Unique SARS-CoV-2 clusters causing a large COVID-19 outbreak in Hong Kong

Authors: Kelvin Kai-Wang To[1,2#], Wan-Mui Chan[1], Jonathan Daniel Ip[1], Allen Wing-Ho Chu[1], Anthony Raymond Tam[3], Raymond Liu[4], Alan Ka-Lun Wu[5], Kwok-Cheung Lung[6], Owen Tak-Yin Tsang[7], Daphne Pui-Ling Lau[7], Wing-Kin To[8], Mike Yat-Wah Kwan[9], Yat-Sun Yau[10], Anthony Chin-Ki Ng[1], Cyril Chik-Yan Yip[2], Kwok-Hung Chan[1], Herman Tse[11], Ivan Fan-Ngai Hung[12], Kwok-Yung Yuen[1,2#]

Author affiliations

[1]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

[2]Department of Microbiology, Queen Mary Hospital, Hong Kong Special Administrative Region, China

[3]Department of Medicine, Queen Mary Hospital, Hong Kong Special Administrative Region, China

[4]Department of Medicine and Geriatrics, Ruttonjee Hospital, Hong Kong Special Administrative Region, China

[5]Department of Microbiology, Pamela Youde Nethersole Eastern Hospital, Hong Kong Special Administrative Region, China

[6]Department of Medicine, Pamela Youde Nethersole Eastern Hospital, Hong Kong Special Administrative Region, China

[7]Department of Medicine and Geriatrics, Princess Margaret Hospital, Hong Kong Special Administrative Region, China

[8]Department of Pathology, Princess Margaret Hospital, Hong Kong Special Administrative Region, China

[9]Department of Paediatrics and Adolescent Medicine, Princess Margaret Hospital, Hong Kong Special Administrative Region, China

[10]Department of Paediatrics, Queen Elizabeth Hospital, Kowloon, Hong Kong Special Administrative Region, China

[11]Department of Pathology, Hong Kong Children's Hospital, Kowloon, Hong Kong Special Administrative Region, China

[12]Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China

Corresponding authors:

Kelvin Kai-Wang To

       Email: kelvinto@hku.hk

       Phone number: (852)-22552413

       Carol Yu Centre for Infection, State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Queen Mary Hospital, Pokfulam, Hong Kong, China

ABSTRACT

After two months of relative quiescence, a large COVID-19 outbreak occurred in Hong Kong in July 2020 after gradual relaxation of social distancing policy. Two unique SARS-CoV-2 phylogenetic clusters have been identified among locally-acquired cases, with most genomes belonging to cluster HK1 which is phylogenetically related to SARS-CoV-2 reported overseas.

INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic has overwhelmed the healthcare system in many parts of the world. To combat the pandemic, social distancing measures have been implemented to reduce community transmission. However, resurgence of COVID-19 cases has been seen in many parts of the world after the relaxation of these social distancing measures.

Hong Kong was one of the first places in the world to report COVID-19 cases [1]. However, the number of COVID-19 cases remained relatively low due to the early implementation of stringent public health measures, including border control, voluntary community-wide wearing of face masks, hand hygiene and social distancing, prompt isolation of suspected cases, and testing and quarantine of close contacts and travelers from epidemic areas [2,3]. The first wave, consisting of imported cases from China and limited local transmission, occurred in January and early March, 2020. The second wave, mainly related to imported cases outside Asia and associated local transmission, occurred in mid-March to May, 2020. After stepping up public control measures, only sporadic local cases were reported in May and June. Hence, most social distancing restrictions were lifted on June 18, 2020. However, a third wave started to occur since early July 2020. Locally-acquired cases re-appeared since July 5, 2020. A total of 617 locally-acquired laboratory-confirmed cases have been reported between July 5 and 21 [4]. In this study, we use whole genome sequencing to investigate the genomic epidemiology of this large summer outbreak.

METHODS

Patient specimens

Archived nasopharyngeal swab or posterior oropharyngeal saliva from COVID-19 patients were retrieved. This study was approved by the the Institutional Review Board of the

University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 13-372, UW 20-292), the Hong Kong East Cluster Research Ethics Committee (REC) (HKECREC-2018-068), the Kowloon West Cluster REC (KW/EX-20-038[144-26]), and the Kowloon Central/Kowloon East Cluster REC (KC/KE-20-0140/ER-1).

Sequencing on clinical specimens

Nanopore sequencing was performed as described previously with modifications [1,5]. RNA was extracted from clinical specimens with Qiagen Viral RNA Mini Kit, and was amplified with sequence-independent single-primer amplification (SISPA) or tiled primers. Nanopore sequencing was performed with the Oxford Nanopore MinION platform. Bioinformatic analysis was performed as described previously with modifications [1]. Sanger sequencing was performed on specimens with low coverage at the cluster-defining mutations. Details on library preparation, bioinformatics analysis, and Sanger sequencing are included in the Supplementary Methods section.

Phylogenetic tree analysis

Genomes from Hong Kong and selected genomes from overseas were included for the phylogenetic tree analysis. Nucleotide position was numbered according to the reference genome Wuhan-Hu-1 (GenBank accession number MN908947.3). Please refer to Supplementary Methods section for details.

RESULTS

In this study, a total of 116 high-quality whole genomes of SARS-CoV-2 were analyzed, including 24 in first wave (January 22 and March 9, 2020), 42 in second wave (March 13 and May 21, 2020), and 50 in third wave (June 22 to July 14, 2020)

(Supplementary Table S1). Ten genomes from the first wave were reported previously [1,5,6]. For the third wave, we included imported cases up to 2 weeks prior to the first locally-acquired case. Spike protein D614G mutation was not found in any genomes during the first wave, which mainly involved travelers from mainland China or other parts of Asia, or the linked local cases. However, D614G mutation was present in 73.8% (31/42) of the genomes in the second wave, which mainly involved travelers returning from Europe or North America (Figure 1A).

In the third wave, 32 specimens collected from locally-acquired cases between July 7 and 14, 2020, and 18 specimens collected from imported cases collected between June 22 and July 14, 2020, were included for analysis. The majority of the locally-acquired cases (29/32) belonged to GISAID clade GR, Nextstrain clade 20B, and Pangolin lineage B.1.1. GISAID clade GR is characterized by C241T, C3037T, A23403G (spike protein D614G), G28882A (nucleoprotein [NP] R203K), and G28883C (NP G204R) [7]. This local cluster formed a single cluster (referred to as cluster HK1) (Figure 1A and 1B) and included 5 unique mutations that are rarely found in other publicly available sequences at GISAID (C2973T [nsp3 A85V], C20312T [nsp15 A231V], C21597T [spike protein S12F], C28308G [NP A12G], C29144T [a synonymous mutation in NP]) (Supplementary Table S2). Phylogenetically, cluster HK1 is most closely related to 4 imported cases from the Philippines but at least 2 of the HK1-defining mutations were not found in these cases.

Another local cluster during the third wave, consisting of 3 members in a household, also belongs to GISAID clade GR, Nextstrain clade 20B, and Pangolin lineage B.1.1. This cluster is characterized by 3 unique mutations, including A19702G (nsp15 N28D), T22020C (spike protein M153T) and C28269T (located between ORF8 and NP gene) (referred to as cluster HK2) (Supplementary Table S2). Cluster HK2 is most closely related to imported

cases from Kazakhstan, but those from Kazakhstan only had A19702G (nsp15 N28D) and T22020C (spike M153T) but not C28269T mutation.

DISCUSSION

In July, 2020, Hong Kong has experienced the third wave of COVID-19, which represented the largest local COVID-19 outbreak since the beginning of the pandemic. In this study, we have identified two unique clusters causing this COVID-19 summer outbreak in Hong Kong. The majority of genomes from locally-acquired cases (91%) during this third wave belong to a cluster HK1, a unique cluster within the GR clade, which is characterized by 4 non-synonymous mutations (nsp3 A85V, nsp15 A231V, spike protein S12F, NP A12G) and 1 synonymous mutation (NP C29144T). Genomes from a local household cluster with 3 members form another unique cluster (cluster HK2), which is characterized by 2 non-synonymous mutations (nsp15 N28D, spike protein M153T) and C28269T. Both clusters, especially cluster HK1, were phylogenetically more closely related to imported cases than the strains collected in Hong Kong during the previous waves before June 2020, suggesting that this summer COVID-19 outbreak is unlikely to be related to silent carriers from previous waves. Instead, our results suggest that the current outbreak may be related to imported cases.

Cluster HK1 is most closely related to genomes of patients traveling from the Philippines, while cluster HK2 is most closely related to those from Kazakhstan. However, there are important differences between our locally-acquired clusters and these imported cases. All the genomes from cases imported from the Philippines did not contain nsp3 A85V or NP A12G that define cluster HK1, while those from Kazakhstan lack the mutation C28269T that define cluster HK2. Hence, our results suggest that there is a missing link between these locally-acquired cases in Hong Kong and those cases from the Philippines or

Kazakhstan. Further investigations are required to investigate the source of this outbreak in Hong Kong.

Both cluster HK1 and HK2 possess D614G, which is now the predominant clade worldwide. D614G, located on the surface of the spike protein promoter, has become a predominant mutation worldwide [8]. D614G mutation is associated with a higher viral load in patients, and is associated with better viral replication in a pseudovirus assay [8].

Several factors may explain this explosive summer outbreak in Hong Kong. First, the sudden increase in social gatherings especially at eateries after the stepping down of public health control measures facilitated person-to-person transmission. Second, this outbreak occurred in the summer, when both the temperature and humidity are high. Whether these climate conditions affect virus transmissibility require further investigations. Third, the unique mutations in these clusters, especially in cluster HK1, may have increased the survival or transmissibility of the virus. The mutations specific to cluster HK1 occur in different proteins, including two nonstructural proteins (nsp3 and nsp15) and two structural proteins (spike protein and nucleoprotein). nsp3 is a putative PL-pro domain. nsp15 is a XendoU: poly(U)-specific endoribonuclease, and has been shown to be a potent interferon antagonist [9]. The surface spike protein is responsible for receptor binding, and the nucleoprotein mainly participates in viral genome transcription, replication and virion assembly. Spike protein S12F is located in the signal sequence, while the nucleoprotein A12G is located in the N1a domain, a linker domain [10]. Cluster HK2 contains two non-synonymous mutations, nsp15 N28D and spike protein M153T. M153T is at the N-terminal domain. This residue is located within a region that is newly found in the SARS-CoV-2 strains when compared with SARS-CoV [11].Whether these mutations affect the function of these proteins remains to be determined.

There are several limitations in this study. First, the current third wave is still ongoing, and there may be other clusters that have not been identified. Second, since whole genome sequencing required samples with relatively high viral load, those with lower viral load could not yield good quality sequence for phylogenetic analysis.

CONCLUSIONS

Two unique SARS-CoV-2 clusters have been identified during this large summer outbreak in Hong Kong shortly after the easing of social distancing policies. Further studies are required to determine how environmental, host or viral factors contribute to this outbreak. Source control at the borders and airport are important to prevent imported cases. Since transmission from asymptomatic or mildly symptomatic patients are common [12,13], community-wide wearing of face mask and social distancing measures, especially at eateries, are required for prevent local spread.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

All authors declare no conflict of interest.

REFERENCES

[1]   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;10.1016/S1473-3099(20)30196-1

[2]   Lam HY, Lam TS, Wong CH, Lam WH, Leung CME, Au KWA, et al. The epidemiology of COVID-19 cases and the successful containment strategy in Hong Kong-January to May 2020. Int J Infect Dis 2020;98:51-8

[3]   Cheng VC, Wong SC, Chuang VW, So SY, Chen JH, Sridhar S, et al. The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. J Infect 2020;10.1016/j.jinf.2020.04.024

[4]   Centre for Health Protection. Latest situation of cases of COVID-19 (as of 15 July 2020). Available at https://www.chp.gov.hk/files/pdf/local_situation_covid19_en.pdf. Accessed on July 16, 2020. 2020

[5]   Chan WM, Ip JD, Chu AW, Yip CC, Lo LS, Chan KH, et al. Identification of nsp1 gene as the target of SARS-CoV-2 real-time RT-PCR using nanopore whole-genome sequencing. J Med Virol 2020;10.1002/jmv.26140

[6]   Chu H, Chan JF, Yuen TT, Shuai H, Yuan S, Wang Y, et al. Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. Lancet Microbe 2020;1:E14-23

[7]   GISAID. Clade and lineage nomenclature aids in genomic epidemiology studies of active hCoV-19 viruses. Available at https://www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/. Accessed on July 19, 2020. 2020

[8]   Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 2020;https://doi.org/10.1016/j.cell.2020.06.043

[9]   Yuen CK, Lam JY, Wong WM, Mak LF, Wang X, Chu H, et al. SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. Emerg Microbes Infect 2020;9:1418-28

[10]   Ye Q, West AMV, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci 2020;10.1002/pro.3909

[11]   Guruprasad L. Evolutionary relationships and sequence-structure determinants in human SARS coronavirus-2 spike proteins for host receptor recognition. Proteins 2020;10.1002/prot.25967

[12]   Huff HV, Singh A. Asymptomatic transmission during the COVID-19 pandemic and implications for public health strategies. Clin Infect Dis 2020;10.1093/cid/ciaa654

[13]   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;10.1016/S1473-3099(20)30364-9

FIGURE LEGEND

Figure 1. Whole genome phylogenetic analysis showing the relationship between the genomes from locally-acquired and imported COVID-19 cases in Hong Kong. The trees were constructed by maximum likelihood method. Clade information as inferred by GISAID, Nextstrain and Pangolin nomenclatures, are shown. The reference genome Wuhan-Hu-1 (GenBank accession number MN908947.3) is used as the root of the tree. (A) The entire phylogenetic tree; (B) A magnified figure focusing on clade GR/20B.

Figure 1A

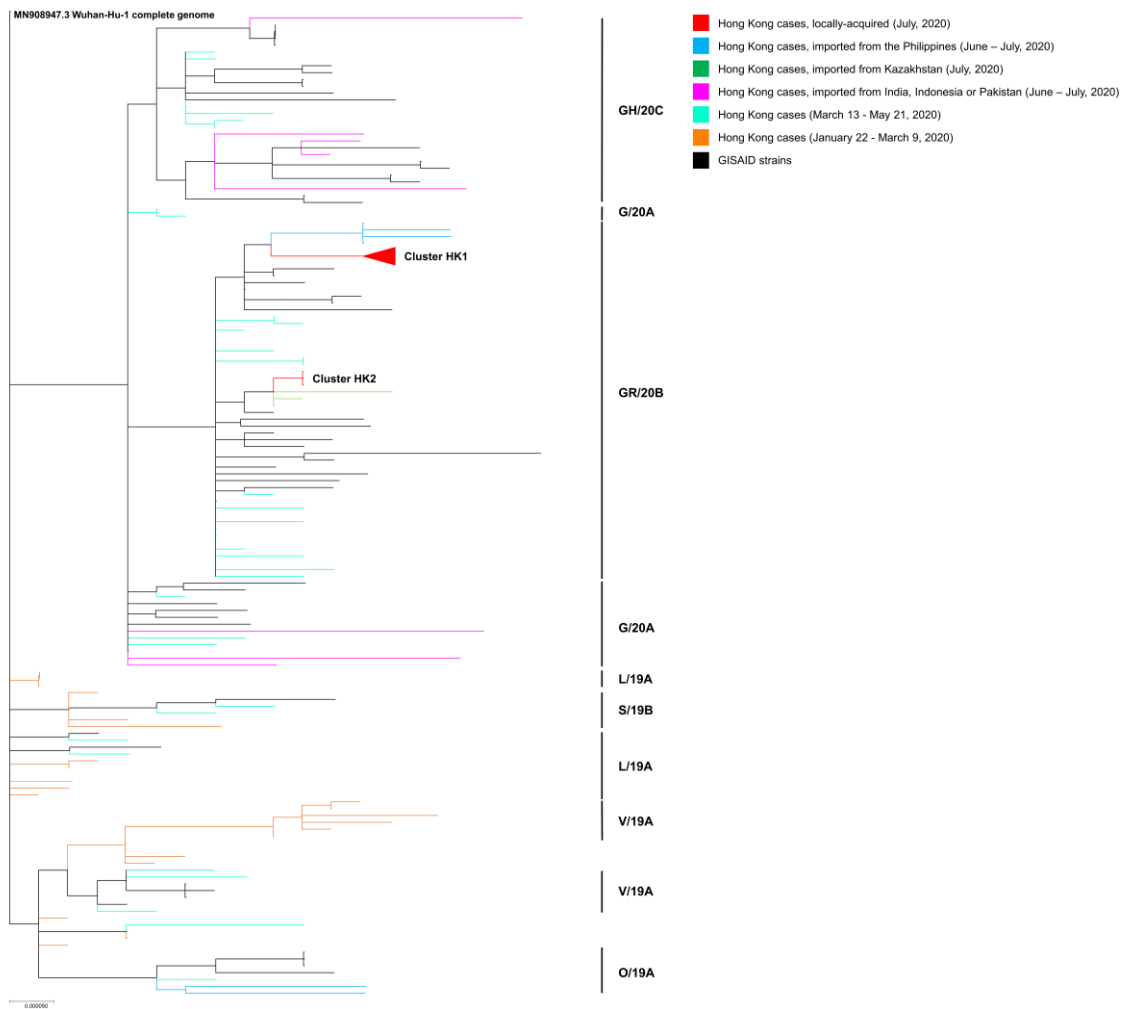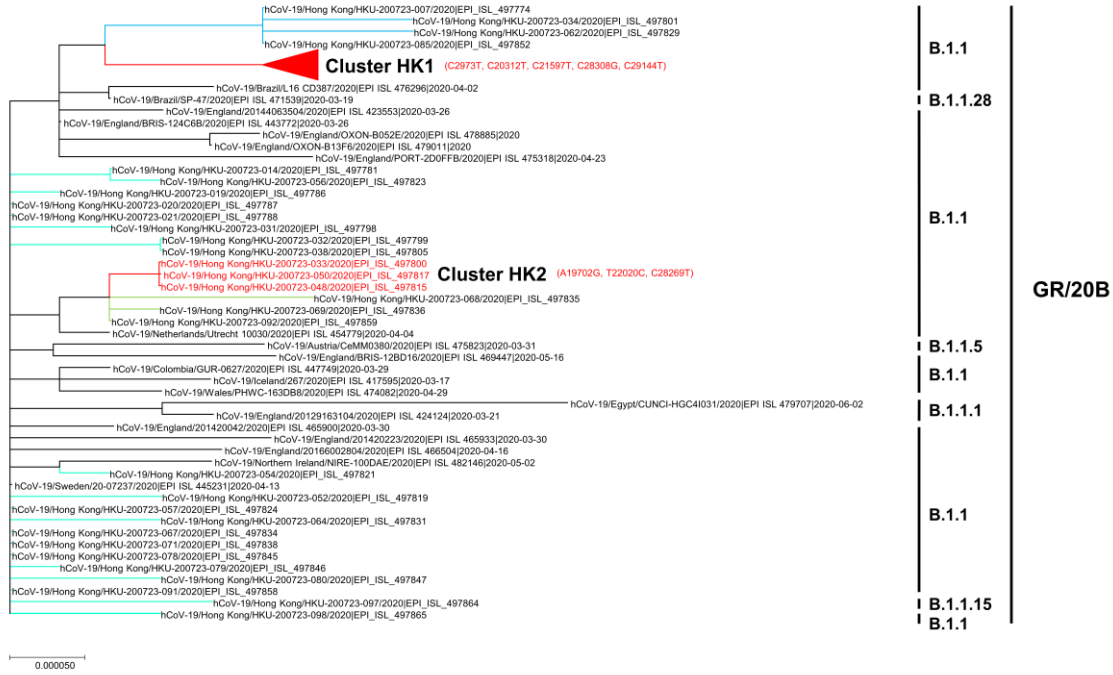# Figure 1B

Hong Kong cases, locally-acquired (July, 2020)
Hong Kong cases, imported from the Philippines (June – July, 2020)
Hong Kong cases, imported from Kazakhstan (July, 2020)
Hong Kong cases (March 13 - May 21, 2020)
GISAID strains