Virologica Sinica 41 (2026) 711–715

Contents lists available at ScienceDirect

Virologica Sinica

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

Letter

# Identification and geographic evolutionary analysis of human pegivirus infection based on metatranscriptomic sequencing

Shan Lu [a,1], Pei-Yu Zhen [a,1], Cai Bian [b,1], Ming-Zhu Zhang [a,c,1], Nan-Nan Yao [b], Xiao-Ling Su [c], Yu-Tong Huang [a], Chen Shan [a], Jia-Yi Pan [a], Lin Zhao [a,e], Yuan-Chun Zheng [b], Jia-Fu Jiang [c], Xiao-Min Zheng [b,*], Run-Ze Ye [d,*], Wu-Chun Cao [a,c,*]

[a] Institute of EcoHealth, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, China
[b] Mudanjiang Forestry Central Hospital, Mudanjiang 157000, China
[c] State Key Laboratory of Pathogen and Biosecurity, Beijing 100071, China
[d] Department of Emergency Medicine, Qilu Hospital of Shandong University, Jinan 250012, China
[e] Shandong Provincial Key Laboratory of Intelligent Monitoring, Early Warning, Prevention and Control for Infectious Diseases, Jinan 250012, China

***Dear Editor,***

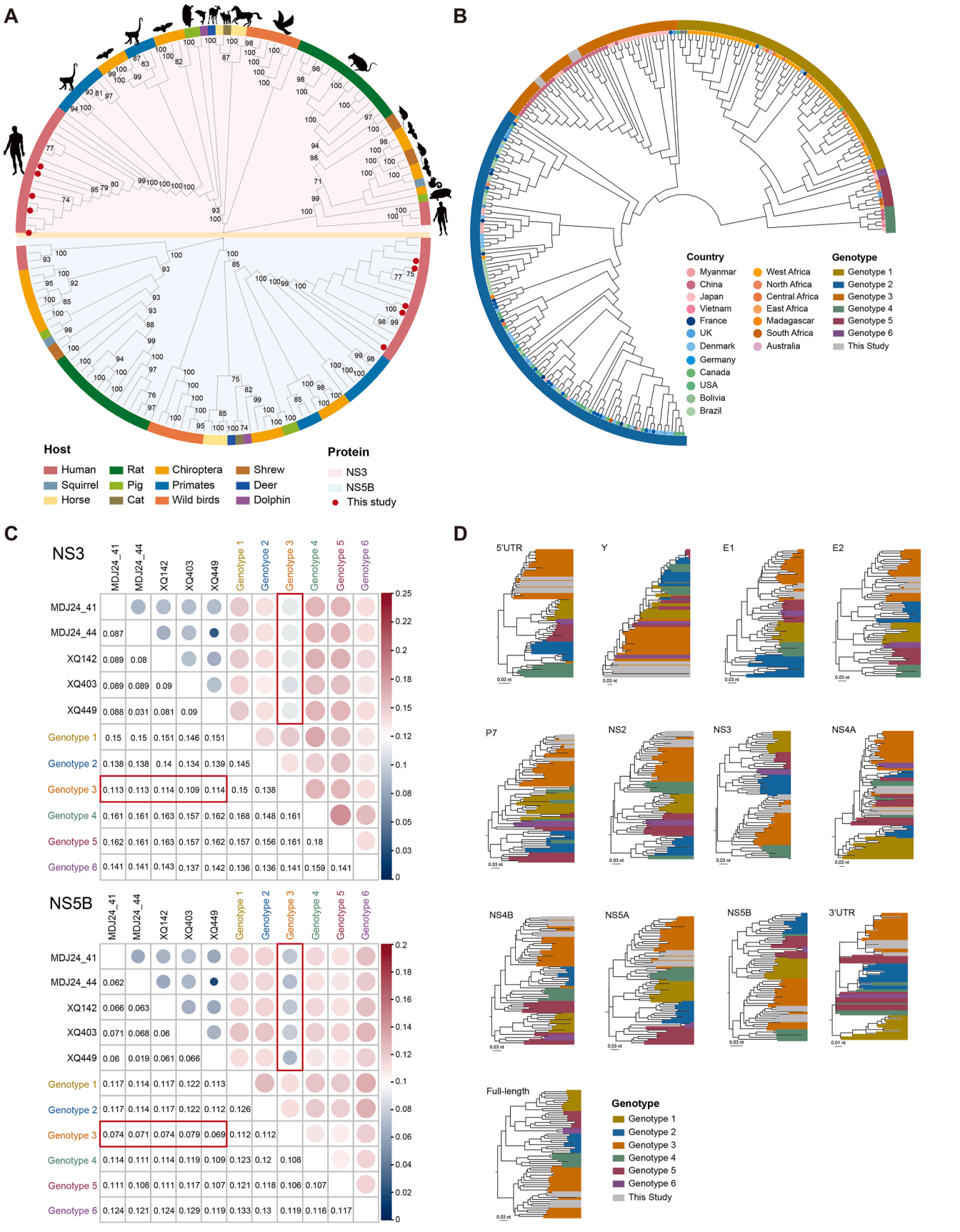
*Pegivirus hominis* (HPgV), originally called GB virus C (GBV-C) or Hepatitis G virus (HGV), is a single-stranded positive-sense RNA virus belonging to the genus *Pegivirus* within the family *Flaviviridae*, with a genome of approximately 9.4 kb in length. HPgV is currently classified into six genotypes (human pegivirus genotype 1–6) (ICTV: https://ictv.global/report/chapter/flaviviridae/flaviviridae/pegivirus). HPgV is primarily transmitted through blood exposure, sexual contact, and vertical transmission (Takayama et al., 1999; Bhanich Supapol et al., 2011; Anggorowati et al., 2013), and is capable of establishing persistent infection in humans (Chivero and Stapleton, 2015). Although it has not been definitively linked to specific clinical symptoms or diseases, numerous studies have indicated that HPgV infection is associated with a slower progression of Acquired Immunodeficiency Syndrome (AIDS) by interfering with the replication of human immunodeficiency virus type 1 (HIV-1) or enhancing host immune competence (Rey et al., 2000; Tillmann and Manns, 2001; Nunnari et al., 2003; Stapleton et al., 2012). However, some studies have proposed that HPgV co-infection may be associated with disease exacerbation in patients with encephalitis or lymphoma (Tuddenham et al., 2020; Fama et al., 2020; He et al., 2025). The prevalence of HPgV exhibits considerable variation across different regions and populations. However, it should be noted that the prevalence data reported in early literature are difficult to interpret directly, as different studies used different detection methods, and even the detection protocols were not standardized (Zhou et al., 1997; Wang et al., 2019). Given its high prevalence and uncertain clinical effects, further investigation is warranted to clarify the clinical significance of HPgV co-infection.

To date, research has primarily focused on the epidemiology of HPgV in different countries and populations, as well as its mechanisms of interaction with other viruses. In contrast, studies on its genetic diversity and geographical evolutionary distribution have been relatively limited, with the transmission and evolutionary dynamics within specific regional populations remaining insufficiently explored. In this study, HPgV was detected in serum samples from patients under active surveillance at a sentinel hospital in Heilongjiang Province, China. This study aims to elucidate the genotypic distribution and genetic variation of HPgV in Heilongjiang Province through phylogenetic analysis of the obtained viral genomes, thereby exploring its phylogeographic characteristics. Our findings can provide scientific data to support further research on the genetic evolution of HPgV.

Participants were recruited from individuals under active surveillance at the sentinel hospital (Mudanjiang Forestry Central Hospital, Heilongjiang Province, northeast China), between May 31 and July 31 in 2023 and 2024. A total of 838 serum samples were collected for metatranscriptomic sequencing to identify HPgV infections. Demographic information, clinical manifestations, underlying diseases, laboratory tests, and treatment outcomes were also recorded (Supplementary Table S1). Five complete viral genomes were recovered from the metatranscriptomic sequencing of serum samples from five distinct subjects. In this study, the seroprevalence of HPgV was 0.6%. Two meta-transcriptome-assembly genomes (MAGs), designated MDJ24-41 and MDJ24-44, were obtained from two participants enrolled in 2024 (Case 4 and Case 5). The remaining three MAGs (XQ142, XQ403, and XQ449) were obtained from three participants enrolled in 2023 (Case

* Corresponding authors.
*E-mail addresses:* hfbzxm@163.com (X.-M. Zheng), runze.ye@mail.sdu.edu.cn (R.-Z. Ye), caowuchun@126.com (W.-C. Cao).
Peer review under the responsibility of editorial board of *Virologica Sinica*.
[1] Shan Lu, Pei-Yu Zhen, Cai Bian, and Ming-Zhu Zhang contributed equally to this work.

https://doi.org/10.1016/j.virs.2026.05.005
Received 9 February 2026; Accepted 21 May 2026
Available online 27 May 2026




*(caption on next page)*

1–3). To verify the HPgV sequences identified by metatranscriptomic sequencing and obtain longer genomic fragments, specific primers were designed for real-time RT-PCR. All primer sequences used in this study are listed in Supplementary Table S2. The amplification products of the positive samples were confirmed by Sanger sequencing. All five viral sequences were deposited in GenBank under accession numbers PX993630–PX993634.

Genus-level phylogenetic analysis based on the conserved NS3 and NS5B regions showed that all five sequences clustered with HPgV. The amino acid p-distances between these strains and known HPgV strains ranged from 0.007 to 0.13 in the NS3 region and 0.006 to 0.175 in the NS5B region. Therefore, all five genomes were classified as HPgV (Fig. 1A). The host range of *Pegivirus* is broad, with the highest prevalence observed in rodents and primates. In contrast, the HPgV species identified in this study infects humans, and it is mainly transmitted through blood exposure, sexual contact, and vertical transmission.

To determine the genotype of the HPgV sequences identified in this study, we performed species-level phylogenetic analysis within HPgV. Whole-genome phylogenetic analysis demonstrated that all sequences from this study clustered with Human pegivirus genotype 3 (Fig. 1B). The nucleotide p-distances between these sequences and the genotype 3 reference sequence (GenBank: D87713) were the lowest in the conserved regions, ranging from 0.109 to 0.114 in the NS3 region and 0.069 to 0.079 in the NS5B region (Fig. 1C). Phylogeographic mapping based on global HPgV complete genome data revealed spatial clustering of genotypes 1–6. Genotype 1 was primarily distributed across African countries; genotype 2 across Europe and the Americas; genotypes 3, 4, and 6 mainly in China and Japan; and genotype 5 in Africa and Europe (Fig. 2A).

Further phylogenetic analysis of individual protein-coding regions showed that, in the NS4A region, the HPgV sequences from this study did not cluster entirely with the genotype 3 reference sequences, which was inconsistent with the patterns observed in other genomic regions and the whole-genome phylogenetic tree (Fig. 1D). To determine whether recombination contributed to this inconsistency, we performed recombination analysis using RDP4. No significant recombination signals were detected, indicating that recombination is unlikely to account for the observed discordance (Supplementary Fig. S1). Human pegivirus genotype 3 sequences that should have clustered together exhibited abnormal clustering in the phylogenetic trees of the 5′UTR, Y, NS4A, and NS4B regions (Fig. 1D), which may be associated with base mutations in these regions. This phenomenon is not uncommon in *Hepacivirus* and *Pestivirus* (Simmonds et al., 2025), yet it is often overlooked in genotyping studies. Our findings highlight the importance of considering potential discrepancies introduced by subgenomic regions when performing phylogenetic classification, particularly for viruses with high genetic diversity. Therefore, whole-genome clustering is considered to more accurately reflect the phylogenetic relationships.

Moreover, base mutations led to amino acid mutations. Therefore, the study sequences were aligned with the amino acid sequences of 50 genotype 3 reference strains, and 10 specific mutation sites were identified at positions aa223, aa411, aa587, aa706, aa734, aa802, aa1887, aa1970, aa2307, and aa2552 (Fig. 2B). In the structural protein region E2, aa223 was Pro in XQ142, and Ser in the remaining sequences. aa411 was Val in XQ449, and Gly in the remaining sequences. aa587 was Gln in MDJ24-41, and His/Asn/Asp in the remaining sequences. In the nonstructural protein region NS2, aa706 in XQ403 was Met, while the others were Val/Leu. aa734 in XQ449 was Arg, while the others were Lys. aa802 in XQ142 was Thr, while the others were Ile/Val. In the nonstructural protein region NS5A, aa1887 was Ile in XQ403, and Leu/Val in the remaining sequences. aa1970 was Ala in XQ403, and Thr/Ile in the remaining sequences. aa2307 was Val in XQ142, and Thr in the remaining sequences. aa2552 was Asp in MDJ24-44 and XQ449, and Glu/Gln in the remaining sequences. It was found that the mutation at position aa587 is present not only in HPgV genotype 3 but also in genotypes 1, 4, and 5. The mutation at aa734 is also observed in genotype 1, and the variation at aa2552 is present in genotypes 1 and 2. These mutations are speculated to be potential evolutionary markers of HPgV strains in northeast China.

Notably, among the five cases, Case 4, who presented with the most severe clinical symptoms, harbored an amino acid variation at position aa587 in the E2 region (His/Asn/Asp→Gln). However, given that HPgV is generally considered non-pathogenic in immunocompetent individuals (Yang et al., 2020) and that no consistent clinical manifestations were observed across the five cases, the relationship between these amino acid variations and clinical symptoms remains unclear. It is also worth noting that whether these cases represented acute or persistent HPgV infection could not be determined due to the lack of longitudinal samples. Further studies with larger sample sizes and longitudinal follow-up are needed to explore the potential clinical significance of these mutations.

In conclusion, we identified and characterized HPgV strains circulating in Heilongjiang Province, China, based on serum samples collected from a sentinel hospital. Phylogenetic analysis based on whole-genome sequences revealed that all obtained HPgV strains clustered within the established Human pegivirus genotype 3 clade, exhibiting high homology with previously reported strains from Asian regions, including China, Japan, and Thailand. These results suggest the potential existence of a regional transmission chain or a shared evolutionary history in East Asia. Notably, some sequences exhibited divergent clustering in the NS4A region, and amino acid alignment identified ten specific mutation sites. This study is the first report of HPgV genotype 3 strains in Heilongjiang Province, China, which expands the known geographic distribution of this virus and reveals its genetic diversity in East Asia. Furthermore, this study demonstrates that subgenomic alignments can lead to different lineage assignments—a finding that underscores the importance of using near-full-length genomes for accurate genotyping. In addition, phylogeographic analysis reveals spatial clustering in the distribution of Human pegivirus genotypes 1–6. The results of this study emphasize the necessity of continuous monitoring and further research on HPgV, particularly regarding its mechanism of action, clinical impacts, and potential for vaccine development.

However, this study also has certain limitations. The HPgV prevalence reported in this study (0.6%) was lower than the prevalence previously reported in healthy Chinese adults (Wang et al., 2019), which is likely due to the lower sensitivity of metagenomic sequencing compared with targeted RT-PCR approaches. Therefore, the true local HPgV prevalence may have been underestimated in this study. Furthermore, this study was limited to surveillance samples from a single sentinel hospital and therefore may not reflect the true prevalence of HPgV in this region. In addition, due to the limited sample size and the lack of longitudinal control samples, the association between HPgV infection and clinical symptoms requires further investigation and validation. Future studies should expand the monitoring scope, conduct in-depth analyses combined with clinical data, and explore the epidemiological and evolutionary dynamics of HPgV in high-risk environments.

**Fig. 1.** **A** Phylogenetic analysis of viruses in the genus *Pegivirus* based on the NS3 protein and NS5B protein. Phylogenetic trees with distinct background colors denoting different regions; viruses detected in this study are marked with red solid circles and the outer circle colors represent the host species of the corresponding sequences. **B** Phylogeny of *Pegivirus hominis* (HPgV) based on the full genome. The color of the inner circle represents the country of origin for each sequence, and the color of the outer circle denotes its genotype. Sequences obtained in this study are marked in gray. **C** Matrix bubble plots of nucleotide p-distance between sequences from this study and sequences of different genotypes of HPgV, based on NS3 peptide segments (up) and NS5B peptide segments (down). Bubble color and size indicate the degree of inter-sequence divergence, and the values in the cells represented the magnitude of nucleotide p-distance from different genotypes. **D** Phylogeny of HPgV based on nucleotide sequences of different peptide segments. Different colors correspond to distinct genotypes. Sequences obtained in this study are marked in gray.

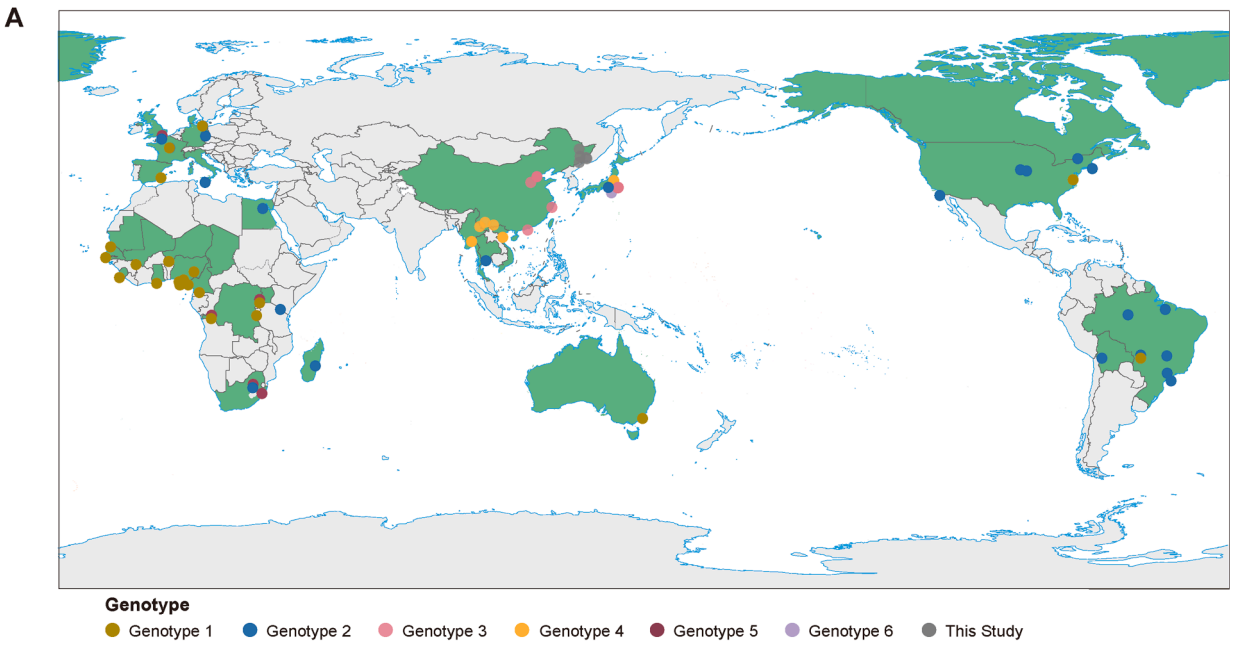




**Fig. 2.** **A** Geographic distribution of complete-genome genotypes of *Pegivirus hominis* (HPgV) worldwide. Countries with detected HPgV sequences are labeled in green. The color of each point indicated the genotypes of sequences. **B** Genomic structure and amino acid variation of the virus in this study. The top part shows the genome organization, including Y, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B regions. The lower part displays the amino acid alignment between the sequences from this study and the reference sequences of Human pegivirus genotype 3. Dots indicates amino acid identity to the reference sequences, red characters representes specific amino acid variations, and the background colors of the peptide segments corresponded to those in the genomic structure above. The reference sequence used for the alignment is K2141/Japan/1996 (GenBank accession: D87713).

## FOOTNOTES

This work was supported by the Natural Science Foundation of China (82103897), the Natural Science Foundation of Shandong Province, China (ZR2024MH182), the Postdoctoral Fellowship Program and China Postdoctoral Science Foundation (BX20240215), the China Postdoctoral Science Foundation (2025M770754), the Young Talent of Lifting engineering for Science and Technology in Shandong, China (SDAST2024QTA079), the Science and Technology Support Plan for Youth Innovation of Colleges and Universities of Shandong Province of China (2022KJ016), and the Cheeloo Young Scholar Program of Shandong University. This study was approved by the ethics committee of Mudanjiang Forest Central Hospital (Ethical Approval Number: 2024 [007]). The authors declare no conflicts of interest.

The HPgV sequences obtained in this study were deposited in GenBank under accession numbers PX993630–PX993634 and also in the Science Data Bank under accession code 10.57760/sciencedb.v.00012. Supplementary data to this article can be found online at https://doi.org/10.1016/j.virs.2026.05.005.

## REFERENCES

Anggorowati, N., Yano, Y., Subronto, Y.W., Utsumi, T., Heriyanto, D.S., Mulya, D.P., Rinonce, H.T., Widasari, D.I., Lusida, M.I., Soetjipto, Hayashi, Y., 2013. GB virus C infection in Indonesian HIV-positive patients. Microbiol. Immunol. 57, 298–308.

Bhanich Supapol, W., Remis, R.S., Raboud, J., Millson, M., Tappero, J., Kaul, R., Kulkarni, P., McConnell, M.S., Mock, P.A., McNicholl, J.M., Roongpisuthipong, A., Chotpitayasunondh, T., Shaffer, N., Butera, S., 2011. Prevalence and correlates of GB virus C infection in HIV-infected and HIV-uninfected pregnant women in Bangkok, Thailand. J. Med. Virol. 83, 33–44.

Chivero, E.T., Stapleton, J.T., 2015. Tropism of human pegivirus (formerly known as GB virus C/hepatitis G virus) and host immunomodulation: insights into a highly successful viral infection. J. Gen. Virol. 96, 1521–1532.

Fama, A., Larson, M.C., Link, B.K., Habermann, T.M., Feldman, A.L., Call, T.G., Ansell, S. M., Liebow, M., Xiang, J., Maurer, M.J., Slager, S.L., Nowakowski, G.S., Stapleton, J. T., Cerhan, J.R., 2020. Human pegivirus infection and lymphoma risk: a systematic review and meta-analysis. Clin. Infect. Dis. 71, 1221–1228.

He, J., Yang, L., Niu, C., 2025. Human pegivirus detected in patient with reversible severe encephalitis and axillary lymphadenopathy: a case report. Diagn. Pathol. 20, 66.

Nunnari, G., Nigro, L., Palermo, F., Attanasio, M., Berger, A., Doerr, H.W., Pomerantz, R. J., Cacopardo, B., 2003. Slower progression of HIV-1 infection in persons with GB virus C co-infection correlates with an intact T-helper 1 cytokine profile. Ann. Intern. Med. 139, 26–30.

Rey, D., Vidinic-Moularde, J., Meyer, P., Schmitt, C., Fritsch, S., Lang, J.M., Stoll-Keller, F., 2000. High prevalence of GB virus C/hepatitis G virus RNA and antibodies in patients infected with human immunodeficiency virus type 1. Eur. J. Clin. Microbiol. Infect. Dis. 19, 721–724.

Simmonds, P., Butković, A., Grove, J., Mayne, R., Mifsud, J.C.O., Beer, M., Bukh, J., Drexler, J.F., Kapoor, A., Lohmann, V., Smith, D.B., Stapleton, J.T., Vasilakis, N., Kuhn, J.H., 2025. Taxonomic expansion and reorganization of *Flaviviridae*. Nat. Microbiol. 10, 3026–3037.

Stapleton, J.T., Chaloner, K., Martenson, J.A., Zhang, J., Klinzman, D., Xiang, J., Sauter, W., Desai, S.N., Landay, A., 2012. GB virus C infection is associated with altered lymphocyte subset distribution and reduced T cell activation and proliferation in HIV-infected individuals. PLoS One 7, e50563.

Takayama, S., Miura, T., Tominaga, T., Taki, M., Matsuo, S., Sugii, S., Shimotohno, K., 1999. Partial nucleotide sequencing of the NS3/helicase region of hepatitis G virus to prove vertical transmission. FEMS Microbiol. Lett. 175, 273–279.

Tillmann, H.L., Manns, M.P., 2001. GB virus-C infection in patients infected with the human immunodeficiency virus. Antivir. Res. 52, 83–90.

Tuddenham, R., Eden, J.S., Gilbey, T., Dwyer, D.E., Jennings, Z., Holmes, E.C., Branley, J.M., 2020. Human pegivirus in brain tissue of a patient with encephalitis. Diagn. Microbiol. Infect. Dis. 96, 114898.

Wang, T., Chen, J., Zhang, Q., Huang, X., Xie, N., Zhang, J., Cai, T., Zhang, Y., Xiong, H., 2019. Prevalence of hepatitis G virus infection among 67,348 blood donors in mainland China. BMC Public Health 19, 685.

Yang, N., Dai, R., Zhang, X., 2020. Global prevalence of human pegivirus-1 in healthy volunteer blood donors: a systematic review and meta-analysis. Vox Sang. 115, 107–119.

Zhou, B., Ma, W., Wang, H., Fu, Y., Xu, L., Lu, L., Jiang, F., Peng, W., 1997. [Investigation on hepatitis G virus (HGV) infection among different populations in Shenzhen]. Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi 11, 348–351.