Virologica Sinica 37 (2022) 307–310

Contents lists available at ScienceDirect

Virologica Sinica

journal homepage: www.keaipublishing.com/en/journals/virologica-sinica
www.virosin.org

Letter

Identification of a novel hepacivirus in Mongolian gerbil (*Meriones unguiculatus*) from Shaanxi, China

Cui-hong An [a,1], Juan Li [b,c,d,1], Yi-ting Wang [c], Shou-min Nie [a], Wen-hui Chang [e], Hong Zhou [b,c,d], Lin Xu [b,c,f], Yang-xin Sun [a,*], Wei-feng Shi [b,c,d,f,*], Ci-xiu Li [b,c,d,*]

[a] Department of Plague and Brucellosis, Shaanxi Center for Disease Control and Prevention, Xi' An 710054, China
[b] Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, 250021, China
[c] Key Laboratory of Etiology and Epidemiology of Emerging Infectious Diseases in Universities of Shandong, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian, 271000, China
[d] Department of Pathogen Biology, School of Basic Medicine, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian, 271000, China
[e] Shaanxi Center for Disease Control and Prevention, Xi'An, 710054, China
[f] School of Public Health, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian, 271000, China

Dear Editor,

Hepaciviruses, members of the family *Flaviviridae*, are enveloped viruses containing a single-stranded positive-sense RNA genome of approximately 8.9–10.5 kb in size (Simmonds et al., 2017). To date, 15 species (*Hepacivirus* A–N, and P) have been documented within the *Hepacivirus* genus that show distinct host ranges, including primates, bats, horses, donkeys, cows, and various rodents (Hartlage et al., 2016). Seven rodent-associated hepaciviruses have been characterized, including *hepacivirus* E, I, G and H infecting rodents of Muridae, *hepacivirus* F and J infecting rodents of Cricetidae (de Souza et al., 2019), and *heapcivirus* P infecting rodents of Xerinae (Li et al., 2019). Additional unclassified rodent hapaciviruses have been described in diverse rodents from Dormouse, Echimyidae, Heteromyidae, and Spalacidae. Mongolian gerbils (*Meriones unguiculatus*) are small rodents belonging to the family Muridae and are widely distributed in the desert grasslands and steppes of northern China, Mongolia, and Russia (Liu et al., 2007). They have been reported as a major host of *Yersinia pestis* causing plagues in China in recent decades (Riehm et al., 2011). Moreover, Mongolian gerbil is known to be susceptible to various viruses and is a commonly used animal model for virus research (Li et al., 2009). Despite this, the natural virome of wild *Meriones unguiculatus* has not been described. Herein, we reported the first hepacivirus detected in Mongolian gerbils captured in Dingbian County of Shaanxi Province, one of the plague zones in China.

A total of 110 Mongolian gerbils were captured during October–November in 2018 from Dingbian County of Shaanxi Province, China. All animals were anesthetized with ether before they were sacrificed. Heart, liver, spleen, lung, kidney and brain tissues were then collected aseptically. Tissues from up to ten individuals were pooled based on groups (SX01 to SX11) and tissue type generating a total of 45 tissue pools (Supplementary Table S1). These samples were then homogenized with steel beads before total RNA extraction using TRIzol reagent (Invitrogen). Quantity and quality were checked before RNA library construction and sequencing. rRNA was first removed using MGIEasy rRNA Depletion Kit (MGI, China) according to manufacturer's instructions followed by library construction using MGIEasy mRNA library Prep Kit (MGI, China). Paired-end (100 bp) sequencing of each RNA library was performed on the BGISEQ-500RS sequencing platform (BGI) (Zhou et al., 2019).

RNA-sequencing generated 489,542–219,228,206 reads per library (Supplementary Table S1), which were then processed using Fastqc v0.20.0 and Trimmomatic v0.36 for quality control (Bolger et al., 2014), before *de novo* assembly using Trinity v2.5.1 (Grabherr et al., 2011). The assembled contigs were annotated by comparison to the NCBI non-redundant nucleotide (nt) and protein (nr) databases using Blastn and Diamond blastx (Buchfink et al., 2015) with e-value thresholds of 1×10^{-10} and 1×10^{-5}, respectively. Examination of the assembled contigs revealed the presence of a novel hepacivirus, here denoted *Gerbil hepacivirus* (GeHV), in multiple libraries (Supplementary Table S1).

To confirm the viral sequence, hepacivirus-associated contigs were extracted and merged by identifying homologous overlapping sequences, followed by iterative read mapping for termini extension (Hu et al., 2021). The trimmed reads were then re-mapped onto the genome using Bowtie2 (Langmead and Salzberg, 2012) and a draft consensus genome was obtained. Overlapping primers (Supplementary Table S2) were then designed based on the viral sequence obtained from NGS for nested RT-PCR and Sanger sequencing (Supplementary Fig. S1). The 3′ terminal sequence was recovered using 3′-RACE (Rapid Amplification of cDNA

* Corresponding authors.
E-mail addresses: sxpco@126.com (Y.-x. Sun), shiwf@ioz.ac.cn (W.-f. Shi), licixiu@sdfmu.edu.cn (C.-x. Li).
[1] Cui-hong An and Juan Li contributed equally to this work.

https://doi.org/10.1016/j.virs.2022.01.016
Received 28 July 2021; Accepted 10 October 2021
Available online 19 January 2022

End) (TaKaRa), while 5′-RACE was attempted unsuccessfully. Open reading frames were predicted using ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) with genes annotated by comparison to the NCBI conserved domain database (CDD). The final near-complete genome of GeHV was 9040 nucleotides (nt) in length and was predicted to encode a single polyprotein of 2826 amino acids (aa). The genes identified were those common to most hepaciviruses including: *5′-C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A,* and *NS5B-3′* (Fig. 1A). Lengths of the partial 5′ and the full 3′ untranslated regions (UTRs) were 325 bp and 234 bp, respectively. Full polyprotein of GeHV exhibited the highest amino acid identity of 64.01% with a hepacivirus (GenBank accession YP_007905733) (Fig. 1B) identified in *Peromyscus maniculatus*, which is from the same family (Muridae) as Mongolian gerbils. Eleven sequences of GeHV with complete CDS generated in this study have been deposited in GenBank under the accession numbers MZ545672–MZ545682. All raw reads generated here are available at the NCBI Sequence Read Archive (SRA) database under the BioProject accession PRJNA749072 with BioSample accessions SAMN20352295–SAMN20352339.

We next performed pairwise alignments based on the aa sequences deduced from the complete polyprotein and the ten viral genes (*C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A,* and *NS5B*) of GeHV with six related rodent hepaciviruses, respectively (Fig. 1B). Amino acid sequence

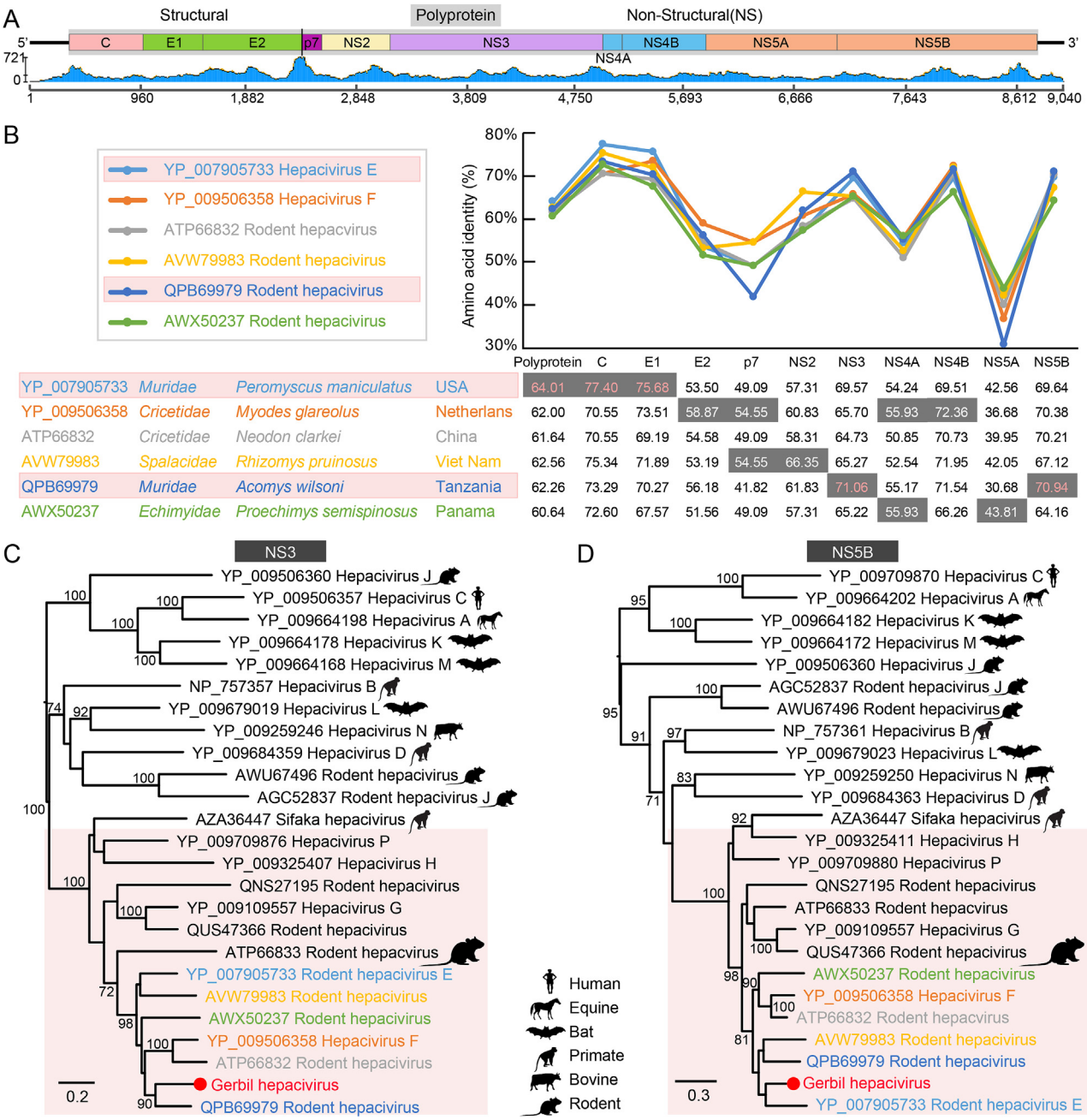

Fig. 1. Genomic characterization and phylogenetic analysis of the *Gerbil hepacivirus* (GeHV) identified in this study. **A** Genomic features and the coverage graph of the DNA reads of GeHV. **B** Comparison of amino acid sequences of the complete polyprotein and the ten viral ORFs between GeHV and representative rodent strains. The gene order (from 5′ to 3′ terminal) of the polyprotein: *C*, *E1*, *E2*, *p7*, *NS2*, *NS3*, *NS4A*, *NS4B*, *NS5A*, and *NS5B*. The two hepaciviruses with the highest amino acid identities in the polyprotein, *NS3* and *NS5B*, with GeHV were highlighted by the light red background. Maximum likelihood phylogenetic tree based on amino acid sequence of the *NS3* protein (**C**) and *NS5B* protein (**D**) from representative hepaciviruses and GeHV. The trees were midpoint rooted for clarity. Bootstrap values of >70% were shown for key nodes. GeHV isolated in this study was labeled with red dot.

Table 1
Pairwise comparisons of amino acid (up) and nucleotide (down) similarity (%) among variants of GeHVs from different libraries.

Strain	SX05/ Spleen	SX05/ Liver	SX06/ Kidney	SX04/ Liver	SX10/ Lung	SX10/ Liver	SX10/ Kidney	SX08/ Lung	SX09/ Lung	SX08/ Liver	SX08/ Kidney
SX05/Spleen		99.96	99.58	99.61	99.72	99.72	99.72	99.75	99.72	99.72	99.65
SX05/Liver	99.97		**99.54**	99.58	99.68	99.68	99.68	99.72	99.68	99.68	99.61
SX06/Kidney	98.89	98.87		99.75	99.79	99.79	99.79	99.82	99.79	99.79	99.76
SX04/Liver	98.86	98.85	99.87		99.82	99.82	99.82	99.86	99.82	99.82	99.76
SX10/Lung	98.91	98.89	99.8	99.82		**100**	**100**	99.96	99.93	99.93	99.86
SX10/Liver	98.9	98.88	99.8	99.82	**99.99**		**100**	99.96	99.93	99.93	99.86
SX10/Kidney	98.86	98.85	99.76	99.78	99.96	99.96		99.96	99.93	99.93	99.86
SX08/Lung	98.91	98.89	99.8	99.82	99.98	99.97	99.93		99.96	99.96	99.9
SX09/Lung	98.91	98.89	99.8	99.82	99.98	99.97	99.93	99.98		99.93	99.86
SX08/Liver	98.91	98.89	99.78	99.8	99.96	99.95	99.91	99.96	99.96		99.93
SX08/Kidney	98.83	**98.81**	99.7	99.71	99.87	99.86	99.82	99.87	99.87	99.87	

Note: numbers in bold means the highest and lowest similarity.

alignment was performed using MAFFT version 7 with the L–INS–i algorithm (Katoh et al., 2002) and the ambiguously aligned regions were removed using TrimAL v1.4 (Capella-Gutierrez et al., 2009). For the two conserved proteins, *NS3* protease and *NS5B* (also known as RNA dependent RNA polymerase, *RdRp*), GeHV shared aa similarities ranging from 64.73% to 71.06% for *NS3* and from 64.16% to 70.94% for *NS5B* with the six representative hepaciviruses, with the highest similarity to one rodent hepacivirus (GenBank accession QPB69979) identified in *Acomys wilsoni* (family Muridae) (Fig. 1B). However, for the *C* and *E1* genes, GeHV shared even higher aa similarities than *NS3* and *NS5B* with *Hepacivirus* E identified in *P. maniculatus* (family: Muridae) at 77.40% and 75.68%, respectively (Fig. 1B). GeHV shared the lowest similarity with other hepaciviruses in *NS5A* and *p7*, with aa identities ranging from 30.68% to 43.81% for *NS5A* and from 41.82% to 54.44% for *p7*, respectively (Fig. 1B).

To determine the evolutionary history of GeHV, we performed phylogenetic analyses based on the protein alignments of *NS3* and *NS5B* with representative species of the *Hepacivirus* genus (*Hepacivirus* A–N, and *Hepacivirus* P), as well as other related strains (Fig. 1C and D). As GeHV was detected in multiple libraries, a representative sequence from the library SX06-kidney was used. Phylogenetic trees were inferred using the maximum likelihood method implemented in IQ-TREE v1.6.1241 (Nguyen et al., 2015), employing the best-fit substitution model with 1000 bootstrap replicates. Consistent with the comparisons of amino acid sequences, GeHV clustered with hepaciviruses infecting various species of rodents, yet most closely with hepaciviruses identified from rodents within the same family (Muridae) (Fig. 1C and D). Specifically, GeHV was grouped with rodent hepacivirus (GenBank accession QPB69979) from *A. wilsoni* (family: Muridae), and they formed a sister clade to the group of *Hepacivirus* F (GenBank accession YP_009506358) and a rodent hepacivirus (GenBank accession ATP66832) in the *NS3* tree (Fig. 1C). In contrast, GeHV was grouped with *Hepacivirus* E (GenBank accession YP_007905733) from *P. maniculatus* (family: Muridae) in the *NS5B* tree (Fig. 1D). Although the topological incongruences were apparent between *NS3* and *NS5B* phylogenies, suggesting recombination events across these gene regions, recombination was not statistically supported by Recombination Detection Program. However, this detection of potential recombination events is confounded by the high level of divergence among these sequences (Fig. 1B). Taken together, our results demonstrated that GeHV represents a novel species.

It was notable that GeHV was very common in *M. unguiculatus*, and was detected in more than half of the sequenced libraries (n = 25/45 including 7 liver, 8 kidney, 6 lung, 3 spleen, and 1 heart libraries), which represented 82% of the gerbil groups (n = 9/11) (SX02, 04, 05, 06, 07, 08, 09, 10, and 11) (Supplementary Fig. S2A and Supplementary Table S1). Regarding virus abundance, estimated as Reads Per Million (RPM), the highest abundance (>100 RPM) was in libraries of kidney, spleen and liver tissues (Supplementary Fig. S2A and Supplementary Table S1), consistent with our expectation for a hepacivirus. Eleven sequences of GeHV with complete coding regions were recovered and used for phylogenetic analyses as well as pairwise comparisons to examine the genetic diversity of GeHV in Mongolian gerbils. Accordingly, GeHVs identified from individual libraries carried unique nucleotide variations (Supplementary Fig. S2B), and sequences roughly formed two distinct phylogenetic groups (Supplementary Fig. S2B). However, a comparison of nucleotide similarity showed that the overall genetic diversity of GeHV in *M. unguiculatus* was low and the viruses differed by a sequence divergence of up to 1.2% (Table 1).

In summary, we report the identification of a novel rodent hepacivirus in Mongolian gerbil (*M. unguiculatus*) from a plague zone in China, increasing our understanding of hepacivirus diversity in rodents. Since Mongolian gerbils have been reported as sensitive animal models for several viruses and a wide range of diseases, they could be potentially used as an animal model for hepacivirus study. Considering the wide distribution of Mongolian gerbils, further investigation of the genetic diversity and pathogenicity of hepaciviruses in Mongolian gerbils is clearly warranted.

Footnotes

This study was funded by the Academic Promotion Programme of Shandong First Medical University (2019QL006), and the Open Research Fund Program of CAS Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology (Grant No. 2021SPCAS002). W.S. was supported by the Taishan Scholars program of Shandong Province. We are grateful to Dr. John-Sebastian Eden (University of Sydney, Australia) for proofreading the manuscript. The authors declare no conflict of interest. This study was approved of the Shandong First Medical University & Shandong Academy of Medical Sciences. All institutional and national guidelines for the care and use of animals were followed.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virs.2022.01.016.

References

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.
Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.
Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
de Souza, W.M., Fumagalli, M.J., Sabino-Santos, G.Jr., Motta Maia, F.G., Modha, S., Teixeira Nunes, M.R., Murcia, P.R., Moraes Figueiredo, L.T., 2019. A novel hepacivirus in wild rodents from South America. Viruses 11, 297.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Hartlage, A.S., Cullen, J.M., Kapoor, A., 2016. The strange, expanding world of animal hepaciviruses. Annu Rev Virol 3, 53–75.
Hu, T., Li, J., Zhou, H., Li, C.X., Holmes, E.C., Shi, W.F., 2021. Bioinformatics resources for SARS-CoV-2 discovery and surveillance. Brief. Bioinform. 22, 631–641.
Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357-U354.
Li, W., Sun, Q., She, R., Wang, D., Duan, X., Yin, J., Ding, Y., 2009. Experimental infection of Mongolian gerbils by a genotype 4 strain of swine hepatitis E virus. J. Med. Virol. 81, 1591–1596.
Li, L.L., Liu, M.M., Shen, S., Zhang, Y.J., Xu, Y.L., Deng, H.Y., Deng, F., Duan, Z.J., 2019. Detection and characterization of a novel hepacivirus in long-tailed ground squirrels (Spermophilus undulatus) in China. Arch. Virol. 164, 2401–2410.
Liu, W., Wan, X., Zhong, W., 2007. Population dynamics of the Mongolian gerbils: seasonal patterns and interactions among density, reproduction and climate. J. Arid Environ. 68, 383–397.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Riehm, J.M., Tserennorov, D., Kiefer, D., Stuermer, I.W., Tomaso, H., Zoller, L., Otgonbaatar, D., Scholz, H.C., 2011. Yersinia pestis in small rodents, Mongolia. Emerg. Infect. Dis. 17, 1320–1322.
Simmonds, P., et al., 2017. ICTV virus taxonomy profile: Flaviviridae. J. Gen. Virol. 98, 2–3.
Zhou, H., Ma, Z., Hu, T., Bi, Y., Mamuti, A., Yu, R., Carr, M.J., Shi, M., Li, J., Sharshov, K., Gao, G.F., Shi, W., 2019. Tamdy virus in ixodid ticks infesting bactrian camels, Xinjiang, China, 2018. Emerg. Infect. Dis. 25, 2136–2138.