



# Genome Sequence of a Human Norovirus GII.4 Hong Kong[P31] Variant in Hong Kong, China

Erin H. Y. Tse,<sup>a</sup> Lin-Yao Zhang,<sup>a</sup> Sin-Leung Lau,<sup>a</sup>  Martin Chi-Wai Chan<sup>a,b,c</sup>

<sup>a</sup>Department of Microbiology, Chinese University of Hong Kong, Hong Kong, China

<sup>b</sup>Stanley Ho Centre for Emerging Infectious Diseases, Chinese University of Hong Kong, Hong Kong, China

<sup>c</sup>Li Ka Shing Institute of Health Sciences, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong, China

**ABSTRACT** We report the nearly complete genome of a norovirus GII.4 Hong Kong[P31] variant (GII strain Hu/HK/2019/GII.4 Hong Kong[P31]/CUHK-NS-2200) that was detected in a patient with gastroenteritis in August 2019. The genome was sequenced by metagenomic next-generation sequencing and was found to have 92.8% nucleotide similarity to the closest GII.4 norovirus sequence in GenBank.

Human noroviruses, belonging to the *Norovirus* genus and the *Caliciviridae* family, are a major cause of gastroenteritis worldwide (1), posing a significant risk to young children, the elderly, and immunocompromised individuals (2). Genetically diverse noroviruses are categorized into 10 genogroups, of which GI, GII, GIV, GVIII, and GIX are known to infect humans (3). Most notably, GII.4 strains have dominated norovirus outbreaks since 2002, with new epidemic variants usually emerging every 2 to 3 years, although the most recent new variant appeared in 2012 (4). Here, we report the nearly complete genome of a norovirus GII.4 variant designated GII.4 Hong Kong[P31].

This norovirus GII strain (Hu/HK/2019/GII.4 Hong Kong[P31]/CUHK-NS-2200) was detected by PCR, as part of our ongoing norovirus surveillance study, in a stool specimen collected in August 2019 from a 42-year-old female patient hospitalized with acute gastroenteritis in Hong Kong, China. A metagenomic next-generation sequencing approach was used to determine the complete viral genome, as we described previously (5). A 0.65- $\mu$ m-filtered stool suspension was digested with DNase, RNase, and benzonase to enrich the virus by eliminating unprotected host and microbial nucleic acids. Viral RNA was extracted using the MagMAX viral RNA isolation kit (Thermo Fisher), and cDNA synthesis was performed with tagged random octamers and SuperScript III reverse transcriptase (Thermo Fisher), followed by second-strand cDNA synthesis using a Klenow fragment (TaKaRa). The cDNA was purified using 1.8 $\times$  AMPure magnetic beads (Beckman Coulter, Inc.) and amplified using Phusion Hot Start II high-fidelity DNA polymerase (Thermo Fisher). A library prepared using the Nextera XT DNA library preparation kit (Illumina) was subjected to paired-end sequencing (2  $\times$  75 bp) on the NextSeq 500 system (Illumina).

A total of 1,226,904 paired-end reads were generated. Reads were trimmed using Trimmomatic 0.36 (6), with a passing rate of 97.0%. Reference sequence mapping was performed using Geneious R9 against the reference strain GII.4 Sydney[P31] (GenBank accession number [JX459908](https://doi.org/10.1128/MRA.01391-19)), of which 8.5% of reads were mapped to norovirus. Taxonomic classification using Taxonomer (7) revealed that norovirus was the only pathogenic mammalian virus detected, with the remaining 7.7% of viral reads being matched to plant viruses. Genome coverage was 95.1%, and the mean sequencing depth was 1,980 $\times$ . The viral sequence was confirmed by *de novo* assembly using SPAdes 3.13.0 (8) in metagenomic mode and by full-length virus capsid Sanger se-

**Citation** Tse EHY, Zhang L-Y, Lau S-L, Chan MC-W. 2020. Genome sequence of a human norovirus GII.4 Hong Kong[P31] variant in Hong Kong, China. *Microbiol Resour Announc* 9:e01391-19. <https://doi.org/10.1128/MRA.01391-19>.

**Editor** Jelle Matthijnssens, KU Leuven

**Copyright** © 2020 Tse et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Martin Chi-Wai Chan, [martin.chan@cuhk.edu.hk](mailto:martin.chan@cuhk.edu.hk).

**Received** 7 November 2019

**Accepted** 20 December 2019

**Published** 16 January 2020

quencing (5). The two methods yielded consistent results. The terminal 5' end of the genome, with no coverage, was determined by amplicon Sanger sequencing (with the first 22 nucleotides corresponding to the forward primer binding region), while the polyadenylated terminal 3' end was sequenced by the rapid amplification of cDNA ends (RACE) PCR method, using a tagged oligo(dT) primer.

The nearly complete genome of our strain is 7,556 nucleotides long (GC content, 49%), excluding the poly(A) tail, and it carries a type P31 polymerase. The best BLAST hit in GenBank was a norovirus GII.4 Osaka[P31] variant from Japan (Hu/GII.4/Osaka3/2007/JP [GenBank accession number [AB541323](https://doi.org/10.1016/S1473-3099(14)70767-4)]), with 92.8% identity at the nucleotide level. Our strain shares 92.6% identity at the amino acid level with the closest GII.4 Osaka variant from Australia (Hu/GII.4/NSW3901/2008/AU [GenBank accession number [GQ845369](https://doi.org/10.1099/jgv.0.001318)]) for the virus capsid, and this capsid was designated GII.4 Hong Kong.

The detection of a GII.4 Hong Kong[P31] strain in a severe case of gastroenteritis is atypical for a middle-aged individual without any known risk factors. Whether this variant emerges in the future to cause a major epidemic remains to be seen.

**Data availability.** The genome sequence of the norovirus GII.4 Hong Kong[P31] strain (Hu/HK/2019/GII.4 Hong Kong[P31]/CUHK-NS-2200) has been deposited in GenBank under accession number [MN400355](https://doi.org/10.1099/jgv.0.001318). The raw Illumina reads are available under BioProject accession number [PRJNA588106](https://doi.org/10.1099/jgv.0.001318).

## ACKNOWLEDGMENTS

We thank the Core Utilities of Cancer Genomics and Pathobiology for providing the next-generation sequencing facilities and assistance in support of this research.

M.C.-W.C. designed and supervised the study. L.-Y.Z. and S.-L.L. performed the experiments. E.H.Y.T. and M.C.-W.C. analyzed the data. E.H.Y.T. drafted the manuscript. All authors critically reviewed the manuscript and approved the final revision. M.C.-W.C. has access to all data and is responsible for the scientific integrity of this work.

This research received no specific grant from any funding agency in the public, commercial, or not-for profit sectors.

## REFERENCES

- Ahmed SM, Hall AJ, Robinson AE, Verhoef L, Premkumar P, Parashar UD, Koopmans M, Lopman BA. 2014. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis* 14:725–730. [https://doi.org/10.1016/S1473-3099\(14\)70767-4](https://doi.org/10.1016/S1473-3099(14)70767-4).
- Hall AJ, Lopman BA, Payne DC, Patel MM, Gastanaduy PA, Vinje J, Parashar UD. 2013. Norovirus disease in the United States. *Emerg Infect Dis* 19:1198–1205. <https://doi.org/10.3201/eid1908.130465>.
- Chhabra P, de Graaf M, Parra GI, Chan MC-W, Green K, Martella V, Wang Q, White PA, Katayama K, Vennema H, Koopmans MPG, Vinje J. 2019. Updated classification of norovirus genogroups and genotypes. *J Gen Virol* 100:1393–1406. <https://doi.org/10.1099/jgv.0.001318>.
- Karst SM, Baric RS. 2015. What is the reservoir of emergent human norovirus strains? *J Virol* 89:5756–5759. <https://doi.org/10.1128/JVI.03063-14>.
- Chan MCW, Kwok K, Hung TN, Chan LY, Chan PKS. 2017. Complete genome sequence of an emergent recombinant GII.4-P16-GII.2 norovirus strain associated with an epidemic spread in the winter of 2016–2017 in Hong Kong, China. *Genome Announc* 5:e00343-17. <https://doi.org/10.1128/genomeA.00343-17>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Flygare S, Simmon K, Miller C, Qiao Y, Kennedy B, Di Sera T, Graf EH, Tardif KD, Kapusta A, Rynearson S, Stockmann C, Queen K, Tong S, Voelkerding KV, Blaschke A, Byington CL, Jain S, Pavia A, Ampofo K, Eilbeck K, Marth G, Yandell M, Schlager R. 2016. Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol* 17:111. <https://doi.org/10.1186/s13059-016-0969-1>.
- 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>.