Diversity and independent evolutionary profiling of rodent-borne pathogens in tropical island, China

Youyou Li[1], Chuanning Tang[1], Yun Zhang[1], Zihan Li[1], Gaoyu Wang[1], Ruoyan Peng[1], Yi Huang[1], Xiaoyuan Hu[1], Henan Xin[2], Boxuan Feng[2], Xuefang Cao[2], Yongpeng He[2], Tonglei Guo[2], Yijun He[2], Haoxiang Su[2], Xiuji Cui[1], Lina Niu[1], Zhiqiang Wu[2], Jian Yang[2], Fan Yang[3], Gang Lu[1], Lei Gao[2], Qi Jin[2], Meifang Xiao[4], Feifei Yin[1], and Jiang Du[5]

[1]Hainan Medical University
[2]Chinese Academy of Medical Sciences & Peking Union Medical College
[3]Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Pathogen Biology
[4]Hainan Women and Children's Medical Center
[5]Affiliation not available

December 2, 2022

Abstract

The risk of emerging infectious diseases (EID) is increasing globally. More than 60% of EIDs worldwide are caused by animal-borne pathogens, and most viral pathogens are rodent-borne. This study aimed to characterise the virome and analyse the phylogenetic evolution and diversity of rodent-borne viruses in Hainan Province, China. We collected 588 anal and throat samples from rodents, combined them into 28 pools according to their species and location, and processed them for next-generation sequencing and bioinformatics analysis. The diverse viral reads closely related to mammals were assigned to 15 viral families. Molecular clues of the important rodent-borne viruses were further identified by polymerase chain reaction for phylogenetic analysis and annotation of genetic characteristics such as coronavirus, arenavirus, picornavirus. We identified a pestivirus in Leopoldoms edwardsi and two bocaviruses in Rattus andamanensis and Leopoldoms edwardsi from the national nature reserves of Jianfengling and Bangxi with low amino acid identity to known pathogens are proposed as the novel species, and their rodent hosts have not been previously reported to carry these viruses. These results expand our knowledge of viral classification and host range and suggest that there are highly diverse, undiscovered viruses that have evolved independently in their unique wildlife hosts in inaccessible areas, which may cause zoonosis if they cross their host barrier. Our virome and phylogenetic analyses of rodent-borne viruses provide basic data for the prevention and control of human infectious diseases caused by rodent-borne viruses in the subtropical area of China.

Introduction

Owing to economic globalisation and environmental changes, the risk of emerging infectious diseases (EID) is increasing globally, posing threats to human health, the economy, and society. Approximately 60.3% of human EIDs are caused by zoonotic pathogens, originating from rodents, bats, birds, and other wildlife [1-3], with viral or prion pathogens accounting for 25.4% of all EID events [1].Coronavirus Disease 2019 (COVID-19) and Ebola hemorrhagic fever (EBHF) bleong to EID events. Rodential is the largest mammalian group on Earth, with 35 families, 389 genera, and approximately 2700 species (˜43% of all mammal species) worldwide [4]. Rodents are widely distributed, diverse in species, have a strong reproductive ability, migrate in groups,

and are mostly concentrated in densely populated and humid and warm places such as Central and Southern China, providing them with many opportunities to interact with humans.

Rodents carry Yersinia pestis, Hantaviruses (HanVs), and Leptospira, which cause plague, haemorrhagic fever, and leptospirosis, respectively [5]. Moreover, several important zoonotic viruses carried by rodents are members of the *Hantaviridae* ,*Flaviviridae* ,*Arenaviridae, Reoviridae* ,*Piconaviridae* , *and Togaviridae* families [4-9]. Many of these family viruses include multiple agents associated with acute gastroenteritis or febrile illnesses that can cause severe diseases in humans. Lassa virus (LASV), lymphocytic choriomeningitis virus (LCMV), Tacaribe, Machupo, Junin, Guanarinto, Sabia, and the recently discovered White-water Arroyo and Lujo viruses are high-impact pathogens belonging to the*Arenaviridae* family [10-13]. HanVs infect a wide range of mammalian hosts; Hanta, Seoul, Puumala, and Sin Nombre viruses are etiological agents of haemorrhagic fever with renal syndrome and hantavirus cardiopulmonary syndrome in humans [7]. Virus-host co-divergence and cross-species transmission have played an important role in the evolution and taxonomy of HanVs [7]. Additionally, rat hepatitis E is the causative agent of hepatitis in humans [14] and tick-borne encephalitis virus [15] causes human meningitis, encephalitis, and meningoencephalitis in the northern European subcontinent. Severe fever with thrombocytopenia syndrome [16], Crimean Congo haemorrhagic fever, and Omsk haemorrhagic fever viruses are transmitted by rodents to humans as vectors [5].

Next-generation sequencing (NGS), viral metagenomic analysis is rapidly evolving in the study of natural environments [17], humans, and animals [18]. Recently, metagenomic and meta-transcriptomic analyses have been increasingly used to investigate the composition of r*odent* virome. For example, the Aichi virus and novel circovirus have been detected in wild rodent faecal samples in the analysis of mouse enterovirus diversity [4]. The detection of multiple infectious viruses in rodent lungs in mainland Southeast Asia [18]. Hainan Island is a tropical island province that has been geographically isolated from mainland China for a long time; therefore, its ecological environment and microorganisms are independent to some degree. Owing to the unique climate and geographical environment of Hainan Island, it is rich in diverse mammalian species, especially rodents. The Hainan Province is an important free trade port in China; therefore, the spread of viral pathogens from this province is expected to increase through economic globalisation. In our previous studies, the Wenzhou virus (WENV) belonging to*Arenaviridae* from *Rattus norvegicus* was detected in Daoke Village, Hainan, and a complete genome sequence was obtained [19]. The present study aimed to . We collected 588 throat and anal swabs from 326 rodents belonging to 6 species and grouped the swabs into 28 pools according to the rodent species, swab types, and sample locations. We sequenced the virome, revealed that coronaviruses, flaviviruses, parvoviruses, astroviruses, and papillomaviruses are carried by rodents, and revealed the viral abundance of each pool. The annotation and phylogenetic description of rodent-borne viruses will help us to explore their origin, spread mode, and evolutionary patterns.

Materials and methods

Ethics statement

This study was conducted in sampling procedures were approved by the Ethics Committee of Hainan Medical University (approval number: HMUEC20180059).

Swabs samples of rodents

We collected 588 throat and anal swab samples from 326 rodents from multiple locations in 9 counties or cities in the Hainan Province, the tropical island province in China, from May 2017 to June 2021. Samples (15-20) were grouped into pools according to their sampling locations and species. Morphology was combined with mitochondrial cytochrome*b* (mt-*cyt b*) to identify murine species and analyse the congruence between viruses and their hosts (Table S1). The collected samples were quickly immersed in the maintenance medium

in the virus-sampling tube (Yocon Biology, Beijing, China) to ensure the sample quality, and transported to the laboratory within 24 h using the low-temperature cold chain [6]. The samples were divided into three parts evenly upon arrival at the laboratory and were stored at -80 °C storage before subsequent experiments, which is consistent with a previous study [6].

Viral nucleic acid library construction and NGS

The 588 samples were combined into 28 pools based on swab type and sample location. Swab samples were passed through 0.45μm filters (MilliporeSigma, Burlington, MA, USA) to remove eukaryotic and bacterium-sized particles. The filtrate was ultracentrifuged at 100,000 $\times g$ and 4 for 3 h. The precipitate collected from 28 pool samples was resuspended in Hank's balanced salt solution and digested with DNase (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) to decompose and remove unprotected nucleic acids. Viral RNA was extracted using a QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany). Viral cDNA was generated using primer K-8N and Superscript III Reverse Transcriptase (Invitrogen, Thermo Fisher Scientific), as previously described [20]. Sequence-independent polymerase chain reaction (PCR) amplification products were purified and subjected to magnetic bead sorting. Viral nucleic acid libraries constructed by pre-processing, such as sonication, were analysed using an Illumina HiSeq2500 sequencer (Illumina Inc., San Diego, CA, USA) for a single read of 150 bp in length. Sequence data were deposited in the National Center for Biotechnology Information (NCBI) sequence reads archive under the accession number PRJNA892773.

Metagenomic sequencing

Quality-controlled reads of each sample were assembled using Trinity V2.5.1 [21]. DIAMOND was used to compare the contigs against the non-redundant protein database from NCBI. All blastx results of contigs were annotated with taxonomy id by an in-house database, which combined accession files (*ftp://ftp.ncbi.nih.gov/pub/taxonomy/accession2taxid/*), and other information obtained from the NCBI Entrez server. Non-redundant viral contigs larger than 500 nucleotides (nt) were selected. Contigs related to bacteriophages, plant viruses, and insect viruses were excluded. Read mapping was performed using BWA-MEM [22]. Low-quality mappings were removed using Converm (v0.6.1, *https://github.com/wwood/CoverM*) with two preset parameters: min-read-percent-identity 0.90 –min-read-aligned-percent 0.75. The relative viral abundance was also generated from Coverm using the transcripts per million (tpm) method. A heatmap for visualising the mouse-related virome profile was plotted using the pheatmap package in R.

Viral genome sequencing

Molecular clues from metagenomic analyses were used to classify sequence reads into viral families or genera using MEGAN. To identify the partial or complete genome, representative viral open reading frames (ORF)-related reads were selected for read-based PCR and sequencing. Reads with accurate genomic locations were used to design specific nested PCR primers and to identify partial genomes, which is consistent with a previous study [6]. cDNA was generated using random primers and Superscript III Reverse Transcriptase (Invitrogen). The remaining genomic sequences were analysed using genome walking and 5'- and 3'- rapid amplification of cDNA ends (Invitrogen; and Takara Bio, Kyoto Japan). The primers used to amplify the sequence obtained in this study are shown in (Table S2).

Genome annotation

The viral basic nucleotide sequences structure of the genomes and the amino acid encoded by effective ORFs and its location were deduced by comparing with the sequence of other virus families. Prediction of conservative protein families and domains using Pfam (http://www.ebi.ac.uk/services/proteins), BLASTP (*https://blast.ncbi.nlm.nih.gov*), and InterProScan 5 (*http://www.ebi.ac.uk/services/proteins*). Routine

sequence alignment was performed using Clustal Omega (*http://www.ebi.ac.uk/Tools/*). The evolution tree beautification was performed using ITOL. (*https://itol.embl.de/*).

Viral prevalence

Nearly complete or partial genomic sequences of the viruses obtained by NGS sequencing were used as templates to design specific, semi-nested primers for the non-structural gene for PCR and screening for viruses in each filtrate sediment and individual sample (Table S2). PCR was performed using Go Taq Colorless Master Mix (Promega). Nest PCR using two microliters of the first-round PCR product as the template of the second round of PCR has high specificity and sensitivity. The thermal cycling conditions for both PCRs were 94 degC for 5 min, followed by 35 cycles at 94 degC for 30 s, 57 degC for 35 s, 72 degC for 30 s, and a final elongation step at 72 degC for 10 min. The PCR products were analysed using 1.5% agarose gel electrophoresis and ultraviolet imaging.

Phylogenetic and data analyses

We used the ClusterW package to align the nucleotide sequences and deduce the amino acid sequences (https://rdrr.io/bioc/muscle/man/muscle-package.html) and default parameters in MEGAX. Relatively conserved viruses were selected to construct a phylogenetic tree using the maximum likelihood method. According to the operation rules, the best substitution model was evaluated using the model selection package function of MEGAX with 1000 bootstrap replicates. The NCBI basic local alignment search tool was used to perform pairwise amino acid alignment between the reference sequences and the new virus.

Results

Sample collection from rodents

In this study, between May 2017 and June 2021, we collected 588 throat and anal swabs of 326 rodents belonging to the order *Rodentia*from Haikou City (CDC, Fucheng, Yanfeng, and Longtang), Baisha County (Yinggeling, Huangjingjiaoling, and Bangxi) Dongfang City, Lingao County, Tunchang County, Lingshui County, Ledong County, Chengmai County, and Sanya City in the Hainan Province, which is the only tropical island province in China (Figure 1 and Table S1). Rodent species were morphologically identified and confirmed by mt-*Cyt b*gene sequences, including *Rattus norvegicus (* 23.53%),*Rattus tanezumi* (38.24%), *Rattus andamanensis*(6.18%)*, Rattus losea* (6.65%), *Niviventer fulvescens* (19.41%), *and Leopoldamys edwardsi* (5.00%). The species from *Rattus* under the *Muridae* are clustered on the evolutionary tree and show a close distance to the *Cricetulus* of the *Circetidae* and a far distance to the *dipus* and*rhizomys* . The 588 samples were combined into 28 pools according to rodent species, swab types, and sample locations.

NGS sequencing

According to species, a total of 136G GB of nucleotide data (7,553,938 valid reads,150 bp long) were obtained from 28 pools. Archaea, bacteria, microbial eukaryotes, and invalid reads with no significant similarity to any amino acid sequences in the viral NR protein database were excluded. A comparison of the total number of contigs and the proportion of viral reads from 6 rodent species and 15 sampling locations is shown in Figure 2a-d. The viral composition of each pooled sample is shown in Figure 2e. The readings were filtered into 57,980 contigs, and the virus-associated contigs that best matched the viral ORF proteins available in each pool were screened.

Metagenomic analysis and virome overview

An overview of collection region, rodent species, and virus-related contigs is shown in Table S3. Mammalian virus-associated contigs were assigned to 15 families (Table S4) and were composed of double-stranded (ds) DNA, dsRNA, single-stranded (ss) DNA, ssRNA viruses and retro-transcribing. Viral reads were assigned to*Flaviviridae* ,*Coronaviridae* , *Arenaviridae* , *Astroviridae* ,*Adenoviridae* , *Parvoviridae* , *Papillomaviridae* ,*Herpesviridae* ,*Picornaviridae* , *Hepeviridae,* and some arthropod viruses, including *Phenoviridae* and *Panviridae* , as well as a large group of unclassified RNA viruses. The heatmap shown in Figure 3 indicates that the viromes hosted by rodents might be related to their geographic distribution. Flaviviruses and picornaviruses are widely distributed in Chengmai, Lingao, and Baisha, with high abundance. Coronaviruses, herpesviruses, and astroviruses are widely distributed throughout the world. Pestivirus and Bocavirus are present only in the mountainous forest area of Baisha. The origins, host, structure information and accession numbers of the viruses identified in this study are shown in Table S5.

Characteristics of negative-stranded RNA viruses

Arenaviridae

Arenaviruses (AreVs; order *Bunyavirales* , family*Arenaviridae*) are a group of enveloped RNA viruses whose genome contains at least two single negative-sense RNA segments totalling approximately 10.5kb. AreVs are currently classified into four genera:*Antennavirus, Hartmanivirus, Mammarenavirus, and Retarenavirus.Mammarenavirus* cause central nervous system disease and haemorrhagic fever in humans in Africa and Latin America [23-24]. Human pathogenic arenaviruses include the Tacaribe Complex (New-World) and LCMV-LASV Complex (Old-World), which include the LASV and LCMV.

To identify partial genomes, representative reads of AreVs were obtained from *Rattus tanezumi* collected from Lingshui (Yelin). Primers were designed for PCR amplification and using Sanger sequencing to cover the gaps. We obtained a WENV partial large segment (L) termed AreV-HMU-2 (Table S5). The preliminary experiment in our study obtained a novel WENV complete sequence from*Rattus norvegicus* collected from Haikou (Daoke) in 2015, termed the HMU virus. AreV-HMU-2 were detected in 9% (2/21) of the throat swabs whereas the HMU virus was detected in 4% (1/22) of the anal swabs. AreV-HMU-2 and HMU viruses showed 95.45% identity with each other and 94.8% and 93.92% identity with the WENV 3, respectively. Both viruses showed [?] 79% identity with other mammalian pathogenic arenaviruses (Table S6). The genome of the HMU virus comprises 7,147 bp and shares high sequence similarity with the WENV isolate Rn-242 [19]. Additionally, AreV-HMU-2 and HMU virus phylogenetic analyses based on partial L showed that both viruses were closely related to WENV and LORV and formed an independent clade within the genus Mammarenavirus (Figure 4).

Characteristics of positive-stranded RNA viruses

Coron*aviridae*

Coronaviruses (CoVs; order*Nidovirales* , suborder *Cornidovirineae* , family*Coronaviridae*) are a group of enveloped RNA viruses with unusually large positive ssRNA genomes of, 26-32kb [25,26]. The 2020 report of the International Committee on the Taxonomy of Viruses (ICTV11https://talk.ictvonline.org/) states that CoVs can be divided into four recognised genera: *Alphacoronavirus* ,*Betacoronavirus* , *Gammacoronavirus* , *and Deltacoronavirus* . CoVs are etiological agents in animals and humans that result in respiratory, hepatic, and enteric diseases such as cold, severe acute respiratory syndrome (SARS) and (COVID-19) [27-29].

To identify partial genomes, representative reads were obtained for CoVs in *Rattus tanezumi* collected from Haikou (Fucheng) and Lingao and Haikou (Yanfeng), and*Niviventer fulvescens*collected from Baisha (Yinggeling). We obtained a novel CoV strain complete genome sequence, named CoV-HMU-1, and three

CoV RNA-dependent RNA polymerase (RdRp) sequences and spike S segments, termed CoV-HMU-2, CoV-HMU-3, and CoV-HMU-4, respectively. CoV-HMU-1 and CoV-HMU-2 were detected in 40% (10/25) and 3% (1/32) of anal swabs, respectively. CoV-HMU-4 and CoV-HMU-3 were both detected in 10% (2/20) of the throat swabs (Table S5). Sequence similarity and phylogenetic analyses of RdRp and S revealed that the identified CoVs could be classified as*Embecovirus* under the genus *Betacoronavirus* , with RdRp and S ORFs showing 96.1-99% and 95.2-97.9% nucleotide (nt) identities, respectively (Table S7). The tree topology showed that the RdRp of CoV-HMU-1 clustered with*Rattus norvegicus* Betacoronavirus (KM349744), CoV-HMU-3 clustered with Betacoronavirus 1 GCCDC4 (MW773844), and CoV-HMU-2 and CoV-HMU-4 clustered with Longquan Rl rat coronavirus (KF294371) (Figure 5a).

The two domains of the S protein of CoVs are related to the invasive ability of the virus into the host cell; the S1 domain is involved in cellular receptor binding and the S2 domain promotes the fusion between the viral capsid and the cell membrane. The putative S1 region was located at residues 493–873 for CoV-HMU-1, 463–775 for CoV-HMU-2 and CoV-HMU-4, and 202–442 for CoV-HMU-3. CoV-HMU-1 S1 shared 94.49% amino acid (aa) identity with Betacoronavirus HKU24, and 59.62–59.26% aa identity with lineages 2 and 3 CoVs (Figure 5b). CoV-HMU-2 and CoV-HMU-4 shared high aa identities with lineage 2 CoVs in the S1 region (88.46–96.31%) and showed 95.19% aa identities with each other. CoV-HMU-3 shared 95.00% aa identity with Betacoronavirus 1 and 66.49-59.26% with CoV-HMU-2 and CoV-HMU-4. The putative S2 region was located at residues 1081–1560 for CoV-HMU-1, 880–1408 for CoV1-HMU-2 and CoV-HMU-4, and 880–1419 for CoV-HMU-3.

Flaviviridae

Flaviviruses (order*Amarillovirales* , family*Flaviviridae*) are a group of enveloped RNA viruses with positive-sense ssRNA genomes of 9-13 kb.*Flaviviridae* can be classified into four genera, *Flavivirus* , *Hepacivirus* ,*Pegivirus* , and *Pestivirus* , which include many important arboviruses and mammalian viruses that are responsible for various mild to severe infectious diseases in humans, primates, bats, horses, and rodents [30-32]. Yellow fever, dengue, and Zika viruses are also important human pathogens [33]. Pestiviruses (PestVs), including the economically important bovine viral diarrhoea virus and classical swine fever virus [34].

We obtained two novel complete PestV genome sequences, PestV-HMU-1 and PestV-HMU-1, using representative reads for pestiviruses in*Leopoldamys edwardsi* collected from Baisha and *Rattus tanezumi* collected from Chengmai County. The Nest-PCR was designed to screen for novel PestVs in the pool with read clues. PestV-HMU-1 and PestV-HMU-2 were detected in 12% (2/17) and 6% (1/18) of throat swabs, respectively (Table S5). Additionally, pairwise alignment revealed that the RdRp (non-structural protein 5, NS5, or NS5B) of PestV-HMU-1 and PestV-HMU-2 showed less than 62.39% and 93.33% aa identity with other PestV members (Table S8), respectively. Phylogenetic analysis based on RdRp aa sequences suggested that the two PestVs could be assigned to two different lineages compared with known PestVs; PestV-HMU-1 clustered with Kioloa rodent PestV (OL452246) and PestV-HMU-2 clustered with Norway rat pestV (NC025677) but appeared to represent a separate evolutionary lineage (Figure 6). Sequence similarity and phylogenetic analysis revealed that both novel sequences could be classified within the genus Pestivirus under *Flaviviridae* .

Astr*oviridae*

Astroviruses (AstroVs; order *Nidovirales* , family*Astroviridae*) are non-enveloped RNA viruses with positive-sense ssRNA genomes of 6.8-7 kb. The AstroVs contained three overlapping ORF.*Mamastrovirus* es and *astroviruses* are members of *the Astroviridae* family. AstroVs infect humans and animals, causing diarrhoea, fever, and vomiting; the major clinical symptom is watery diarrhoea. AstroVs may be associated with meningitis, aseptic encephalitis, and meningoencephalomyelitis in humans and animals [35].

We obtained a novel AstroV strain complete genome sequence termed AstroV-HMU-4, and two AstroVs RdRp and capsid proteins, termed AstroV-HMU-3 and AstroV-HMU-5, using representative reads for AstroVs in*Rattus tanezumi* collected from Lingao and Chengmai. AstroV-HMU-3, AstroV-HMU-4, and

AstroV-HMU-5 were detected in 11% (2/18), 22% (4/18), and 96% (24/25) of the anal swabs, respectively, and AstroV-HMU-5 was detected in 68% (17/25) of throat swabs (Table S5). The identity analysis of RdRp showed that the three identified AstroVs had [?]97.59% aa identity with other known AstroVs (Table S9). Moreover, evolutionary trees were constructed for complete protein sequences of RdRp and CP. Phylogenetic analysis of RdRp and CP showed that the three AstroVs clustered with Civet astrovirus (OM451116) and rodent astrovirus (KT946730) within the genus *Mamastrovirus* ; however, AstroV-HMU-3 appeared to represent a separate evolutionary lineage (Figure 7).

Calici*viridae*

Calicivirus virions are positive-sense, non-enveloped ssRNA viruses belonging to the family *Caliciviridae* . Partial Calicivirus RdRp (CaliV-HMU-1) was obtained from *Rattus norvegicus* in Sanya (Table S5). However, the sequence was difficult to compare at the nucleotide level and showed aa identity with other known caliciviruses. Therefore, we suggest that CaliV-HMU-1 belongs to a new species under *Sapovirus* , showing 68.06% aa identity with the human Sapovirus GI.

Hepeviridae

Hepeviruses include enterically transmitted small quasi-enveloped viruses with positive-sense RNA genomes. Hepatitis E is an important causative agent of acute sporadic viral hepatitis worldwide. We obtained a partial RdRp sequence of picornavirus from Haikou*Rattus norvegicus* and named it HepeV-HMU-1 (Table S5). Sequence similarity of partial RdRp indicated that HepeV-HMU-1 showed 92.16% nt identity with Paslahepevirus isolated from Guangdong *Rattus norvegicus* .

*Picornavir*idae

Picornaviruses are members of the Picornaviridae family and are small, non-enveloped, and positive ssRNA viruses. Picornavirus can cause skin, mucous membrane, intracranial, heart, liver, nervous, and respiratory diseases in many vertebrate hosts [36]. We obtained a partial RdRp sequence of the picornavirus from Chengmai Rattus tanezumi and named it PicoV-HMU-1 (Table S5). Sequence similarity of partial RdRp indicated that PicoV-HMU-1 shared 93.14% nt identity with kobuvirus 1 isolated from *Guangdong Rattus losea.*

Characteristics of DNA viruses

Parvoviridae

Parvoviruses (ParVs; order *Piccovirales* , family*Parvoviridae*) are small, non-enveloped, non-segmented, ssDNA viruses with an average genome size of 4-6 kb [37]. ParVs are classified into three subfamilies: *Densovirinae* ,*Hamaparvovirinae* , *and Parvovirinae* . ParVs infect many different animal hosts, including bovines, canines, bats, rodents, and nonhuman primates. Cross-species transmission of Parvovirinae among carnivores has been previously reported [38]. ParV B19 causes hydrops fetalis in fetuses, erythema infectiosum in children, arthritis in adults, and aplastic crisis in patients with hemoglobinopathies.

To identify the genomes, we used representative reads for ParVs in*Rattus tanezumi* collected from Chengmai and Haikou (Fucheng),*Rattus andamanensis*collected from Baisha (Bangxi), and *Leopoldamys edwardsi*collected from Baisha (Huangjingjiaoling) and obtained two novel ParV complete genome sequences, ParV-HMU-1 and ParV-HMU-2. Moreover, three ParV partial NS1 strains were obtained and named ParV-HMU-3, ParV-HMU-4, and ParV-HMU-5, respectively. The analysis revealed that 14% (3/21), 11% (2/18), and 20% (4/20) of the anal swabs were positive for ParV-HMU-1, ParV-HMU-4, and ParV-HMU-5, respectively. Moreover, 8% (3/25) and 11% (2/18) of throat swabs were positive for ParV-HMU-2 and ParV-HMU-3, respectively (Table S5). Genome identity analysis of NS1 revealed that ParV-HMU-4 and ParV-HMU-5 share [?]99.09% aa identity with other known ParVs and with each other (Table S10). Pairwise alignment of NS1 indicated that ParV-HMU-1, ParV-HMU-2, and ParV-HMU-3 showed high diversity (66.36 %, 68.79 %,

and 75.58% aa identity with other known ParVs, respectively). Evolutionary trees constructed for NS showed that the five ParVs were clustered with rat bocavirus (KT454512) under *bocavirus* ; however, ParV-HMU-1, ParV-HMU-2, and ParV-HMU-3 appeared to represent a separate evolutionary cluster (Figure 8a).

Papillomaviridae

Papillomaviruses (PVs; order *Zurhausenvirale* s, family *Papillomaviridae* , subfamily *Firstpapillomavirinae* , and *Secondpapillomavirinae*) are a diverse group of small, non-enveloped viruses with dsDNA genomes of 5-8 kb. However, ancestral PVs are composed of four major ORFs that encode early (E1, E2) and late (L1 and L2) proteins. PVs can cause persistent infections in the skin and mucosal membranes in humans and mammals and may also cause epidermal proliferative lesions.

Representative reads for PVs in *Rattus norvegicus* collected from Tunchang (Nandian), *Rattus tanezumi* collected from Haikou (Yanfeng), and *Niviventer* collected from Baisha (Yinggeling) were used to obtain two novel PV strains with complete genome sequences, named PV-HMU-5 and PV-HMU-6, and two PVs L1 sequences, named PV-HMU-3 and PV-HMU-4. PV-HMU-3 was detected in 45% (9/20) of throat swabs whereas PV-HMU-5 was detected in 61% (11/18) of throat swabs and 55% (10/18) of anal swabs. PV-HMU-6 was detected in 50% (9/18) of throat swabs and 38% (7/18) of anal swabs whereas PV-HMU-4 was detected in 10% (2/20) of throat swabs (Table S5). Pairwise alignment revealed that the L segments of PV-HMU-3 and PV-HMU-4 showed 98.98% and 100.00% aa identity with known PVs (Table S11), respectively. Moreover, genome identity analysis of L revealed that PV-HMU-5 and PV-HMU-6 showed 89.96% and 89.63% nt identities with all other known PVs, respectively. Phylogenetic analysis based on the L nucleotide sequences revealed that four PVs were assigned to the genus *Firstpapillomavirinae* ; however, PV-HMU-3, PV-HMU-4, and PV-HMU-5 clustered with each other; PV-HMU-6 clustered with Mastomys natalensis papillomavirus 1 (MRU01834) and separated from the members of the PVs (Figure 8b).

Characteristics of unclassified RNA viruses

In the present study, we identified seven unclassified RNA viruses in swabs of different rodent species from different sites on Hainan Island (Table S4). The seven unclassified virus families mainly belonged to the unclassified *Picornavirales* under *Pisuvricota* . After annotating the complete or partial ORF or RdRp coding regions of these fully or partially sequenced viruses, many unclassified *Picobirnaviridae* and *Picornaviridae* reads belonged to pools from the hosts of *Rattus* and *Niviventer* . Moreover, reads of *Rhabdoviridae* , *Dicistroviridae* , and *Chuviridae* appeared in pools from Rattus hosts.Data analysis showed that rodents carry a larger proportion of viral RNA and a wider range of phylogenetic diversity. Therefore, further research on rodent-borne RNA will have important implications for expanding our understanding of the viral spectrum and pathogenicity of rodent-borne viruses.

Discussion

With economic globalisation and environmental changes, the scope of human activities continues to expand, which increases the possibility of pathogens jumping from rodent hosts to humans and domestic animals. The risk of zoonotic EID is elevated in tropical forest areas experiencing land-use change and wildlife biodiversity (mammal species richness), based on the prediction of demographic, environmental, and biological correlates associated with EID events [39]. Traditional virus detection methods rely on cell culture based on visible cytopathic effects and molecular detection of known viral genome sequences. However, the application of NGS technology enables the sequencing of pathogenic genomes in one sample, including unknown pathogens. Hainan Island is an important trade hub and tropical monsoon area with large animal species diversity. Therefore, virome and phylogenetic analyses of rodent-borne viruses in Hainan Island can provide insights into the prevention and control of rodent-borne diseases in China.

The virome data obtained in this study revealed that the risk of rodent-borne pathogens on Hainan Island should not be underestimated. In this study, we collected swab samples from six species and three genera of Muridae, covering nine districts of Hainan Province, to analyse the viral diversity carried by rodents living next to humans and living in natural reserves. The novel CoVs, PestVs, AstVs, ParVs, and PVs identified in this study increase our knowledge of the viral classification of each viral family. Many rodent-borne viruses detected in this study were not only found on Hainan Island but also in the southwestern (Yunnan, Hunan) and northern (Jilin, Ningxia, Inner Mongolia, Xinjiang) regions of China [20], Thailand, Lao PDR, Cambodia, and other parts of the world [18]. Moreover, we also identified widely distributed respiratory tract or enteric RNA viruses, such as AstroV, pestivirus, papillomavirus, picornaviruses, rotavirus, and picobirnavirus, which is consistent with a previous study [20]. However, the abundance of these viruses in the metagenomic analysis in the present study was different from that previously reported in rodent lung samples [18]. Our data was not abundant in viruses that are transmitted through blood or bodily fluids, such as phleboviruses, hepaciviruses, Hantaan viruses, and arteriviruses. This phenomenon may be due to the different tissue tropisms of these viruses. The results of the viral positivity rate suggest that these viruses infect hosts primarily through the respiratory tract, such as CoVs and PestVs. Moreover, some viruses were positive in anal swabs but negative in throat swabs. These results may be attributed to insufficient collection of virus-infected exfoliated host cells during the collection of swabs or to the different replication characteristics of faecal/oral- or respiratory-transmitted viruses, which may be detected at different locations in different cycles [19]. Some viruses, such as AstroV-HMU-5, were detected in both throat and anal swabs, indicating that the host may carry the virus for a long time and may transmit it via the faecal-oral route (food- or water-borne routes) [40]. Additionally, PV-HMU-5 was detected with PV-HMU-6 in the same host, and throat and anal swabs showed PV positivity in multiple hosts, indicating that this host group carried the PVs for a long time, which increases the probability of the virus becoming a potential pathogen.

The murine hepatitis virus in murine coronaviruses was first isolated in 1949 [41], and a variant named sialodacryoadenitis coronavirus was detected in rats in 1970 [42]. Coronavirus recombination events are highly common, such as feline CoV type I and canine coronavirus recombination events [43]. Human CoVs recombine to produce three genotypes [44] and SARS-CoV constantly undergoes evolutionary recombination [45]. The results of the phylogenetic analysis of the RdRp and S protein strains of CoVs were consistent. The four CoVs identified in this study shared high RdRp and S protein identities with known murine CoVs and clustered with murine CoVs under the subgenus Embecovirus in the evolutionary tree. CoV-HMU-4, CoV-HMU-2, CoV-HMU-3, and mouse hepatitis virus showed a relatively close distance in the evolutionary tree, especially CoV-HMU-3 and mouse hepatitis virus, under the same branch. Based on the rodent hosts of known and newly discovered CoVs, our results suggest that the same virus can be detected in different regions and rodents, indicating that rodents carrying CoVs may not exhibir geographic and host specificity. The high similarity of CoVs observed in our study suggests that each lineage shares a common ancestor. Additionally, beta-CoVs were detected in the throat and anal swabs of rodents and only a few clues were found in rodent lungs, suggesting that these viruses primarily infect the upper respiratory tract whereas the absence of alpha-CoVs in the two types of samples indicates that alpha-CoVs may not be the main cause of lower respiratory tract infections in rodents.

Pestiviruses may cause clinical illnesses, such as acute diarrhoea, acute haemorrhagic syndrome, wasting disease, and transplacental infection, leading to fetal death [46], and posing serious health and economic burden. PestV genomic sequences have been detected in bat and rat samples using NGS [8, 47]. In the present study, pairwise alignment revealed that the RdRp of PestV-HMU-1 and PestV-HMU-2 showed low aa identities with known pathogens, suggesting that they may be novel species, and rodents have not been previously reported to carry these viruses. PestV-HMU-1 was carried by the *Leopoldamys edwardsi* collected from the Huanjinjiaoling reserve and PestV-HMU-2 was carried by the *Rattus tanezumi* collected from the Chengmai. Moreover, PestV-HMU-1 was also present with ParV-HMU-2 in *Leop* oldamys edwardsi from the Huanjinjiaoling reserve of Hainan. The abundance and diversity of viruses are high in the central mountainous areas of Hainan.

Multiple infectious ParVs have been detected in animals; porcine ParV causes reproductive failure in pigs

[48] whereas canine and feline panleukopenia ParVs are pathogenic to dogs and cats, respectively [49]. ParV infection can also cause serious harm to humans; toad virus, tusavirus, and cutavirus are mainly detected in children with diarrhoea [50]. ParV-HMU-1 and ParV-HMU-2, carried by *Rattus andamanensi* and *Leopoldamys edwardsi*collected from the two nature reserves in Baisha, showed low identities with known pathogens. Overall, the virome reported in this study is highly diverse. These results extend our knowledge of viral taxonomy and host range and show that in inaccessible areas, there are still highly diverse viruses that have evolved independently in their unique wildlife hosts. If these viruses cross the host barrier, they are highly likely to cause zoonosis.

Our results provide a profile of the composition of the rodent-borne virome and a baseline for rodent-borne viruses on Hainan Island, China. We detected PestVs and bocaviruses with low identities that were not previously reported to be carried by rodents. The associated viral genome sequences obtained in our study were corroborated, allowing us to further understand the phylogenetic relationship of the viruses on Hainan Island to other regions and obtain data on the diversity and independent evolutionary profiling of rodent-borne pathogens. This study suggests that mammals still harbour a large number of uncharacterised viruses, which require further investigation. This study provides basic data for the prevention and control of rodent-borne infectious diseases, evaluates the potential pathogenicity of newly discovered microorganisms, and provides a baseline for future virome studies and identification of potential pathogens in EIDs.

Conclusions

In this study, we report multiple novel viruses from different viral families in nasal and anal swabs collected from rodents in the Hainan Province. These findings greatly increase our knowledge of the rodent-borne virus community on Hainan Island and provide basic data for analysing the evolutionary independence of viruses. The pathogenicity and the associated impact of these novel viruses on humans and animals should be evaluated in further studies.

Author Contributions

FY, LG and QJ designed the study. YL, CT, YZ, ZL, GW, PR, YH, XH, XC, LN, HS, MX and GL collected the specimens and performed the experiments. HX, BF, XFC, ZW, JY, YPH, TG, YJH and FY analysed the data. YL and JD wrote the manuscript. All authors read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Key Research and Development Plan of Hainan Province (No. ZDYF2020150, and ZDYF2022SHFZ085), the National Natural Science Foundation of China (Nos. 32060015, 81860367 and 82060378), CAMS Innovation Fund for Medical Sciences (CIFMS)" (grant 2021-I2M-1-037), the Major Science and Technology Program of Hainan Province (No. ZDKJ202003), the Hainan Provincial Natural Science Foundation of China (Nos. 820RC628 and 2019RC218), the Scientific Research Foundation of the Higher Education Institutions of Hainan Province, China (No. Hnky2021-40), the Research Project of the Hainan Academician Innovation Platform (No. YSPTZX202004), the Hainan Talent Development Project (No. SRC200003), Graduate student innovation grant of Hainan Medical University (No. Hys2020-33, and Qhys2021-356).

Data availability statement

Sequence data were deposited in the National Center for Biotechnology Information (NCBI) sequence reads archive under the accession number PRJNA892773.

References

[1] Jones KE, Patel NG, Levy MA, et al. Global trends in emerging infectious diseases. Nature, 2008;451(7181):990-3.

[2] Taylor LH, Latham SM, Woolhouse ME. Risk factors for human disease emergence. Philos Trans R Soc Lond B Biol Sci. 2001;356(1411):983-9.

[3] Wu T, Perrings C, Kinzig A, et al. Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: A review . Ambio. 2017;46(1):18-29.

[4] Phan TG, Kapusinszky B, Wang C, et al. The fecal viral flora of wild rodents. PLoS Pathog. 2011;7(9):e1002218.

[5] Han BA, Schmidt JP, Bowden SE, et al. Rodent reservoirs of future zoonotic diseases. Proc Natl Acad Sci U S A. 2015;112(22):7039-44.

[6] Wu Z, Du J, Lu L, et al. Detection of Hantaviruses and Arenaviruzses in three-toed jerboas from the Inner Mongolia Autonomous Region, China. Emerg Microbes Infect. 2018;7(1):35.

[7] Guo P, Lin XD, Wang W, et al. Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. PLoS Pathog. 2013;9(2):e1003159.

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

[9] Du J, Lu L, Liu F, et al. Distribution and characteristics of rodent picornaviruses in China. Sci Rep. 2016;6:34381.

[10] Ishii A, Thomas Y, Moonga L, et al. Novel arenavirus, Zambia. Emerg Infect Dis. 2011;17(10):1921-4.

[11] Briese T, Paweska JT, Mcmullan LK, et al. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. PLoS Pathog. 2009;5(5):e1000455.

[12] Knust B, Stroher U, Edison L, et al. Lymphocytic choriomeningitis virus in employees and mice at multipremises feeder-rodent operation, United States, 2012. Emerg Infect Dis/ 2014;20(2):240-7.

[13] Palacios G, Druce J, Du L, et al. A new arenavirus in a cluster of fatal transplant-associated diseases. N Engl J Med. 2008;358(10):991-998.

[14] Sridhar S, Yip CCY, Wu S, et al. Rat Hepatitis E Virus as Cause of Persistent Hepatitis after Liver Transplant. Emerg Infect Dis. 2018;24(12):2241-50.

[15] Pulkkinen LIA, Butcher SJ, Anastasina M. Tick-Borne Encephalitis Virus: A Structural View. Viruses.2018;10(7):350.

[16] Liu Q, He B, Huang SY, et al. Severe fever with thrombocytopenia syndrome, an emerging tick-borne zoonosis. Lancet Infect Dis. 2014;14(8):763-72.

[17] Djikeng A, Kuzmickas R, Anderson NG, et al. Metagenomic analysis of RNA viruses in a fresh water lake. PLoS One. 2009;4(9):e7264.

[18] Wu Z, Han Y, Liu B, 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(1):18.

[19] Zhang Y, Rao L, Lv G, et al. Complete Genome Sequence of a Mammarenavirus Harbored by Rodents on Hainan Island, China. Genome Announc. 2018;6(10):e00129-18.

[20] Wu Z, Lu L, Du J, et al. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases . Microbiome. 2018; 6(1):178.

[21] Haas BJ, Papanicolaou A, Yassour M, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8):1494-1512.

[22] Li H, Durbin RR. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.

[23] McCormick JB, Webb PA, Krebs JW, et al. A prospective study of the epidemiology and ecology of Lassa fever. J Infect Dis. 1987;155(3):437-444.

[24] Goni SE, Iserte JA, Stephan BI, et al. Molecular analysis of the virulence attenuation process in Junin virus vaccine genealogy. Virus Genes. 2010;40(3):320-8.

[25] Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. Virology. 2001;279(2):371-4.

[26] Gonzalez JM, Gomez-Puertas P, Cavanagh D, et al. A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. Arch Virol. 2003;148(11):2207-35.

[27] Ren LL, Wang M, Wu ZQ, et al. Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. Chin Med J (Engl). 2020;133(9):1015-24.

[28] Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin . Nature. 2020;579(7798):270-3.

[29] Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 2019;17(3):181-92.

[30] Drexler JF, Corman VM, Muller MA, et al. Evidence for novel hepaciviruses in rodents. PLoS Pathog. 2013;9(6):e1003438.

[31] Quan PL, Firth C, Conte JM, et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses. Proc Natl Acad Sci U S A. 2013;110(20):8194-9.

[32] Smith DB, Becher P, Bukh J, et al. Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family. J Gen Virol. 2016;97(11):2894-907.

[33] Thiel HJ, Collett MS, Gould EA, et al. Family Flaviviridae [J]. Virus Taxonomy: Eight Report of the International Committee on the Taxonomy Of Viruses. Elsevier, Amsterdaml; 2005, pp. 981-998.

[34] Liu L, Larska M, Xia H, et al. Atypical pestivirus and severe respiratory disease in calves, Europe. Emerg Infect Dis, 2012;18(11):1917-8.

[35] Calistri A, Palu G. Editorial commentary: Unbiased next-generation sequencing and new pathogen discovery: undeniable advantages and still-existing drawbacks. Clin Infect Dis. 2015;60(6):889-91.

[36] Zell R, Delwart E, Gorbalenya AE, et al. ICTV virus taxonomy profile: Picornaviridae. J Gen Virol. 2017;98(10):2421-2.

[37] Mietzsch M, Penzes JJ, Agbandje-Mckenna M. Twenty-five years of structural Parvovirology. Viruses. 2019;11(4):362.

[38] Allison AB, Kohler DJ, Fox KA, et al. Frequent cross-species transmission of parvoviruses among diverse carnivore hosts. J Virol. 2013;87(4):2342-7.

[39] Allen T, Murray KA, Zambrana-Torrelio C, et al. Global hotspots and correlates of emerging zoonotic diseases. Nat Commun. 2017;8(1):1124.

[40] Walter JE, Mitchell DK. Role of astroviruses in childhood diarrhea. Curr Opin Pediatr. 2000;12(3):275-279.

[41] Bailey OT, Pappenheimer AM, Cheever FS, et al. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin: II. pathology. J Exp Med.1949;90(3):195-212.

[42] Parker JC, Cross SS, Rowe WP. Rat coronavirus (RCV): a prevalent, naturally occurring pneumotropic virus of rats. Arch Gesamte Virusforsch. 1970, 31(3): 293-302.

[43] Herrewegh AA, Smeenk I, Horzinek MC, et al. Feline coronavirus type II strains 79-1683 and 79-1146 originate from a double recombination between feline coronavirus type I and canine coronavirus. J Virol. 1998;72(5):4508-14.

[44] Woo P C, Lau SK, Huang Y, et al. Coronavirus diversity, phylogeny and interspecies jumping. Exp Biol Med (Maywood). 2009;234(10):1117-27.

[45] Graham R L, Baric RS. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J Virol. 2010;84(7): 3134-46.

[46] Moennig V, Plagemann PG. The pestiviruses. Adv Virus Res. 1992,41:53-98.

[47] Wu Z, Ren X, Yang L, et al. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. J Virol. 2012;86(20):999-1012.

[48] Schirtzinger EE, Suddith AW, Hause BM, et al. First identification of porcine parvovirus 6 in North America by viral metagenomic sequencing of serum from pigs infected with porcine reproductive and respiratory syndrome virus. Virol J. 2015;12:170.

[49] Parker J S, Murphy WJ, Wang D, et al. Canine and feline parvoviruses can use human or feline transferrin receptors to bind, enter, and infect cells. J Virol. 2001;75(8): 3896-902.

[50] Vaisanen E, Fu Y, Hedman K, et al. Human Protoparvoviruses. Viruses. 2017; 9(11).

Figure 1. Sample collection in the Hainan Island and metagenomics of rodents according to species and locations. The maps of Haikou City, Baisha County, Dongfang City, Lingao County, Tunchang County, Lingshui County, Ledong County, Chengmai County, and Sanya City in Hainan Island, showing the locations of the samples. Rodent numbers and species are also shown on the map with different shapes, colours, and sizes.

Figure 2. (a-d**).** Comparison of the number of total contigs and proportions of viral reads from the six rodent species. Comparison of the numbers of total reads and proportions of viral reads from the fifteen sampling locations. (e**).** The proportions of the number of reads assigned to each viral family in each of the 28 pools are shown in the bar graph.

Figure 3. Rodent-related virome profile in the transcriptomic data from the 28 pools. Expression is reported in log2(TPM), where TPM refers to transcripts per million.

Figure 4. Phylogenetic tree of arenaviruses (AreVs). The phylogenetic tree is based on the partial L protein (RdRp) acid sequences of AreVs. Phylogenetic trees were constructed by the maximum likelihood method using the best-fit models (LG+G+I). AreVs identified in this study are labelled in red. Host genus and location of each virus are labelled by the 5-point stars and dots of different colours. The outer colour rings represent additional taxonomic information about these viruses.

Figure 5 . Phylogenetic tree of coronaviruses (CoVs). (a) The phylogenetic tree is based on the amino acid sequences of complete RNA-dependent RNA polymerase (RdRp) genes and S (spike) genes of CoVs. (b)Phylogenetic trees were constructed by the maximum likelihood method using the best-fit models (WAG+G). All CoVs identified in this study are labelled in red. Host genus and location of each virus are labelled by the 5-point stars and dots of different colours. The outer colour rings represent additional taxonomic information about these viruses.

Figure 6 . Phylogenetic tree of pestiviruses (PestVs). The phylogenetic tree is based on the amino acid sequences of complete RNA-dependent RNA polymerase (RdRp) (NS5) genes of PestVs. Phylogenetic trees

were constructed by the maximum likelihood method using the best-fit models (rtREV+G). All PestVs identified in this study are labelled in red. Host genus and location of each virus are labelled by the 5-point stars and dots of different colours. The outer colour rings represent additional taxonomic information about these viruses.

Figure 7 . Phylogenetic tree of astroviruses (AstroVs). (a-b) The phylogenetic tree is based on the amino acid sequences of complete RNA-dependent RNA polymerase (RdRp) genes and CP genes. Phylogenetic trees were constructed by the maximum likelihood method using the best-fit models (mtREV+G+I). All AstroVs identified in this study are labelled in red. Host genus and location of each virus are labelled by the 5-point stars and dots of different colours. The outer colour rings represent additional taxonomic information about these viruses.

Figure 8 . Phylogenetic tree of parvoviruses (ParVs) and papillomaviruses (PVs). (a) Phylogenetic tree based on the amino acid sequences of complete NS5 genes of ParVs. Phylogenetic trees were constructed by the maximum likelihood method using the best-fit models (rtREV+G+I). (b) Phylogenetic trees are based on the amino acid sequences of complete L genes of PVs and constructed by the maximum likelihood method using the best-fit models (LG+G). All ParVs and PVs identified in this study are labelled in red. Host genus and location of each virus are labelled by the 5-point stars and dots of different colours. The outer colour rings represent additional taxonomic information about these viruses.