



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Original article

## Intra-host non-synonymous diversity at a neutralizing antibody epitope of SARS-CoV-2 spike protein N-terminal domain

Jonathan Daniel Ip<sup>1</sup>, Kin-Hang Kok<sup>1</sup>, Wan-Mui Chan<sup>1</sup>, Allen Wing-Ho Chu<sup>1</sup>,  
Wai-Lan Wu<sup>1</sup>, Cyril Chik-Yan Yip<sup>2</sup>, Wing-Kin To<sup>3</sup>, Owen Tak-Yin Tsang<sup>4</sup>,  
Wai-Shing Leung<sup>4</sup>, Thomas Shiu-Hong Chik<sup>4</sup>, Kwok-Hung Chan<sup>1</sup>, Ivan Fan-Ngai Hung<sup>5</sup>,  
Kwok-Yung Yuen<sup>1,2</sup>, Kelvin Kai-Wang To<sup>1,2,\*</sup>

<sup>1</sup> State Key Laboratory for Emerging Infectious Diseases, Carol Yu Centre for Infection, Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China

<sup>2</sup> Department of Microbiology, Queen Mary Hospital, Hong Kong Special Administrative Region, China

<sup>3</sup> Department of Pathology, Princess Margaret Hospital, Hong Kong Special Administrative Region, China

<sup>4</sup> Department of Medicine and Geriatrics, Princess Margaret Hospital, Hong Kong Special Administrative Region, China

<sup>5</sup> Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China

## ARTICLE INFO

## Article history:

Received 18 July 2020

Received in revised form

22 September 2020

Accepted 25 October 2020

Available online 2 November 2020

Editor: M. Cevik

## Keywords:

COVID-19

Illumina sequencing

Intra-host diversity

Nanopore sequencing

Neutralizing antibody

Non-synonymous mutation

Spike

## ABSTRACT

**Objectives:** SARS-CoV-2 has evolved rapidly into several genetic clusters. However, data on mutations during the course of infection are scarce. This study aims to determine viral genome diversity in serial samples of COVID-19 patients.

**Methods:** Targeted deep sequencing of the spike gene was performed on serial respiratory specimens from COVID-19 patients using nanopore and Illumina sequencing. Sanger sequencing was then performed to confirm the single nucleotide polymorphisms.

**Results:** A total of 28 serial respiratory specimens from 12 patients were successfully sequenced using nanopore and Illumina sequencing. A 75-year-old patient with severe disease had a mutation, G22017T, identified in the second specimen. The frequency of G22017T increased from  $\leq 5\%$  (nanopore: 3.8%; Illumina: 5%) from the first respiratory tract specimen (sputum) to  $\geq 60\%$  (nanopore: 67.7%; Illumina: 60.4%) in the second specimen (saliva; collected 2 days after the first specimen). The difference in G22017T frequency was also confirmed by Sanger sequencing. G22017T corresponds to W152L amino acid mutation in the spike protein which was only found in  $<0.03\%$  of the sequences deposited into a public database. Spike amino acid residue 152 is located within the N-terminal domain, which mediates the binding of a neutralizing antibody.

**Discussion:** A spike protein amino acid mutation W152L located within a neutralizing epitope has appeared naturally in a patient. Our study demonstrated that monitoring of serial specimens is important in identifying hotspots of mutations, especially those occurring at neutralizing epitopes which may affect the therapeutic efficacy of monoclonal antibodies. **Jonathan Daniel Ip, Clin Microbiol Infect 2021;27:1350.e1–1350.e5**

© 2020 European Society of Clinical Microbiology and Infectious Diseases. Published by Elsevier Ltd. All rights reserved.

## Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread rapidly, resulting in more than 28 million laboratory-

confirmed COVID-19 cases globally as of 14 September 2020. SARS-CoV-2 mainly causes respiratory tract infection, although extrapulmonary manifestations have been reported [1]. The efficient person-to-person transmission may be related to the high viral load shortly after symptom onset and the large number of asymptomatic individuals [2,3]. Prolonged persistence of RNA in body fluids is common despite the development of neutralizing antibodies [2].

\* Corresponding author. Kelvin Kai-Wang To, Department of Microbiology, 19th Floor, Block T, Queen Mary Hospital, Pokfulam Road, Pokfulam, Hong Kong Special Administrative Region.

E-mail address: [kelvinto@hku.hk](mailto:kelvinto@hku.hk) (K.K.-W. To).

As an RNA virus, the genome replication of SARS-CoV-2 is prone to error, and gene mutations arise frequently. Whole genome sequencing showed that the viral genomes may differ between family members [4]. Major genetic diversity has already been seen [5]. Phylogenetic analysis has demonstrated that even patients within the same geographical region were infected with genetically diverse SARS-CoV-2 strains [6].

Most studies on viral genome focus on the population level. However, understanding viral mutations on an individual patient level can have important implications for pathogenesis and treatment of virus infection. For A(H1N1)pdm09 virus, the D225G substitution of haemagglutinin was more enriched in the lower respiratory tract specimens because this substitution enables the virus to bind more efficiently to  $\alpha$ 2,3 sialic acid receptors that are predominantly found in the alveoli [7]. Furthermore, analysis of serial specimens allows the detection of resistant mutants arising during antiviral treatment [8]. A recent analysis of bronchoalveolar fluid from COVID-19 patients showed that intra-host viral genome variation is a rare event, with a median of 1 nucleotide variant with minor allele frequency of  $\geq 20\%$  [9], but whether mutations arise during the course of illness was not examined.

In this study, we analysed serial specimens from individual patients to search for nucleotide and amino acid variations that develop over the course of illness. We have specifically chosen the spike protein for targeted sequencing because the spike protein is responsible for the binding of the virus to the host cell angiotensin-converting enzyme 2 (ACE2) receptor and membrane fusion [10]. Our recent study has demonstrated that the level of IgG against spike protein RBD has a high correlation with neutralizing antibody titre [11].

## Patients and methods

### Patients

This study included archived respiratory specimens from COVID-19 patients with laboratory confirmation by real-time reverse transcription-polymerase chain reaction (RT-PCR) targeting the E or RdRp-Hel gene as described previously [12,13]. The respiratory specimens included nasopharyngeal swab, saliva, endotracheal aspirate and sputum. Serial specimens were defined as specimens collected from the same patient at least 1 day apart. Some patients in this study were enrolled in our previous randomized controlled trial on triple therapy with interferon- $\beta$ 1b, lopinavir-ritonavir and ribavirin [14], or in our previous study on viral load and serological profile [2]. This study has been approved by the Institutional Review Board of The University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 13-372).

### SARS-CoV-2 spike gene RT-PCR

Viral RNA was extracted from respiratory specimens using Qiagen Viral RNA Mini Kit (Hilden, Germany) or BioMérieux NucliSENS easyMag (Marcy-l'Étoile, France) as described previously [4]. Full-length spike gene was amplified using SuperScript™ III one-step RT-PCR system with Platinum™ Taq High Fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, MA, USA) using primer set 1, or primer set 2, and 3 (Table S1) under the following RT-PCR conditions: 50°C for 30 min, 94°C for 2 min; then 40 cycles of 94°C for 30 s, 55°C for 30 s and 68°C for 250 s. Amplified PCR products were either purified by 0.5x AMPure XP bead (Beckman Coulter, California, CA, USA) or gel purified by QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Purified PCR products were then subjected to nanopore or Illumina library preparations.

Since RT-PCR may lead to bias in assessing frequency of nucleotide variants, we performed separate RT-PCRs for nanopore and Illumina sequencing.

### Nanopore targeted sequencing

DNA libraries were prepared using the Ligation Sequencing Kit (SQK-LSK109, Oxford Nanopore Technologies) according to the manufacturer's instructions with modifications [14]. Briefly, amplified and purified spike gene PCR products were subjected to PCR barcoding using PCR Barcoding Expansion 1-96 kit (EXP-PBC096). Barcoded samples were then pooled at equal molar ratio prior to end repair, sequencing adapter ligation and clean-up. The end repair incubation time was prolonged from the manufacturer's recommendation of 5 min at 20°C and 5 min at 65°C, to 30 min at 20°C and 30 min at 65°C. The Oxford Nanopore MinION platform and R9.4.1 flow cell were used for sequencing.

Guppy v3.4.5 was used in converting the raw signal data into FASTQ format, demultiplexing and removing nanopore adaptor sequences. Only reads with a minimum Q score of 10 were included for subsequent analysis. The sequencing run was quality checked using MinIONQC [15]. Quality-checked reads were filtered by length to remove primer dimer and chimera. The filtered reads were then aligned with the SARS-CoV-2 reference genome Wuhan-Hu-1 (GenBank accession number MN908947.3) using the Burrows–Wheeler Aligner (BWA) [16]. SAMtools v1.10 and BCftools mpileup were used in creating a variant file [17,18]. Only reads with mapping quality  $\geq 30$  and base quality  $\geq 20$  were used in piling. BCftools [17] call, vcfutils.pl [18] and SEQTK [19] were used in generating the FASTA consensus sequence. The variation frequencies were obtained using SAMtools v1.10 [18] and VarScan2 [20] with reads of mapping quality score  $\geq 30$ , basecalling quality score  $\geq 20$ . Only single nucleotide polymorphisms (SNPs) with a minimum of 250  $\times$  coverage,  $p < 0.001$ , minimum variant frequency of 0.5%, minimum average basecalling quality score of 30 and minimum read support for variant of 10 reads were reported. The raw data have been deposited in the NCBI Sequence Read Archive with accession number PRJNA664839.

### Targeted sequencing of spike gene using Illumina sequencing

For Illumina sequencing, DNA libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina) following the manufacturer's protocol with 1 ng of the spike gene cDNA as the starting material. In brief, DNA fragments with adaptors were generated by fragmentation reaction at 55°C for 5 min. The fragmented DNA were then indexed and amplified in a 50  $\mu$ L reaction volume with 12 cycles of PCR, followed by AMPure XP bead clean-up. The quality of the enriched libraries was then validated using Agilent Bioanalyzer, Qubit and real-time RT-qPCR.

The libraries were pooled at equal molar ratio, denatured and diluted to optimal concentration prior to sequencing. The Illumina NovaSeq 6000 was used for sequencing to generate pair-end 151 bp reads. The whole dataset was deposited in the NCBI Sequence Read Archive with accession number PRJNA664541.

Illumina adapters were removed from the reads. Any reads with length of at least 100 bp and at least 90% of bases with quality score of  $\geq 30$  were retained during the quality filtering process using FASTP [21]. Pair-end reads were aligned with the reference genome SARS-CoV-2 Wuhan-Hu-1 (MN908947.3) using the BWA [16]. Sorting and read deduplication were performed using SAMtools v1.10 [18] and Picard v2.22.3. The consensus sequence was generated using SEQTK [19] and BCFTools [17] with reads of mapping quality and basecalling quality score  $\geq 30$ . The coverage depth is generated by SAMtools depth command. The variation frequencies

were obtained using SAMtools v1.10 [18] and VarScan2 [20] with reads of mapping quality score  $\geq 30$  and basecalling quality score  $\geq 35$ . Only SNPs with a minimum of  $250 \times$  coverage  $p < 0.001$ , minimum variant frequency of 0.5%, minimum average basecalling quality score of 30 and minimum read support for variant of 10 reads were reported.

### Sanger sequencing

Sanger sequencing was performed as described previously [22]. Both strands were sequenced twice with an ABI Prism 3730xl DNA Analyser (Applied Biosystems, Foster City, CA, USA) using forward and reverse primers.

### Consensus sequence

The consensus sequences were deposited into GISAID (Accession number EPI\_ISL\_538524–538551).

## Results

This study included 21 patients with at least two serial respiratory tract specimens available. A total of 98 serial respiratory tract specimens, collected between January 23 and March 16, 2020, were retrieved. SARS-CoV-2 spike gene RT-PCR was positive in 19 patients and 55 specimens. Of these 19 patients, 12 patients had at least two serial samples positive on different days (Table S2). Hence, nanopore and Illumina sequencing were performed for 12 patients on 28 specimens. These 12 patients included five males and seven females. The median age was 62 years, with a range between 30 and 75 years. Out of the 12 patients, four (33.3%) required oxygen supplementation, two (16.7%) were admitted to the intensive care unit, one (8.3%) was intubated and one (8.3%) died.

For these 12 patients, the mean filtered coverage was  $36\,393 \times$  for nanopore sequencing and  $40\,990 \times$  for Illumina sequencing runs (Tables S3 and S4). Out of these 12 patients, only one patient had nucleotide differences between samples collected on different days and with a difference in variant frequency between initial and subsequent samples exceeding 10% in both nanopore and Illumina sequencing. This patient was a 75-year-old man who required oxygen supplementation during hospitalization. For nanopore sequencing, the G22017T (guanine to thymine) was found in 3.8% (158/4116) of the reads in the sputum specimen collected on day 7 (first specimen) after symptom onset, while it accounts for 67.8% (12060/17795) of reads in the saliva specimen collected on day 9 after symptom onset (second specimen) (Fig. 1A). For Illumina sequencing, G22017T was found in 5.0% (356/7082) of the reads in the first specimen, while this mutation was found in 60.4% (4578/7574) in the second specimen (Fig. 1B). Sanger sequencing was performed to verify the mixed population at the position 22 017 in the second specimen (Fig. 1C). He had received lopinavir–ritonavir and ribavirin one day before the collection of the second specimen. G22017T resulted in the non-synonymous mutation of W152L (tryptophan to leucine) in the N-terminal domain of the S1 subunit. This mutation is within the N3 loop of the N-terminal domain, which mediates the binding of a neutralizing antibody, 4A8 [23].

Next, we determined the prevalence of W152L mutation among SARS-CoV-2 isolates deposited in GISAID up to 15 September 2020 (Table S5). Of 92942 sequences available, 21 (0.023%) had W152L mutation.

## Discussion

### Summary of principal findings

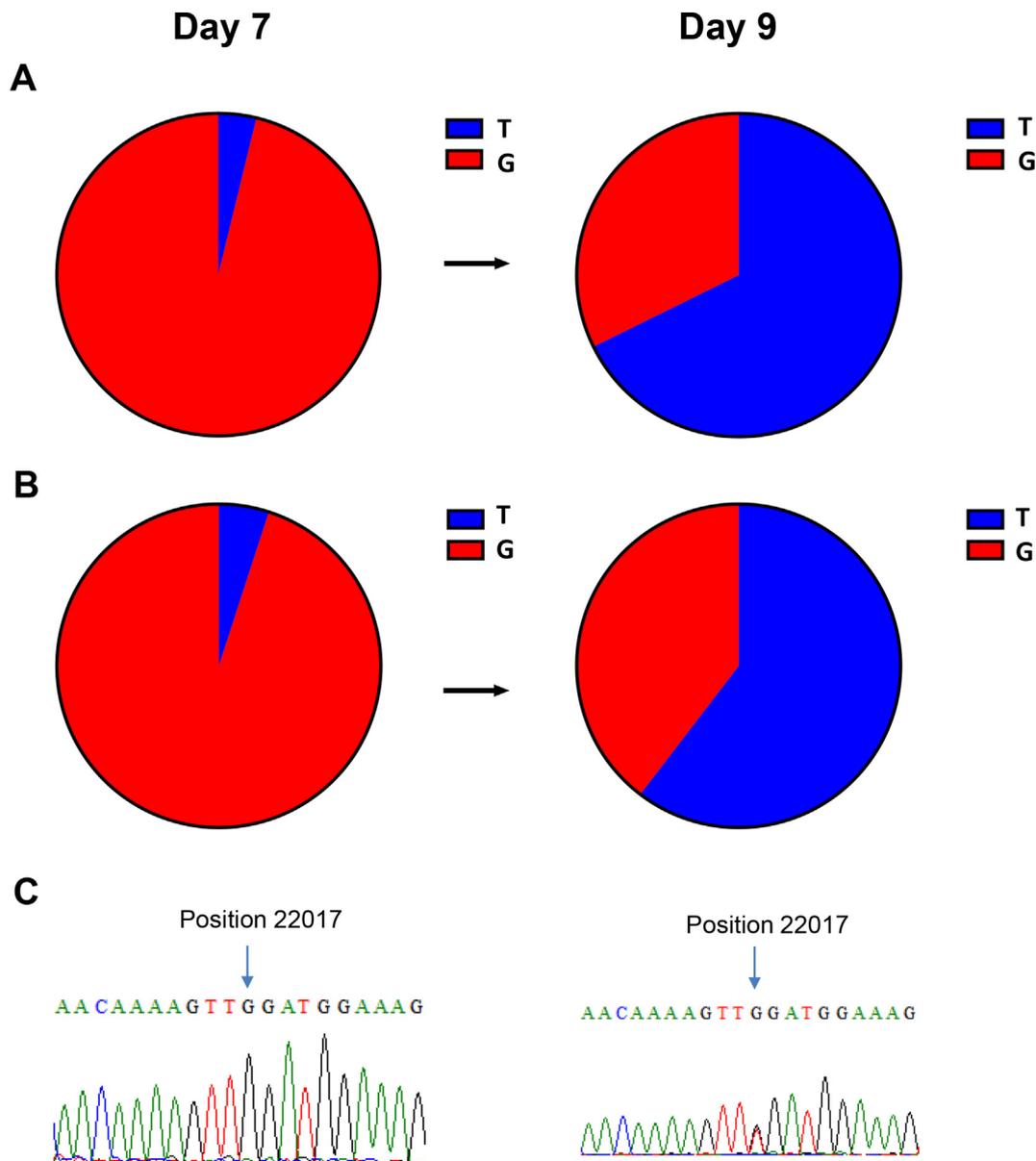
This study assessed the SARS-CoV-2 spike gene mutations in serial specimens from COVID-19 patients using a combination of nanopore, Illumina and Sanger sequencing. In one patient with severe disease, an important non-synonymous mutation G22017T, which results in W152L (tryptophan to leucine) mutation in the N-terminal domain of the spike gene, was present at a low frequency of  $\leq 5\%$  in the sputum specimen but represented the predominant population ( $\geq 60\%$ ) in the saliva specimen collected 2 days later. W152L is located within the binding site of a recently identified neutralizing antibody 4A8. Spike protein amino acid residue 152, together with residue 145 in the N3 loop of N-terminal domain, interacts with the 4A8 antibody via hydrophobic interactions [23]. Since tryptophan has an aromatic side chain while leucine has an aliphatic side chain, the W152L mutation might change the structure of the N3 loop in the N-terminal domain, hence affecting the binding of neutralizing antibodies. Such observation may be an example of microevolution induced by host antibody response in an attempt to evade neutralizing antibody response which arise during the course of illness.

### Comparison with other studies

W152L mutation is located at the N-terminal domain of the spike protein. The presence of W152L from a minor population to become the predominant population suggests that this site may be under immune selection pressure in this patient. Neutralizing antibodies against N-terminal domain could be found in COVID-19 patients [23–25], and may neutralize virus infection for several reasons. First, Chi et al. postulated that the antibody may prevent the conformational change of the S protein which is necessary for fusion [23]. Second, antibody against N-terminal domain may block the spike protein that binds to an unidentified receptor. The N-terminal domain of S1 subunit has been shown to participate in binding for other coronaviruses. The human coronaviruses OC43 and HKU1 has been shown to bind to 9-O-acetylated sialic acids, while MERS-CoV-2 binds to non-acetylated sialic acid, via the N-terminal domain [26,27]. The mouse hepatitis virus binds to host cell receptor carcinoembryonic antigen-related cell adhesion molecule via the S1 N-terminal domain [28].

Although nucleotide changes are frequent, amino acid mutation is actually an infrequent event for the spike gene [29], except for D614G which now accounts for the majority of the SARS-CoV-2 reported worldwide [5]. Mutations at spike protein amino acid residue 152 are rarely found, accounting for  $< 0.03\%$  of the strains deposited in GISAID as of 15 September 2020. The low prevalence of W152L mutation suggest that this mutation may confer reduced fitness or transmissibility. However, since usually only one viral sequence from each patient would be deposited into public databases, the prevalence of W152L mutation may have been underestimated.

Our patient with W152L mutation was treated with ribavirin and lopinavir–ritonavir which started between the collection of the first and second specimen. Our previous clinical trial showed that triple combination, which includes ribavirin, lopinavir–ritonavir and interferon- $\beta$  1b, can shorten the duration of illness in COVID-19 patients [30]. Ribavirin is known to induce mutations in RNA viruses, and this may have promoted the mutation in this patient. However, out of 4000 bp sequenced, only W152L was found to be the predominant mutant. Therefore, it is unlikely that the mutation is related to the antivirals given.



**Fig. 1.** G22017T spike gene mutation in patient HKU-IHCE0511-006. (A) Nanopore sequencing. (B) Illumina sequencing. (C) Sanger sequence tracing, demonstrating the double peak at position 22017 in the day 9 specimen.

#### Limitations of this study

There are several limitations in this study. First, we can only determine the SNPs among specimens with successful RT-PCR of the spike region. Second, the serial specimens of each patient that were successfully sequenced were collected <14 days apart. This is because the viral load reduces substantially during the second week of infection [2,31]. Therefore, we were not able to determine the variants that are present in patients with prolonged viral shedding. Viral load can also be affected by other factors. Older age has been associated with a higher viral load [2]. The type of specimen can also affect the viral load [31]. The duration of viral shedding can be longer in fecal than in respiratory specimens [32]. Third, the amplification of the spike gene may cause distortion in the variant frequencies. We have addressed this problem by

performing a separate RT-PCR reaction for nanopore and Illumina sequencing of the same specimen. Furthermore, we only focused on variants that exceeded 10% variation from the same individual. Fourth, for the patient with W152L mutation in the saliva specimen collected on day 9 after symptom onset, only sputum specimen was available on day 7 after symptom onset. It is possible that W152L mutation is already present as a predominant population in the saliva of the patient on day 7 after symptom onset.

#### Conclusions and implications for clinical practice and research studies

Our study demonstrates that mutations may arise spontaneously at neutralizing antibody sites during the course of COVID-19. Although these mutations may not be sustained during person-to-

person transmission, these amino acid mutations may affect the antiviral activity of neutralizing antibodies. Monitoring for intra-host mutational hotspots are warranted. Serial monitoring of mutations is needed for clinical trials on monoclonal antibody therapeutic trials.

### Transparency declaration

All authors declare no conflict of interest.

### Funding

This work was supported by the Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Diseases and Research Capability on Antimicrobial Resistance for the Department of Health of the HKSAR Government, and donations of Richard Yu and Carol Yu, May Tam Mak Mei Yin, the Shaw Foundation Hong Kong, Michael Seak-Kan Tong, Respiratory Viral Research Foundation Limited, Hui Ming, Hui Hoy and Chow Sin Lan Charity Fund Limited, Chan Yin Chuen Memorial Charitable Foundation, Marina Man-Wai Lee, the Hong Kong Hainan Commercial Association South China Microbiology Research Fund, the Jessie & George Ho Charitable Foundation, Perfect Shape Medical Limited, Kai Chong Tong, Tse Kam Ming Laurence, Betty Hing-Chu Lee, and Ping Cham So.

### Author's contribution

J.D.I., K.H.K., W.M.C., K.Y.Y. and K.K.W.T. designed the study. J.D.I., W.M.C., C.C.Y.Y., A.W.H.C. and K.H.C. acquired the data. O.T.Y.T., W.S.L., T.S.H.C. and I.F.N.H. were involved in patient recruitment. All authors interpreted the data, revised the manuscript critically for important intellectual content and approved the final report.

### Acknowledgement

We are grateful to the Centre for PanorOmic Sciences (CPOS) of the University of Hong Kong for providing Illumina sequencing. We gratefully acknowledge the originating and submitting laboratories who contributed sequences to GISAID (Table S5).

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cmi.2020.10.030>.

### References

- [1] Cheung KS, Hung IF, Chan PP, Lung KC, Tso E, Liu R, et al. Gastrointestinal manifestations of SARS-CoV-2 infection and virus load in fecal samples from the Hong Kong cohort and systematic review and meta-analysis. *Gastroenterology* 2020;159:81–95.
- [2] To KK, Tsang OT, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020;20:565–74.
- [3] Hung IF, Cheng VC, Li X, Tam AR, Hung DL, Chiu KH, et al. SARS-CoV-2 shedding and seroconversion among passengers quarantined after disembarking a cruise ship: a case series. *Lancet Infect Dis* 2020;20:1051–60.
- [4] Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020;395:514–23.
- [5] To KK, Chan WM, Ip JD, Chu AW, Tam AR, Liu R, et al. Unique SARS-CoV-2 clusters causing a large COVID-19 outbreak in Hong Kong. *Clin Infect Dis* 2020:ciaa1119. <https://doi.org/10.1093/cid/ciaa1119>.
- [6] Fauver JR, Petrone ME, Hodcroft EB, Shioda K, Ehrlich HY, Watts AG, et al. Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the United States. *Cell* 2020;181:990–6.
- [7] Chen H, Wen X, To KK, Wang P, Tse H, Chan JF, et al. Quasispecies of the D225G substitution in the hemagglutinin of pandemic influenza A(H1N1) 2009 virus from patients with severe disease in Hong Kong, China. *J Infect Dis* 2010;201:1517–21.
- [8] Hung IFN, To KKW, Chan JFW, Cheng VCC, Liu KSH, Tam A, et al. Efficacy of clarithromycin-naproxen-oseltamivir combination in the treatment of patients hospitalized for influenza A(H3N2) infection: an open-label randomized, controlled, Phase IIb/III Trial. *Chest* 2017;151:1069–80.
- [9] Shen Z, Xiao Y, Kang L, Ma W, Shi L, Zhang L, et al. Genomic diversity of severe acute respiratory syndrome-coronavirus 2 in patients with coronavirus disease 2019. *Clin Infect Dis* 2020;71:713–20.
- [10] Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 2020;181:281–92. e6.
- [11] To KK, Cheng VC, Cai JP, Chan KH, Chen LL, Wong LH, et al. Seroprevalence of SARS-CoV-2 in Hong Kong Special Administrative Region and our returnees evacuated from Hubei province of China: a multi-cohort study. *Lancet Microbe* 2020;21:6595.
- [12] Chan JF, Yip CC, To KK, Tang TH, Wong SC, Leung KH, et al. Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/HEP real-time reverse transcription-PCR assay validated in vitro and with clinical specimens. *J Clin Microbiol* 2020;58:e00310–20.
- [13] Yip CC, Sridhar S, Cheng AK, Leung KH, Choi GK, Chen JH, et al. Evaluation of the commercially available LightMix® Modular E-gene kit using clinical and proficiency testing specimens for SARS-CoV-2 detection. *J Clin Virol* 2020;129:104476.
- [14] Hung IF, Lung KC, Tso EY, Liu R, Chung TW, Chu MY, et al. Triple combination of interferon beta-1 b, lopinavir-ritonavir, and ribavirin in the treatment of hospitalized COVID-19 patients: a phase 2 open-label randomized controlled trial. *Lancet* 2020;395:1695–704. [https://doi.org/10.1016/S0140-6736\(20\)31042-4](https://doi.org/10.1016/S0140-6736(20)31042-4).
- [15] Lanfear R, Schalamun M, Kainer D, Wang W, Schwesinger B. MinIONQC: fast and simple quality control for MinION sequencing data. *Bioinformatics* 2019;35:523–5.
- [16] Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589–95.
- [17] Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics* 2016;32:1749–51.
- [18] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;25:2078–9.
- [19] GitHub. Available at: <https://github.com/lh3/seqtk>. [Accessed 12 May 2020].
- [20] Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22:568–76.
- [21] Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–90.
- [22] To KKW, Chan WM, Li KSM, Lam CSF, Chen Z, Tse H, et al. High prevalence of four novel astrovirus genotype species identified from rodents in China. *J Gen Virol* 2017;98:1004–15.
- [23] Chi X, Yan R, Zhang J, Zhang G, Zhang Y, Hao M, et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* 2020;369:650–5.
- [24] Wee AZ, Wrapp D, Herbert AS, Maurer DP, Haslwanter D, Sakharkar M, et al. Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science* 2020;369:731–6.
- [25] Barnes CO, West Jr AP, Huey-Tubman KE, Hoffmann MAG, Sharaf NG, Hoffman PR, et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* 2020;182:828–42. <https://doi.org/10.1016/j.cell.2020.06.025>.
- [26] Hulswit RJG, Lang Y, Bakkers MJG, Li W, Li Z, Schouten A, et al. Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc Natl Acad Sci USA* 2019;116:2681–90.
- [27] Park YJ, Walls AC, Wang Z, Sauer MM, Li W, Tortorici MA, et al. Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. *Nat Struct Mol Biol* 2019;26:1151–7.
- [28] Shang J, Wan Y, Liu C, Yount B, Gully K, Yang Y, et al. Structure of mouse coronavirus spike protein complexed with receptor reveals mechanism for viral entry. *PLoS Pathog* 2020;16:e1008392.
- [29] Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Next-strain: real-time tracking of pathogen evolution. *Bioinformatics* 2018;34:4121–3.
- [30] Hung IF, Lung KC, Tso EY, Liu R, Chung TW, Chu MY, et al. 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* 2020;395:1695–704.
- [31] Wylie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, et al. Saliva or nasopharyngeal swab specimens for detection of SARS-CoV-2. *N Engl J Med* 2020;383:1283–6.
- [32] Zheng S, Fan J, Yu F, Feng B, Lou B, Zou Q, et al. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January–March 2020: retrospective cohort study. *BMJ* 2020;369:m1443.