


RESEARCH

Open Access



Decoding the RNA viromes in rodent lungs provides new insight into the origin and evolutionary patterns of rodent-borne pathogens in Mainland Southeast Asia

Zhiqiang Wu^{1,2*}, Yelin Han^{1†}, Bo Liu^{1†}, Hongying Li³, Guangjian Zhu³, Alice Latinne^{3,4,5}, Jie Dong¹, Lilin Sun¹, Haoxiang Su¹, Liguang Liu¹, Jiang Du¹, Siyu Zhou¹, Mingxing Chen¹, Anamika Kritiyakan⁶, Sathaporn Jittapalpong⁶, Kittipong Chaisiri⁷, Phillipe Buchy⁸, Veasna Duong⁹, Jian Yang¹, Jinyong Jiang¹⁰, Xiang Xu¹⁰, Hongning Zhou¹⁰, Fan Yang¹, David M. Irwin¹¹, Serge Morand⁶, Peter Daszak³, Jianwei Wang¹ and Qi Jin^{1*} 

Abstract

Background: As the largest group of mammalian species, which are also widely distributed all over the world, rodents are the natural reservoirs for many diverse zoonotic viruses. A comprehensive understanding of the core virome of diverse rodents should therefore assist in efforts to reduce the risk of future emergence or re-emergence of rodent-borne zoonotic pathogens.

Results: This study aimed to describe the viral range that could be detected in the lungs of rodents from Mainland Southeast Asia. Lung samples were collected from 3284 rodents and insectivores of the orders Rodentia, Scandentia, and Eulipotyphla in eighteen provinces of Thailand, Lao PDR, and Cambodia throughout 2006–2018. Meta-transcriptomic analysis was used to outline the unique spectral characteristics of the mammalian viruses within these lungs and the ecological and genetic imprints of the novel viruses. Many mammalian- or arthropod-related viruses from distinct evolutionary lineages were reported for the first time in these species, and viruses related to known pathogens were characterized for their genomic and evolutionary characteristics, host species, and locations.

Conclusions: These results expand our understanding of the core viromes of rodents and insectivores from Mainland Southeast Asia and suggest that a high diversity of viruses remains to be found in rodent species of this area. These findings, combined with our previous virome data from China, increase our knowledge of the viral community in wildlife and arthropod vectors in emerging disease hotspots of East and Southeast Asia.

Keywords: Mainland Southeast Asia, Rodent lungs, Core virome, Viral evolution, Emerging infectious diseases

* Correspondence: wuzq2009@ipbcams.ac.cn; zdsys@vip.sina.com

†Zhiqiang Wu, Yelin Han and Bo Liu contributed equally to this work.

¹NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Most human infectious viral diseases are of zoonotic origin, such as Coronavirus Disease 2019 (COVID-19) and Ebola [1–4]. There are multiple animal hosts for these zoonotic viruses, including Rodentia. Rodentia is the most diverse mammalian order and is comprised of at least 33 families and 2277 species that occur over a large variety of terrestrial habitats including both natural and human-made environments [5, 6]. Rodents live in close contact with arthropods (e.g., fleas, ticks, and mites), wildlife, and domestic animals and are a critical link in the transmission of various zoonotic pathogens between humans and animals [7, 8]. Several high-impact zoonotic viruses from the families *Arenaviridae*, *Flaviviridae*, *Hantaviridae*, *Nairoviridae*, *Reoviridae*, and *Togaviridae* are of rodent origin [8–15]. Lassa virus (LASV), White-water Arroyo virus, Machupo virus, Junin virus, Guanarito virus, Tacaribe virus, and Sabia virus are examples of viruses from the family *Arenaviridae* that can cause severe human hemorrhagic fever and related diseases [16–19]. Another arenavirus (AreV) is the lymphocytic choriomeningitis virus (LCMV), which is the causative agent for human aseptic meningitis, encephalitis, and meningoencephalitis [20]. Many species of hantaviruses (HanVs), such as Hantaan virus, Seoul virus, Puumala virus, and Sin Nombre Virus, are the causative agents for hemorrhagic fever with renal syndrome (HFRS) or hantavirus pulmonary syndrome (HPS) in humans [13]. Rat hepatitis E virus (HEV) causes persistent hepatitis in humans [21]. Moreover, rodents play important roles in the natural circulation of vector-borne (especially tick-borne) viruses such as the Crimean-Congo hemorrhagic fever virus (CCHFV) of the family *Nairoviridae* and the tick-borne encephalitis virus (TBEV) and Omsk hemorrhagic fever virus (OHFV) of the family *Flaviviridae* [7, 8].

Recently, meta-genomic and meta-transcriptomic analyses have been increasingly used to survey the virome in diverse animals over a wide geographical area [22–28]. Instead of identifying the pathogens for diseases that have already emerged, virome surveillance provides abundant data to help understand the existing viral population and their ecology in ecosystems before they emerge into human populations causing large-scale outbreaks, to eventually prevent the zoonotic emerging infectious diseases (EIDs) [1, 3]. Previously, we conducted virome surveys in China that have illustrated the ecological characteristics and evolution of viruses in rodents, bats, and other small animals [29–31]. In addition to the discovery of novel viruses, many other known viruses were found in new hosts or locations that allowed a better understanding of the origin and evolutionary patterns of pathogens such as HanV, AreV, coronavirus (CoV), and arterivirus (ArteV) [10, 32–36]. These results suggest that there is still a rich as-yet-to-be-discovered

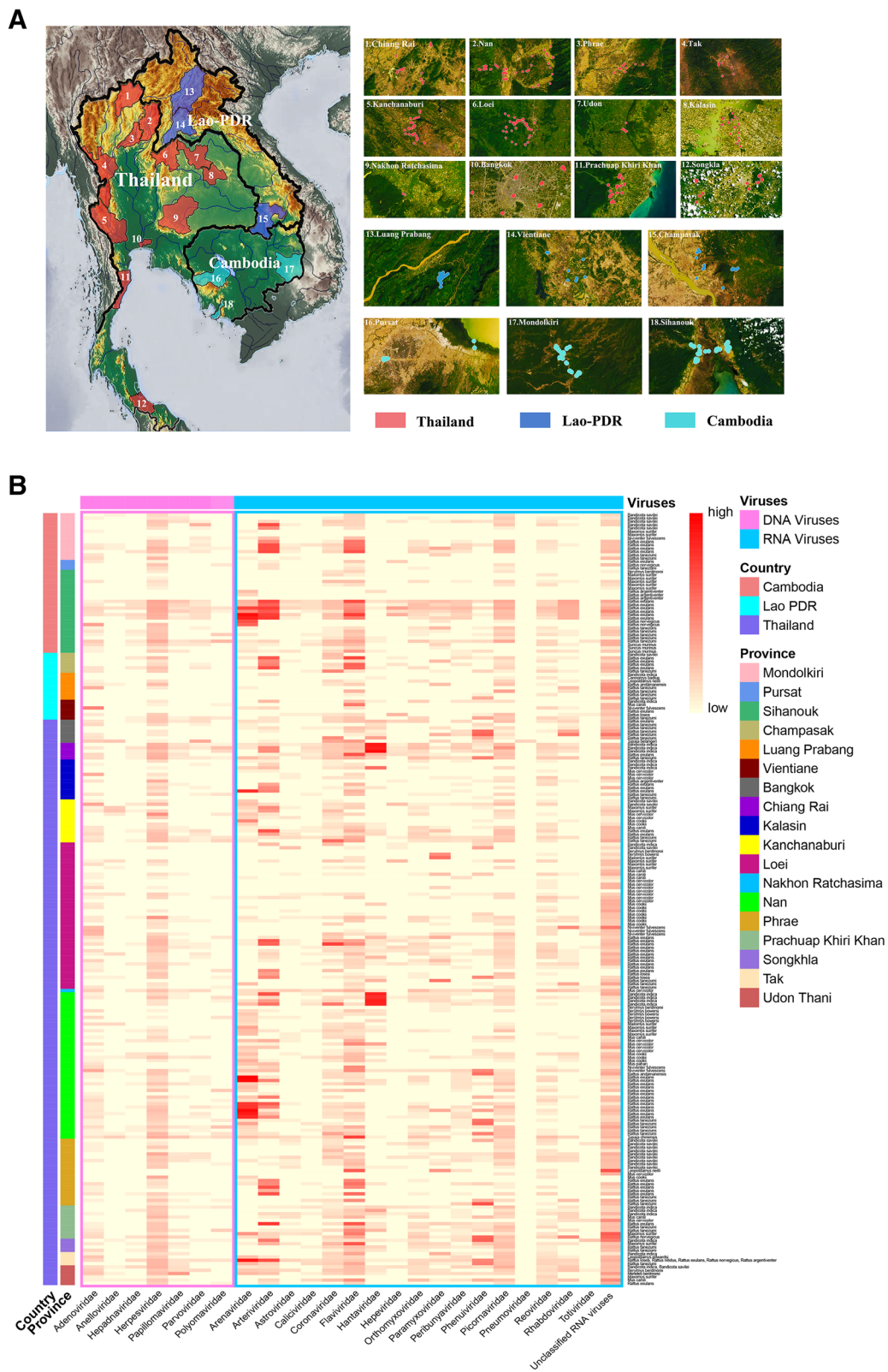
viral diversity in small mammalian species and that the risk of EIDs originating from these rodents should not be underestimated.

Southeast Asia, characterized by a great diversity of mammals, a hotspot of rodent diversification, and a hotspot of biodiversity at threat coupled to cultural traditions of wildlife consumption in many regions, is identified as a hotspot for EIDs [37–41]. A number of zoonotic diseases, including viral hemorrhagic fever, severe acute respiratory syndrome, and Nipah virus-related brain or respiratory diseases, already emerged in this area [42, 43]. Accelerating environmental change and highly connected globalization, leading to increased risks of EID being shared across borders, suggests that joint efforts in surveillance, early detection, prevention, and response to EIDs are needed. For this current study, we collected 3284 lung samples from 30 rodent and insectivore (shrew, tree shrew, and short-tailed gymnure) species in 12 provinces of Thailand, three provinces of Lao PDR, and three provinces of Cambodia over a 12-year period (2006–2018) for a meta-transcriptomic analysis. The virome data obtained revealed an overview of the viral community present in the lungs of these small mammalian species across the Indochina Peninsula. The description of the host and geographical bias of these diverse viral populations and characterization of novel viruses in multiple viral communities, allowed us to identify the ecological and genetic imprints of known human and animal viral pathogens, including HanVs, phleboviruses (PhleVs), AreVs, hepaciviruses, pegiviruses, and arteriviruses (ArteVs) in these regions. This study provides a baseline that can be used for future studies to identify changes in viromes and potentially identify novel EIDs before they emerge.

Results

Animal sampling

A total of 3284 lung samples of small mammal individuals were collected from different sites in the three countries between April 2006 and November 2018. These comprised 3191 lung samples from rodents, 66 lung samples from shrews and two lung samples from short-tailed gymnure under the order Eulipotyphla, and 25 lung samples from tree shrews under the order Scandentia. Samples were obtained from a total of 18 provinces, including the Bangkok urban area, Kanchanaburi, Chiang Rai, Tak, Prachuap Khiri Khan, Loei, Nan, Songkla, Udon Thani, Kalasin, Phrae, and Nakhon Ratchasima in Thailand, Vientiane, Champasak, and Luang Prabang in Lao PDR, and Preah Sihanouk, Monduliri, and Pursat in Cambodia (Fig. 1a). The samples included 25 species of rodents from the families Muridae, Spalacidae, and Sciuridae, three shrew species from the family Soricidae, and two species of tree shrew from the family Tupaiidae that reside in



(See figure on previous page.)

Fig. 1 a Sampling sites from the various provinces of Southeast Asia. Map of Southeast Asia showing the eighteen provinces in the three countries where animals were sampled. Sampling sites are indicated in different color, with red color representing Thailand, blue representing Lao PDR, and cyan representing Cambodia. Sampling sites within the eighteen provinces are shown on the right, with dots indicating sampling sites with each province. The number of samples from the species and the provinces and dates of collection are detailed in Table S1. **b** Heatmap based on the normalized numbers of sequence reads for 24 families of mammalian viruses and a group of unclassified RNA viruses in each pooled sample. Mammalian host species are listed on the right as a text column. Location information is indicated on the left as color code, with colors defined on the right. Names of the mammalian viral families are indicated at the bottom. Relative abundance of the viruses in each species in each location are indicated as a heat map ranging from low (yellow) to high (red) based on the normalized average viral genome size and total sequencing reads in each pool

urban, rural, and wild areas throughout the Indochina Peninsula. The most commonly sampled species were *Rattus exulans*, *R. tanezumi*, *Bandicota savilei*, *B. indica*, *Maxomys surifer*, and *Mus cervicolor* (Table S1). *Rattus* was the dominant genus, accounting for 49.66% of the total samples. The genus *Bandicota* accounted for 18.30% and *Mus* accounted for 14.92% of the total samples. These rodent species are ecologically closely associated with humans, arthropod vectors, and other animals in these areas.

Meta-transcriptomic analysis and overview of the virome

Due to the repeated sampling of some species in the same location, the 3284 isolated lung total RNA samples were combined into 233 pools with equal quantity. To determine whether viral RNA was present in each pool, an rRNA excluded RNA library was constructed and processed for NGS-based meta-transcriptomic analysis. A total of 262.38 GB of nucleotide data was obtained. Reads that were classified as cellular organisms (including bacteria, archaea, and eukaryotes) and reads with no significant homology to any amino acid sequence in the NCBI NR database were removed. The remaining 495,579 reads were used to identify their best-matched hit with viral proteins available in the NCBI NR database (Table S2). Due to the presence of numerous transcripts from the hosts and other cellular organisms, most pools had low levels of viral RNA.

To analyze the virome, virus-associated reads were classified into a group of unclassified RNA viruses and 98 known families under the double-stranded (ds) RNA viruses, reverse-transcribing viruses, single-stranded (ss) RNA viruses, dsDNA viruses, and ssDNA viruses. By further characterization of dietary habits and other host traits, non-vertebrate-associated viral reads and retrotransposon-related sequence reads that have previously been described [23, 31] were removed. The remaining 406,869 viral reads (approximately 82.1% of the total viral hits) were then assigned into 24 mammal-related viral families and a group of unclassified RNA viruses. The reads for each viral family in each pool were normalized by the viral genome size and the proportion of total viral reads, and the prevalence of each viral family in each province and animal species is shown in Fig. 1b and Table S3. Viral reads from the families

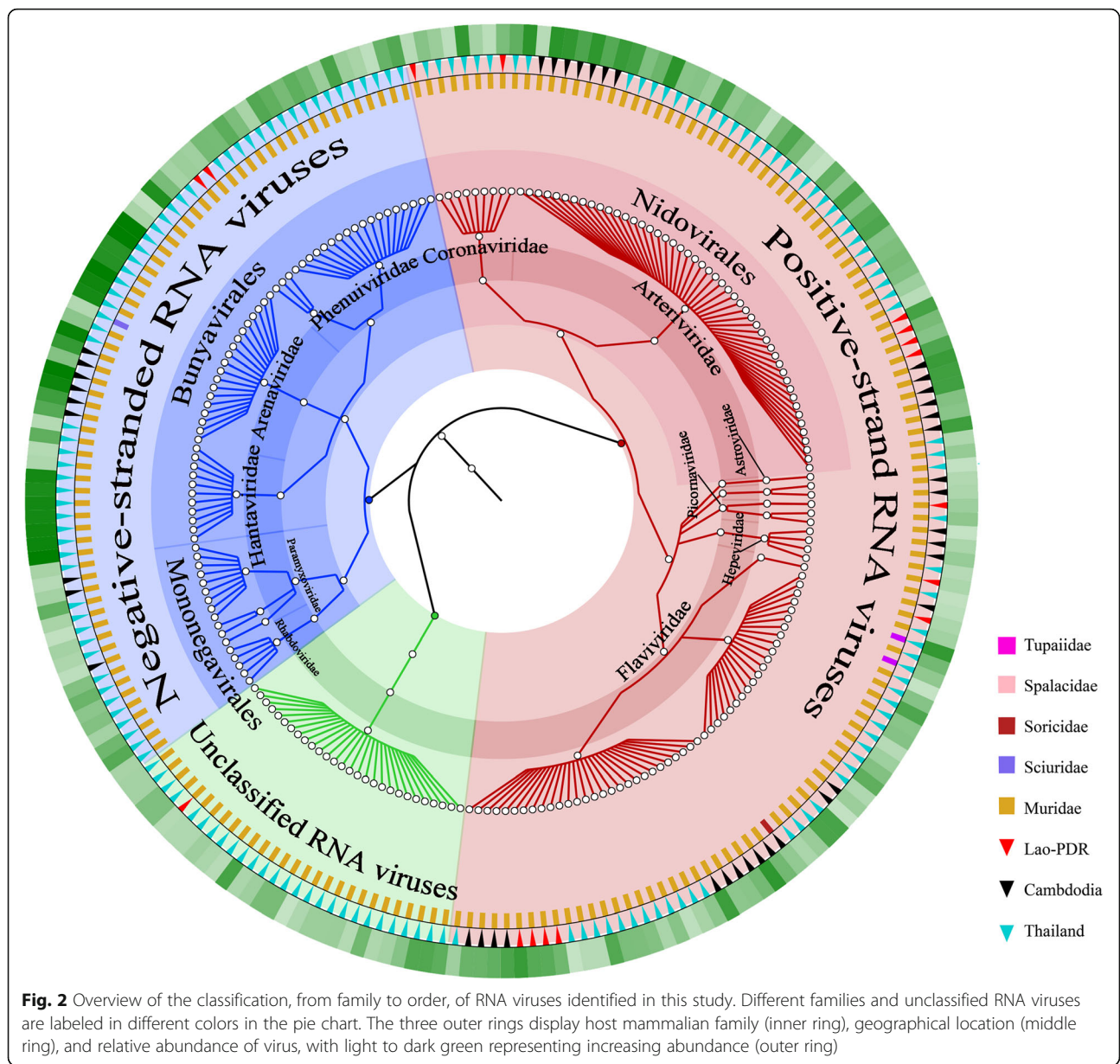
Arteriviridae, *Arenaviridae*, *Flaviviridae*, *Hantaviridae*, *Herpesviridae*, and *Phenuiviridae* were widely distributed, in differing abundances, in a variety of rodent and insectivore species from the different regions. Virus families *Adenoviridae*, *Astroviridae*, *Anelloviridae*, *Coronaviridae*, *Caliciviridae*, *Hepeviridae*, *Herpesviridae*, *Paramyxoviridae*, *Picornaviridae*, *Peribunyaviridae*, *Rhabdoviridae*, and *Reoviridae* were found in fewer species in the different regions. Although sequence reads related to the families *Hepadnaviridae* and *Orthomyxoviridae* were occasionally present in some of the virome, when we used RT-PCR to amplify genomic sequences of these viruses, we failed to amplify any sequences. This might suggest that the *Hepadnaviridae* and *Orthomyxoviridae* viruses were of low viral load or spurious sequence similarities. In addition to the family assigned reads, a substantial number of viral reads were for unclassified RNA viruses in the realm *Riboviria*, including diverse *Chuviridae*-, *Nodaviridae*-, or *Totiviridae*-related viruses. Although most viral hits (98.64%) were assigned into the RNA virus group, a small number of DNA viral reads were found in sequence data due to their corresponding RNA transcripts being detected. Due to their low number, we did not perform any further analyses of these DNA viruses.

Based on the virome data provided by these meta-transcriptomes, the prevalence and diversity of viruses in families including pathogens that are known to cause human and animal infection or are novel in rodent were then confirmed by PCR screening of individual lung samples. In total, 216 representative virus strains were selected for genomic or partial genomic sequencing (Fig. 2 and Table S4). Below we outline the characteristics of these different types of viruses.

Characteristics of negative-stranded RNA viruses

HanVs

Rodent HanVs are segmented negative-stranded RNA viruses with three genome segments (L, M, and S). These viruses of the genus *Orthohantavirus*, family *Hantaviridae* (under the order *Bunyavirales*), can be divided into two groups: the Murinae-related phylogroup III HanVs and the Sigmodontinae-, Neotominae-, and Arvicolinae-related phylogroup IV HanVs [10, 13, 44].



Here, we identified eleven HanV strains that were found in three species (*B. indica*, *R. exulans*, and *R. tanezumi*) of Murinae from five provinces (Bangkok, Chiang Rai, Kalasin, Loei, and Nan) of Thailand in multiple years (Table S4). *B. indica* is the main host for these HanVs. Eight strains were determined for genome sequences, and three strains were selected for sequencing of partial L, M, and S. The complete genome sequence (including non-coding regions (NCRs)) of these viruses showed more than 93.5% nucleotide (nt) identity with each other, suggesting that they all belonged to the same viral species. The open reading frame (ORF) for L of this species showed 79–81.4% nt identity with those of Anjozoro-robe HanVs detected in *R. rattus* and *Eliurus majori*

from Madagascar in 2014 [45], the M and S ORFs of this species showed 95–96% and 96.1–99.4% nt identities with those of Thailand virus strains found in *B. indica* of Thailand in 1994 and 2004 [46, 47] (Table S5). Phylogenetic trees based on the complete M segment-encoded glycoprotein precursor (GPC), L segment-encoded RNA-dependent RNA polymerase (RdRp), and S segment-encoded nucleocapsid protein (N) amino acid sequences were constructed (Fig. 3). All HanVs identified here were assigned to the Murinae-related phylogroup III, clustered together, and formed a separate clade associated with the clade of Anjozoro-robe HanVs that are closely related to the lineage of the Seoul virus strains. Although the exact relationship between these

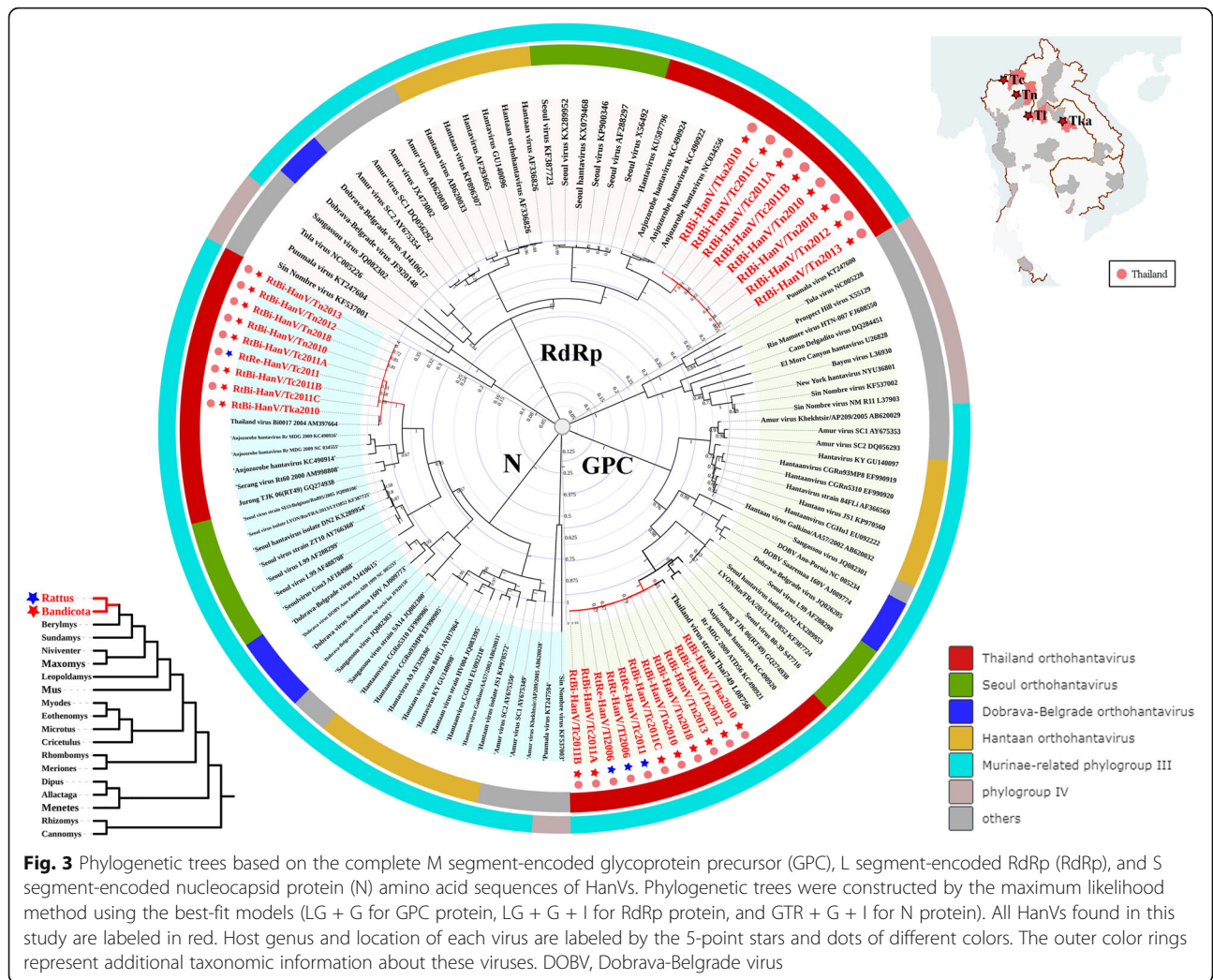


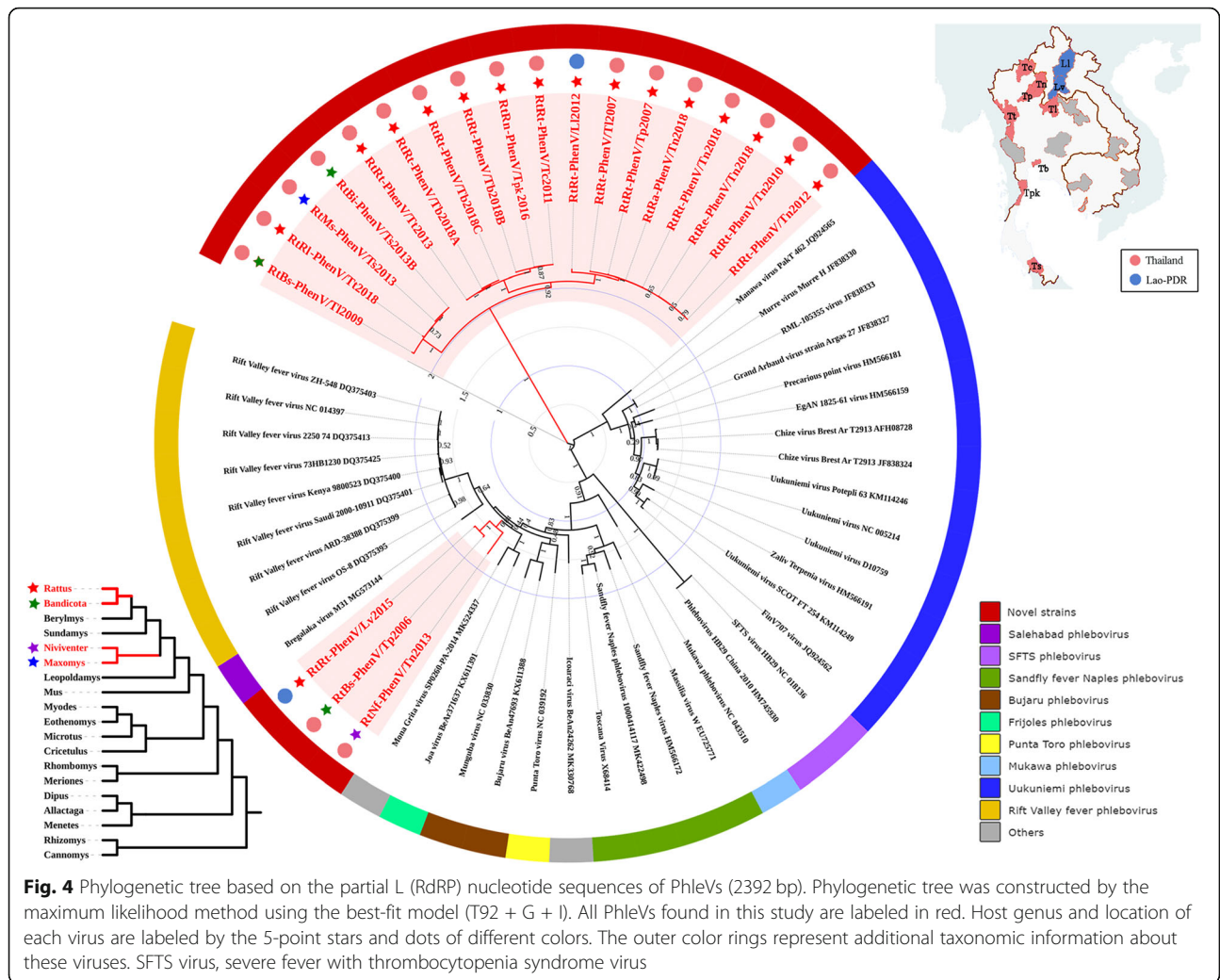
Fig. 3 Phylogenetic trees based on the complete M segment-encoded glycoprotein precursor (GPC), L segment-encoded RdRp (RdRp), and S segment-encoded nucleocapsid protein (N) amino acid sequences of HanVs. Phylogenetic trees were constructed by the maximum likelihood method using the best-fit models (LG + G for GPC protein, LG + G + I for RdRp protein, and GTR + G + I for N protein). All HanVs found in this study are labeled in red. Host genus and location of each virus are labeled by the 5-point stars and dots of different colors. The outer color rings represent additional taxonomic information about these viruses. DOBV, Dobrava-Belgrade virus

HanVs and Thailand virus cannot be resolved, due to the absence of complete L sequences from the Thailand viruses in GenBank, we propose that a single lineage of the species *Thailand orthohantavirus* [48] includes all of the HanV strains detected in Thailand from the M, S, and partial L-based alignments and our phylogenetic analysis. This suggested that these HanVs have circulated in diverse Thai provinces for several decades.

PhleVs

Similar to HanVs, viruses of the genus *Phlebovirus* are also linear segmented negative-stranded RNA viruses that belong to the family *Phenuiviridae* of the order *Bunyvirales* [49]. This genus contains many highly virulent viruses such as Rift Valley fever virus (RVFV), Toscana virus, and severe fever with thrombocytopenia syndrome virus. These viruses are arthropod-borne, naturally harbored by ruminant or camel reservoirs, and are transmitted by mosquitoes, sandflies, or ticks, and cause severe diseases in humans and animals [50–52]. After

mapping the sequencing reads, a total of 21 rodent PhleVs’ genome sequences were completely or partially confirmed in the lung samples from *B. savilei*, *M. surifer*, *Niviventer fulvescens*, and several diverse *Rattus* species from eight Thai provinces and two Laotian provinces (Table S4). Species of *Rattus* species were the main hosts of the PhleVs. Pairwise alignment revealed that the partially sequenced L segments of these viruses showed less than 71.2% nt identity with all other known PhleVs, which suggested that these rodent PhleVs represent novel species (Table S6). Phylogenetic analysis of the partial L nucleotide sequences revealed two distinct lineages of rodent PhleVs in the genus *Phlebovirus*, lineage 1 related to Uukuniemi PhleVs, and lineage 2 related to RVFV and Salehabad PhleVs (Fig. 4). The different PhleVs identified in species of the *Rattus* genus from different locations showed very close genetic relationships and can be further divided into two clades within lineage 1. However, some of the PhleVs from diverse rodent species shared high sequence identities and close genetic

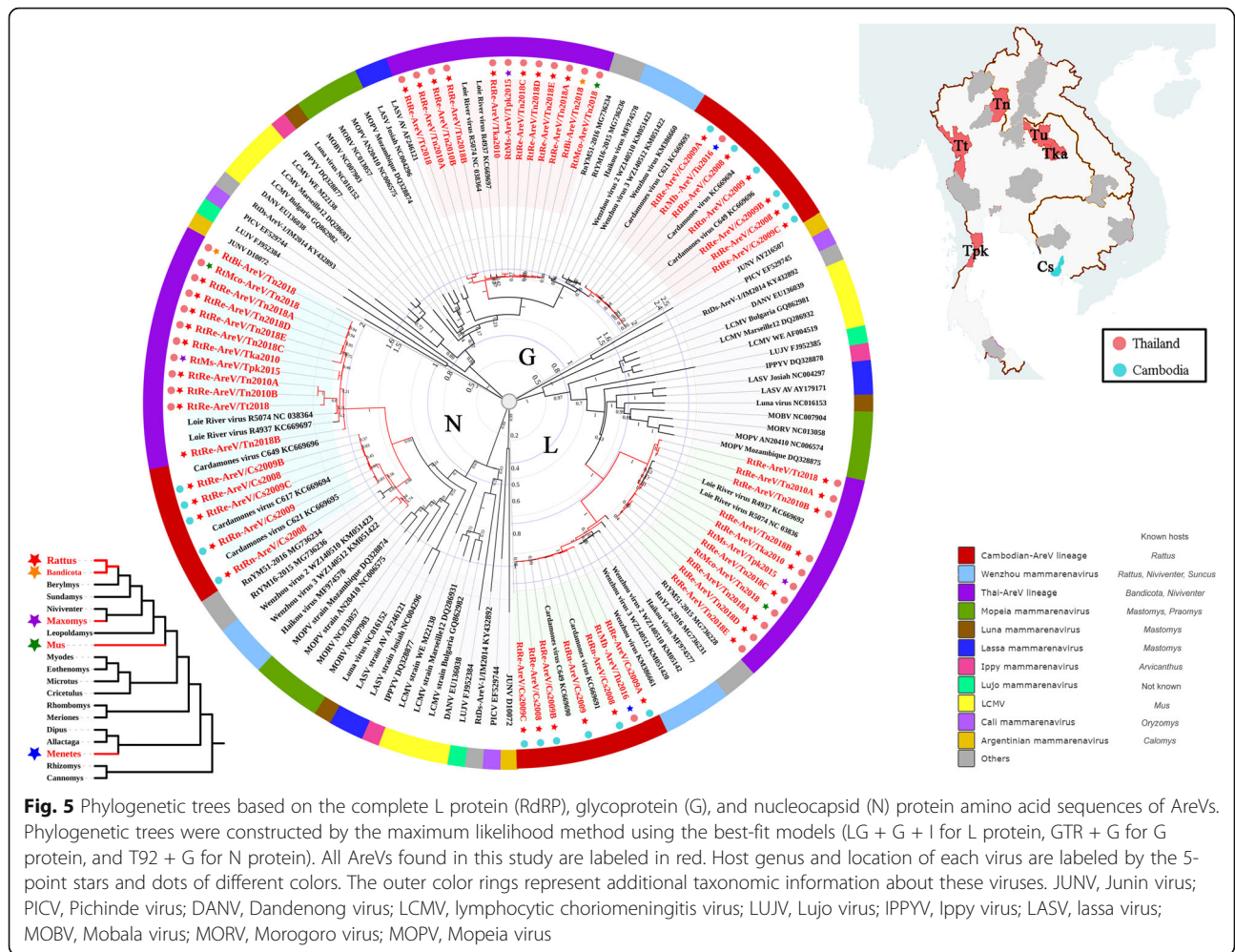


relationships. For example, RtRI-PhenV/Tt2018, RtMs-PhenV/Ts2013, and RtBs-PhenV/Tl2009 identified in three different rodent genera are closely related to each other within lineage 1, and as well, a close relationship was seen between RtBs-PhenV/Tp2006 and RtRt-PhenV/Lv2015 within lineage 2.

AreVs

Rodent AreVs of the genus *Mammarenavirus*, family *Arenaviridae*, can be divided into the Old-World and New-World complexes [16]. These viruses are a group of linear segmented negative-stranded RNA viruses with two genome segments (L and S). A total of nineteen Old-World AreVs strains were identified from lung samples from *B. indica*, *M. cookii*, *M. surifer*, *Menetes berdmorei*, and *Rattus* species from five Thai provinces and the Cambodian Sihanouk province (Table S4). Species of *Rattus* are the main hosts for these AreVs. Fourteen strains were characterized for genome sequences, and five strains were selected for sequencing of partial L and S. Pairwise

alignment of the complete genome sequences (including NCRs) revealed that seven strains shared high sequence similarity and were closely related to Cardamones virus found in Veal Renh, Cambodia, in 2009, with 96.1–100% nt identities, and 12 strains shared high sequence similarity and were closely related to the Loei River virus found in Loei province, Thailand, in 2008, with 88.2–95.1% nt identities [53] (Table S7). In accordance with sequence alignment results, phylogenetic analyses based on the RdRP (L), glycoprotein (G), and nucleocapsid (N) proteins suggested that these AreVs could be assigned into two different lineages related to AreVs reported in China (Fig. 5). We designated them as Thai-AreV lineage and Cambodian-AreV lineage. These results revealed that the Thai-AreVs have circulated in five Thai provinces for at least 9 years (2010–2018) and that the Cambodian-AreV circulated in the Sihanouk region of Cambodian between 2008 and 2009. The only exception to this was RtMb-AreV/Tu2016, which is a member of the Cambodian-AreV lineage but was found in Udon Thani province of Thailand in 2016.

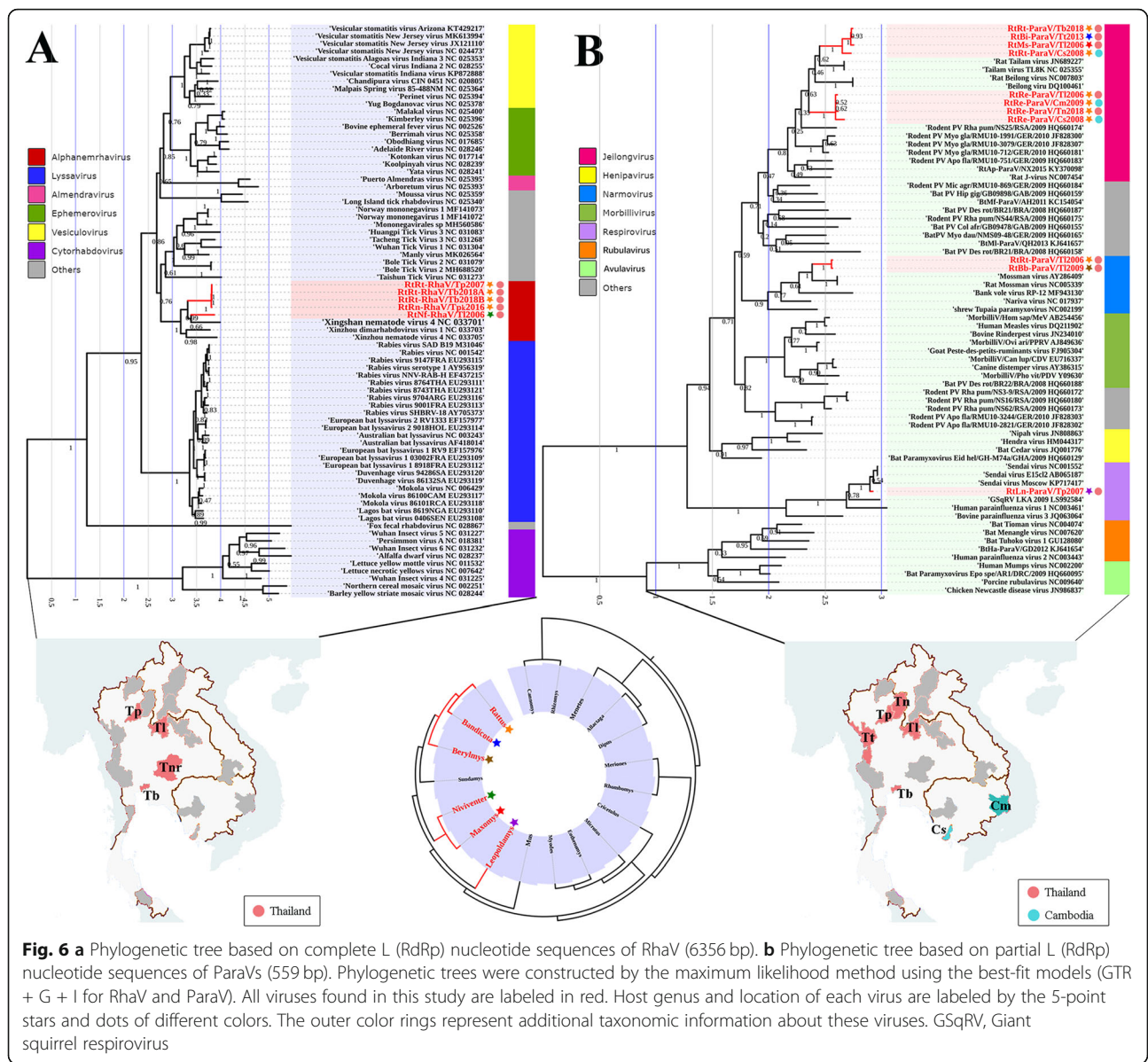


Rhabdoviruses (RhaVs)

RhaVs are a large group of linear negative-stranded RNA viruses of the family *Rhabdoviridae*, which currently has 20 genera [54, 55]. These viruses infect diverse vertebrates, invertebrates, and plants, and some of them can cause mild-to-severe diseases such as vesicular stomatitis virus and rabies virus [56, 57]. Here, we identified five RhaVs in *N. fulvescens* and *Rattus* species from four Thai provinces, and the complete genome sequences of two viruses were confirmed (Table S4). Unlike the previously reported rodent lyssavirus, mokola virus, and murine feces-associated rhabdovirus (MuFARV) [58, 59], these four RhaVs showed less than 66.5% nt homology with known *Rhabdoviridae* members (Table S8). The most closely related virus was the Xingshan nematode virus 4, a newly identified RhaV in *Spirurian nematodes* [23]. Phylogenetic analysis based on the complete L nucleotide sequences suggested that these novel rodent RhaVs clustered with the Xingshan nematode virus 4 of genus *Alphanemrhavirus* (Fig. 6a).

Paramyxoviruses (ParaVs)

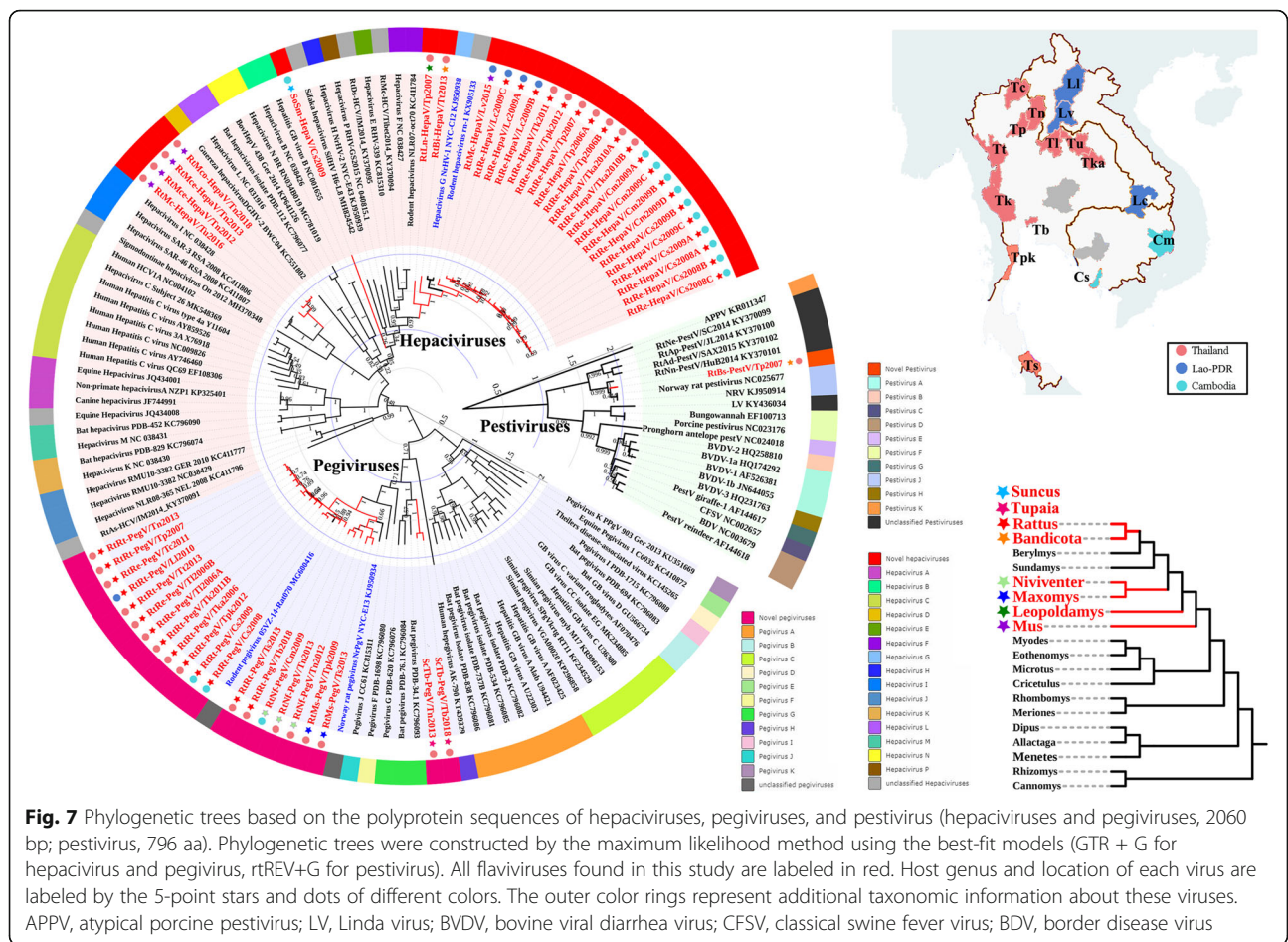
The family *Paramyxoviridae* is a group of enveloped viruses with negative-stranded RNA genomes that are responsible for many mild-to-severe human and animal diseases [35, 60–66]. Twelve rodent ParaVs’ sequences were identified in the lung samples from *B. indica*, *Berylmys bowersi*, *Leopoldamys neilli*, *M. surifer*, and diverse species of *Rattus* from five Thai provinces and two Cambodian provinces (Table S4). We obtained full-length sequences of ten virus strains. Of these, nine ParaVs were closely related to members of the genus *Jeilongvirus* with high sequence similarity (75.3–77% nt identities for L ORF), two were Mossman virus related (75.8–76.5% nt identities for L ORF), and one was Sendai virus related (91.9% nt identities for partial sequenced L) (Table S9). Phylogenetic analysis based on the partial L nucleotide sequences revealed that these rodent ParaVs were assigned into the genera *Narmovirus*, *Jeilongvirus*, and *Respirovirus* (Fig. 6b).



Characteristics of positive-stranded RNA viruses
Hepaciviruses, pegiviruses, and pestiviruses

The genera *Hepacivirus*, *Pegivirus*, and *Pestivirus* are within the family *Flaviviridae* and are positive, single-stranded RNA viruses. These viruses infect a variety of mammalian hosts, including primates, bats, horses, and rodents [67–69]. The hepatitis C virus of the genus *Hepacivirus* is an important causative agent of hepatitis and hepatocellular carcinoma in humans [70], and the two classic types of pestiviruses, bovine viral diarrhea virus and classical swine fever virus, are important causative agents of mild-to-severe diseases in cattle and pigs [71, 72]. Here, we found a total of 51 viral members of the family *Flaviviridae* within the diverse rodent and shrew species lung samples from almost all sampling

sites across Thailand, Lao PDR, and Cambodia (Table S4). Thirty-four strains underwent genome sequencing and seventeen strains were selected for sequencing of partial polyproteins. Pairwise alignment of the complete or partial genome sequences suggested that twenty-eight of them were hepaciviruses, with 41.4–100% nt identities with each other, twenty-two were pegiviruses, with 42.1–96.1% nt identities with each other, and one was pestivirus, with 75.2% nt identity to a known rodent pestivirus (Table S10). Phylogenetic analysis based on the partial polyproteins revealed that the 51 novel viruses could be assigned to distinct novel lineages within the genera *Hepacivirus*, *Pegivirus*, and *Pestivirus* (Fig. 7). Several host-specific lineages, including a *Rattus exulans*-related lineage for *Hepacivirus*, and *N. fulvescens*-

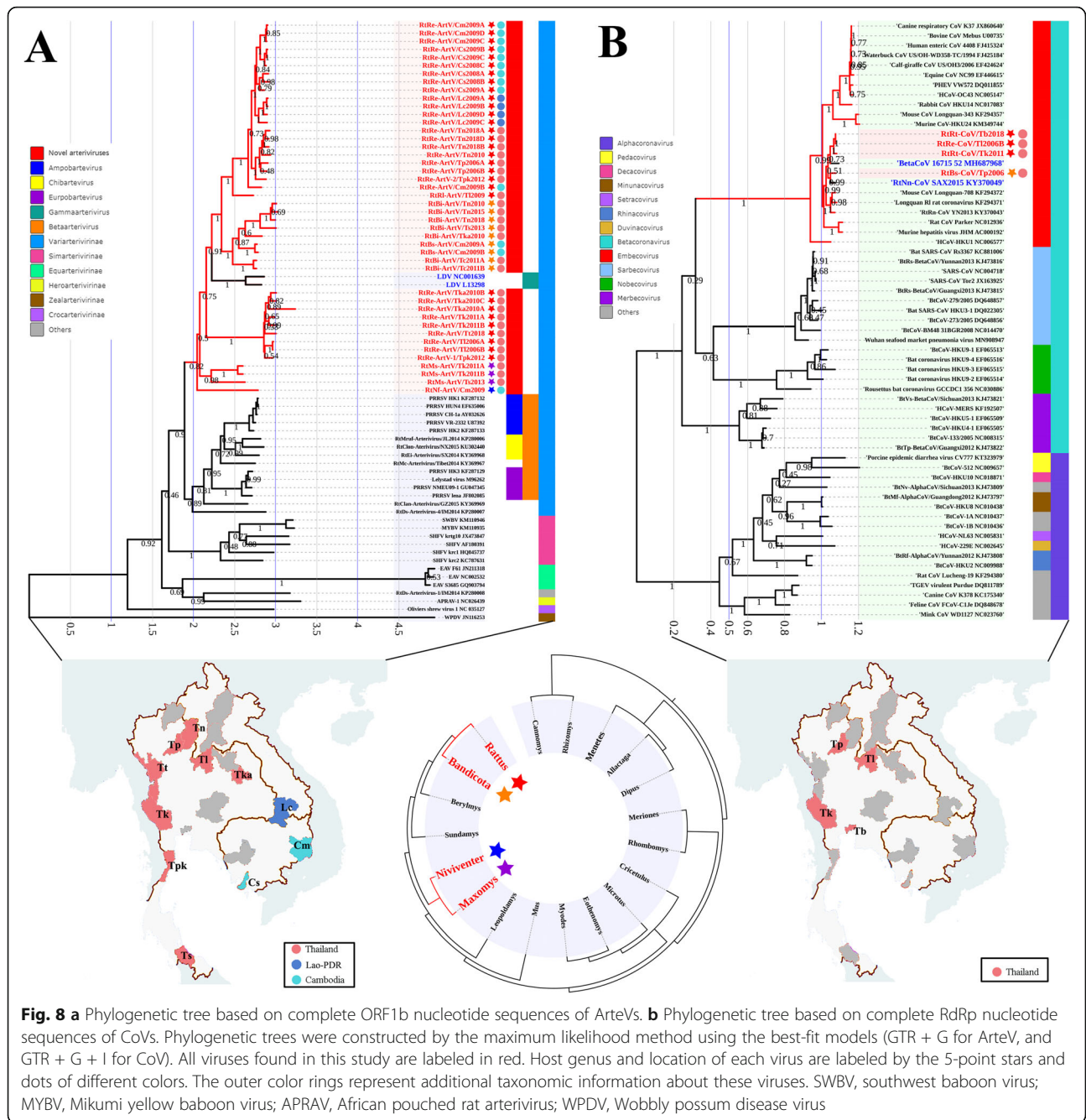


related and *Rattus*-related lineages for *Pegivirus*, were detected that suggested that the phylogenies of most of these viruses were strictly congruent with the relationships of their rodent or insectivore hosts. For the first time, we reported hepacivirus and pegivirus (SoSm-HepaV/Cs2009, ScTb -PegV/Tb2018, and ScTb-PegV/Tn2013) for insectivores (shrew and tree shrew). The virus SoSm-HepaV/Cs2009 represented a separate hepacivirus clade with less than 44.4% nt identity with any known virus. ScTb -PegV/Tb2018 and ScTb-PegV/Tn2013 represent a separate pegivirus clade with less than 57.6% nt identity with known viruses. These viruses, together with previously reported bat viruses, formed the main evolutionary frames for these two genera.

ArteVs

ArteVs, of the family *Arteriviridae*, are a group of enveloped viruses with positive single-stranded RNA genomes that are responsible for a variety of mild to severe diseases in horses, simians, and swine, such as equine arteritis virus (EAV), simian hemorrhagic fever virus (SHFV), and porcine reproductive and respiratory syndrome virus (PRRSV) [73–76]. A total of 49 ArteVs’ genomic or

partial genomic sequences were identified in diverse rodent tissues from almost all sampling sites throughout Thailand, Lao PDR, and Cambodia (Table S4). Pairwise alignment of ORF1b revealed that these viruses shared 55.7–98.6% nt identities with each other, and less than 66% nt identity with known ArteVs (Table S11). These sequences are most similar to previously reported *Betaarterivirus* and *Gammaarterivirus* members of the genera, PRRSVs, lactate dehydrogenase-elevating virus (LDV), and unclassified rodent ArteVs. However, unlike the results of our previous study of rodent pharyngeal and anal samples from China that showed a diverse phylogenetic scattering of ArteVs throughout the family *Arteriviridae*, here, phylogenetic analysis based on ORF1b and ORF5 revealed that all ArteVs found here in lung tissues clustered with each other as a separate clade within the subfamily *Variarterivirinae* with different host-specific lineages (such as *Rattus*-related, *Maxomys*-related, and *Bandicota*-related lineages) (Fig. 8a). These lineages represented distinct viral classifications that differ from the previously identified genera *Betaarterivirus* and *Gammaarterivirus* of the subfamily *Variarterivirinae*.



CoVs

CoV is a group of enveloped viruses with a large positive single-stranded RNA genome within the subfamily *Coronavirinae* and include viruses that result in human diseases such as colds, severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and COVID-19 [1, 3, 77]. The subfamily *Coronavirinae* is divided into four recognized genera, *Alphacoronavirus*, *Betacoronavirus*, *Deltacoronavirus*, and *Gammacoronavirus* [78–81]. Previously, we had found a large number of rodent CoVs in the diverse rodent species in China that were assigned

into two separate lineages within *Alphacoronavirus* and *Betacoronavirus* [29]. Here, however, we only found nine CoVs within lung samples from species of *Bandicota* and *Rattus* from five Thai provinces and the Laoian Champasak province, and two of them were identified for genome sequencing (Table S4). Sequence similarity and phylogenetic analysis of RdRp revealed that all of these CoVs could be classified within *Embecovirus* under the genus *Betacoronavirus*, with nt sequence identities between 93.7 and 100% (Fig. 8b and Table S12). Despite our large sample size, alpha-CoV was not found in our samples.

Hepatitis E viruses (HEVs)

HEVs of the family *Hepeviridae* are a group of small, non-enveloped, positive single-stranded RNA viruses. Members of the species *Orthohepevirus A* are one of the most common causative agents of hepatitis in humans, and rodent-borne *Orthohepevirus C* was recently reported to be zoonotic and cause persistent hepatitis in humans [21, 27, 29, 42, 82]. The partial genome sequences of four HEVs were identified in *M. surifer* of Thai Loei province, *R. losea* of Laotian Vientiane province, and *R. exulans* of Cambodian Sihanouk province (Table S4). All of these viruses were closely related to previously reported rodent HEV, strains Vietnam-105, and human patient HEV strain LCK-3110, with 77.7–80.7% nt identities in ORF1, but less than 58.9% nt identity in ORF1 compared to HEVs from other hosts (Table S13). Phylogenetic analysis based on ORF1 assigned these HEVs into the species *Orthohepevirus C*, and they were closely related to the Vietnam-105 and LCK-3110 lineage, which was suspected to be a causative agent of human persistent hepatitis (Fig. 9a).

Picornaviruses (PicoVs)

Members of the family *Picornaviridae* are small, non-enveloped, positive single-stranded RNA viruses. Diverse PicoVs cause mucocutaneous, encephalic, cardiac, hepatic, neurological, and respiratory diseases in a wide variety of vertebrate hosts [83]. In our samples, three PicoVs were identified in *B. indica* of Thai Chiang Rai province, *R. tanezumi* of Laotian Luang Prabang province, and *R. tanezumi* of Cambodian Sihanouk province (Table S4). Sequence similarity and phylogenetic analysis of complete RdRp indicated that all of these PicoVs were closely related to known rodent PicoVs with nt sequence identities between 42 and 88.4% (Table S14 and Figure S10).

Astrovirus (AstroV)

AstroVs comprise positive single-stranded RNA viruses and are members of the genus *Mamastrovirus* within the family *Astroviridae* and infect many mammals and cause gastroenteritis [84]. We identified only one AstroV in *B. savilei* of the Thai Loei province. This virus shared 82.05 % nt identity in complete RdRp with an AstroV previously reported in China (Table S16 and Figure S11).

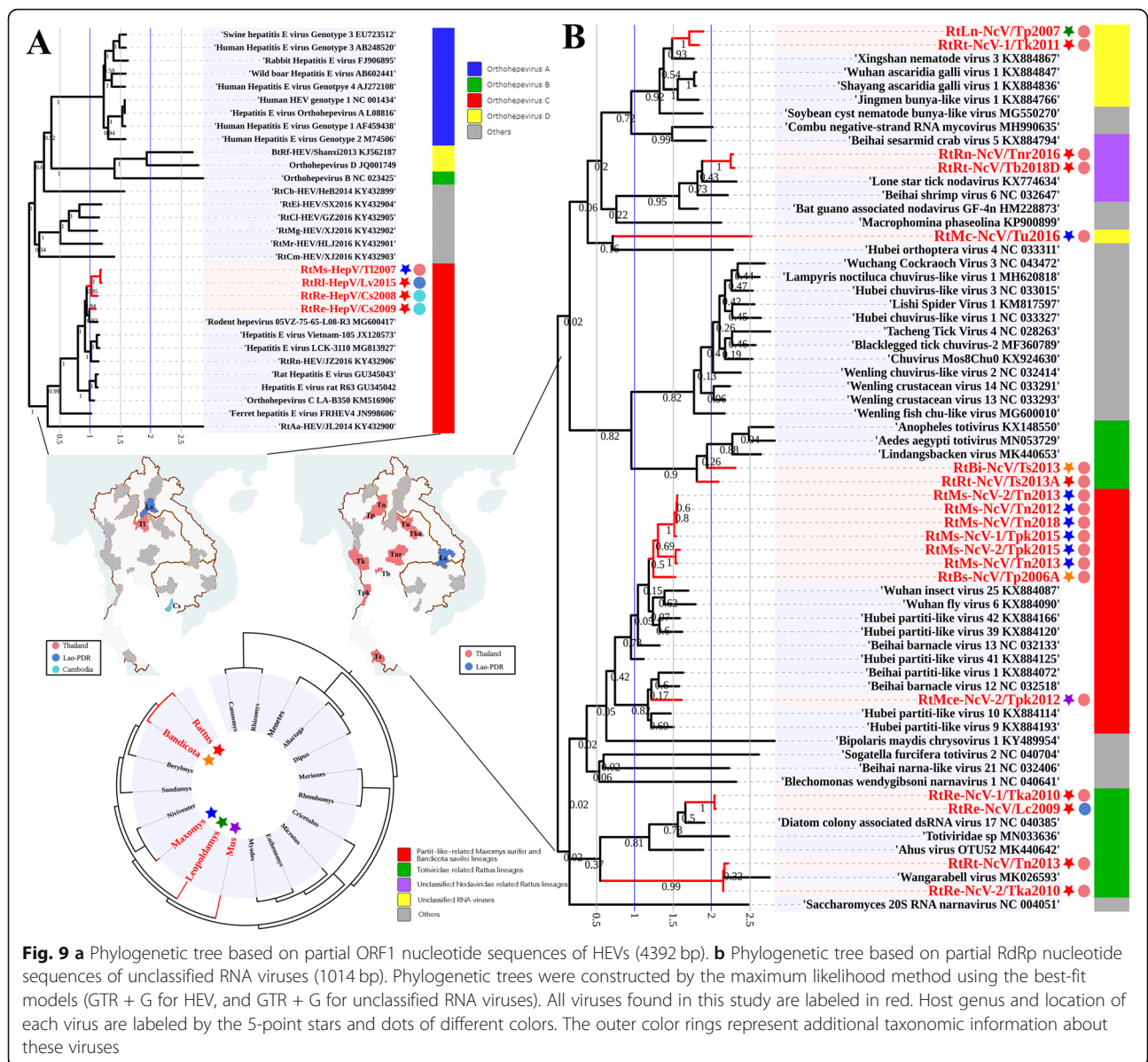
Characteristics of unclassified RNA viruses

Recent viral surveillance studies in invertebrates, amphibians, reptiles, and fishes have revealed a new view of the RNA virosphere that is more diverse than the current taxonomy [23, 26]. Here, we identified a total of 28 unclassified RNA viruses that were found in the lungs of diverse rodent species from different Thai and Laotian provinces (Table S4). After ORF annotation for these

complete or partial sequenced viruses, pairwise alignment revealed that the code regions of partial RdRp for these viruses showed 22.7–99.9% nt identities with each other, and less than 66.7% nt identity with other known RNA viruses (Table S16), suggesting that these newly identified viruses are highly diverse and distinct from known and undefined viral families. To further examine the evolutionary relationships between these viruses, phylogenetic trees were constructed based on the partial RdRp proteins of viral genomes from all related known families, genera, and unclassified taxa of invertebrates, amphibians, reptiles, and fish. The unclassified RNA viruses found here formed at least 10 distinct lineages (Fig. 9b). Most of the viruses tended to form different lineages such as the partiti-like-related viruses found in the *M. surifer* and *B. savilei* lineages, the *Rhabdoviridae*-related viruses found in the *Rattus* lineage, and the *Totiviridae*-related viruses found in the *Rattus* lineage. These data suggested that RNA viruses in rodents occupy a broader range of phylogenetic diversity and are similar to the RNA viral spectrum observed in invertebrates.

Discussion

As many animal-borne pathogens are transmitted through the respiratory tract, oral cavity, or enteric canal, previous studies have focused on systemically deciphering the viral communities found in respiratory or fecal samples (which are easy to get through non-invasive sampling) from natural animal hosts [25, 29–31, 58]. But identification of novel viruses found in open environmental samples raises question of whether these viruses are harbored by the animal host and cause actual infections, or they are just viruses that infect insect, fungi, and plants (and other environmental sources) associated with the life habits of this animal host. Although our previous reports provided evolutionary evidence for the zoonotic origin of viruses such as henipavirus, HanV, AreV, SARS-CoV, and MERS-CoV [10, 32–36, 85, 86], the presence of numerous vertebrate-irrelevant viral sequences in oral and rectal virome data complicated the definition of the host-virus relationships. Hence, organ- or tissue-based virome analysis, designated as “core virome” analysis, is also needed to accurately describe the viral infection status in natural reservoirs. This effort should give us a better understanding of which threats are directly coming from the infections of natural hosts and which threats just originate from the environment. Recently, a metatranscriptomic-based virome analysis of the internal organs of amphibians, reptiles, and fish conducted by Shi et al identified extremely distant vertebrate-associated viruses and defined the long evolutionary history of vertebrate RNA viruses [26]. However, to date, large-scale core virome surveillance has not been performed to



understand the ecological diversity of the viral community in the internal organs or tissues of rodents, the largest mammalian group and one of the most important natural reservoirs of many zoonotic viruses with severe impacts on human populations [7].

For the first time, this study details the core viromes residing in the lungs of rodents and other small insectivores throughout the three main countries of Mainland Southeast Asia. Muridae and Spalacidae rodents represent a large part of the rodent diversity in Mainland Southeast Asia and were the main focus of this study. With 10 genera and 24 species from the families Muridae and Spalacidae represented in our samples, our study includes most rodent species living in human-modified and natural habitats in the region [87–89].

Sampling of species belonging to the families Sciuridae, Erinaceidae, Soricidae, and Tupaiidae was opportunistic and represent only a small fraction of the total taxonomic diversity within these families. The lung was used for this meta-transcriptomic analysis as the lung is the organ where many viral infections happen in animals, as well as the place where viral spillover initiates between species [7, 8]. By greatly reducing the interference by irrelevant environmental microorganism, the actual abundances of RNA viruses and the transcriptional levels of DNA viruses in the lung organs was revealed. We identified new viral families, genera and species with the characterization of 216 diverse unclassified RNA viruses, HanVs, PhleVs, AreVs, RhaVs, ParaVs, hepaciviruses, pegiviruses, ArteVs, CoVs, HEVs, and PicoVs, extending

our knowledge of the viral biodiversity of rodents in the Indochinese Peninsula, albeit the roles of some of these viruses in disease remain unclear. The characterization of a number of the unclassified RNA viruses that are different from all known viral families suggests that our knowledge of rodent viral richness is still very limited and that there is an important number of “missing viruses” needing further investigation.

The core virome harbored by the rodent internal organs is quite different from the virome found in nasopharyngeal or rectal swabs previously [24, 29, 58, 85, 86]. The widely distributed RNA viruses, such as various PheVs, hepaciviruses, pegiviruses, and ArteVs detected in this study, were not or only occasionally found in the oral and anal samples previously described. In contrast, RNA viruses that mainly infect their hosts through the upper respiratory tract or enteric canal, such as pestivirus, PicoV, HEV, AstroV, sapovirus, norovirus, rotavirus, morbillivirus, and picobirnavirus, were seldomly or never detected in the internal organ samples screened in this study [27, 28, 66]. This difference may be due to the different tissue tropisms of these viruses: viruses such as PheVs, hepaciviruses, and pegiviruses are blood- or body fluid-borne pathogens while viruses such as pestivirus, PicoV, HEV, and morbillivirus are fecal-oral or respiratory-transmitted pathogens. The blood- or body fluid-borne viruses could be detected in various internal organs with different viral loads because of their different infection cycles, viremia, or inevitable cross-contamination of tissue and blood during sampling, while the fecal-oral or respiratory-transmitted viruses could also be occasionally detected in internal organs because of their diverse replication features or viremia. Compared to the widely distributed beta-CoVs identified in rodent oral and rectal samples, only a few beta-CoVs were detected in the lung specimens, suggesting that beta-CoV infection mainly occurs in the upper respiratory tract of rodents. Furthermore, the absence of rodent alpha-CoVs in any of the lung samples suggests that this CoV group may not be responsible for lower respiratory tract infection in these rodent species. Similarly, all lung ArteVs show a separate evolutionary footprint differing from the previously reported PRRSV-related rodent ArteVs identified in the oral and anal specimens, also suggesting two different types of ArteV infection in rodents.

By showing that species of small mammals that live in the Indochinese Peninsula carry greater numbers of viruses than previously thought, we extend the known host and geographical ranges of many viral families. Most of the viruses identified in this study show no strict geographical or host restrictions. Here, we first report some viruses in specific hosts, such as shrew hepacivirus and pegivirus, tree shrew pegivirus, and a large

group of novel rodent PheVs, which indicates that these species may harbor a greater diversity of both known and unknown viruses, some of which may have zoonotic potential. A large number of viruses were identified in rodents belonging to the genera *Rattus* and *Bandicota*. This suggests that these two genera might act as major reservoirs for diverse mammalian viruses throughout the Mainland of Southeast Asia. Species members of these two genera are characterized by their synanthropic behavior and have been implicated in the emergence and spread of infectious diseases of public health importance, such as plague, murine typhus, scrub typhus, and leptospirosis, in addition to diseases caused by viruses [90].

HanV and AreV of rodent origin are important causative agents of human diseases. Almost all previously reported phylogroup III HanVs and Old-World AreVs are Murinae borne, with the exceptions of Jerboa hantavirus and Jerboa arenavirus we identified recently [91]. The Thailand virus from *B. indica* is the only reported hantavirus species carried by rodents in Thailand [46, 47]. Although this virus is suspected to be the causative agent of HFRS, only a few fragmentary pieces of sequence information for this virus are available so far. We identified Thailand virus-related HanVs in *B. indica* and two *Rattus* species in five Thai provinces between 2006 and 2018, indicating that this viral species of phylogroup III HanV is continuously circulating in different Murinae species in Thailand. The first confirmed genomic sequences of this virus obtained in this study enable us to further understand the phylogenetic relationship between the Thailand virus and other HanVs. Similarly, the identification of Thai- and Cambodian-AreV lineages here reveals the continual existence of AreVs in Indochinese Peninsula, and the relatively strict geographical restrictions of AreVs in this region. However, we found a Cambodian-AreV, RtMb-AreV/Tu2016, in *M. berdmorei* of the Thai Udon Thani province, which is geographically distant from the Cambodian Sihanouk province, suggesting that this virus was disseminated recently to additional host species over a wider geographical range. And furthermore, the identification of RtMb-AreV/Tu2016 in *Menetes* species of the family Sciuridae indicate again that Old-World AreV are able to infect more mammalian hosts than previously thought and that Murinae species are not their only hosts. Since this AreV is confirmed to be involved in human infections in Southeastern Asia [53], the potential for disease emergence from rodent-origin AreV in this region should not be underestimated.

Rodents can serve as potential sources for vector-borne disease outbreaks and many arboviruses such as CCHFV of the family *Nairoviridae*, TBEV and OHFV of the family *Flaviviridae*, and Venezuelan equine encephalitis virus of the family *Togaviridae* are harbored by

rodents and are transmitted to humans and other animals via vectors such as the tick, mosquito, or sand-fly [8]. Here, we found diverse rodent viruses in the lung that are related to arthropod viruses, including the families *Phenuiviridae*, *Rhabdoviridae*, *Totiviridae*, and *Chuviridae*, and a large group of unclassified RNA viruses. These viruses show both close and distant phylogenetic relationships with known arthropod viruses, suggesting that rodents may act as one of the mammal reservoirs for these arthropod-related viruses in the environment, and play roles in the arthropod-cycle and the conservation of such viruses. Since Mainland Southeast Asia is a hotspot for many arbovirus-related tropical diseases [42, 43], the threat from rodent reservoirs should not be ignored.

Conclusions

These findings, combined with our previous viral surveys from rodents, bats, and mosquitoes, and the online viral databases for these animals (DBatVir, DRodVir, and DMosVir, <http://www.mgc.ac.cn/>) [29–31, 92–94], greatly increase our knowledge of the viral community in wildlife and arthropod vectors in densely populated countries of East and Southeast Asia. Continued efforts in viral surveillance among wildlife hosts will reveal greater diversity of viral lineages, as recently hypothesized globally [37] and provide evidence to mitigate the risk of potential zoonotic disease emergence, as well as build the regional and global capacity for effective prevention and response to EIDs [22].

Methods

Animal samples

Rodents were trapped once or twice (in the wet season between June and October, and in the dry season between December and March) over a 12-year period between 2006 to 2018. Trapping sessions were conducted once or twice annually for each locality, with the trapping conducted over a 4-night period, with 30 lines of 10 traps placed in three different habitats, namely, (1) forest and mature plantations, (2) non-flooded lands or fields (shrubby waste land, young plantations, orchards), and (3) rain-fed lowland paddy rice fields (cultivated floodplain) (with the exception of Bangkok Metropolitan recreational parks). This yielded a total of 1200 night-traps per trapping session. Locally made live cage traps were used. At each locality, the 30 lines were placed within a 10-km² area. Each line was separated from others by at least 1 km (different habitats) or 3 km (same habitat). Villages and isolated houses, which correspond to a fourth habitat category, human settlement, were also sampled opportunistically using cage traps distributed to residents [43]. The locations of sampling sites were recorded by place name and GPS coordinates with

latitude and longitude. The identity of the captured species was initially determined by morphology by experienced field biologists and subsequently confirmed by barcoding of mitochondrial cytochrome b, mitochondrial cytochrome oxidase I, or 16sRNA sequences [40, 95]. Complete data on the animals used as reference specimens for the barcoding assignment is available on the “Barcoding Tool/RodentSEA” section of the CERoPath project website (www.ceropath.org). Captured animals were humanely euthanized and lung organ was collected. Each lung sample was placed in a cryogenic vial and temporarily stored in liquid nitrogen before being transported to the laboratory and stored at –80 °C.

RNA library construction and next-generation sequencing

Lung samples were washed with Hank’s balanced salt solution twice and then homogenized using a TissueLyser (Qiagen) in the presence of lysis buffer (Buffer RLT Plus) and stainless-steel beads (5 mm mean diameter). Total RNA was extracted from each sample using a RNeasy Plus Mini Kit (Qiagen) and quantified by a Qubit Fluorometer (Thermo Fisher Scientific). For library construction, RNAs from the same or related rodent species and collected at the same or related sites were pooled in equal quantity and then checked for integrity by an Agilent 2100 Bioanalyzer (Agilent Technologies), rRNA was removed from each pool using NEBNext rRNA Depletion Kit (NEB), and the final libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB). For sequencing, each pool was labeled with a specific index using NEBNext Multiplex Oligos for Illumina (NEB) and then processed for next-generation sequencing (NGS) using an Illumina HiSeq2500 sequencer (Illumina) using the 125 bp paired-end method. Raw sequence reads were filtered using previously described criteria to obtain valid sequences, and reads with no call sites, sequencing adaptor reads, duplicate reads, and low-complexity reads were removed [31, 96].

Taxonomic assignment and viral RNA quantification

Sequence similarity-based taxonomic assignment was conducted as previously described [31]. Briefly, sequence reads were evaluated for viral origin by conducting alignments with the NCBI protein database (NR), using BLASTx (-e 1e-5 -F T). To detect potential viruses with remote similarities, reads with no hits in the NCBI NR database were further assembled using the software Trinity (v.2.4.0), and the resulting contigs were aligned again to NCBI NR to identify any viral-like sequences. The taxonomic identity of the aligned reads and contigs with the best BLAST scores (E-value < 10⁻⁵) were then parsed and exported using a MEGAN 6 - MetaGenome Analyzer. To estimate the abundance of the viral RNAs

in each pool, reads were assigned to each viral taxon at the level family and were normalized by the total number of sequencing reads (or using previously described VTMK-index [96]) and viral genome size. While maintaining a low false-positive rate, any viral taxon supported by less than three unique reads was excluded from the quantitative analysis.

Calculation of viral prevalence and the full-genome sequencing of each virus in the positive samples

Sequencing reads and assembled contigs for the same family or genus of each virus were extracted from MEGAN 6 - MetaGenome Analyzer. The accurate locations of these reads and contigs and the relative distances between them were determined using the alignment results from BLASTx or BLASTn. A draft genome, which contained the single nucleotide polymorphisms from each virus, was obtained. Specific primers were designed from the located sequence reads and contigs to screen for each virus in the individual samples from each mammalian species. Representative positive samples for each virus were selected for genome sequencing as a viral quasi-species. Fragments between the reads and contigs were amplified with overlapping, nested specific primers, and then sequenced. The majority of the viral genomes were obtained by merging sequence reads, assembled contigs, and amplified fragments using the SeqMan program (Lasergene). Any remaining gaps and the termini of genomes were determined using genome walking, inverse PCR, and 5'- and 3'-rapid amplification of cDNA ends.

Genome annotation

The ORFs in the complete and partial genomes of the sequenced viruses were predicted with the NCBI ORF Finder tool (<https://www.ncbi.nlm.nih.gov/orffinder/>). Changes in the nucleotide sequences of the genomes and the amino acid sequences of ORFs were deduced by comparing the sequences with those of related viruses available in GenBank. Conserved protein families and domains were predicted using InterProScan 5 (available at: <http://www.ebi.ac.uk/services/proteins>). Pairwise and multiple sequence alignments were performed by using Clustal Omega, Needle (<http://www.ebi.ac.uk/Tools/>), MegAlign (Lasergene), and T-coffee (<http://tcoffee.org/cat/apps/tcoffee/index.html>) with manual curation.

Phylogenetic analysis

All related reference viral sequences were downloaded from GenBank. We used MEGA6.0 to align the nucleotide and deduced amino acid sequences, using the MUSCLE program with default parameters [97]. The best substitution model was then evaluated with the Model Selection package. We constructed a maximum

likelihood phylogenetic tree in MEGA6.0, using the appropriate substitution model, with 1000 bootstrap replicates. Phylogenetic trees were displayed, manipulated, and annotated using The Interactive Tree Of Life (<https://itol.embl.de>) [98].

Data accessibility

All genome sequences were submitted to GenBank. Accession numbers for the viruses are MT085080 to MT085341 (Table S4). Illumina HiSeq2500 sequence data was deposited into NCBI BioProject under accession number PRJNA605875.

Description of Supplementary Information (SI)

Supplementary Tables and Figures are available with the online version of this paper.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-020-00965-z>.

Additional file 1. Table S1. Samples of the 30 animal species used in this study and the countries (provinces) and dates of collection. **Table S2.** Overview of virus-associated reads. **Table S3.** Overview of sequence reads from mammal related viral families and unclassified RNA viruses. **Table S4.** Origin and accession number of viruses identified in this study. **Table S5.** Nucleotide sequence identity of novel rodent hantaviruses and known hantaviruses within the L, M, and S segment. **Table S6.** Nucleotide sequence identity of novel rodent PhleVs and known PhleVs within the L open reading frame. **Table S7.** Nucleotide sequence identity of novel rodent arenaviruses and known arenaviruses (AreV) in the L-, G- and N-encoding regions. **Table S8.** Nucleotide sequence identity of novel rodent rhabdoviruses and known rhabdoviruses (RhaV) in the L protein-encoding regions. **Table S9.** Nucleotide sequence identity of novel rodent paramyxoviruses and known paramyxoviruses (ParaV) in the L protein-encoding regions. **Table S10.** Nucleotide sequence identity of novel rodent and known hepa-, pegi- and pestviruses. **Table S11.** ORF1b nucleotide sequence identity of novel rodent and known arteriviruses. **Table S12.** Nucleotide sequence identity of novel rodent and known coronaviruses (CoVs) in the RdRP-encoding region. **Table S13.** Nucleotide sequence identity of novel rodent and known hepeviruses (HEVs) in the ORF1 region. **Table S14.** Nucleotide sequence identity of novel rodent and known picornaviruses (PicoVs) in the RdRP-encoding region. **Table S15.** Nucleotide sequence identity of novel rodent and known astroviruses (AstroVs) in the RdRP-encoding region. **Table S16.** Nucleotide sequence identity of novel rodent unclassified RNA viruses and known viruses in the RdRP-encoding region.

Additional file 2. Figure S10. Phylogenetic tree based on the complete amino acid sequence of the RNA-dependent RNA polymerase of picornaviruses. The viruses found in this study are labeled in red font. **Figure S11.** Phylogenetic tree based on the complete amino acid sequence of the RNA-dependent RNA polymerase of astroviruses (AstroVs). The virus found in this study is labelled in red font.

Acknowledgements

Not applicable.

Authors' contributions

Conceived and designed the experiments: ZW, PD, QJ. Performed the experiments: ZW, YH, JD, LS, HS, LL, JD. Analyzed the data: ZW, YH, BL, MC, SZ, JY. Contributed reagents/materials/analysis tools: ZW, HL, GZ, AL, AK, SJ, KC, PB, VD, JJ, XX, HZ, FY, SM, PD, JW. Wrote the paper: ZW, HL, AL, DM.I, SM. The authors read and approved the final manuscript.

Funding

This work was supported by the National S&T Major Project “China Mega-Project for Infectious Disease” (Grant No. 2018ZX10101001 and 2018ZX10711001), the National Natural Science Foundation of China (Grant No. 81772228), the CAMS Innovation Fund for Medical Sciences (Grant Nos. 2020-I2M-CoV19-010, 2016-I2M-1-014, and 2017-I2M-B&R-12), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2019PT310029), and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (Award No. R01AI110964). All work conducted by EcoHealth Alliance staff after April 24th 2020 was also supported by generous funding from The Samuel Freeman Charitable Trust, Pamela Thye, The Wallace Fund, and an anonymous Donor c/o Schwab Charitable. Field work studies were supported by ANR 07 BDIV 012 CERoPath project (Community ecology of rodents and their pathogens in a changing environment in Southeast Asia), ANR 11 CPEL 002 Biodiv-HealthSEAproject (Local impacts and perceptions of global changes: biodiversity, health, and zoonoses in Southeast Asia), and ANR-17-CE35-0003-02 FutureHealthSEA (Predictive scenarios of health in Southeast Asia).

Availability of data and materials

Datasets generated and analyzed during the current study are available in this published article (and its supplementary information files) and the NCBI sequence reads archive (SRA) under accession number PRJNA605875.

Ethics approval and consent to participate

Animals were treated according to the guidelines of the Regulations for the Administration of Laboratory Animals (Decree No. 2 of the State Science and Technology Commission of the People’s Republic of China, 1988), the American Society of Mammalogists, and the European Union legislation ((Directive 86/609/EEC). Sampling procedures were approved by the Ethics Committee of the Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College (Approval number: IPB EC20100415). Each trapping campaign was approved by national, regional, and local health authorities. Approval notices for trapping and investigation of rodents were provided by the Ministry of Health Council of Medical Sciences, National Ethics Committee for Health Research (NHCHR) Lao PDR, number 51/NECHR, and by the Ethical Committee of Mahidol University, Bangkok, Thailand, number 0517.1116/661 [43]. None of the species are considered to be endangered and none are included on the CITES list or the Red List (IUCN).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Author details

¹NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China. ²Key Laboratory of Respiratory Disease Pathogenomics, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China. ³EcoHealth Alliance, New York, NY, USA. ⁴Wildlife Conservation Society, Viet Nam Country Program, Ha Noi, Vietnam. ⁵Wildlife Conservation Society, Health Program, Bronx, NY, USA. ⁶Faculty of Veterinary Technology, Kasetsart University, Bangkok, Thailand. ⁷Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. ⁸GlaxoSmithKline Vaccines, Singapore City, Singapore. ⁹Virology Unit, Institut Pasteur in Cambodia, Phnom Penh, Cambodia. ¹⁰Yunnan Institute of Parasitic Diseases, Pu’er, PR China. ¹¹Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada.

Received: 16 September 2020 Accepted: 6 December 2020

Published online: 21 January 2021

References

- Ren LL, Wang YM, Wu ZQ, Xiang ZC, Guo L, Xu T, Jiang YZ, Xiong Y, Li YJ, Li XW, et al. Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. *Chin Med J (Engl)*. 2020;133:1015–24.

2. Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, Delicat A, Paweska JT, Gonzalez JP, Swanepoel R. Fruit bats as reservoirs of Ebola virus. *Nature*. 2005;438:575–6.
3. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270–3.
4. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579:265–9.
5. Blanga-Kanfi S, Miranda H, Penn O, Pupko T, DeBry RW, Huchon D. Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evol Biol*. 2009;9:71.
6. Wilson DE, Reeder DM. *Mammal species of the world: a taxonomic and geographic reference*. 3rd ed. Baltimore: Johns Hopkins University Press; 2005.
7. Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. *Proc Natl Acad Sci U S A*. 2015;112:7039–44.
8. Meerburg BG, Singleton GR, Kijlstra A. Rodent-borne diseases and their risks for public health. *Crit Rev Microbiol*. 2009;35:221–70.
9. Charrel RN, de Lamballerie X. Zoonotic aspects of arenavirus infections. *Vet Microbiol*. 2010;140:213–20.
10. Wu Z, Du J, Lu L, Yang L, Dong J, Sun L, Zhu Y, Liu Q, Jin Q. Detection of Hantaviruses and Arenaviruses in three-toed jerboas from the Inner Mongolia Autonomous Region, China. *Emerg Microbes Infect*. 2018;7:35.
11. Cao S, Ma J, Cheng C, Ju W, Wang Y. Genetic characterization of hantaviruses isolated from rodents in the port cities of Heilongjiang, China, in 2014. *BMC Vet Res*. 2016;12:69.
12. Zhang S, Wang S, Yin W, Liang M, Li J, Zhang Q, Feng Z, Li D. Epidemic characteristics of hemorrhagic fever with renal syndrome in China, 2006–2012. *BMC Infect Dis*. 2014;14:384.
13. Manigold T, Vial P. Human hantavirus infections: epidemiology, clinical features, pathogenesis and immunology. *Swiss Med Wkly*. 2014;144:w13937.
14. Ergonul O. Crimean-Congo haemorrhagic fever. *Lancet Infect Dis*. 2006;6:203–14.
15. Yoshii K, Song JY, Park SB, Yang J, Schmitt HJ. Tick-borne encephalitis in Japan, Republic of Korea and China. *Emerg Microbes Infect*. 2017;6:e82.
16. Ishii A, Thomas Y, Moonga L, Nakamura I, Ohnuma A, Hang’ombe B, Takada A, Mweene A, Sawa H. Novel arenavirus, Zambia. *Emerg Infect Dis*. 2011;17:1921–4.
17. Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, Palacios G, Khristova ML, Weyer J, Swanepoel R, Egholm M, et al. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathog*. 2009;5:e1000455.
18. Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, Conlan S, Quan PL, Hui J, Marshall J, et al. A new arenavirus in a cluster of fatal transplant-associated diseases. *N Engl J Med*. 2008;358:991–8.
19. Vieth S, Drosten C, Lenz O, Vincent M, Omilabu S, Hass M, Becker-Ziaja B, ter Meulen J, Nichol ST, Schmitz H, Gunther S. RT-PCR assay for detection of Lassa virus and related Old World arenaviruses targeting the L gene. *Trans R Soc Trop Med Hyg*. 2007;101:1253–64.
20. Knust B, Stroher U, Edison L, Albarino CG, Lovejoy J, Armeanu E, House J, Cory D, Horton C, Fowler KL, et al. Lymphocytic choriomeningitis virus in employees and mice at multipremises feeder-rodent operation, United States, 2012. *Emerg Infect Dis*. 2014;20:240–7.
21. Sridhar S, Yip CCY, Wu S, Cai J, Zhang AJ, Leung KH, Chung TWH, Chan JFW, Chan WM, Teng JLL, et al. Rat hepatitis E virus as cause of persistent hepatitis after liver transplant. *Emerg Infect Dis*. 2018;24:2241–50.
22. Carroll D, Daszak P, Wolfe ND, Gao GF, Morel CM, Morzaria S, Pablos-Mendez A, Tomori O, Mazet JAK. The Global Virome Project. *Science*. 2018;359:872–4.
23. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, et al. Redefining the invertebrate RNA virosphere. *Nature*. 2016;540:539–43.
24. Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, Delwart EL. The fecal viral flora of wild rodents. *PLoS Pathog*. 2011;7:e1002218.
25. Li L, Victoria JG, Wang C, Jones M, Fellers GM, Kunz TH, Delwart E. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J Virol*. 2010;84:6955–65.
26. Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, Wang W, Eden JS, Shen JJ, Liu L, et al. The evolutionary history of vertebrate RNA viruses. *Nature*. 2018;556:197–202.

27. Firth C, Bhat M, Firth MA, Williams SH, Frye MJ, Simmonds P, Conte JM, Ng J, Garcia J, Bhuvu NP, et al. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal *Rattus norvegicus* in New York City. *MBio*. 2014;5:e01933–14.
28. Sachsenroder J, Braun A, Machnowska P, Ng TFF, Deng X, Guenther S, Bernstein S, Ulrich RG, Delwart E, Johne R. Metagenomic identification of novel enteric viruses in urban wild rats and genome characterization of a group A rotavirus. *J Gen Virol*. 2014;95:2734–47.
29. Wu Z, Lu L, Du J, Yang L, Ren X, Liu B, Jiang J, Yang J, Dong J, Sun L, et al. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. *Microbiome*. 2018;6:178.
30. Wu Z, Yang L, Ren X, He G, Zhang J, Yang J, Qian Z, Dong J, Sun L, Zhu Y, et al. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME J*. 2016;10:609–20.
31. Wu Z, Ren X, Hu Y, Yang J, He G, Zhang J, Dong J, Sun L, Du J, et al. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. *J Virol*. 2012;86:10999–1012.
32. Du J, Yang L, Ren X, Zhang J, Dong J, Sun L, Zhu Y, Yang F, Zhang S, Wu Z, Jin Q. Genetic diversity of coronaviruses in *Miniopterus fuliginosus* bats. *Sci China Life Sci*. 2016;59:604–14.
33. Wu Z, Yang L, Ren X, Zhang J, Yang F, Zhang S, Jin Q. ORF8-related genetic evidence for Chinese horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. *J Infect Dis*. 2016;213:579–83.
34. Yang L, Wu Z, Ren X, Yang F, Zhang J, He G, Dong J, Sun L, Zhu Y, Zhang S, Jin Q. MERS-related betacoronavirus in *Vespertilio superans* bats, China. *Emerg Infect Dis*. 2014;20:1260–2.
35. Wu Z, Yang L, Yang F, Ren X, Jiang J, Dong J, Sun L, Zhu Y, Zhou H, Jin Q. Novel Henipa-like virus, Mojjiang Paramyxovirus, in rats, China, 2012. *Emerg Infect Dis*. 2014;20:1064–6.
36. Yang L, Wu Z, Ren X, Yang F, He G, Zhang J, Dong J, Sun L, Zhu Y, Du J, et al. Novel SARS-like betacoronaviruses in bats, China, 2011. *Emerg Infect Dis*. 2013;19:989–91.
37. Olival KJ, Hosseini PR, Zambrana-Torrel C, Ross N, Bogich TL, Daszak P. Host and viral traits predict zoonotic spillover from mammals. *Nature*. 2017;546:646–50.
38. Allen T, Murray KA, Zambrana-Torrel C, Morse SS, Rondinini C, Di Marco M, Breit N, Olival KJ, Daszak P. Global hotspots and correlates of emerging zoonotic diseases. *Nat Commun*. 2017;8:1124.
39. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. Global trends in emerging infectious diseases. *Nature*. 2008;451:990–3.
40. Pages M, Chaval Y, Herbreteau V, Waengsothorn S, Cosson JF, Hugot JP, Morand S, Michaux J. Revisiting the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. *BMC Evol Biol*. 2010;10:184.
41. Wilcove DS, Giam X, Edwards DP, Fisher B, Koh LP. Navjot's nightmare revisited: logging, agriculture, and biodiversity in Southeast Asia. *Trends Ecol Evol*. 2013;28:531–40.
42. Grace D, Gilbert J, Lapor ML, Unger F, Fevre S, Nguyen-Viet H, Schelling E. Zoonotic emerging infectious disease in selected countries in Southeast Asia: insights from ecohealth. *Ecohealth*. 2011;8:55–62.
43. Morand S, Blasdel K, Bordes F, Buchy P, Carcy B, Chaisiri K, Chaval Y, Claude J, Cosson JF, Desquesnes M, et al. Changing landscapes of Southeast Asia and rodent-borne diseases: decreased diversity but increased transmission risks. *Ecol Appl*. 2019;29:e01886.
44. Guo WP, Lin XD, Wang W, Tian JH, Cong ML, Zhang HL, Wang MR, Zhou RH, Wang JB, Li MH, et al. Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. *PLoS Pathog*. 2013;9:e1003159.
45. Reynes JM, Razafindralambo NK, Lacoste V, Olive MM, Barivelo TA, Soarimalala V, Heraud JM, Lavergne A. Anjozorobe hantavirus, a new genetic variant of Thailand virus detected in rodents from Madagascar. *Vector Borne Zoonotic Dis*. 2014;14:212–9.
46. Pattamadilok S, Lee BH, Kumperasart S, Yoshimatsu K, Okumura M, Nakamura I, Araki K, Khoprasert Y, Dangsupa P, Panlar P, et al. Geographical distribution of hantaviruses in Thailand and potential human health significance of Thailand virus. *Am J Trop Med Hyg*. 2006;75:994–1002.
47. Xiao SY, Leduc JW, Chu YK, Schmaljohn CS. Phylogenetic analyses of virus isolates in the genus *Hantavirus*, family *Bunyaviridae*. *Virology*. 1994;198:205–17.
48. Laenen L, Vergote V, Calisher CH, Klempa B, Klingstrom J, Kuhn JH, Maes P. *Hantaviridae*: current classification and future perspectives. *Viruses*. 2019;11:788.
49. Noronha LE, Wilson WC. Comparison of two zoonotic viruses from the order *Bunyavirales*. *Curr Opin Virol*. 2017;27:36–41.
50. Clark MHA, Warimwe GM, Di Nardo A, Lyons NA, Gubbins S. Systematic literature review of Rift Valley fever virus seroprevalence in livestock, wildlife and humans in Africa from 1968 to 2016. *PLoS Negl Trop Dis*. 2018;12:e0006627.
51. Charrel RN, Bichaud L, de Lamballerie X. Emergence of Toscana virus in the mediterranean area. *World J Virol*. 2012;1:135–41.
52. Silvas JA, Aguilar PV. The emergence of severe fever with thrombocytopenia syndrome virus. *Am J Trop Med Hyg*. 2017;97:992–6.
53. Blasdel KR, Duong V, Eloit M, Chretien F, Ly S, Hul V, Deubel V, Morand S, Buchy P. Evidence of human infection by a new mammarenavirus endemic to Southeastern Asia. *Elife*. 2016;5:e13135.
54. Dietzgen RG, Kondo H, Goodin MM, Kurath G, Vasilakis N. The family *Rhabdoviridae*: mono- and bipartite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. *Virus Res*. 2017;227:158–70.
55. Walker PJ, Blasdel KR, Calisher CH, Dietzgen RG, Kondo H, Kurath G, Longdon B, Stone DM, Tesh RB, Tordo N, et al. ICTV virus taxonomy profile: *Rhabdoviridae*. *J Gen Virol*. 2018;99:447–8.
56. Whitfield AE, Huot OB, Martin KM, Kondo H, Dietzgen RG. Plant rhabdoviruses—their origins and vector interactions. *Curr Opin Virol*. 2018;33:198–207.
57. Bruner K, Nadin-Davis S, Biek R. Genomic sequencing, evolution and molecular epidemiology of rabies virus. *Rev Sci Tech*. 2018;37:401–8.
58. Williams SH, Che X, Garcia JA, Klena JD, Lee B, Muller D, Ulrich W, Corrigan RM, Nichol S, Jain K, Lipkin WI. Viral diversity of house mice in New York City. *MBio*. 2014;5:e01933–14.
59. Delmas O, Holmes EC, Talbi C, Larrous F, Dacheux L, Bouchier C, Bourhy H. Genomic diversity and evolution of the lyssaviruses. *PLoS One*. 2008;3:e2057.
60. Noh JY, Jeong DG, Yoon SW, Kim JH, Choi YG, Kang SY, Kim HK. Isolation and characterization of novel bat paramyxovirus B16-40 potentially belonging to the proposed genus *Shaanvirus*. *Sci Rep*. 2018;8:12533.
61. Clayton BA. Nipah virus: transmission of a zoonotic paramyxovirus. *Curr Opin Virol*. 2017;22:97–104.
62. Woo PC, Lau SK, Wong BH, Wong AY, Poon RW, Yuen KY. Complete genome sequence of a novel paramyxovirus, *Tailam virus*, discovered in Sikkim rats. *J Virol*. 2011;85:13473–4.
63. Li Z, Yu M, Zhang H, Magoffin DE, Jack PJ, Hyatt A, Wang HY, Wang LF. Beilong virus, a novel paramyxovirus with the largest genome of non-segmented negative-stranded RNA viruses. *Virology*. 2006;346:219–28.
64. Jack PJ, Boyle DB, Eaton BT, Wang LF. The complete genome sequence of *J* virus reveals a unique genome structure in the family *Paramyxoviridae*. *J Virol*. 2005;79:10690–700.
65. Forth LF, Konrath A, Klose K, Schlottau K, Hoffmann K, Ulrich RG, Hoper D, Pohlmann A, Beer M. A novel squirrel respirovirus with putative zoonotic potential. *Viruses*. 2018;10:373.
66. Drexler JF, Corman VM, Muller MA, Maganga GD, Vallo P, Binger T, Gloza-Rausch F, Cottontail VM, Rasche A, Yordanov S, et al. Bats host major mammalian paramyxoviruses. *Nat Commun*. 2012;3:796.
67. Quan PL, Firth C, Conte JM, Williams SH, Zambrana-Torrel CM, Anthony SJ, Ellison JA, Gilbert AT, Kuzmin IV, Niezgodna M, et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses. *Proc Natl Acad Sci U S A*. 2013;110:8194–9.
68. Drexler JF, Corman VM, Muller MA, Lukashev AN, Gmyl A, Coutard B, Adam A, Ritz D, Leijten LM, van Riel D, et al. Evidence for novel hepaciviruses in rodents. *PLoS Pathog*. 2013;9:e1003438.
69. Smith DB, Becher P, Bukh J, Gould EA, Meyers G, Monath T, Muerhoff AS, Pletnev A, Rico-Hesse R, Stapleton JT, Simmonds P. Proposed update to the taxonomy of the genera *Hepacivirus* and *Pegivirus* within the *Flaviviridae* family. *J Gen Virol*. 2016;97:2894–907.
70. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol*. 2019;16:589–604.
71. Liu L, Larska M, Xia H, Utenthal A, Polak MP, Stahl K, Alenius S, Shan H, Yin H, Belak S. Atypical pestivirus and severe respiratory disease in calves, Europe. *Emerg Infect Dis*. 2012;18:1917–8.

72. Chander V, Nandi S, Ravishankar C, Upmanyu V, Verma R. Classical swine fever in pigs: recent developments and future perspectives. *Anim Health Res Rev.* 2014;15:87–101.
73. Snijder EJ, Kikkert M, Fang Y. Arterivirus molecular biology and pathogenesis. *J Gen Virol.* 2013;94:2141–63.
74. Snijder EJ, Meulenberg JJ. The molecular biology of arteriviruses. *J Gen Virol.* 1998;79(Pt 5):961–79.
75. Nilubol D, Tripipat T, Hoonsuan T, Kortheerakul K. Porcine reproductive and respiratory syndrome virus, Thailand, 2010–2011. *Emerg Infect Dis.* 2012;18:2039–43.
76. Nelsen CJ, Murtaugh MP, Faaberg KS. Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. *J Virol.* 1999;73:270–80.
77. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019;17:181–92.
78. Forni D, Cagliani R, Clerici M, Sironi M. Molecular evolution of human coronavirus genomes. *Trends Microbiol.* 2017;25:35–48.
79. Lau SK, Woo PC, Li KS, Tsang AK, Fan RY, Luk HK, Cai JP, Chan KH, Zheng BJ, Wang M, Yuen KY. 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.* 2015;89:3076–92.
80. Woo PC, Lau SK, Huang Y, Yuen KY. Coronavirus diversity, phylogeny and interspecies jumping. *Exp Biol Med (Maywood).* 2009;234:1117–27.
81. Anthony SJ, Johnson CK, Greig DJ, Kramer S, Che X, Wells H, Hicks AL, Joly DO, Wolfe ND, Daszak P, et al. Global patterns in coronavirus diversity. *Virus Evol.* 2017;3:vev012.
82. Pavio N, Meng XJ, Renou C. Zoonotic hepatitis E: animal reservoirs and emerging risks. *Vet Res.* 2010;41:46.
83. Zell R, Delwart E, Gorbalenya AE, Hovi T, King AMQ, Knowles NJ, Lindberg AM, Pallansch MA, Palmenberg AC, Reuter G, et al. ICTV Virus Taxonomy Profile: Picornaviridae. *J Gen Virol.* 2017;98:2421–2.
84. Cortez V, Meliopoulos VA, Karlsson EA, Hargest V, Johnson C, Schultz-Cherry S. Astrovirus biology and pathogenesis. *Annu Rev Virol.* 2017;4:327–48.
85. Du J, Li Y, Lu L, Zheng D, Liu B, Yang L, Su H, Dong J, Sun L, Zhu Y, et al. Biodiversity of rodent anelloviruses in China. *Emerg Microbes Infect.* 2018;7:38.
86. Du J, Lu L, Liu F, Su H, Dong J, Sun L, Zhu Y, Ren X, Yang F, Guo F, et al. Distribution and characteristics of rodent picornaviruses in China. *Sci Rep.* 2016;6:34381.
87. Latinne A, Waengsothorn S, Rojanadilok P, Eiamampai K, Sribuarod K, Michaux JR. Diversity and endemism of Murinae rodents in Thai limestone karsts. *Syst Biodivers.* 2013;11:323–44.
88. Morand S, Bordes F, Blasdel K, Pilosof S, Cornu JF, Chaisiri K, Chaval Y, Cosson JF, Claude J, Feyfant T, et al. Assessing the distribution of disease-bearing rodents in human-modified tropical landscapes. *J Appl Ecol.* 2015;52:784–94.
89. Francis CM. *A Field Guide to the Mammals of South-East Asia.* London: New Holland; 2008.
90. Kosoy M, Khlyap L, Cosson JF, Morand S. Aboriginal and invasive rats of genus *Rattus* as hosts of infectious agents. *Vector Borne Zoonotic Dis.* 2015;15:3–12.
91. Wu Z, Du J, Lu L, Yang L, Dong J, Sun L, Zhu Y, Liu Q, Jin Q. Detection of Hantaviruses and Arenaviruses in three-toed jerboas from the Inner Mongolia Autonomous Region, China. *Emerg Microbes Infect.* 2018;7:35.
92. Chen L, Liu B, Wu Z, Jin Q, Yang J. DRodVir: A resource for exploring the virome diversity in rodents. *J Genet Genomics.* 2017;44:259–64.
93. Chen L, Liu B, Yang J, Jin Q. DBatVir: the database of bat-associated viruses. *Database (Oxford).* 2014;2014:bau021.
94. Du J, Li F, Han Y, Fu S, Liu B, Shao N, Su H, Zhang W, Zheng D, Lei W, et al. Characterization of viromes within mosquito species in China. *Sci China Life Sci.* 2019;63:1089–92.
95. Galan M, Pages M, Cosson JF. Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples. *PLoS One.* 2012;7:e48374.
96. Yang J, Yang F, Ren L, Xiong Z, Wu Z, Dong J, Sun L, Zhang T, Hu Y, Du J, et al. Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. *J Clin Microbiol.* 2011;49:3463–9.
97. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 2013;30:2725–9.
98. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47:W256–9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

