversity and distribution in different species, is essential to establish an early warning and traceable system for rodent derived EIDs [24]. To this end, the current study captured rodents in a representative area of Guangdong, China, and conducted a viral metagenomic analysis after grouping them according to species and organs. However, simply generating a viral genome by metagenomics does not allow for easy validation of the zoonotic potential of the new virus, which would require recreating the intact virion and then testing it in cell culture or animal models, complicated steps that cannot be done in bulk. It is critical to evaluate the zoonotic potential of this new or unknown virus, because once these viruses are spilled into the human population, pandemics are likely to occur, such as Nipah virus encephalitis and Ebola [25,26]. Hence, a machine learning model was used to prioritize viruses of zoonotic potential only based on their genomic signatures. It was found that >70% of the viruses were judged to have a high or very high likelihood of zoonotic disease [27]. This study provides a new perspective on preventing spillover events by focusing on high-risk viruses through rapid zoonotic risk assessment of viral genome sequences.

2. Materials and methods

2.1. Rodent trapping

Five cities in Guangdong Province, China were selected as sampling sites for this study, including Guangzhou, Shenzhen, Chaozhou, Zhanjiang, and Shaoguan (Table S1). To include rodents in various habitats of the city, sampling was conducted in locations such as hospitals, airports, wharfs, stations, farms, and forest parks (Fig. 1). Rodents were captured in cages using cooked food as bait, and anesthesia was administered. The information regarding sex, body length, tail length and hair colour was recorded for the initial species identification. Sterile cardiac blood sampling was performed using a syringe, and serum was separated after centrifugation at 3000 r/min for 15 min and frozen. The animals were sacrificed by neck clamping under anesthesia to minimize suffering. Subsequently, they were packaged, frozen, and transported to the laboratory of Sun Yat-sen University. Dissection was performed in a biosafety cabinet, and intestinal contents and tissue samples were removed and placed into EP tubes and numbered. The procedure was performed in a sterile environment, and the samples were subsequently placed in cryopreservation tubes and stored in a − 80 refrigerator. Animal species were further identified using PCR targeting the mammalian cytochrome B gene [28]. In sum, a total of 194 rodents from four species were collected for analysis. Serum (Se), brain (B), liver (G), spleen (Sp), lung(L) and intestinal contents(F) were collected from *Rattus norvegicus* (Rn), *Rattus flavipectus* (Rf), *Rattus losea* (Rl) and *Mus musculus* (Mm) for study (Table S2).

2.2. RNA library construction

To assess the viral diversity, the collected rodent specimens were divided into 24 pools according to species and organs. Firstly, the organs were broken up and homogenized, and then the tissue and cellular debris were removed through low-speed centrifugation. The supernatant obtained was transferred to the ultracentrifugation tube and centrifuged at 4 °C with HIMAC CP 100wx ultracentrifuge (Hitachi). After removing the supernatant, resuspended the pellet in 200ul SB buffer, added EMB and EM proportionally, and incubated at 37 °C for 60 min; added SS solution (2ul) proportionally, mix well, and inactivate the enzyme

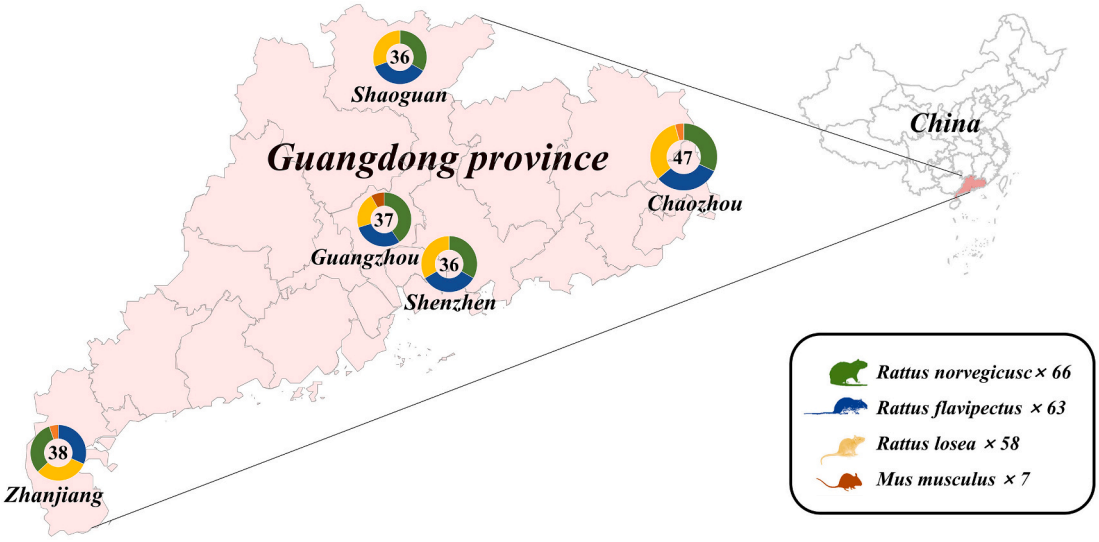

Fig. 1. The sampling sites and the numbers of rodent samples in Guangdong Province, China. The numbers of the 194 samples belonging to the 4 rodent species are indicated by a pie chart for each city.

reaction at 65–75 °C for 10 min, centrifuge at 2000r for 5 min, and stored 200ul of the supernatant at-20 °C for subsequent experiments. Next, the Hipure Universal RNA Mini Kit (Magen) was used to isolate high-purity RNA from the sample collections. After being at room temperature for 3 min, the samples were spun at 4 °C for 15 min at a speed of 12,000 x g to eliminate any precipitates. The products were reverse transcribed to Double-Stranded DNA using REPLI-g Cell WGA & WTA Kit (QIAGEN). Finally, the concentration and length of the amplified products were qualitatively evaluated using Qubit 4 Fluorometer (Thermo Fisher Scientific), and Qsep 1 DNA Analyzer (BiOptic), respectively.

2.3. Next-generation sequencing

Qualified DNA samples from the previous step were fragmented using a sonicator to generate fragments of ≤800 bp. High-quality sequencing libraries were constructed using the NEBNext Ultra™ II DNA Library Prep Kit for Illumina (New England Biolabs) instructions. Briefly, the end of the DNA fragment is repaired and A-tail is added, followed by ligation of the adaptor and PCR amplification product. Qubit 4 Fluorometer (Thermo Fisher Scientific), and Qsep 1 DNA Analyzer (BiOptic) were again used to qualitatively assess the concentration and length of the sequencing libraries. The qualified libraries were sequenced by Novaseq 6000 (Illumina) platform with 150 bp paired-end reads. The sequencing results were stored in FASTQ file format, which contained the sequencing sequences (Reads) and the corresponding quality information.

2.4. Library quality control

Given the presence of low-quality reads in the raw data obtained from sequencing, we use Trimmomatic (v0.39) for quality control, and the specific processing steps are as follows:(I) removing paired reads with adapter; (II) when the number of low-quality bases in a single-end read exceeded 20% of the total number of bases in the read, the paired reads were removed. (III) removing the repetitive reads generated by PCR amplification; (IV) removing the ployX sequence [29]. To eliminate ribosomal and host sequence contamination, the clean reads obtained in the previous step were mapped to the ribosomal and host databases, respectively, using BWA (v0.7.17), and the next data analysis was performed after the mapped sequences were removed.

2.5. Assembly and genome annotation

Megahit (v1.2.9) was used to conduct *De novo* assembly of the high-quality reads into contigs, and sequences that matched the host sequence database were removed. The assembled contigs were clustered using CD-HIT (v4.7). Then, the unique contigs were compared to the Virus-NT (extracted from the NCBI nucleotide database, containing phages) and Virus-NR (extracted from the NCBI Protein database, containing phages) databases using BLASTN and BLASTX, respectively. The resulting contigs were filtered by meeting $e < 1 \times 10^{-5}$ and $e < 1 \times 10^{-3}$ as candidate virus contigs. Additionally, MetaGeneMark (v3.38) and hmmsearch (v3.2.1) were used to predict gene sequences and compare the protein sequences with HMM (VPFs and vFam), respectively. The unique contigs that satisfied the screening requirements ($e \leq 1 \times 10^{-5}$) were also taken into consideration as potential viral sequences [30,31]. After using BLAST (v2.9.0+) to match the potential virus sequences with the NT database, the findings were filtered using a significance threshold of $e \leq 1 \times 10^{-10}$. Diamond (v0.9.10) was used to compare the sequences to the NR database, and the outcomes were selected based on a significance threshold of $e \leq 1 \times 10^{-3}$. Finally, NCBI Taxonomy library was used to annotate the above alignments, and if >20% of the top 50 alignments were nonviral, the sequence was considered nonviral and the remaining sequences were considered viral.

2.6. Abundance and diversity analyses

The reads were aligned to the identified viral contigs, and the RPKM value of each contig was calculated. The sum of the relative abundance of all contigs corresponding to each family was the relative abundance of that family. Principal co-ordinates analysis (PCoA) is a visualization method to study the similarity or difference of data, which can observe the differences between individuals or groups. PCoA uses a dimensionality reduction analysis method, utilizing Bray-Curtis distance, to select the most contributing combination for analysis. In the PCoA results, the closer the two samples are, the higher the similarity between the samples.

2.7. Inferring zoonotic potential

The zoonotic potential of the viruses was ranked using machine learning models that infer the likelihood of infecting humans based on the host range features encoded in the viral genomes, with a cut-off value of 0.293 as reported in a recent study [32]. The features contain a measure of bias in genomic composition of viruses. For instance, the relative frequencies of each codon, amino acid, and dinucleotide. The characteristics were likewise matched to human RNA transcripts for calculating similarity between human and virus sequences. The specific operation procedure involved specifying the viral genomes in FASTA format and manually defining the open reading frames (ORFs). Using the NCBI ORF Finder program (https://www.ncbi.nlm.nih.gov/orffinder/), the full and incomplete viral genome ORFs were inferred. The complete and partial viral genome ORFs were predicted using the NCBI ORF Finder tool (https://www.ncbi.nlm.nih.gov/orffinder/). The ranking of the viruses was generated by inputting the specified sequence format, providing the paths to the genome and metadata files, and naming the output files.

2.8. Phylogenetic analysis

Deduced amino acid sequences of the families *Hantaviridae*, *Reoviridae*, *Picobirnavirdae*, and *Astroviridae* were selected for phylogenetic analysis in this study. The NCBI-nr database was used to obtain all reference sequences. Sequence alignment was performed using MEGA11 and CLUSTW program with default parameters. The Model Selection package was utilized to evaluate the best substitution model. The appropriate substitution model was then used to construct a maximum likelihood phylogenetic tree in MEGA11, with 1000 bootstrap replicates as a measure of tree robustness.

3. Result

3.1. Metagenomic analysis and sequence assembly

In this study, 24 libraries were constructed according to species and organs of rodents. The raw data has been uploaded to the Sequence Read Archive (SRA) library (BioProject accession no. PRJNA939277). We generated 1,311,238,528 raw reads (a total of 393.5 GB of data) and after removing low-quality sequences, a total of 707,438,887 clean reads were retained for further analysis. The 189,796,947 high-quality reads obtained after removal of host and ribosome were assembled with Megahit, resulting in 298, 231 contigs (Table S3). After removing the host, ribosome, and clustering, the 199,756 unique contigs were multiple times aligned with the database before being used for species annotation. Contigs annotated to DNA viruses were deleted, leaving only RNA viruses to be used for further research as we only constructed the sequencing library on RNA (Fig. S1).

3.2. Viral communities analysis

Our study resulted in 2163 viral contigs belonging to 25 families that

infect a wide range of hosts, including vertebrates, invertebrates, amoebas, and plants. Since viruses from *Rhabdoviridae* and *Flaviviridae* families can affect both vertebrate and invertebrate hosts, it is suggested that they are possibly vector-borne viruses. Viral abundance was represented by calculating RPKM, with *Retroviridae*, *Picobirnaviridae* and *Arenaviridae* being the most abundant families (Fig. 2A). *Arenaviridae, Hantaviridae, Paramyxoviridae, Picobirnaviridae, Picornaviridae, Astroviridae, Retroviridae, Hepeviridae, Reoviridae, Flaviviridae, Phenuiviridae* and *Virgaviridae* were detected in all rodent species (Fig. 2C). *Arteriviridae, Astroviridae, Retroviridae, Nyamiviridae, Phenuiviridae* and *Reoviridae* are present in every organ of rodents (Fig. 2B). In contrast, *Hantaviridae, Picobirnaviridae, Arteriviridae* and *Picornaviridae* have been detected in organ samples other than serum and brain. *Flaviviridae* are mainly found in serum samples while *Nyamiviridae, Phenuiviridae* and *Reoviridae* have been detected in organ samples other than serum.

The comparison of viral communities through principal-coordinate analysis (PCoA) suggested that there was no statistically significant difference between four species ($P > 0.05$) (Fig. S2B), and the viral communities may be similar. However, it varies considerably between six organs ($P < 0.05$) (Fig. S2A). Specifically, serum and intestinal contents were distinguished from the other four organs, thus having a different virome composition.

3.3. Zoonotic prediction

Zoonotic risk of viral sequences after high-throughput sequencing was evaluated as previously reported to predict the probability of virus transmission to humans. We excluded viral sequences from Amoeba, Bacteria, and Plants, as well as ORFs shorter than 450 bp to improve the accuracy of our predictions [32]. In the end, we evaluated 92 vertebrates, invertebrates, and unclassified viruses in 14 families (Fig. 3A). Among the 21 viruses considered high risk, *Seoul orthohantavirus* was found to have the highest zoonotic risk as previously reported (Fig. 3B) [33]. *Picobirnaviridae* had eight viral species that were considered as high zoonotic risk. Some viruses, such as *Mammarenavirus, Astroviruses* and *Pestivirus* were both high risk and abundant.

3.4. Phylogenetic relationships of selected viruses

3.4.1. Seoul orthohantavirus

Seoul orthohantavirus, which can cause Hantavirus hemorrhagic fever with renal syndrome belongs to *Hantaviridae* family (under the order *Bunyavirales*) [34]. The main reservoir and source of infection are small rodents prevalent in Eurasia, which shed the virus in their blood, saliva, urine, and feces. *Seoul orthohantavirus* is a negative-strand RNA virus with three different segments including S (small), M (medium), and L (large).

A total of 4 contigs were identified in this study, mainly from lung of *Rattus losea*. They covered 11,435 nucleotides including three genome segments (L, M, and S). The ORF of L showed 99.81% aa identity with the *Seoul orthohantavirus* in *Rattus norvegicus* and *Rattus tanezumi* from southern China in 2021 and S ORFs of this species showed nearly 100% aa identity [35]. Phylogenetic tree based on the complete L protein sequences-encoded RNA-dependent RNA polymerase (RdRp) of *Hantaviruses* including representatives of the main strains. HV-RlL01 was closely related to the isolate YN45, which has the possibility of cross-species spread between *Rattus norvegicus* and *Rattus tanezumi* (Fig. 4A).

3.4.2. Rotaviruses

Rotaviruses that cause severe diarrheal illness in children globally are members of the family *Reoviridae*. Their genome contains segments 1, 2, 3, 4, 6, and 9 coding for 6 structural (VP1-VP4, VP6, VP7) proteins and segments 5, 8, 7, 10, and 11 coding for 6 non-structural (NSP1-NSP6) proteins [36]. *Rotaviruses* are classified into ten different groups (species A-J and possibly K, L) based on the genetic variability of the viral protein 6 (VP6). Recently, there have been increasing reports of cross-species transmission of *Rotaviruses* [37–39].

In total, 49 contigs were identified as *Rotavirus*, mainly from intestinal contents (F) of rodent samples. The ORF of RVA-RlF01 shows 96.32% aa identity with RdRp of *Human rotavirus A* detected in patients with diarrhea in China [40]. The ORFs of RVB-MmF01 and RVB-RfF01 show 91.11% and 91.89% aa identity with RdRp of *Human rotavirus B* detected in human and *Rattus norvegicus* in 1994 [41]. Phylogenetic tree based on the RdRp of *Rotaviruses*. Contigs obtained from the intestinal content of *Rattus losea* belong to the G3 type of human rotavirus A (Fig. 4B). Meanwhile, contigs obtained from the intestinal contents of *Rattus flavipectus* and *Mus musculus* belong to *Rotavirus B*.

3.4.3. Picobirnavirus

Picobirnaviruses, which were initially discovered in humans and rodents are non-enveloped, ds-RNA viruses with two genomic segments [42]. Segment 1 (2.2–2.7 kb) includes ORF1 and ORF2 which encode the capsid protein and hydrophilic protein, respectively, whereas segment 2 (1.2–1.9 kb) contains an ORF encoding RdRp [43]. *Picobirnaviruses* have been found in the feces of various animals, including humans, rabbits, dogs, pigs, rodents, and birds [44]. Their presence has been associated with diarrhea, especially in immunodeficient patients [45].

A total of 1608 contigs were identified as *Picobirnaviruses*, from

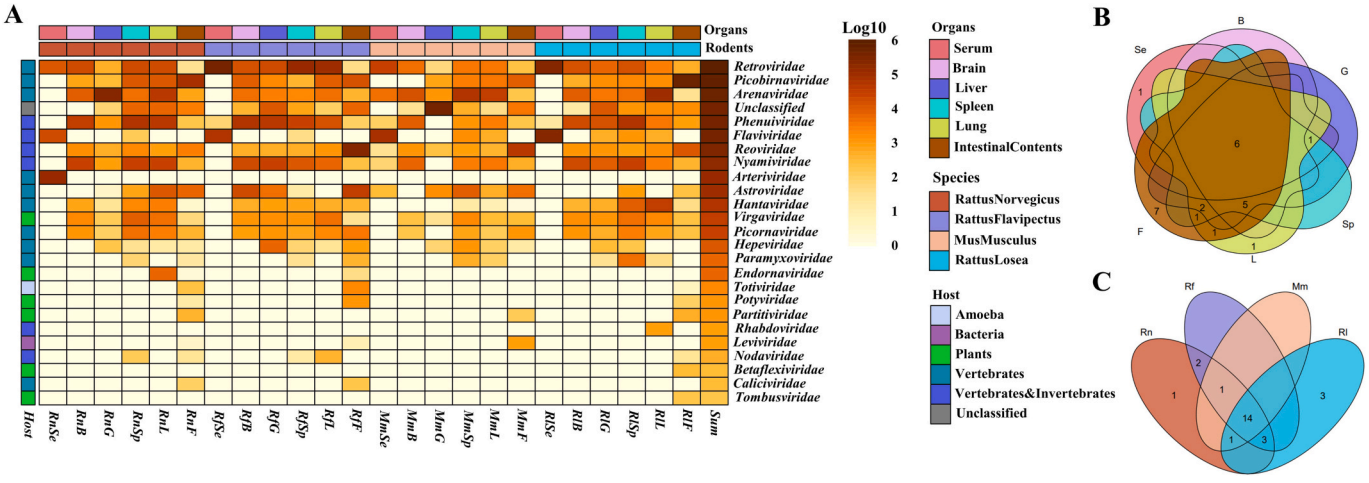

Fig. 2. Viral communities analyses at the family level. (A) Relative abundance (RPKM) of the viral family in each pool are indicated as a heat map ranging from low (0) to high (>5). Names of the viral families are listed on the right as a text column. Organs, rodents, and hosts information is indicated on the left as colour code, with colors defined on the right. (B)(C) Venn plot showing the numbers of viral families found in each species or organs.

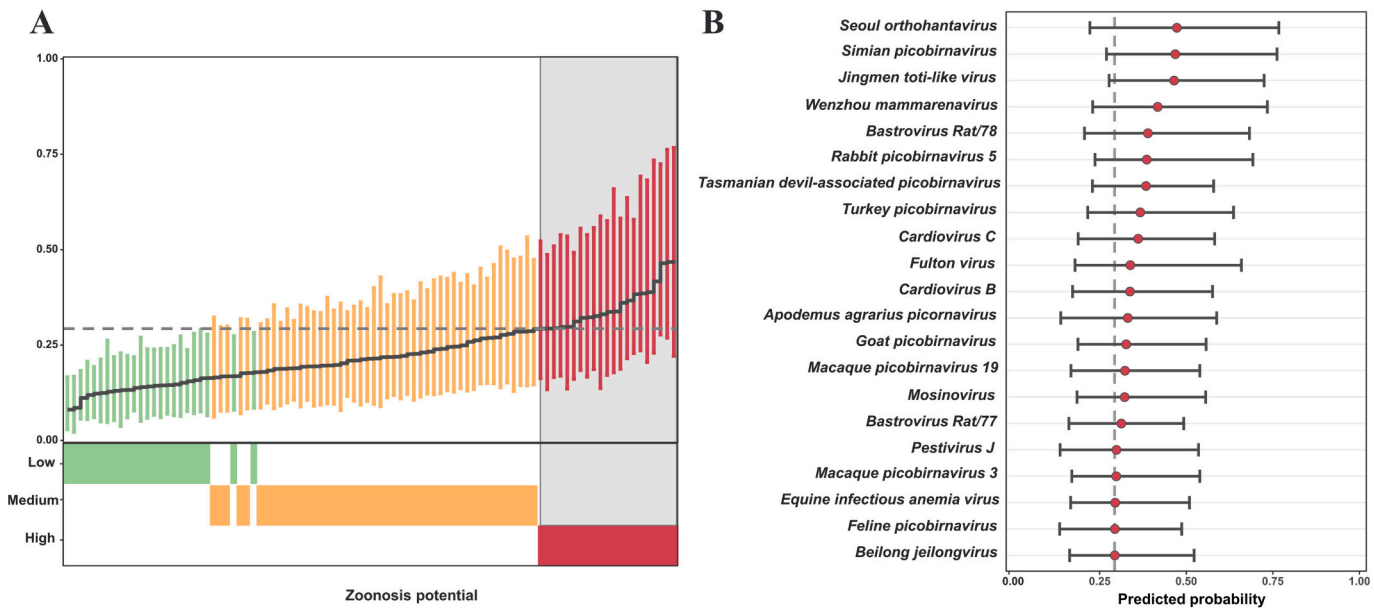

Fig. 3. Zoonotic risk prediction: After excluding the identified Amoeba, Bacteria and Plants viruses, and the ORFs that shorter than 450 bp, 92 viruses were assessed for zoonotic risk (A) Probability of human infection based on viral genomes. Points reveal the mean calibrated score, with lines indicating 95% confidence intervals. The black line indicates a cutoff at 0.293 and Colors show the assigned zoonotic potential categories. Green represents low risk, yellow represents medium risk and red represents high risk. (B) 21 viruses were classified as high risk. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

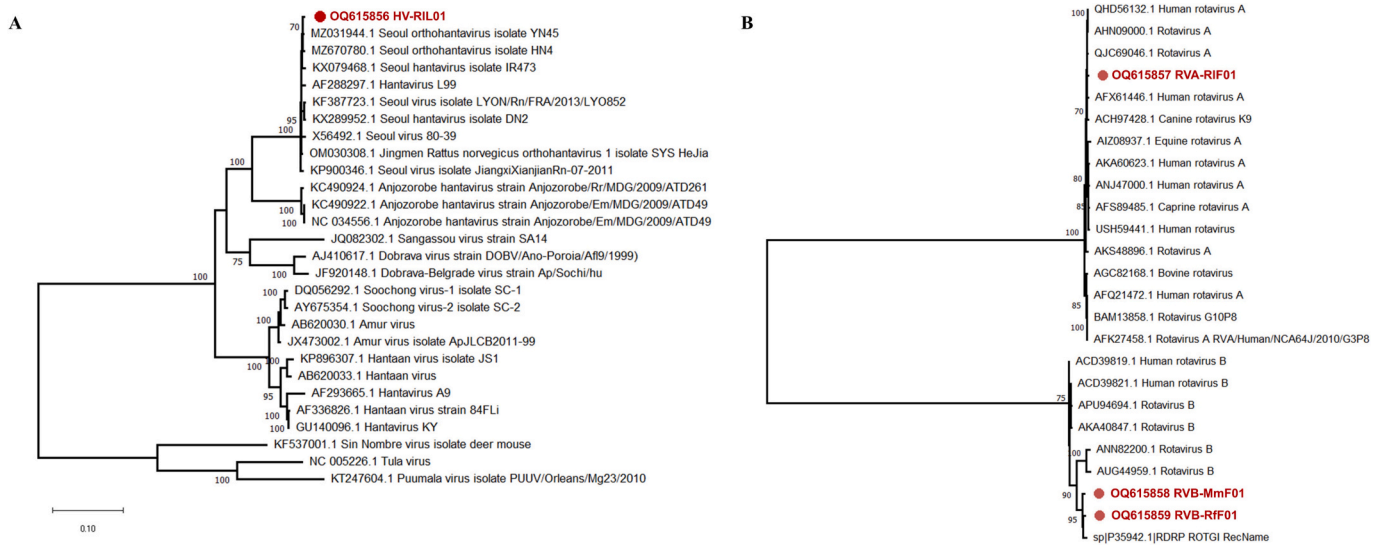

Fig. 4. (A) Phylogenetic tree based on the complete L segment-encoded RdRp amino acid sequences of *Seoul orthohantavirus*. Phylogenetic trees were constructed by the maximum likelihood method with 1000 bootstrap replicates using the best-fit models (LG + G + I for RdRp protein). *Seoul orthohantavirus* found in this study is labeled in red circle. Clade labeling indicates specific genera. Bootstrap support values >0.7 are shown, which are sized according to degree of support (bootstrap support value of 1 has the largest size). (B) Phylogenetic tree based on VP1 (RdRp) protein amino acid sequences of *Rotaviruses.* Phylogenetic trees were constructed by the maximum likelihood method with 1000 bootstrap replicates using the best-fit models (LG + G). *Rotaviruses* found in this study are labeled in red circle. Bootstrap support values >0.7 are shown, which are sized according to degree of support (bootstrap support value of 1 has the largest size). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

organs other than serum in all rodent samples. Here, we identified seven RdRp sequences and the amino acid identities of RdRp with known *Picobirnaviruses* were <74.20%, suggesting that these viruses may be members of potential new species or genera in family *Picobirnavirus* (Fig. 5A). Phylogenetic trees built from amino acid sequences of RdRp were generated to study the evolutionary position of these *Picobirnaviruses*. Besides known genotypes, these RdRp sequences of *Picobirnaviruses* create several new genogroups (Fig. 5B).

3.4.4. *Astroviruses*

Astroviruses, positive single-stranded RNA viruses of *Astroviridae* family, can infect both mammals and birds, resulting in gastroenteritis and other extraintestinal diseases [46,47]. According to the latest suggestion from the International Committee on the Taxonomy of Viruses (ICTV), the taxonomy of *Astroviruses* depends on the amino acid genome distance of the capsid protein (between 0.338 and 0.783). The *Astroviruses* genome comprises ORF1a, ORF1b, and ORF2. ORF1 encodes both protease and RdRp, whereas ORF2 encodes capsid protein.

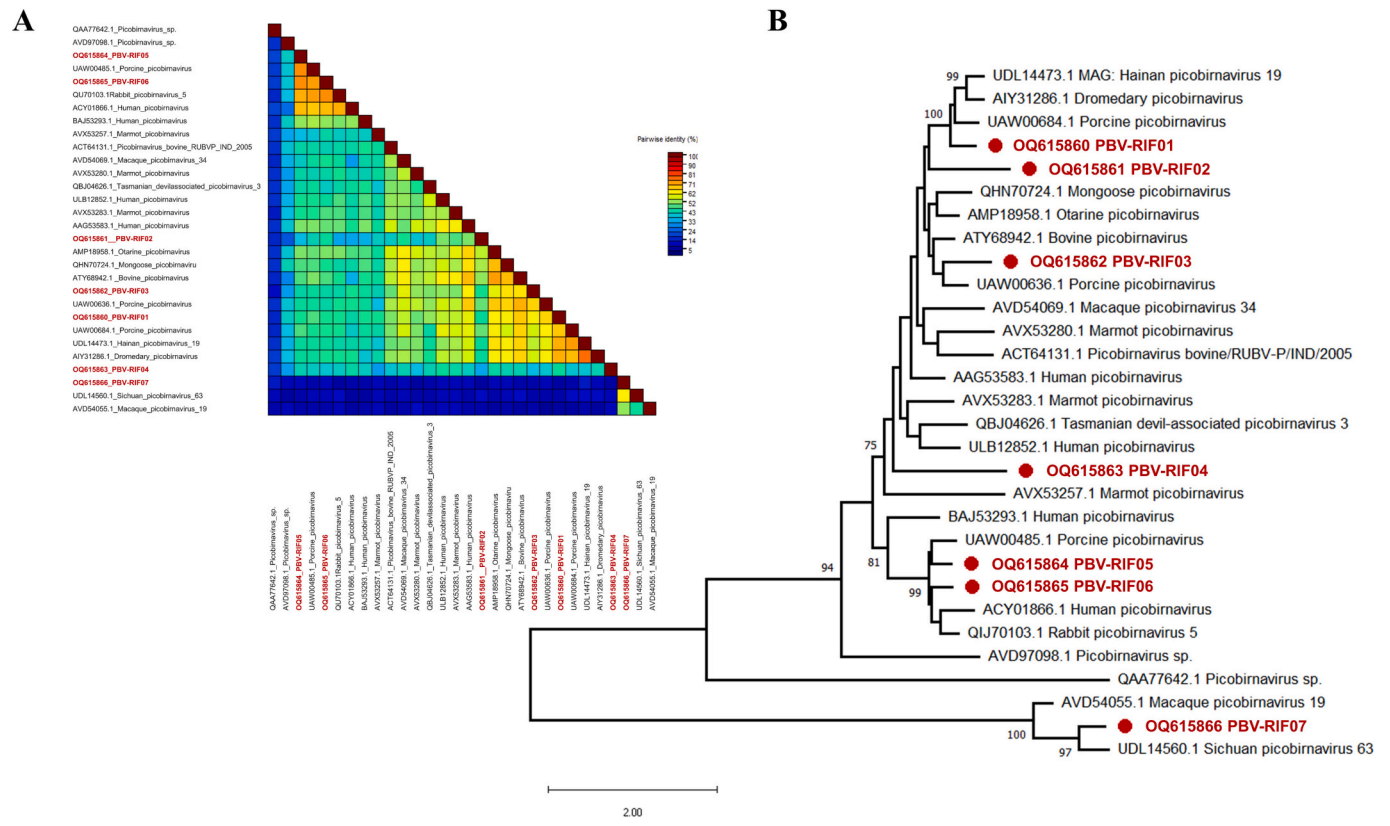

Fig. 5. Phylogenetic relationship of *Picobirnavirus*. (A) Pairwise genetic distance heatmap of *Picobirnavirus* RdRp domain protein sequences. (B) Phylogenetic tree based on RdRp amino acid sequences of *Picobirnavirus* and were constructed by the maximum likelihood method with 1000 bootstrap replicates using the best-fit models (LG + G). *Picobirnaviruses* found in this study are labeled in red circle. Bootstrap support values >0.7 are shown, which are sized according to degree of support (bootstrap support value of 1 has the largest size). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In total, 84 contigs were identified as *Astroviruses* in all rodent samples and five nearly complete astrovirus genomes were assembled. Phylogenetic trees based on capsid amino acid sequences were generated to study the evolutionary position of *Astroviruses* (Fig. 6B). The three ORF2 (AstV-RfF01, AstV-RfF02, AstV-RfF03) of astrovirus genomes for capsid protein showed 94.2–94.6% aa identity with Rat astrovirus Rn/S510/Guangzhou detected in *Rattus norvegicus* from China in 2017 (Fig. 6A). The ORFs of AstV-RfF04 and AstV-RnF01 show 87.7% and 98.0% aa identity with capsid protein of Rodent astrovirus in *Rattus norvegicus* and *Rattus losea*, respectively.

4. Discussion

The pandemic of emerging and re-emerging infectious diseases highlights the necessity of unbiased research on both known and unknown zoonotic viruses, which is a major issue of One Health. This international strategy refers to the prevention of emerging infectious diseases and the protection of human, animal, and environmental health through interdisciplinary, cross-sectoral, and cross-regional collaboration [48,49]. Rodents serve as common hosts for numerous zoonotic viruses, and their habitats overlap constantly with those of humans, posing a significant challenge to public health and safety [50]. They present a serious threat to the prevention and control of infectious illnesses because of their close interaction with wildlife and humans [51]. Additionally, rodent-borne viruses can be indirectly transmitted by vectors involving ticks, fleas, and mites [52]. Further investigation of rodent-borne viruses will not only expand our knowledge of RNA virus diversity in rodents but also reveal a considerable number of unknown rodent viruses.

In order to identify viral genomes and analyze the viral infection status in rodents, the serum, brain, liver, spleen, lung, and intestinal contents representing *Rattus norvegicus, Rattus flavipectus, Rattus losea* and *Mus musculus* were combined into 24 pools according to organs and species. This study resulted in 2163 viral contigs assigned to 25 families that attack a broad range of hosts. Some viral families from vertebrates and invertebrates predominated in the viromes, with a smaller presence from plants, bacteria, and amoeba. Viral abundance varied among rodent species, which may be related to their ecology and behavior [53].

The viral communities do not differ statistically between species while they vary considerably between organs. The statistically no difference between species may be due to the small number of samples [6]. The brain, liver, spleen, and lung had widespread RNA viruses, including *Arenaviridae*, *Phenuiviridae*, and *Nyamiviridae*, while in serum or intestinal contents these viruses were either not found or only found in small amounts. Conversely, RNA viruses that primarily infect their hosts through the enteric canal, such as *Picobirnaviridae* and *Reoviridae*, were seldom or never found in internal organs and serum. This difference in detection may be attributed to the different organ tropisms of these viruses, with body fluid-borne pathogens such as *Arenaviridae*, *Phenuiviridae*, and *Nyamiviridae*, and fecal-oral-transmitted pathogens such as *Picobirnaviridae* and *Reoviridae*. *Flaviviridae* and *Arteriviridae* were only detected in serum samples because they are transmitted through blood. In comparison to detection in fecal samples, the examination of organ samples represents a critical step in minimizing the possibility for mistakes in the taxonomic classification of novel viruses. This approach is expected to provide a robust illustration of viruses harbored by rodents [24].

Given the exponential pace at which new viruses are being discovered with many unknown viruses lacking sufficient pathogenic information, it is imperative to focus finite research and surveillance efforts

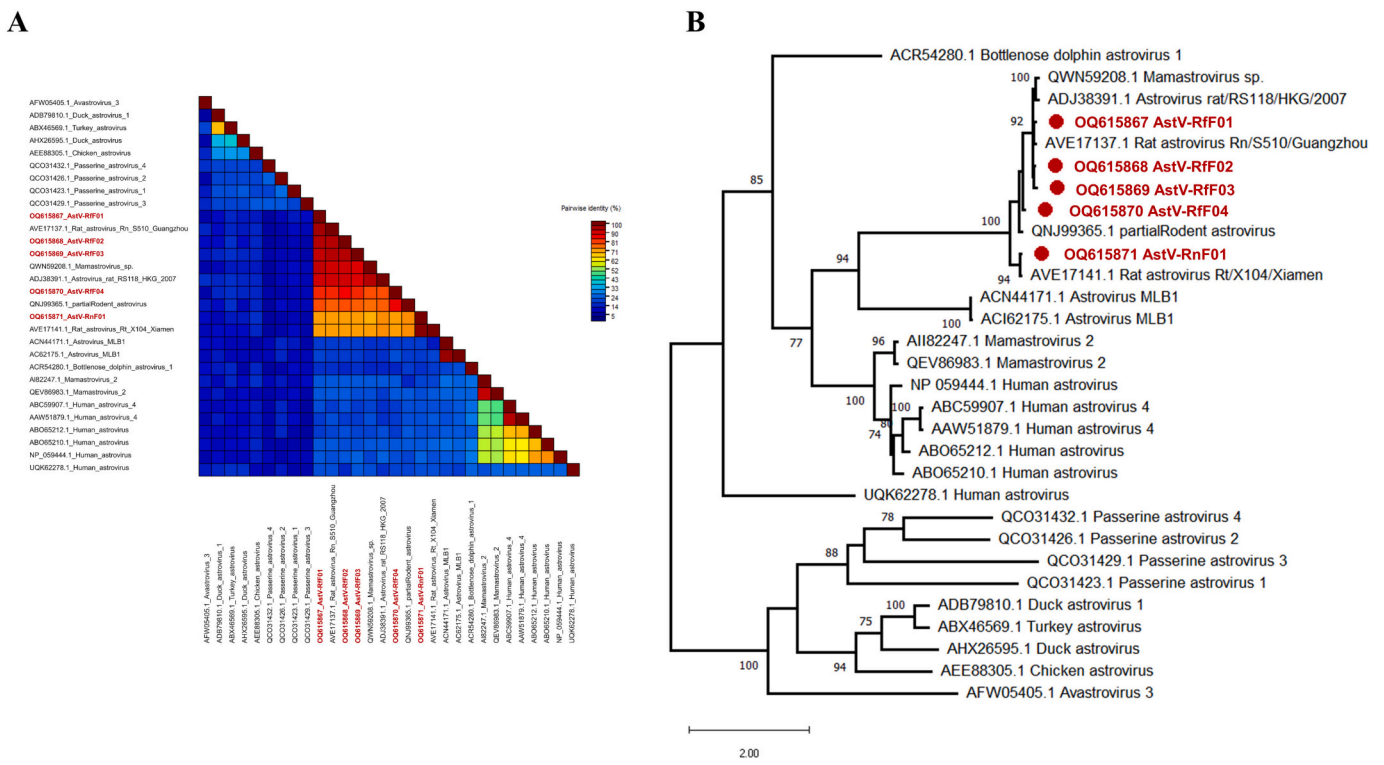

Fig. 6. Phylogenetic relationship of *Astroviridae* (A) Pairwise genetic distance heatmap of *Astroviridae* capsid protein sequences. (B) Phylogenetic tree based on capsid protein amino acid sequences of *Astroviridae* and were constructed by the maximum likelihood method with 1000 bootstrap replicates using the best-fit models (LG + G + F). *Astroviridae* found in this study are labeled in red circle. Bootstrap support values >0.7 are shown, which are sized according to degree of support (bootstrap support value of 1 has the largest size). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

on the viruses most likely to endanger humanity. To identify high-risk viruses, machine learning models prioritize viruses with zoonotic potential based on genomic features, which have been shown to predict zoonotic potential more accurately than models based on relatedness [32]. Despite the potential influence of compositional similarity on predictions, there is no single or specific feature of the viral genome that determines its potential role in zoonoses [27]. Due to the intricate and non-linear nature of the associations between various traits, a genome-based ranking approach offers a dependable and consistent method for zoonotic prediction. After zoonotic prediction, species of *Hantaviridae, Picobirnavirdae, Astroviruses, Flaviviridae, Retroviridae* and *Arenavirus* are divided into high zoonotic risk and abundance, and this prediction is consistent with previous reports. Despite being based solely on genomic signatures, this genome-based ranking approach provides valuable insights for further studies.

Hantaviruses, which are harbored by multiple small mammal species, pose a worldwide public health threat [54]. Based on phylogenetic trees built using the L segments, it was determined that *Rattus losea* carried the *Seoul orthohantavirus,* which is almost identical to the virus sequence found in southern China [35]. *Rattus norvegicus* and *Rattus tanezumi* are major hosts of the *Seoul orthohantavirus* genotype in previous studies from China, indicating that this genotype is prevalent in the *Rattus* and has a spreading tendency.

Rotaviruses, which are primarily transmitted by fecal oral route, act as a significant zoonotic virus [55]. Sequences of *Rotavirus A* detected in *Rattus losea* were most closely related to the *Human rotavirus A* detected in China with diarrhea, suggesting that an overflow event may be occurring. The *Human rotavirus B* discovered in *Rattus norvegicus* and Humans in 1994 has the most genetic similarities with the *Rotavirus B* discovered in *Mus musculus* and *Rattus flavipectus* of this study. Also, the high abundance of *Rotavirus A* and *Rotavirus B* in intestinal contents might suggest *Rotavirus* replication in intestinal tract of the rodents. Our identification of a related argues for the continued circulation of *Rotaviruses* between rodents and human.

Picobirnaviruses can infect hosts and cause diarrhea, and approximately 20% of human fecal diarrhea samples with unclear causes tested positive in the Netherlands [56]. Many *Picobirnaviruses* were identified that matched two distinct viral segments, including segment I and segment II. At the same time, phylogenetic trees based on the RdRp show that there are three complete sequences belonging to a different branch from known viruses, which means they were identified as new viruses. This study further confirms that *Picobirnaviruses* are highly genetically diverse, indicating that rodents are significant wildlife hosts [24,57].

Astroviruses are recently discovered viruses that infect humans and poultry in China [58]. It could lead to gastrointestinal symptoms in humans, and several cases of neurological impairment have been reported [59]. *Astroviruses* were frequently detected in the present study, and the majority of the identified *Astroviruses* sequences can be allocated to three separate clusters along with reference viral sequences from rodents. A new strain of *Astrovirus* has been discovered and has the highest similarity with the *Rodent astrovirus*.

The present study has several limitations: The representativeness of rodent species samples may be insufficient due to limited catches and a non-dominant population of some species in the sampling area. Additionally, the study is constrained to autumn, which hinders our ability to investigate the influence of seasons on the viral composition carried by rodents. Future investigations should be conducted at various spatio-temporal scales to assess the importance of virus surveillance at the human-animal-environment interface. Lastly, the detection and validation of the virus in predicting zoonotic diseases have not been performed yet, and it is essential to employ experimental methods to assess the effectiveness of the prediction model in the future.

5. Conclusions

This study characterizes the RNA viromes of rodents and assesses the

risk of transmission to humans, providing a new insight into the viral diversity in Guangdong (south of China). Because of the global diversity and distribution of rodents, it is essential to increase attention to their role in viral diseases. Further attempts to conduct virus surveillance in rodents will identify more diverse virus lineages with early warning to minimize the risk of future zoonotic illnesses. As a result of these efforts, regional and international capacity for EIDs prevention and response will be strengthened.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.onehlt.2023.100631.

Ethics statement

The animal study was reviewed and approved by Ethical Review Committee for Biomedical Research, School of Public Health, Sun Yat-sen University (Approval number: SYSU-PHME-2022005).

Author contributions

BY, QD, JL, XW, and CG contributed to conception and designed of the study. BY drafted and revised the manuscript. BY and CG analyzed the data and built the figures in both original and revised versions. QD, XL, RZ, and QW performed the experiments. All authors read and approved the final manuscript.

Funding

This work was funded by the National Key Research and Development Program of China (2018YFE0208000), the Key-Area Research and Development Program of Guangdong Province (2018B020241002), the Shenzhen Science and Technology Program (JSGG20220606142207017), the Guangdong Province Drug Administration Science and Technology Innovation Project (2022ZDZ12) and the Central Government Guides Local Science and Technology Development Funds to Freely Explore Basic Research Project (2021Szvup171).

Declaration of Competing Interest

We claim that there is no conflict of interest associated with the paper entitled “Evaluating the zoonotic potential of RNA viromes of rodents provides new insight into rodent-borne zoonotic pathogens in Guangdong, China”.

Data availability

Metagenomic sequencing raw data were deposited in the NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject) with BioProject accession no. PRJNA939277. Virus sequences were deposited in GenBank (https://www.ncbi.nlm.nih.gov/genbank) with accession numbers “OQ615856–OQ615871”.

Acknowledgments

We acknowledge and appreciate our colleagues for their valuable suggestions and technical assistance for this study.

References

[1] K.E. Jones, et al., Global trends in emerging infectious diseases, Nature 451 (7181) (2008) 990–993.
[2] W.B. Karesh, et al., Ecology of zoonoses: natural and unnatural histories, Lancet 380 (9857) (2012) 1936–1945.
[3] M.E. Woolhouse, S. Gowtage-Sequeria, Host range and emerging and reemerging pathogens, Emerg. Infect. Dis. 11 (12) (2005) 1842–1847.
[4] D.E. Wilson, D.M. Reeder, Mammal Species of the World: A Taxonomic and Geographic Reference vol. 1, JHU Press, 2005.
[5] B.G. Meerburg, G.R. Singleton, A. Kijlstra, Rodent-borne diseases and their risks for public health, Crit. Rev. Microbiol. 35 (3) (2009) 221–270.
[6] Z. Wu, et al., Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases, Microbiome 6 (1) (2018) 178.
[7] B.A. Han, et al., Rodent reservoirs of future zoonotic diseases, Proc. Natl. Acad. Sci. 112 (22) (2015) 7039–7044.
[8] H.C. Yin, D.C. Wan, H.Y. Chen, Metagenomic analysis of viral diversity and a novel astroviruse of forest rodent, Virol. J. 19 (1) (2022) 138.
[9] Z. Wu, 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 9 (1) (2021) 18.
[10] S.K. Lau, et al., Discovery of a novel coronavirus, China Rattus coronavirus HKU24, from Norway rats supports the murine origin of Betacoronavirus 1 and has implications for the ancestor of Betacoronavirus lineage a, J. Virol. 89 (6) (2015) 3076–3092.
[11] B.A. Han, A.M. Kramer, J.M. Drake, Global patterns of zoonotic disease in mammals, Trends Parasitol. 32 (7) (2016) 565–577.
[12] A.D. Luis, et al., A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proc. Biol. Sci. 280 (1756) (2013), 20122753.
[13] Y.X. Wang, A Complete Checklist of Mammal Species and Subspecies in China: A Taxonomic and Geographic Reference, China Forestry Publishing House, Beijing, 2003, pp. 41–46.
[14] E.E. Puckett, D. Orton, J. Munshi-South, Commensal rats and humans: integrating rodent Phylogeography and Zooarchaeology to highlight connections between human societies, Bioessays 42 (5) (2020), e1900160.
[15] T. Allen, et al., Global hotspots and correlates of emerging zoonotic diseases, Nat. Commun. 8 (1) (2017) 1124.
[16] T.G. Ksiazek, et al., A novel coronavirus associated with severe acute respiratory syndrome, N. Engl. J. Med. 348 (20) (2003) 1953–1966.
[17] Y. Yang, et al., Clinical indices and mortality of hospitalized avian influenza a (H7N9) patients in Guangdong, China, Chin. Med. J. 132 (3) (2019) 302–310.
[18] Z. Pan, et al., Analysis on surveillance of plague host animals in Guangzhou city from 2008 to 2009, J. Trop. Med. (Guangzhou) 10 (11) (2010) 1336–1360.
[19] B. Wang, et al., Detection and characterization of three zoonotic viruses in wild rodents and shrews from Shenzhen city, China, Virol. Sin. 32 (4) (2017) 290–297.
[20] Q. Su, et al., Epidemiology and genetic diversity of zoonotic pathogens in urban rats (Rattus spp.) from a subtropical city, Guangzhou, southern China, Zoonoses Public Health 67 (5) (2020) 534–545.
[21] J. Wu, et al., Genomic characterization of Wenzhou mammarenavirus detected in wild rodents in Guangzhou City, China, One Health 13 (2021), 100273.
[22] X.A. Zhang, et al., A zoonotic Henipavirus in febrile patients in China, N. Engl. J. Med. 387 (5) (2022) 470–472.
[23] R. Lu, et al., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet 395 (10224) (2020) 565–574.
[24] H. Du, et al., Metagenome-assembled viral genomes analysis reveals diversity and infectivity of the RNA virome of Gerbillinae species, Viruses 14 (2) (2022).
[25] K.B. Chua, et al., Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia, Lancet 354 (9186) (1999) 1257–1259.
[26] D. Malvy, et al., Ebola virus disease, Lancet 393 (10174) (2019) 936–948.
[27] J.T. Ladner, Genomic signatures for predicting the zoonotic potential of novel viruses, PLoS Biol. 19 (9) (2021), e3001403.
[28] T.D. Kocher, et al., Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers, Proc. Natl. Acad. Sci. U. S. A. 86 (16) (1989) 6196–6200.
[29] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (15) (2014) 2114–2120.
[30] D. Paez-Espino, et al., Uncovering Earth’s virome, Nature 536 (7617) (2016) 425–430.
[31] P. Skewes-Cox, et al., Profile hidden Markov models for the detection of viruses within metagenomic sequence data, PLoS One 9 (8) (2014), e105067.
[32] N. Mollentze, S.A. Babayan, D.G. Streicker, Identifying and prioritizing potential human-infecting viruses from their genome sequences, PLoS Biol. 19 (9) (2021), e3001390.
[33] E. Krautkrämer, M. Zeier, A. Plyusnin, Hantavirus infection: an emerging infectious disease causing acute renal failure, Kidney Int. 83 (1) (2013) 23–27.
[34] M. Goeijenbier, et al., Seoul hantavirus in brown rats in the Netherlands: implications for physicians–epidemiology, clinical aspects, treatment and diagnostics, Neth. J. Med. 73 (4) (2015) 155–160.
[35] W. He, et al., Detection and genetic characterization of Seoul virus in liver tissue samples from Rattus norvegicus and Rattus tanezumi in urban areas of southern China, Front. Vet. Sci. 8 (2021), 748232.
[36] L. Hu, et al., Rotavirus non-structural proteins: structure and function, Curr. Opin. Virol. 2 (4) (2012) 380–388.
[37] R. Johne, et al., Distantly related rotaviruses in common shrews, Germany, 2004-2014, Emerg. Infect. Dis. 25 (12) (2019) 2310–2314.
[38] V. Martella, et al., Zoonotic aspects of rotaviruses, Vet. Microbiol. 140 (3–4) (2010) 246–255.
[39] K. Bányai, et al., Candidate new rotavirus species in Schreiber's bats, Serbia, Infect. Genet. Evol. 48 (2017) 19–26.
[40] Y.H. Wang, et al., Complex evolutionary patterns of two rare human G3P[9] rotavirus strains possessing a feline/canine-like H6 genotype on an AU-1-like genotype constellation, Infect. Genet. Evol. 16 (2013) 103–112.

[41] J.J. Eiden, C. Hirshon, Sequence analysis of group B rotavirus gene 1 and definition of a rotavirus-specific sequence motif within the RNA polymerase gene, Virology 192 (1) (1993) 154–160.
[42] H.G. Pereira, et al., A virus with a bisegmented double-stranded RNA genome in rat (Oryzomys nigripes) intestines, J. Gen. Virol. 69 (Pt 11) (1988) 2749–2754.
[43] X.L. Luo, et al., Marmota himalayana in the Qinghai-Tibetan plateau as a special host for bi-segmented and unsegmented picobirnaviruses, Emerg. Microbes Infect. 7 (1) (2018) 20.
[44] A.Y. Kashnikov, N.V. Epifanova, N.A. Novikova, Picobirnaviruses: prevalence, genetic diversity, detection methods, Vavilovskii Zhurnal Genet Selektsii 24 (6) (2020) 661–672.
[45] I.R. Haga, et al., Identification of a bisegmented double-stranded RNA virus (Picobirnavirus) in faeces of giant anteaters (Myrmecophaga tridactyla), Vet. J. 158 (3) (1999) 234–236.
[46] V. Cortez, et al., Astrovirus biology and pathogenesis, Annu. Rev. Virol. 4 (1) (2017) 327–348.
[47] S.N. Roach, R.A. Langlois, Intra- and cross-species transmission of astroviruses, Viruses 13 (6) (2021).
[48] G.F. Albery, et al., The science of the host-virus network, Nat. Microbiol. 6 (12) (2021) 1483–1492.
[49] P. Overgaauw, et al., A one health perspective on the human-companion animal relationship with emphasis on zoonotic aspects, Int. J. Environ. Res. Public Health 17 (11) (2020).
[50] K.J. Olival, et al., Host and viral traits predict zoonotic spillover from mammals, Nature 546 (7660) (2017) 646–650.
[51] F. Bordes, K. Blasdell, S. Morand, Transmission ecology of rodent-borne diseases: new frontiers, Integr. Zool. 10 (5) (2015) 424–435.
[52] J.O. Lloyd-Smith, et al., Epidemic dynamics at the human-animal interface, Science 326 (5958) (2009) 1362–1367.
[53] J. Lu, et al., Gut Virome of the World's highest-elevation lizard species (Phrynocephalus erythrurus and Phrynocephalus theobaldi) reveals versatile commensal viruses, Microbiol. Spectr. 10 (1) (2022), e0187221.
[54] X.D. Lin, et al., Migration of Norway rats resulted in the worldwide distribution of Seoul hantavirus today, J. Virol. 86 (2) (2012) 972–981.
[55] C.K. Yinda, et al., Cameroonian fruit bats harbor divergent viruses, including rotavirus H, bastroviruses, and picobirnaviruses using an alternative genetic code, Virus Evol. 4 (1) (2018) vey008.
[56] M. van Leeuwen, et al., Human picobirnaviruses identified by molecular screening of diarrhea samples, J. Clin. Microbiol. 48 (5) (2010) 1787–1794.
[57] X. He, et al., Metagenomic analysis of viromes in tissues of wild Qinghai vole from the eastern Tibetan plateau, Sci. Rep. 12 (1) (2022) 17239.
[58] K. To, et al., High prevalence of four novel astrovirus genotype species identified from rodents in China, J. Gen. Virol. 98 (5) (2017) 1004–1015.
[59] C. Donato, D. Vijaykrishna, The broad host range and genetic diversity of mammalian and avian astroviruses, Viruses 9 (5) (2017).