



Near-Complete Genome Sequence of a Hepatitis A Subgenotype IB Virus Isolated from Frozen Raspberries

Haifeng Chen,^a Weimin Wang,^a Shiliang Wang,^b Yuan Hu^c

^aDivision of Molecular Biology, Office of Applied Research and Safety Assessment, Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, Laurel, Maryland, USA

^bTranslational Bioinformatics, Bristol-Myers Squibb, Princeton, New Jersey, USA

^cOffice of Regulatory Affairs, Northeast Food and Feed Laboratory, Food and Drug Administration, Jamaica, New York, USA

ABSTRACT Hepatitis A virus is one of the most common causes of foodborne viral illness. Here, we report the nearly complete genome sequence of a hepatitis A virus (subgenotype IB) isolated from frozen raspberries using RNA sequencing-based metagenomics.

Hepatitis A virus (HAV) is the most frequently identified pathogen of acute viral hepatitis worldwide and is recognized as one of the most common foodborne viruses (1, 2). HAV is a small nonenveloped virus classified into the *Hepatovirus* genus within the *Picornaviridae* family. The viral genome is an approximately 7,500-nucleotide (nt) positive-sense, single-stranded RNA. The HAV genome is heterogeneous, enabling classification of the virus into six different genotypes, designated I to VI, where I, II, and III genotypes, which are divided into subgenotypes A and B, have been associated with human disease (3–5). IA is the most prevalent subgenotype in the United States, followed by IB and IIIA (6). HAV is generally acquired through the fecal-oral route, either by direct person-to-person contact or by ingestion of contaminated water or food, particularly shellfish, soft fruit, and raw vegetables (7). Although HAV is considered to be one of the leading foodborne viruses, direct detection of HAV in implicated foods is complicated by the complexity of the food matrix and low contamination levels of the virus. Foodborne HAV infection is not often reported except when triggered by an outbreak. Here, we report the nearly complete genome sequence of a subgenotype IB HAV from frozen raspberries that were collected during an infection event which occurred in New Hampshire in 2013 using a metagenomic technique.

RNA was prepared from frozen raspberries following the U.S. Food and Drug Administration Bacteriological Analytical Manual 26B protocol (8), with the exception of not using murine norovirus 1. The RNA extract was then concentrated with an Amicon ultra 0.5 centrifugal filter (Millipore, MA). RNA amplification, sequencing library preparation, and paired-end sequencing were performed as described previously (9, 10). Briefly, the prepared RNA was reverse transcribed and amplified using the Ovation RNA sequencing (RNA-seq) system v2 (NuGEN, CA), in accordance with the manufacturer's instructions. Library preparation was conducted with 1 ng of amplified products using the Nextera XT library prep kit (Illumina, CA). The prepared library was loaded in the flow cell of a 500-cycle reagent kit v2, and paired-end sequencing was run on the MiSeq platform (Illumina). The total number of sequence reads was 5,053,188, and the average read length was 194.02 bp. All reads were quality trimmed, and adapter sequences were removed using CLC Genomics Workbench 9.0 (Qiagen, MD). The trimmed reads were assembled *de novo* using SPAdes 3.8.1 with default parameters (11). We used BLASTn to compare the generated *de novo* assembled contigs against the NCBI nucleotide sequence database to identify related foodborne viruses. Contigs with

Citation Chen H, Wang W, Wang S, Hu Y. 2019. Near-complete genome sequence of a hepatitis A subgenotype IB virus isolated from frozen raspberries. *Microbiol Resour Announc* 8:e00522-19. <https://doi.org/10.1128/MRA.00522-19>.

Editor Simon Roux, DOE Joint Genome Institute

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Haifeng Chen, haifeng.chen@fda.hhs.gov.

Received 13 May 2019

Accepted 11 June 2019

Published 3 July 2019

sequence similarity to related foodborne viruses were selected for mapping assembly using Burrows-Wheeler aligner (12). A fully assembled genome of HAV was identified as being 7,475 bp long, with a G+C content of 37.86%, and the mean read depth was 60. The genome comprised a 5' untranslated region (UTR) (734 bp), a single open reading frame (ORF) (6,678 bp), and a 3' UTR (63 bp). Phylogenetic analysis of the nearly complete genome sequence using Hepatitis A Virus Genotyping Tool version 1.0 (<https://www.rivm.nl/mpf/typingtool/hav/>) revealed that the HAV associated with the contaminated frozen raspberries belonged to subgenotype IB.

Data availability. The genome sequence generated in this study was deposited in GenBank under the accession number [MK829707](https://www.ncbi.nlm.nih.gov/nuclseq/MK829707). Raw sequence data were deposited in the NCBI Sequence Read Archive under the accession number [PRJNA535492](https://www.ncbi.nlm.nih.gov/sra/PRJNA535492).

ACKNOWLEDGMENTS

We thank New Hampshire Public Health Laboratories for providing the food samples.

The findings and conclusions of this article are those of the authors and do not necessarily represent the official views or policies of the U.S. Food and Drug Administration.

REFERENCES

1. Wasley A, Fiore A, Bell BP. 2006. Hepatitis A in the era of vaccination. *Epidemiol Rev* 28:101–111. <https://doi.org/10.1093/epirev/mxj012>.
2. Vaughan G, Goncalves Rossi LM, Forbi JC, de Paula VS, Purdy MA, Xia G, Khudyakov YE. 2014. Hepatitis A virus: host interactions, molecular epidemiology and evolution. *Infect Genet Evol* 21:227–243. <https://doi.org/10.1016/j.meegid.2013.10.023>.
3. Robertson BH, Jansen RW, Khanna B, Totsuka A, Nainan OV, Siegl G, Widell A, Margolis HS, Isomura S, Ito K, Ishizu T, Moritsugu Y, Lemon SM. 1992. Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. *J Gen Virol* 73:1365–1377. <https://doi.org/10.1099/0022-1317-73-6-1365>.
4. Lu L, Ching KZ, de PV, Nakano T, Siegl G, Weitz M, Robertson BH. 2004. Characterization of the complete genomic sequence of genotype II hepatitis A virus (CF53/Berne isolate). *J Gen Virol* 85:2943–2952. <https://doi.org/10.1099/vir.0.80304-0>.
5. Costa-Mattioli M, Cristina J, Romero H, Perez-Bercoff R, Casane D, Colina R, Garcia L, Vega I, Glikman G, Romanowsky V, Castello A, Nicand E, Bassin M, Billaudel S, Ferre V. 2002. Molecular evolution of hepatitis A virus: a new classification based on the complete VP1 protein. *J Virol* 76:9515–9525. <https://doi.org/10.1128/JVI.76.18.9516-9525.2002>.
6. Nainan OV, Xia G, Vaughan G, Margolis HS. 2006. Diagnosis of hepatitis A virus infection: a molecular approach. *Clin Microbiol Rev* 19:63–79. <https://doi.org/10.1128/CMR.19.1.63-79.2006>.
7. Bosch A, Sánchez G, Abbaszadegan M, Carducci A, Guix S, Le Guyader FS, Netshikweta R, Pinto RM, van der Poel WHM, Rutjes S, Sano D, Taylor MB, van Zyl WB, Rodríguez-Lázaro D, Kovač K, Sellwood J. 2011. Analytical methods for virus detection in water and food. *Food Anal Methods* 4:4–12. <https://doi.org/10.1007/s12161-010-9161-5>.
8. U.S. Food and Drug Administration. BAM 26B appendices: multi-laboratory validation of hepatitis A virus concentration and detection protocols. <https://www.fda.gov/food/laboratory-methods-food-safety/bam-26b-appendices-multi-laboratory-validation-hepatitis-virus-concentration-and-detection-protocols>.
9. Chen H, Wang S, Wang W. 2018. Complete genome sequence of a human norovirus strain from the United States classified as genotype GII.P6_GII.6. *Genome Announc* 6:e00489-18. <https://doi.org/10.1128/genomeA.00489-18>.
10. Hu Y, Chen H. 2018. Whole-genome sequences of two human norovirus GII.4 variants isolated in the United States. *Microbiol Resour Announc* 7:e00840-18. <https://doi.org/10.1128/MRA.00840-18>.
11. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
12. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.