




Genome Sequences of SARS-CoV-2 Strains Detected in Hong Kong

 Chun Hang Au,^a Wai Sing Chan,^a Ho Yin Lam,^a Dona N. Ho,^a Simon Y. M. Lam,^a Jonpaul S. T. Zee,^a Tsun Leung Chan,^a Edmond S. K. Ma^a

^aMolecular Pathology Division, Department of Pathology, Hong Kong Sanatorium & Hospital, Hong Kong

ABSTRACT We sequenced severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genomes from deep throat saliva samples of three imported cases in Hong Kong by Nanopore sequencing. Epidemiological and clinical features of these coronavirus disease 2019 (COVID-19) cases were presented for genomic epidemiology studies.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; family *Coronaviridae* and genus *Betacoronavirus*) is causing an ongoing coronavirus disease 2019 (COVID-19) pandemic. There were 1,109 COVID-19 cases in Hong Kong as of 12 June 2020 (https://www.chp.gov.hk/files/pdf/local_situation_covid19_en.pdf), with most cases reported in March to April 2020. We report here the genome sequences of SARS-CoV-2 strains detected from three imported cases in Hong Kong.

Case 1 (58-year-old man, HKSH0003) and case 2 (61-year-old female, HKSH0004) are a couple who had symptom onset on 10 March 2020 after returning from the United States to Hong Kong and tested positive on 14 March. Case 1 was relatively asymptomatic apart from fever, but case 2 showed symptoms and signs of pneumonia during hospitalization. They were enrolled in a clinical trial of interferon beta-1b, lopinavir-ritonavir, and ribavirin (1) and were discharged after two negative PCR tests. Case 3 (59-year-old man, HKSH0007) had symptom onset on 21 March after returning from Germany to Hong Kong and tested positive on 1 April. He showed very mild symptoms and denied treatment during isolation. The patient was discharged after two negative PCR tests. Informed consent was obtained from every case. This study was reviewed and approved by the Research Committee of Hong Kong Sanatorium & Hospital (reference number RC-2020-11).

Deep throat saliva samples treated with Sputasol (Thermo Scientific, USA) were subjected to nucleic acid extraction using eMAG (bioMérieux, France). The full genome was amplified following the ARTIC nCoV-2019 V1 protocol (<https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w>). Briefly, cDNA was synthesized using a SuperScript III first-strand synthesis system (Invitrogen, USA) with random hexamers. Two multiplex PCRs of 49 amplicons (V1 panel) were performed using AmpliTaq Gold DNA polymerase (Applied Biosystems, USA), combined, purified using AMPure XP beads (Beckman Coulter, USA), and quantified using a Qubit double-stranded DNA (dsDNA) high-sensitivity assay (Invitrogen, USA). Each genome was sequenced separately. For each sample, a separate Nanopore sequencing library was prepared using a ligation sequencing kit (SQK-LSK109; Oxford Nanopore Technologies, UK), sequenced using a new MinION R9.4.1 flow cell (FLO-MIN106D) with MinKNOW live base calling version 19.12.2 (fast model), and analyzed following the ARTIC nCoV-2019 bioinformatics protocol. Briefly, filtered reads (mean quality score, ≥ 7 ; length, 400 to 700 bp) were mapped to the Wuhan-Hu-1 reference strain (GenBank accession number [MN908947.3](https://www.ncbi.nlm.nih.gov/nuccore/MN908947.3)) using Minimap2 version 2.17 (parameter -x map-ont) and subjected to primer soft-

Citation Au CH, Chan WS, Lam HY, Ho DN, Lam SYM, Zee JST, Chan TL, Ma ESK. 2020. Genome sequences of SARS-CoV-2 strains detected in Hong Kong. *Microbiol Resour Announc* 9:e00697-20. <https://doi.org/10.1128/MRA.00697-20>.

Editor John J. Dennehy, Queens College

Copyright © 2020 Au 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 Edmond S. K. Ma, eskma@hksh.com.

Received 15 June 2020

Accepted 13 July 2020

Published 30 July 2020

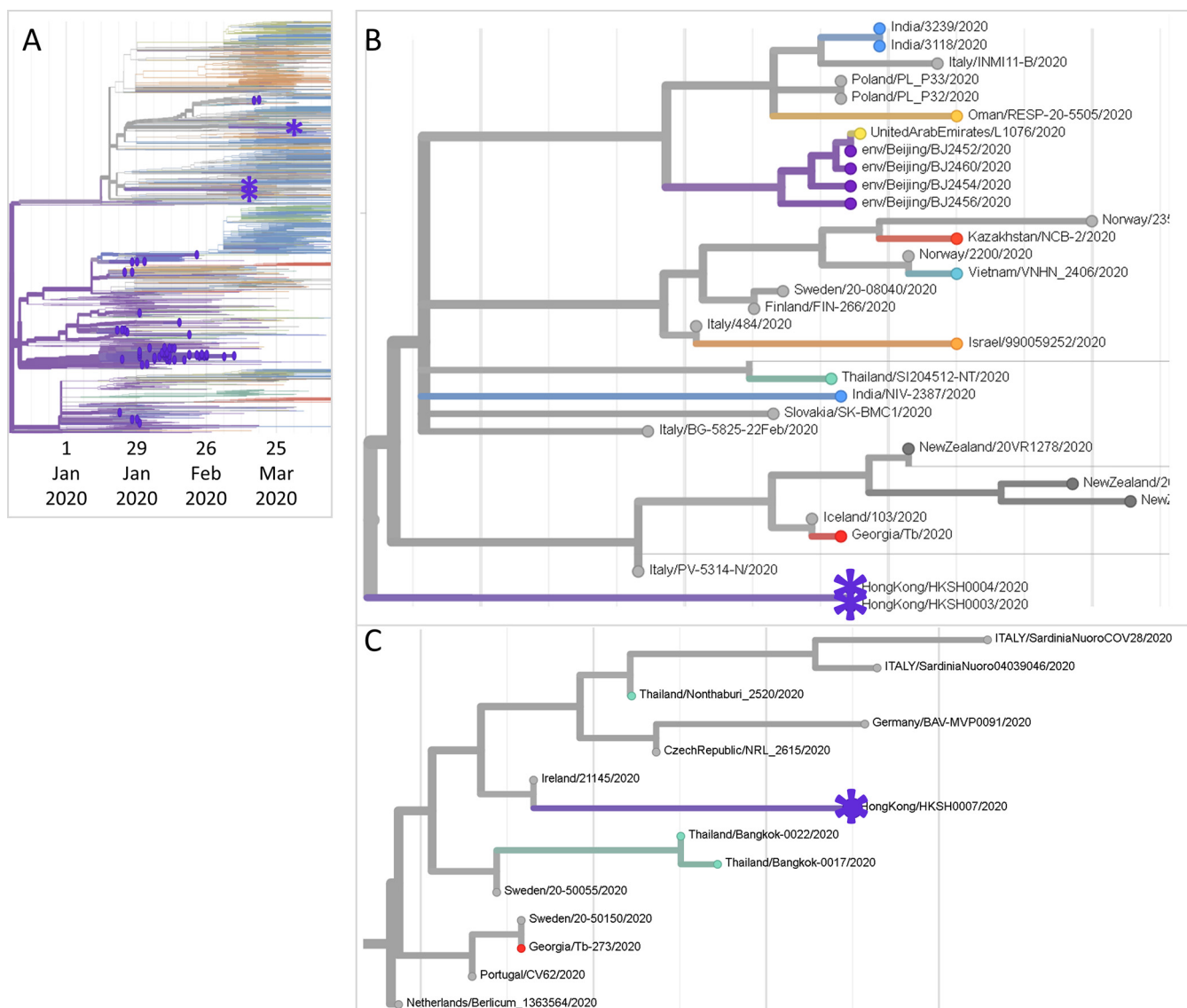


FIG 1 (A) Maximum-likelihood time-resolved phylogenetic tree of 2,949 SARS-CoV-2 genome sequences sampled up to 16 April 2020 in the GISAID database on 12 June 2020 (<https://nextstrain.org/ncov/global?dmax=2020-04-16>). Genome sequences sampled in Hong Kong are indicated as purple asterisks for the three sequences reported in this announcement and purple circles for the other 72 sequences. Phylogeny was rooted relative to early sequences from Wuhan, China. All tree branches are scaled according to sampling times (terminal nodes) and the most likely time of divergence (internal nodes). (B) Close-up view of the phylogeny (<https://nextstrain.org/ncov/asia?dmax=2020-04-16>) of case 1 (HKSH0003, purple asterisk), case 2 (HKSH0004, purple asterisk), and 29 closely related sequences (gray for non-Asian samples and other colors for Asian samples) (late February 2020 to mid-April 2020). (C) Close-up view of the phylogeny (<https://nextstrain.org/ncov/asia?dmax=2020-04-16>) of case 3 (HKSH0007, purple asterisk) and 13 closely related sequences (gray for non-Asian samples and other colors for Asian samples) (late February 2020 to mid-April 2020).

clipping using align_trim (version 1.1.0) (<https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html>) or BAMClipper version 1.1.1 (parameter -u 40 -d 40) (2). Consensus genomes were built from Nanopolish version 0.12.5 (3) variant calls. Maximum-likelihood time-resolved phylogenetic analysis was conducted using Tree-Time (4) and Nextstrain (5) (<https://github.com/nextstrain/ncov>) on 12 June 2020 (default GTR substitution model and nucleotide substitution rate of 0.0008 per site per year). Default parameters were used for all software unless otherwise specified.

Nanopore sequencing runs yielded 0.497-Gbp reads for HKSH0003 (853,213 reads; mean length, 582 bp), 0.534-Gbp reads for HKSH0004 (781,048 reads; mean length, 684 bp), and 7.740-Gbp reads for HKSH0007 (16,122,660 reads; mean length, 480 bp). All three SARS-CoV-2 genomes were sequenced at a mean depth of >3,500×, reference coverage of 92% to 96%, length of 27,384 to 28,723 nucleotides, and GC content of

38%. Phylogenetic analysis showed that they descended from European samples (88%, 88%, and 100% confidence for cases 1, 2, and 3, respectively) and were distinct from most other Hong Kong sequences (Fig. 1A). Although the disease severities of cases 1 and 2 were different, there was no detectable difference between their SARS-CoV-2 genomes (Fig. 1B). This suggests that host factors played a more important role in clinical outcome than viral genetic variation, as shown by emerging clinical studies (6). The phylogeny of genomes related to case 3 was also consistent with a pattern of international dissemination by air travel (Fig. 1C).

Data availability. These sequences were deposited in the Global Initiative on Sharing All Influenza Data (GISAID) database (strain identifiers EPI_ISL_430018, EPI_ISL_430063, and EPI_ISL_451957) and in NCBI GenBank under the accession numbers [MT510744](#), [MT510745](#), and [MT517419](#), BioProject number [PRJNA634965](#), BioSample numbers [SAMN15013847](#), [SAMN15013848](#), and [SAMN15028143](#), and SRA numbers [SRR11844878](#), [SRR11844877](#), and [SRR11852109](#).

ACKNOWLEDGMENTS

We thank all the institutions that shared their SARS-CoV-2 sequences with the GISAID database (a full list of institutions is available from <https://github.com/tommyau/sars-cov-2>). We thank our colleagues at the Molecular Pathology Division, Department of Pathology, Hong Kong Sanatorium & Hospital, for their dedicated and professional work on routine laboratory diagnostics.

REFERENCES

- Hung IF, Lung KC, Tso EY, Liu R, Chung TW, Chu MY, Ng YY, Lo J, Chan J, Tam AR, Shum HP, Chan V, Wu AK, Sin KM, Leung WS, Law WL, Lung DC, Sin S, Yeung P, Yip CC, Zhang RR, Fung AY, Yan EY, Leung KH, Ip JD, Chu AW, Chan WM, Ng AC, Lee R, Fung K, Yeung A, Wu TC, Chan JW, Yan WW, Chan WM, Chan JF, Lie AK, Tsang OT, Cheng VC, Que TL, Lau CS, Chan KH, To KK, Yuen KY. 2020. Triple combination of interferon beta-1b, lopinavir-ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *Lancet* 395: 1695–1704. [https://doi.org/10.1016/S0140-6736\(20\)31042-4](https://doi.org/10.1016/S0140-6736(20)31042-4).
- Au CH, Ho DN, Kwong A, Chan TL, Ma ESK. 2017. BAMClipper: removing primers from alignments to minimize false-negative mutations in amplicon next-generation sequencing. *Sci Rep* 7:1567. <https://doi.org/10.1038/s41598-017-01703-6>.
- Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods* 12:733–735. <https://doi.org/10.1038/nmeth.3444>.
- Sagunenko P, Puller V, Neher RA. 2018. TreeTime: maximum-likelihood phylodynamic analysis. *Virus Evol* 4:vex042. <https://doi.org/10.1093/ve/vex042>.
- Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagunenko P, Bedford T, Neher RA. 2018. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34:4121–4123. <https://doi.org/10.1093/bioinformatics/bty407>.
- Zhang X, Tan Y, Ling Y, Lu G, Liu F, Yi Z, Jia X, Wu M, Shi B, Xu S, Chen J, Wang W, Chen B, Jiang L, Yu S, Lu J, Wang J, Xu M, Yuan Z, Zhang Q, Zhang X, Zhao G, Wang S, Chen S, Lu H. 2020. Viral and host factors related to the clinical outcome of COVID-19. *Nature* 583:437–440. <https://doi.org/10.1038/s41586-020-2355-0>.