

Article

Open Access

Identification of novel mammalian viruses in tree shrews (*Tupaia belangeri chinensis*)

Hong Zhou^{1,2,#}, Ren-Rong Tian^{3,#}, Xiu-Rong Wang^{4,#}, Jin-Xuan Yang³, Yun-Xiao Wang⁴, Ming-Liang Zhao³, Xu-Dong Zhang⁴, Yu-Hua Ma⁵, Long-Bao Lv⁵, Edward C. Holmes⁶, Yong-Tang Zheng^{3,5,*}, Wei-Feng Shi^{7,8,*}

¹ School of Clinical and Basic Medical Sciences, Shandong First Medical University & Shandong Academy of Medical Sciences, Ji'nan, Shandong 250117, China

² Key Laboratory of Emerging Infectious Diseases in Universities of Shandong, Shandong First Medical University and Shandong Academy of Medical Sciences, Taian, Shandong 271000, China

³ Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences/Key Laboratory of Bioactive Peptides of Yunnan Province, KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Center for Biosafety Mega-Science, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

⁴ School of Public Health, Shandong First Medical University & Shandong Academy of Medical Sciences, Ji'nan, Shandong 250117, China

⁵ National Resource Center for Non-Human Primates, Kunming Primate Research Center, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650107, China

⁶ Sydney Institute for Infectious Diseases, School of Medical Sciences, The University of Sydney, Sydney, New South Wales 2006, Australia

⁷ Department of Infectious Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

⁸ Shanghai Institute of Virology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

ABSTRACT

The Chinese tree shrew (*Tupaia belangeri chinensis*), a member of the mammalian order Scandentia, exhibits considerable similarities with primates, including humans, in aspects of its nervous, immune, and metabolic systems. These similarities have established the tree shrew as a promising experimental model for biomedical research on cancer, infectious diseases, metabolic disorders, and mental health conditions. Herein, we used meta-transcriptomic sequencing to analyze plasma, as well as oral and anal swab samples, from 105 healthy asymptomatic tree shrews to identify the presence of potential zoonotic viruses. In total, eight mammalian viruses with complete genomes were identified, belonging to six viral families, including *Flaviviridae*, *Hepeviridae*, *Parvovirinae*, *Picornaviridae*, *Sedoreoviridae*, and *Spinareoviridae*. Notably, the presence of rotavirus was recorded in tree shrews for the first time. Three viruses — hepacivirus 1, parvovirus, and picornavirus — exhibited low genetic similarity (<70%) with previously reported viruses at the whole-genome scale, indicating novelty. Conversely, three other viruses — hepacivirus 2,

hepatovirus A and hepevirus — exhibited high similarity (>94%) to known viral strains. Phylogenetic analyses also revealed that the rotavirus and mammalian orthoreovirus identified in this study may be novel reassortants. These findings provide insights into the diverse viral spectrum present in captive Chinese tree shrews, highlighting the necessity for further research into their potential for cross-species transmission.

Keywords: Tree shrew (*Tupaia belangeri chinensis*); Meta-transcriptomic sequencing; Mammalian viruses; Genomic analysis

INTRODUCTION

Tree shrews, rat-sized mammals classified within the genus *Tupaia* (family Tupaiidae, order Scandentia; Wang, 1987), are found extensively across South Asia, Southeast Asia, and Southwest China (Fan et al., 2013; Ye et al., 2021). The Chinese tree shrew (*Tupaia belangeri chinensis*) and northern tree shrew (*Tupaia belangeri*) are both found in China,

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2024 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

Received: 14 December 2023; Accepted: 25 December 2023; Online: 26 December 2023

Foundation items: This work was supported by the National Key R & D Program of China (2021YFC2300900, 2021YFC2301300), Academic Promotion Programme of Shandong First Medical University (2019QL006), Natural Science Foundation of Shandong Province (ZR2020QH274), Yunnan Key Research and Development Program (202103AQ100001, 202102AA310055), and Key Program of Chinese Academy of Sciences (KJZD-SW-L11)

*Authors contributed equally to this work

*Corresponding authors, E-mail: zhengyt@mail.kiz.ac.cn; shiwf@ioz.ac.cn

inhabiting moist, tropical, and subtropical forests. Despite diverging from primates approximately 85 million years ago, the taxonomic position of tree shrews has proven contentious, including potential classification as a “low-level primate” mammalian order. Given their relatively close phylogenetic relationship with humans, tree shrews are considered valuable animal models for biomedical research, facilitating the study of viral infections, human diseases, drug testing, and safety evaluations (Lu et al., 2021; Xiao et al., 2017). Indeed, tree shrews have been successfully infected with human hepatitis B virus, hepatitis C virus, influenza virus, and SARS-CoV-2 (Amako et al., 2010; Guo et al., 2018; Li et al., 2018; Xu et al., 2020). Furthermore, investigations have identified the presence of various viruses in wild populations, including adenovirus (Schöndorf et al., 2003), paramyxovirus (Tidona et al., 1999), circovirus (Okamoto et al., 2001), hepacivirus (GenBank accession No. OX394189), papillomavirus, and polyomavirus (Liu et al., 2019), emphasizing the potential importance of these animals as reservoirs for a broad range of viruses.

In this study, through meta-transcriptomic sequencing of *T. belangeri chinensis* from Kunming, Yunnan Province, China, we identified eight viruses with near-complete genomes related to human viruses. These comprised three novel viruses, including hepacivirus, picornavirus, and parvovirus, and five known viruses, including three hepaciviruses, one rotavirus, and one mammalian orthoreovirus, each exhibiting novel genetic characteristics.

MATERIALS AND METHODS

A total of 236 samples, including plasma samples ($n=60$), anal swabs ($n=81$), and oral swabs ($n=95$), were collected from 105 apparently healthy Chinese tree shrews (*T. belangeri chinensis*) from Kunming, Yunnan Province, China in April 2019 and 2022. The samples were mixed to generate 25 pools, each containing 7–11 samples. The swab samples were homogenized with steel beads before total RNA extraction using TRIzol reagent (Takara, Japan). After total RNA extraction, RNA quantity and quality were checked, followed by RNA library construction. Ribosomal RNA (rRNA) was removed using a MGIEasy rRNA Depletion Kit (Human-Mouse-Rat) (MGI, China) according to the manufacturer's instructions. Meta-transcriptomic libraries were subsequently constructed using a MGIEasy mRNA Library Prep Kit (MGI, China). Paired-end (100 bp) sequencing was performed using the BGISEQ-500RS and MGISEQ-2000 sequencing platforms (BGI, China).

Quality control of the raw reads was performed, using Fastp v.0.20.0 and Trimmomatic v.0.36 (Bolger et al., 2014) to remove adaptor and low-quality reads. Clean reads were *de novo* assembled using Trinity v.2.5.1 (Grabherr et al., 2011) with default settings. To identify potential viral sequences, the assembled contigs were compared against the NCBI non-redundant nucleotide (nt) and protein (nr) databases using BLASTn and Diamond BLASTx (Buchfink et al., 2015) with an e-value cutoff of 1×10^{-10} and 1×10^{-5} , respectively. Associated viral contigs were validated by read mapping using Bowtie2 and checked with Geneious v.2021.0.1. Relative abundance of identified viruses was estimated using reads per million (RPM). Assessment of completeness was conducted using CheckV (Nayfach et al., 2021) for non-segmented virus genomes.

The assembled genome sequences of potential novel

viruses were further confirmed by quantitative real-time polymerase chain reaction (qPCR), PCR amplification, and Sanger sequencing. TaqMan-based qPCR assays were established to test the positive samples: cDNA was synthesized using a ReverTra Ace qPCR RT Kit (TOYOBO, Japan), followed by qPCR using a set of probe and primer pairs (Supplementary Table S1) with a ProTaq HS Premix Probe qPCR Kit (AG, China) performed on the LightCycler 96 Real-Time PCR System (Roche, Switzerland). Complete viral genomes were also confirmed by Sanger sequencing. After cDNA was synthesized, PCR amplification was performed using specific overlapping primers (Supplementary Table S1) designed based on the genome sequences from meta-transcriptomic sequencing. The PCR products were then subjected to Sanger sequencing. Results showed that the genome sequences derived by Sanger sequencing were consistent with those obtained from meta-transcriptomic sequencing.

Open reading frames (ORFs) were predicted using ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) with genes annotated against the NCBI conserved domain database (CDD). Viral amino acid or nucleotide sequences identified in this study were aligned with representative reference viruses using MAFFT v.7 (Kato & Standley, 2013), employing the E-INS-i algorithm in Geneious v.2021.0.1. Phylogenetic trees were estimated using the maximum-likelihood (ML) method in IQ-TREE v.1.6.12, employing the best-fit substitution model and 1 000 bootstrap replicates (Nguyen et al., 2015). Phylogenetic analysis at the genus level was performed based on amino acid sequences of reference viral genomes sharing <90% identities with the study-obtained sequences. Phylogenetic analysis at the species level was performed using nucleotide sequences. All phylogenetic trees were mid-point rooted for clarity, and only bootstrap values >70% were shown adjacent to the nodes.

RESULTS

Meta-transcriptomic analysis

Total RNA sequencing generated 21 197 600–113 227 379 raw reads per library (Table 1). Analyses of the assembled viral contigs from the raw data of each library revealed the presence of seven RNA viruses and one DNA virus, including Tupaia hepacivirus 1 (TuHV1), Tupaia hepacivirus 2 (TuHV2), Tupaia hepatovirus A (TuHAV), Tupaia hepevirus (TuHEV), Tupaia parvovirus (TuPaV), Tupaia picornavirus (TuPiV), Tupaia rotavirus A (TuRVA), and Tupaia mammalian orthoreovirus (TuMRV) (Table 1). Raw reads were then re-mapped to estimate viral abundance (RPM), and full-length or near-complete genome sequences were obtained (Table 1).

BLASTn analysis revealed that TuPaV shared the highest nucleotide identity (66.5%) with a chipmunk parvovirus (GenBank accession No. NC_038543). The complete polyprotein of TuPiV showed the highest (51.98%) amino acid similarity to a canine picornavirus (GenBank accession No. APY24210), followed by *Miniopterus schreibersii* picornavirus 1 (47.6% identity). The complete polyprotein of TuHV1 exhibited the highest amino acid identity of 52.49% with a hepacivirus (GenBank accession No. YP_009553586) identified in a ground squirrel (*Citellus dauricus*, order Rodentia). The remaining five viruses exhibited >90% nucleotide identities with their closest relatives available in GenBank (Table 2).

Table 1 Summary of meta-transcriptomic sequencing libraries and sequence coverage of viruses identified in this study

No.	Group ID	Sample type	Individuals	Raw reads (PE)	Mapped reads (RPM)							
					TuHV1	TuHV2	TuHAV	TuHEV	TuPaV	TuPiV	TuRVA	TuMRV
1	P01	Plasma	8	83 806 456	94 068 (561.2)	119 088 (710.5)			196 (1.2)			
2	P02	Plasma	8	87 248 835	63 872 (366.0)	149 318 (855.7)						
3	P03	Plasma	8	66 502 579	342 590 (2 576.8)	496 802 (3 735.2)			970 (7.3)			
4	OS01	Oral swab	8	82 816 867								
5	OS02	Oral swab	7	78 026 732								
6	P04	Plasma	9	39 693 452	1 378 540 (17 364.8)	2 036 404 (25 651.6)						
7	P05	Plasma	9	31 092 753	1 283 290 (20 636.5)	1 564 892 (25 164.9)						
8	P06	Plasma	9	23 401 546	1 343 912 (28 714.2)	2 201 202 (47 031.1)						
9	P07	Plasma	9	21 197 600	1 532 200 (36 144.9)	2 086 674 (49 219.6)						
10	AS01	Anal swab	10	26 755 978			316 (5.9)					
11	AS02	Anal swab	10	30 816 120								
12	AS03	Anal swab	10	42 759 907					96 (1.1)			
13	AS04	Anal swab	10	47 434 795					396 (4.2)			
14	AS05	Anal swab	10	29 762 700					136 (2.3)			
15	AS06	Anal swab	10	63 577 215					79 588 (625.9)			
16	AS07	Anal swab	10	113 227 379			4 462 (19.7)		27 768 (122.6)			
17	AS08	Anal swab	11	22 835 621					2 298 (50.3)	592-5 266 (13.0-115.3)		
18	OS03	Oral swab	10	30 610 074								
19	OS04	Oral swab	10	28 742 971								
20	OS05	Oral swab	10	46 447 759								
21	OS06	Oral swab	10	25 824 833								
22	OS07	Oral swab	10	33 912 415								
23	OS08	Oral swab	10	81 859 028			196 (1.2)					170-1 196 (1.0-7.3)
24	OS09	Oral swab	10	25 188 720			1 002 (19.9)		138 (2.7)			44-280 (0.9-5.6)
25	OS10	Oral swab	10	31 305 599					2 434 (39.9)	34-164 (0.5-2.6)		20-108 (0.3-1.9)

RPM: Reads Per Million.

The TuHV1 and TuHV2 hepaciviruses were detected in seven plasma libraries, with viral abundances (RPM) ranging from 366.0 to 36 144.9 and 710.5 to 49 219.6, respectively. TuHEV was only detected in library AS01, with an RPM of 5.9. A novel parvovirus TuPaV was identified in two plasma libraries, with RPM values of 1.2 and 7.3, respectively. TuPiV was detected in six anal swab libraries and two oral swab libraries, while TuHAV was detected in one anal swab library and two oral swab libraries. The highest abundances of TuPiV and TuHAV were found in the libraries AS06 and OS09, with RPM values of 625.9 and 19.9, respectively. TuMRV was identified in three oral swab libraries, and virus abundance (RPM) in the 10 segments ranged from 1.0 to 7.3 (library OS08), 0.9 to 5.6 (library OS09), and 0.3 to 1.9 (library OS10), respectively. A rotavirus (TuRVA) was also identified in the anal swab library AS08 and oral swab library OS10 (Table 1), with viral abundance in the 11 segments ranging from 13.0 to 115.3 (library AS08) and 0.5 to 2.6 (library OS10), respectively.

Tupaia hepatitis viruses

Hepaciviruses, belonging to the family *Flaviviridae*, are enveloped viruses with a single-stranded positive-sense RNA

genome approximately 8.9–10.5 kb in length (Simmonds et al., 2017). Here, two hepaciviruses, TuHV1 and TuHV2, were identified in the tree shrew plasma (libraries P01–P07). To confirm their prevalence, TaqMan-based qPCR assays were used to screen all plasma samples ($n=60$). Results showed that TuHV1 and TuHV2 were both detected in 50% of the samples, with cycle threshold (Ct) values ranging from 20.08 to 34.95 and 16.64 to 29.33, respectively.

The final near-complete genomes of hepaciviruses TuHV1 and TuHV2 were 9 143 and 9 202 nucleotides in length (>94% completeness estimated by CheckV), predicted to encode single polyproteins of 2 859 and 2 865 amino acids, respectively. The complete polyprotein of TuHV1 exhibited the highest amino acid identity (52.49%) with *Hepacivirus* P isolate RHV-GS2015 (GenBank accession No. YP_009553586). The full-length polyprotein of TuHV2 showed the highest nucleotide and amino acid identities (95.7% and 99.51%, respectively) with a northern tree shrew hepacivirus (GenBank accession No. OX394189) (Table 2) identified in *T. belangeri*. The two viruses exhibited a typical hepacivirus genome structure, including: 5'-C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B-3' (Figure 1). Pairwise comparisons of amino acid and nucleotide similarities among

Table 2 Viruses identified in this study

Virus name	Family	Genus	Closest relative (Accession ID Name)	Coverage (%)	Identity (%)	Encoded protein	No. ^a
Hepacivirus TuHV1	<i>Flaviviridae</i>	<i>Hepacivirus</i>	YP_009553586 Hepacivirus P isolate RHV-GS2015	97	52.49		7
Hepacivirus TuHV2	<i>Flaviviridae</i>	<i>Hepacivirus</i>	OX394189 Northern treeshrew hepacivirus	100	95.70		7
Hepatovirus TuHAV	<i>Picornaviridae</i>	<i>Hepatovirus</i>	NC_028981 Tupaia hepatovirus A isolate TN1	99	94.43		2
Hepevirus TuHEV	<i>Hepeviridae</i>	unclassified	KR905549 Hepeviridae isolate Yunnan-2013	99	95.37		1
Parvovirus TuPaV	<i>Parvovirinae</i>	<i>Erythroparvovirus</i>	NC_038543 Chipmunk parvovirus	80	66.50		1
Picornavirus TuPIV	<i>Picornaviridae</i>	<i>Mischivirus</i>	APY24210 Canine picornavirus isolate A128thr	80	51.98		5
Rotavirus TuRVA	<i>Sedoreoviridae</i>	<i>Rotavirus</i>	OK651080 RVA/Horse-wt/IND/ERV4/2017/G3P[3]	99	91.89	VP1	2
			ON221572 RVA/Bat-tc/CHN/CHYNC82/2016/G3P3	99	92.86	VP2	
			LC576610 RVA/Human-wt/THA/MS2015-1-0001/G3P[10]	100	94.35	VP3	
			AB055967 RVA/Goat/KOR/GRV/1998/G3P5[3]	99	87.24	VP4	
			LC328216 RVA/Cat-tc/JPN/FRV303/1993/G3P[3]	100	81.04	VP6	
			MK161358 RVA/Human-wt/THA/2CR23/2016/G3P[8]	99	80.93	VP7	
			ON221577 RVA/Bat-tc/CHN/CHYNC82/2016/G3P3	100	86.07	NSP1	
			JX036372 RVA/Horse-wt/ARG/E3198/2008/G3P[3]	100	92.88	NSP2	
			KU597752 RVA/Human-wt/CHN/M2-102/2014/G3P[3]	100	91.35	NSP3	
			OP963645 RVA/Bat/2018/S18CXBatR24	100	93.57	NSP4	
			KU597754 RVA/Human-wt/CHN/M2-102/2014/G3P[3]	100	98.77	NSP5	
Mammalian orthoreovirus TuMRV	<i>Spinareoviridae</i>	<i>Orthoreovirus</i>	KM087105 MRV2/Bat/CHN/RpMRV-YN2012/2012	100	96.45	lambda-3	2
			JN799427 MRV/Pig/AUS/729/1998	100	92.03	lambda-2	
			MG451073 MRV3/TS/China/2012	100	96.29	lambda-1	
			MG451065 MRV1/TS/China/2011	100	98.31	mu-2	
			MG451064 MRV1/TS/China/2011	100	96.06	mu-1	
			OP057406 MRV1 40/Bat/Kansas USA/2018	100	91.18	mu-NS	
			KP185123 MRV1/T1/T28/KM/2013	99	94.05	sigma-1	
			OP963631 MRV/Bat/2017/S17BSBatR47	99	94.90	sigma-2	
			MG451079 MRV3/TS/China/2012	99	95.51	sigma-NS	
			MG451080 MRV3/TS/China/2012	99	94.57	sigma-3	

^aNumber of near-complete genomes of the virus obtained in this study.

variants of TuHV1 and TuHV2 from different libraries indicated low genetic diversity in the tree shrews investigated here, with the two viruses only differing 2.7%–4.4% and 0.48%–0.58% at the nucleotide and protein levels, respectively (Supplementary Table S2).

To determine the evolutionary histories of TuHV1 and TuHV2, we performed phylogenetic analyses of the polyprotein sequences of representative *Hepacivirus* species (*Hepacivirus A*–*Hepacivirus N* and *Hepacivirus P*), as well as related viruses (Figure 1). As the two hepaciviruses were detected in multiple libraries, a representative sequence of TuHV1 from the library P01 and a representative sequence of TuHV2 from the library P02 were used. Consistent with sequence comparisons, TuHV1 and TuHV2 were clustered with hepaciviruses infecting tree shrews, monkeys, sloths, and various species of rodents. TuHV2 was most closely related to the hepacivirus from the northern tree shrew, with this pair then clustered with TuHV1 (Figure 1). The three tree shrew hepaciviruses were then grouped with a rodent hepacivirus (GenBank accession No. CAI5760841), forming a sister clade to hepaciviruses containing *Hepacivirus P* (GenBank accession No. YP_009553586) from *Citellus dauricus* (family: Sciuridae) and other host species (Figure 1).

Hepevirus TuHEV was identified in an anal swab library (AS01). Hepeviruses (family *Hepeviridae*) are single-stranded positive-sense RNA viruses, comprising at least eight genotypes, four of which can infect humans. Genotypes G3 and G4 are zoonotic, while genotypes G1 and G2 have only been found in primates, including humans (Meng, 2011). A number of novel HEV-like viruses have recently been identified in a variety of animal hosts, including pigs, deer, rabbits, chickens, camels, rats, ferrets, mink, and fish (Li et al., 2016; Meng et al., 2022; Smith et al., 2016; Zhang et al., 2022). The near-complete genome of TuHEV was 6 538 nucleotides in length with 94.01% completeness (estimated by CheckV). TuHEV shared the highest similarity with a tree shrew hepevirus (95.37% nucleotide identity, GenBank accession No. KR905549). From the phylogenetic tree based on the complete genomes of representative hepeviruses (G1–G4) and other related HEV strains, TuHEV was grouped with the tree shrew hepevirus isolate Yunnan-2013 (GenBank accession No. KR905549), forming a sister clade to a group of animal and human hepeviruses (Figure 2).

Novel Tupaia parvovirus

Parvoviruses (family *Parvoviridae*) are small non-enveloped

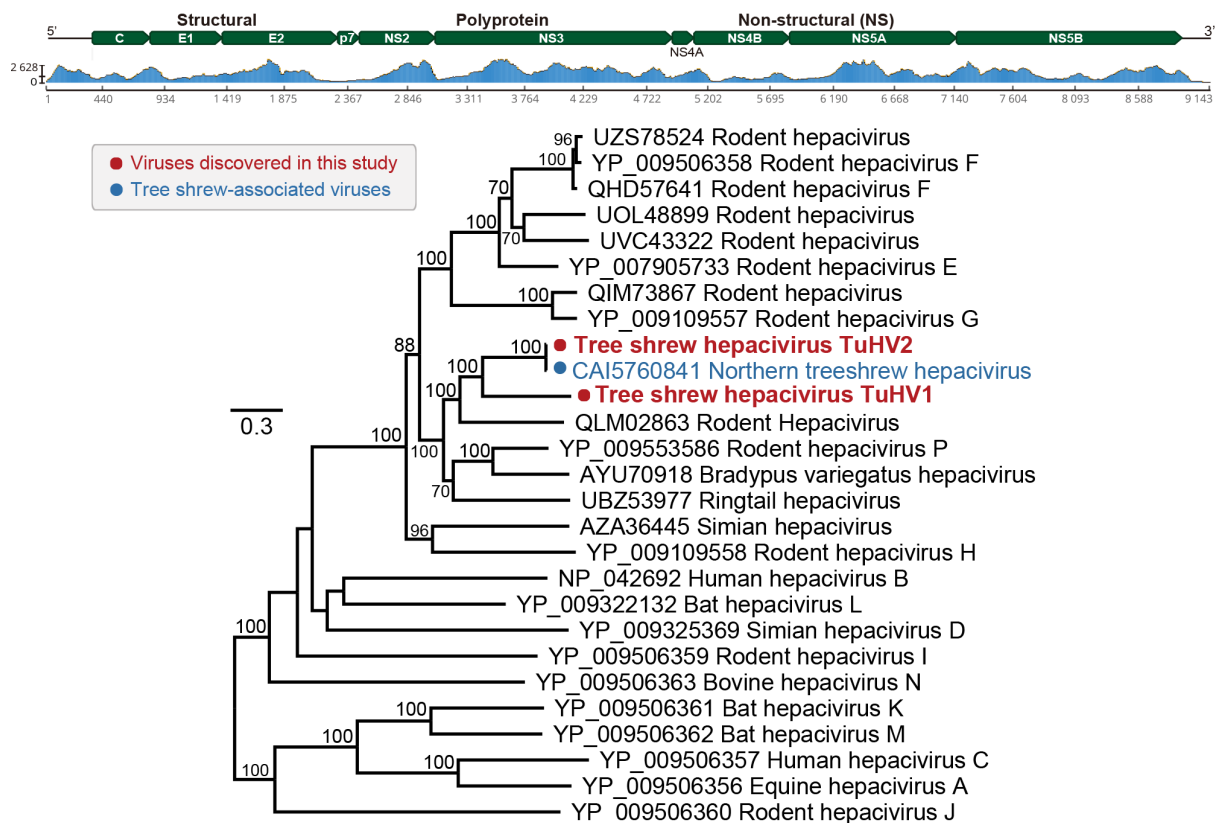


Figure 1 Genomic characterization and evolutionary relationships of tree shrew hepatitis virus identified in this study

Genomic features and sequence coverage (read abundance) of TuHV1 were analyzed and visualized using Geneious. Phylogenetic analysis of viruses TuHV1 and TuHV2 (marked with red dots) and representative hepatitis viruses was performed based on amino acid sequences. Scale bar represents number of amino acid substitutions per site and bootstrap values >70% are shown for key nodes. Tree was mid-point rooted.

viruses, with a linear positive-sense single-stranded DNA genome of ~5 kb in length. Eight genera in the subfamily *Parvovirinae* infect vertebrates, of which human parvovirus B19, non-human primate simian parvoviruses, and chipmunk parvovirus (ChpPV) belong to the genus *Erythroparvovirus* (Chen et al., 2010). Certain parvoviruses are noted for their broad host range, indicative of cross-species virus transmission (Hueffer & Parrish, 2003).

The complete genome of TuPaV was 5 283 bp in length (100% completeness), possessing a typical genome organization similar to that of other members of the family *Parvoviridae* (Yoo et al., 1999), with the first ORF (positions 246–2 558) encoding a putative nonstructural protein NP1/NP2 of 770 amino acids and a second ORF (positions 2 559–5 018) encoding a putative capsid protein VP1/VP2 of 819 amino acids (Figure 3). This novel TuPaV shared 69.88% and 53.08% amino acid identities with the chipmunk parvovirus ChpPV (GenBank accession No. NC_038543) in the VP1 and NS1 proteins, respectively. Phylogenetic analyses revealed that the NS1 of TuPaV fell within the genus *Erythroparvovirus* and formed a subclade containing chipmunk and tree shrew-associated parvoviruses (Figure 3A). Similar phylogenetic relationships were also observed in VP1 (Figure 3B). Taken together, these results indicate that TuPaV may represent a novel member within the genus *Erythroparvovirus*, closely related to ChpPV but distinct from primate parvoviruses.

Tupaia picornaviruses

The family *Picornaviridae* includes a large group of non-enveloped positive single-stranded RNA viruses with a

genome of 7–9 kb in length encoding a single polyprotein. They can cause mucocutaneous, encephalic, cardiac, hepatic, neurological, and respiratory diseases in a variety of vertebrates (Tracy et al., 2006; Wu et al., 2016). We identified two picornaviruses, TuPiV and TuHAV, in this study. Four complete and one near-complete genome sequence of TuPiV were recovered. Pairwise nucleotide comparisons revealed >99.59% nucleotide sequence similarity with each other. The genome size of TuPiV ranged from 8 037 to 8 163 bp in length (100% completeness). The TuPiV genome sequence from the library AS06 was selected as the representative strain to determine evolutionary relationships. Notably, this novel virus shared the highest nucleotide identity (51.98%) with canine picornavirus (GenBank accession No. APY24210). Phylogenetic analysis of the polyprotein sequences revealed that TuPiV clustered with the canine picornavirus, with this cluster forming a sister clade to bat picornavirus 1 (GenBank accession No. YP_009361319) and *Suncus murinus* picornavirus (GenBank accession No. AWK02669) within the genus *Mischivirus* (Figure 4).

In addition, two complete genome sequences of TuHAV, 7 385 and 7 316 bp in length, respectively (>99% completeness), were recovered and showed 99.88% nucleotide similarity. TuHAV-positive libraries AS07 and OS09 included anal and oral swabs from tree shrews Tu85 to Tu94 (Supplementary Table S3). Four and five positive samples were detected using qPCR in libraries AS07 and OS09, respectively, with four individual tree shrews (Tu91, Tu92, Tu93, and Tu94) found to be TuHAV-positive for both sample types (Supplementary Table S3). The assembled TuHAV

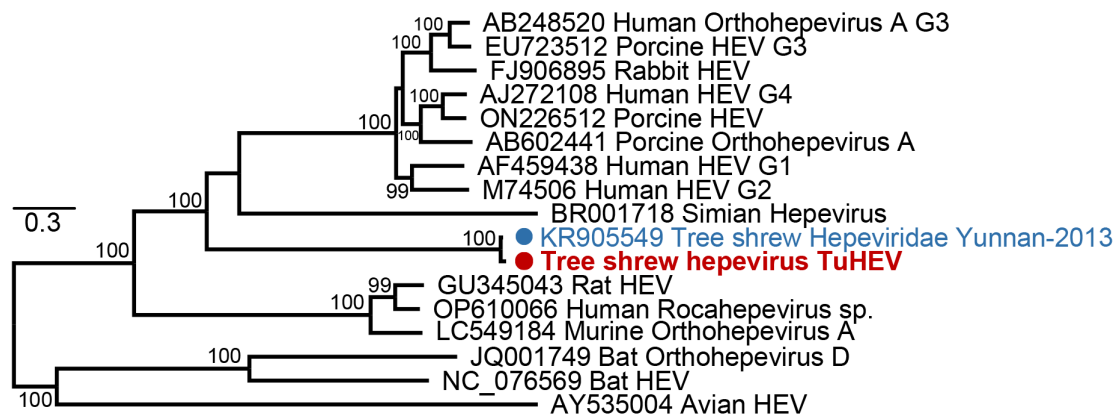


Figure 2 Phylogenetic analysis of nucleotide sequences from tree shrew virus TuHEV and representative hepeviruses

Strains are highlighted with a red dot. Scale bar represents number of amino acid substitutions per site and bootstrap values >70% are shown for key nodes. Tree was mid-point rooted.

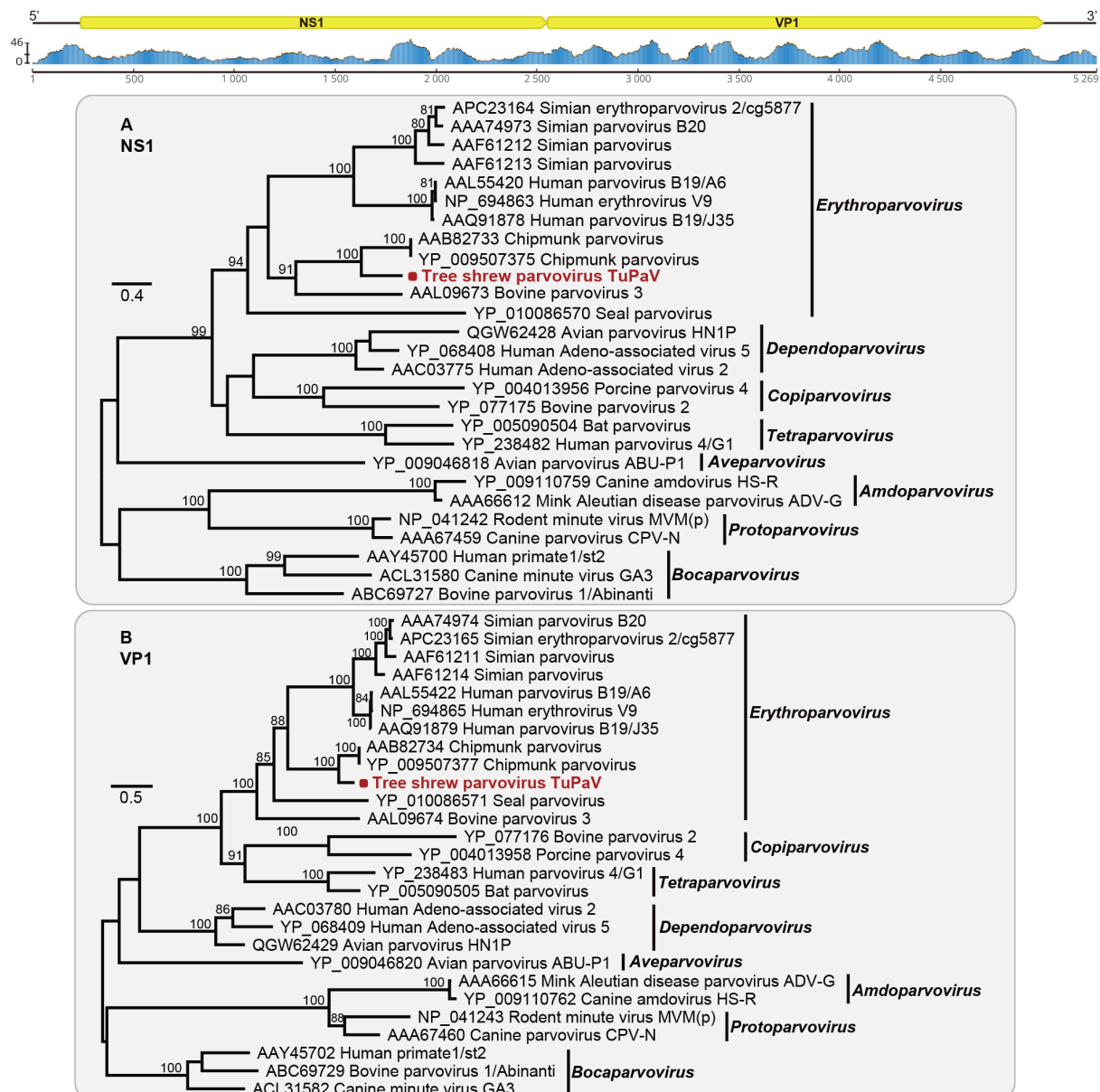


Figure 3 Genomic characterization and evolutionary relationships of tree shrew parvovirus identified in this study

Genomic features and sequence coverage (read abundance) of TuPaV were analyzed and visualized using Geneious. Phylogenetic analysis of TuPaV (highlighted with a red dot) and representative parvoviruses was performed based on NS1 (A) and VP1 (B) amino acid sequences. Scale bars represent number of amino acid substitutions per site and bootstrap values >70% are shown for key nodes. Trees were mid-point rooted.

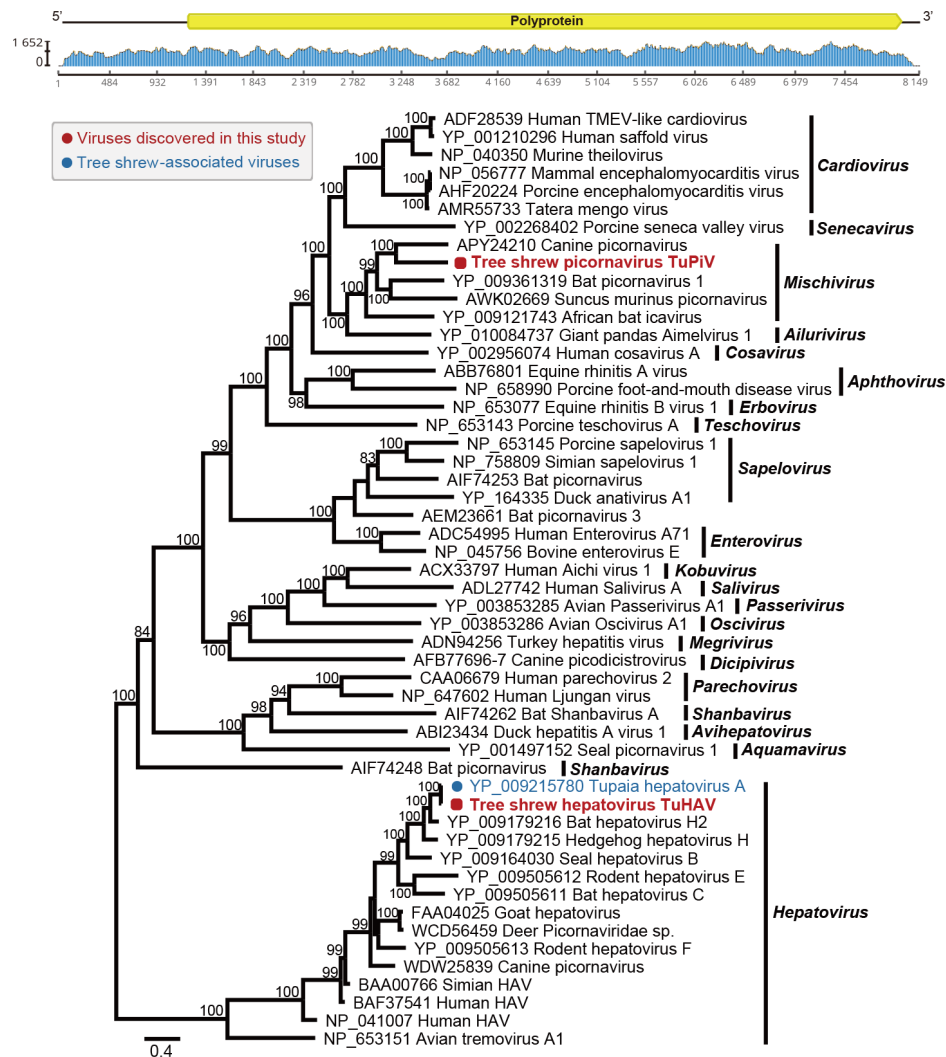


Figure 4 Genomic characterization and evolutionary relationships of tree shrew viruses TuPiV and TuHAV identified in this study

Genomic features and sequence coverage (read abundance) of novel virus TuPiV (highlighted with a red dot) were predicted and visualized using Geneious. Phylogenetic analysis of TuPiV and TuHAV and representative picornaviruses was performed based on amino acid sequences. Scale bar represents number of amino acid substitutions per site and bootstrap values >70% are shown for key nodes. The tree was mid-point rooted.

genome sequence from library AS07 was most closely related to the *Tupaia* hepatovirus A isolate TN1 (GenBank accession No. YP_009215780), showing the highest amino acid identity (99.06%). Phylogenetic analysis revealed that they clustered together within the genus *Hepatovirus*, grouping with strains known to infect tree shrews, bats, hedgehogs, and seals (Figure 4).

Tupaia viruses TuMRV and TuRVA

Mammalian orthoreoviruses (MRVs) are non-enveloped icosahedral viruses, which contain a segmented genome of 10 double-stranded RNAs, including three large (L1–L3), three medium (M1–M3), and four small (S1–S4) genome segments. The S1 gene segment defines the MRV serotype (Gaillard & Joklik, 1980). MRVs are associated with gastroenteritis, respiratory illness, encephalitis, and other diseases in humans and various animal species (Dermody et al., 2013). At least four tree shrew MRVs have been described in recent years (Li et al., 2020; Xu et al., 2013). Genetic reassortment and intragenic rearrangement events have also been detected in MRVs in a wide range of hosts, including bats, humans, pigs, and minks (Duncan, 1999; Lo et al., 2022).

Here, two complete MRV genome sequences, TuMRV01

and TuMRV02, were recovered from libraries OS08 and OS09, with nucleotide sequence identities ranging from 99.73% to 100% across the 10 separate segments. The nucleotide sequences of the S1, S2, S3, S4, L1, and L2 segments were used to estimate the phylogenetic trees (Figure 5). The S1 segment showed that TuMRV had the highest sequence identity (94.05%) with MRV1/T1/T28/KM/2013 (GenBank accession No. KP185123), also identified in a tree shrew. Phylogenetic analysis of the S1 gene revealed that the two TuMRV strains fell within the MRV serotype 1 and were grouped with two other tree shrew MRVs: MRV1/T1/T28/KM/2013 and MRV1/TS/China/2011 (Figure 5A). In the phylogenetic trees of the S3 and S4 gene segments, TuMRV clustered with tree shrew MRV1/TS/China/2011 and MRV3/TS/China/2012 (Figure 5C, E). However, in the S2 phylogenetic tree, the two novel sequences did not cluster with MRV1/TS/China/2011 and MRV3/TS/China/2012 directly, with the two latter sequences clustering with bat MRV/Bat/2017/S17BSBatR47 (Figure 5B). In the L1 phylogenetic tree, TuMRV01 and TuMRV02 formed a clade with bat MRV2/RpMRV-YN2012 and MRV3/TS/China/2012 (Figure 5D). The L2 segment phylogeny

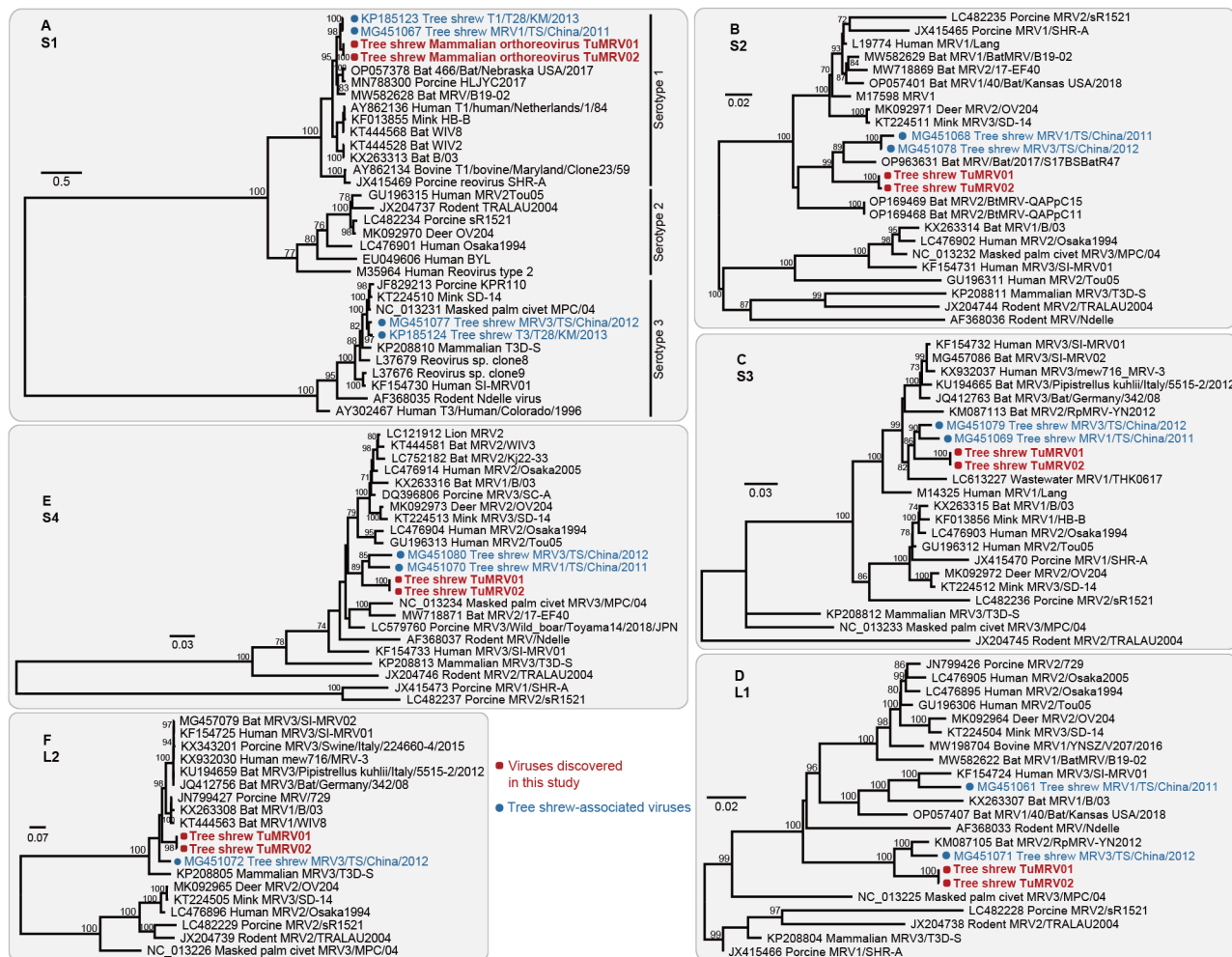


Figure 5 Phylogenetic analysis of nucleotide sequences from tree shrew virus TuMRV and representative MRV viruses

A: S1 gene; B: S2 gene; C: S3 gene; D: L1 gene; E: S4 gene; F: L2 gene. Scale bars represent number of amino acid substitutions per site and bootstrap values >70% are shown for key nodes. Novel TuMRV strains are highlighted with a red dot, and trees were mid-point rooted.

revealed a close relationship between TuMRV and porcine and bat MRV strains (Figure 5F). Taken together, these findings indicate that the MRV isolates described here may be reassortants.

Rotaviruses are important pathogens causing severe gastroenteritis in children and animals worldwide. Rotaviruses have been classified into species A–I and tentative species J (Bányai et al., 2017; Johne et al., 2019; Matthijnsens & Van Ranst, 2012). Rotavirus A (RVA) is responsible for most seasonal endemic diarrheal diseases in young children (Troeger et al., 2018). The rotavirus genome comprises 11 double-stranded RNA segments (Estes & Greenberg, 2013). In this study, we identified a rotavirus from libraries AS08 and OS10, and all 11 genome segments were recovered. The VP6 protein phylogenetic tree (Figure 6A) indicated a consistent branching of this rotavirus with other RVA species, hence its name TuRVA. To the best of our knowledge, RVA has not been reported previously in tree shrews, although rotavirus antigens have been detected in wild Chinese tree shrews (Wang et al., 2011). Anal and oral swabs from tree shrews Tu95 to Tu105 were pooled into TuRVA-positive libraries AS08 and OS10, respectively (Supplementary Table S4). Seven anal samples were TuRVA-positive based on qPCR analysis of library AS08, and nine oral samples were positive for TuRVA in the library OS10. Among these, six individual

tree shrews (Tu95, Tu96, Tu99, Tu100, Tu101, and Tu102) were TuRVA-positive for both sample types (Supplementary Table S4). The nucleotide identities of the two TuRVA strains ranged from 99.85% to 100% across segments. Phylogenetic analyses revealed that seven TuRVA gene segments (NSP2, NSP3, NSP4, VP1–3, and VP6) fell within the same lineages as human rotaviruses RVA/Human-tc/JPN/AU-1/1982/G3P3[9] and RVA/Human-tc/THA/T152/1998/G12P[9] (Supplementary Figure S1 and Table S5). However, TuRVA clustered with sequences of A9, H6, P[3], and G3 in the NSP1, NSP5, VP4, and VP7 gene trees, respectively, differing from those of RVA/Human-tc/JPN/AU-1/1982/G3P3[9] and RVA/Human-tc/THA/T152/1998/G12P[9] (Figure 6B–6E). Taken together, the full genome constellation of TuRVA was G3-P[3]-I3-R3-C3-M3-A9-N3-T3-E3-H6, suggesting that TuRVA may be a reassortant and novel RVA member.

DISCUSSION

Zoonotic viruses, including adenovirus, paramyxovirus, papillomavirus, and polyomavirus, have been identified previously in tree shrews (Liu et al., 2019; Schöndorf et al., 2003; Tidona et al., 1999). In the current study, we conducted a meta-transcriptomic study of 236 samples from 105 healthy asymptomatic Chinese tree shrews in Kunming, Yunnan, China. Analysis of the transcriptomes led to the discovery of

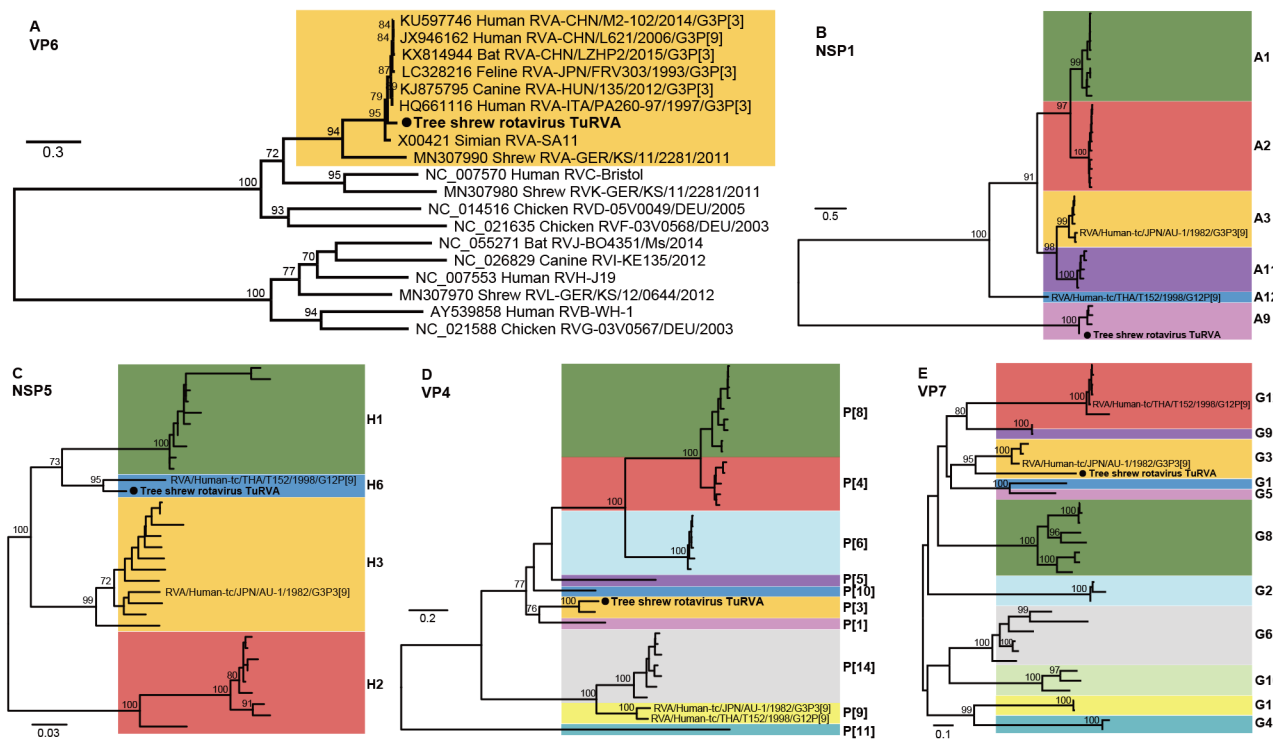


Figure 6 Phylogenetic analysis of sequences from tree shrew virus TuRVA and representative rotaviruses

A: Phylogenetic analysis of VP6 proteins of representative RVA to RVJ sequences; B–E: Phylogenetic analyses of nucleotide sequences of gene segments NSP1 (B), NSP5 (C), VP4 (D), and VP7 (E) of representatives of RVA. Scale bars represent number of amino acid or nucleotide substitutions per site and bootstrap values >70% are shown for key nodes. Trees were mid-point rooted. Novel TuRVA strains are highlighted with a red dot, and lineages of each gene are marked on the right side.

seven RNA viruses and one DNA virus, including three novel viruses, which shared 50%–70% nucleotide or amino acid similarity to the most closely related viruses. In addition, the presence of an RVA family member (TuRVA) was identified in tree shrews for the first time. These results highlight the capacity of tree shrews to harbor a wide spectrum of RNA and DNA viruses. As such, the presence of these viruses should be considered when tree shrews are used as an animal model for biomedical research.

To date, 15 *Hepacivirus* species (*Hepacivirus* A–N and P) have been documented to infect a wide range of hosts, including primates, bats, horses, donkeys, cows, and various rodents (Hartlage et al., 2016). The two hepaciviruses (TuHV1 and TuHV2) detected in blood samples of tree shrews were grouped together with another tree shrew hepacivirus (GenBank accession No. CAI5760841), which then formed a sister clade to the *Hepacivirus* P group. In addition, hepatovirus TuHAV and hepevirus TuHEV — members of the families *Picornaviridae* and *Hepeviridae*, respectively — were also detected in swab samples. These findings have important implications for the use of tree shrews as laboratory infection models of human hepatitis-associated viruses.

We also identified a novel parvovirus (TuPaV) in the tree shrews, which shared the highest nucleotide identity (66.5%) with a chipmunk parvovirus (GenBank accession No. NC_038543) but was distinct from other members of the genus *Erythroparvovirus* found in humans and non-human primates. Previous research has shown that the large non-structural protein (NS1) of ChpPV, initially isolated from Manchurian chipmunks in Korea, induces apoptosis in B19V semipermissive UT7/Epo-S1 cells (Chen et al., 2010). The potential of the novel TuPaV strain to cross the species barrier

and infect other mammals, including humans, clearly warrants further study. Regarding picornaviruses, we also identified a novel TuPIV in the tree shrews, which exhibited only ~52% amino acid identity to a canine picornavirus (GenBank accession No. APY24210). As many picornaviruses can cause disease (Tracy et al., 2006), the pathogenicity of this novel picornavirus should be further investigated.

The two remaining segmented tree shrew viruses, TuMRV and TuRVA, were from the family *Reoviridae*. The two TuMRV strains (both from oral swabs) shared 99.8%–100% nucleotide identities, suggesting that this mammalian orthoreovirus may have been circulating in our captive tree shrew population. Phylogenetic analyses indicated that they belonged to serotype 1. Of note, we believe that TuRVA is the first identification of rotavirus in tree shrews, with phylogenetic analyses revealing a broad range of mammalian hosts of this virus.

In conclusion, we discovered four hepatitis viruses, one parvovirus, one picornavirus, one rotavirus, and one mammalian orthoreovirus from captive tree shrews in China, three of which represent potential novel viruses. These findings suggest that tree shrews may be an important reservoir for viruses that can infect a broad range of animal hosts, including humans.

DATA AVAILABILITY

Sequencing data supporting the findings of this study were deposited in the Genome Sequence Archive (GSA) database (<https://ngdc.cnpc.ac.cn/gsa/>) under accession number CRA015269, Science Data Bank (doi: <https://10.57760/sciencedb.j00139.00120>), and NCBI under BioProjectID PRJNA1019067.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y.T.Z. and W.F.S. designed the study. R.R.T., H.Z., Y.H.M., and L.B.L. collected the living tree shrew samples. X.R.W., Y.X.W., X.D.Z., J.X.Y., and M.L.Z. performed the laboratory work including RNA extraction, meta-transcriptomic sequencing, and qPCR. H.Z. and W.F.S. performed data analysis. H.Z., E.C.H., Y.T.Z., and W.F.S. drafted and revised the paper. All authors read and approved the final version of the manuscript.

REFERENCES

- Amako Y, Tsukiyama-Kohara K, Katsume A, et al. 2010. Pathogenesis of hepatitis C virus infection in *Tupaia belangeri*. *Journal of Virology*, **84**(1): 303–311.
- Bányai K, Kemenesi G, Budinski I, et al. 2017. Candidate new rotavirus species in Schreiber's bats, Serbia. *Infection, Genetics and Evolution*, **48**: 19–26.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15): 2114–2120.
- Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, **12**(1): 59–60.
- Chen ZJ, Chen AY, Cheng F, et al. 2010. Chipmunk parvovirus is distinct from members in the genus *Erythrovirus* of the family *Parvoviridae*. *PLoS One*, **5**(12): e15113.
- Dermody TS, Parker JSL, Sherry B. 2013. Orthoreoviruses. In: Knipe DM, Howley PM. *Fields Virology*. 6th ed. Philadelphia: Lippincott Williams & Wilkins, 1304–1346.
- Duncan R. 1999. Extensive sequence divergence and phylogenetic relationships between the fusogenic and nonfusogenic orthoreoviruses: a species proposal. *Virology*, **260**(2): 316–328.
- Estes MK, Greenberg HB. 2013. Rotaviruses. In: Knipe DM, Howley PM. *Fields Virology*. 6th ed. Philadelphia: Lippincott Williams & Wilkins, 1347–1401.
- Fan Y, Huang ZY, Cao CC, et al. 2013. Genome of the Chinese tree shrew. *Nature Communications*, **4**: 1426.
- Gaillard RK, Joklik WK. 1980. The antigenic determinants of most of the proteins coded by the three serotypes of reovirus are highly conserved during evolution. *Virology*, **107**(2): 533–536.
- Grabherr MG, Haas BJ, Yassour M, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**(7): 644–652.
- Guo WN, Zhu B, Ai L, et al. 2018. Animal models for the study of hepatitis B virus infection. *Zoological Research*, **39**(1): 25–31.
- Hartlage AS, Cullen JM, Kapoor A. 2016. The strange, expanding world of animal hepaciviruses. *Annual Review of Virology*, **3**: 53–75.
- Hueffer K, Parish CR. 2003. Parvovirus host range, cell tropism and evolution. *Current Opinion in Microbiology*, **6**(4): 392–398.
- Johne R, Tausch SH, Grütze J, et al. 2019. Distantly related rotaviruses in common shrews, Germany, 2004–2014. *Emerging Infectious Diseases*, **25**(12): 2310–2314.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**(4): 772–780.
- Li RF, Yuan B, Xia XS, et al. 2018. Tree shrew as a new animal model to study the pathogenesis of avian influenza (H9N2) virus infection. *Emerging Microbes & Infections*, **7**(1): 166.
- Li TC, Yang TT, Yoshizaki S, et al. 2016. Ferret hepatitis E virus infection induces acute hepatitis and persistent infection in ferrets. *Veterinary Microbiology*, **183**: 30–36.
- Li XF, Sun XM, Lu CX, et al. 2020. Isolation and identification of two new strains of mammalian orthoreovirus from Chinese tree shrews. *Archives of Virology*, **165**(7): 1541–1550.
- Liu P, Qiu Y, Xing C, et al. 2019. Detection and genome characterization of two novel papillomaviruses and a novel polyomavirus in tree shrew (*Tupaia belangeri chinensis*) in China. *Virology Journal*, **16**(1): 35.
- Lo VT, Yoon SW, Noh JY, et al. 2022. Characterization of replication and variations in genome segments of a bat reovirus, BatMRV/B19–02, by RNA-seq in infected Vero-E6 cells. *Archives of Virology*, **167**(11): 2133–2142.
- Lu T, Peng HM, Zhong LP, et al. 2021. The tree shrew as a model for cancer research. *Frontiers in Oncology*, **11**: 653236.
- Matthijnssens J, Van Ranst M. 2012. Genotype constellation and evolution of group A rotaviruses infecting humans. *Current Opinion in Virology*, **2**(4): 426–433.
- Meng K, Liu MC, Zhang YX, et al. 2022. A genetically novel avian Hepatitis E virus in China. *Virus Genes*, **58**(6): 589–593.
- Meng XJ. 2011. From barnyard to food table: the omnipresence of hepatitis E virus and risk for zoonotic infection and food safety. *Virus Research*, **161**(1): 23–30.
- Nayfach S, Camargo AP, Schulz F, et al. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature Biotechnology*, **39**(5): 578–585.
- Nguyen LT, Schmidt HA, von Haeseler A, et al. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**(1): 268–274.
- Okamoto H, Nishizawa T, Takahashi M, et al. 2001. Genomic and evolutionary characterization of TT virus (TTV) in tupaia and comparison with species-specific TTVs in humans and non-human primates. *Journal of General Virology*, **82**(9): 2041–2050.
- Schöndorf E, Bahr U, Handermann M, et al. 2003. Characterization of the complete genome of the Tupaia (tree shrew) adenovirus. *Journal of Virology*, **77**(7): 4345–4356.
- Simmonds P, Becher P, Bukh J, et al. 2017. ICTV virus taxonomy profile: *Flaviviridae*. *Journal of General Virology*, **98**(1): 2–3.
- Smith DB, Simmonds P, Izopet J, et al. 2016. Proposed reference sequences for hepatitis E virus subtypes. *Journal of General Virology*, **97**(3): 537–542.
- Tidona CA, Kurz HW, Gelderblom HR, et al. 1999. Isolation and molecular characterization of a novel cytopathogenic paramyxovirus from tree shrews. *Virology*, **258**(2): 425–434.
- Tracy S, Chapman NM, Drescher KM, et al. 2006. Evolution of virulence in picornaviruses. In: Domingo E. *Quasispecies: Concept and Implications for Virology*. Berlin, Heidelberg: Springer, 193–209.
- Troeger C, Khalil IA, Rao PC, et al. 2018. Rotavirus vaccination and the global burden of rotavirus diarrhea among children younger than 5 years. *JAMA Pediatrics*, **172**(10): 958–965.
- Wang XX, Li JX, Wang WG, et al. 2011. Preliminary investigation of viruses to the wild tree shrews (*Tupaia belangeri Chinese*). *Zoological Research*, **32**(1): 66–69. (in Chinese)
- Wang YX. 1987. Taxonomic research on Burma-Chinese tree shrew, *Tupaia belangeri* (Wagner), from southern China. *Zoological Research*, **8**(3): 213–230. (in Chinese)
- Wu ZQ, Yang L, Ren XW, et al. 2016. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *The ISME Journal*, **10**(3): 609–620.
- Xiao J, Liu R, Chen CS. 2017. Tree shrew (*Tupaia belangeri*) as a novel laboratory disease animal model. *Zoological Research*, **38**(3): 127–137.
- Xu J, Huang XY, Li XF, et al. 2013. Isolation and identification of *Tupaia* orthoreovirus. *Zoological Research*, **34**(2): 116–120. (in Chinese)
- Xu L, Yu DD, Ma YH, et al. 2020. COVID-19-like symptoms observed in Chinese tree shrews infected with SARS-CoV-2. *Zoological Research*, **41**(5): 517–526.
- Ye MS, Zhang JY, Yu DD, et al. 2021. Comprehensive annotation of the Chinese tree shrew genome by large-scale RNA sequencing and long-read isoform sequencing. *Zoological Research*, **42**(6): 692–709.
- Yoo BC, Lee DH, Park SM, et al. 1999. A novel parvovirus isolated from Manchurian chipmunks. *Virology*, **253**(2): 250–258.
- Zhang WJ, Ami Y, Suzuki Y, et al. 2022. Mongolia gerbils are broadly susceptible to hepatitis E virus. *Viruses*, **14**(6): 1125.