

# The surveillance of spike protein for patients with COVID-19 detected in Hong Kong in 2020

Gannon C. K. Mak  | Angela W. L. Lau | Andy M. Y. Chan  | Edman T. K. Lam | Rickjason C. W. Chan | Dominic N. C. Tsang

All from Microbiology Division, Department of Health, Public Health Laboratory Services Branch, Centre for Health Protection, Hong Kong Special Administrative Region, China

## Correspondence

Gannon C.K. Mak, 9/F, Public Health Laboratory Centre, 382 Nam Cheong Street, Shek Kip Mei, Kowloon, Hong Kong Special Administrative Region, China.  
Email: [so\\_phls10@dh.gov.hk](mailto:so_phls10@dh.gov.hk)

## Abstract

In 2020, numerous fast-spreading severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants have been reported. These variants had unusually high genetic changes in the spike (S) protein. In an attempt to understand the genetic background of SARS-CoV-2 viruses in Hong Kong, especially before vaccination, the purpose of this study is to summarize the S protein mutations detected among coronavirus disease 2019 (COVID-19) patients in Hong Kong in 2020. COVID-19 cases were selected every month in 2020. One virus from each case was analyzed. The full encoding region of the S proteins was sequenced. From January 2020 to December 2020, a total of 340 COVID-19 viruses were sequenced. The amino acids of the S protein for 44 (12.9%) were identical to the reference sequence, WIV04 (GenBank accession MN996528). For the remaining 296 sequences (87.1%), a total of 43 nonsynonymous substitution patterns were found. Of the nonsynonymous substitutions found, some of them were only detected at specific time intervals and then they disappeared. The ongoing genetic surveillance system is important. It would facilitate early detection of mutations that can increase infectivity as well as mutations that are selected for the virus to escape immunological restraint.

## KEYWORDS

coronavirus, evolution, genetics, genetic variability, mutation, virus classification

## 1 | INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been spreading around the world since 2019 and variants harboring various mutations have been increasingly identified (henceforth: "variant" stands for virus with one or more mutations that are different from the original virus<sup>1</sup>). Several of the mutations, especially those in the receptor-binding domain, were associated with enhanced affinity to the receptor in humans. Preliminary findings showed that they were fast-spreading variants. These variants

shared common features, notably high genetic changes in the spike (S) protein,<sup>2-6</sup> which is a key target for vaccine, virus entry, and infectivity.<sup>7</sup>

In Hong Kong, the SARS-CoV-2 vaccine was available by end of February 2021 for the public free of charge.<sup>8</sup> Genetic surveillance would facilitate early detection of mutations that can increase infectivity as well as mutations that are selected for the virus to escape immunological restraint, especially when the vaccine is available. The purpose of this study is to summarize the S protein mutations detected among COVID-19 patients in Hong Kong in 2020.

## 2 | METHODS

We focused our study on the SARS-CoV-2 S protein to characterize viral mutations that are potentially associated with neutralization and immunity. It also served the purpose of knowing the genetic background of SARS-CoV-2 viruses in Hong Kong before the launch of vaccination.

With the exception of May and June 2020 when five and six COVID-19 cases were selected respectively due to a low number of COVID-19 cases, at least ten COVID-19 cases were selected every month in 2020. COVID-19 cases were selected in different time points to achieve a balanced representation and to avoid selection bias. The protocols were based on the previous study.<sup>9</sup> Only one virus from each case was analyzed. The S protein consists of 1273 amino acids; the full encoding region of 3822 nucleotides (from start to stop codons) was analyzed for each virus. Epidemiological information was collected as described.<sup>10</sup>

## 3 | RESULTS

From January 2020 to December 2020, a total of 340 SARS-CoV-2 viruses were sequenced for the S protein. Among those COVID-19 cases selected, 95 were imported cases (27.9%), 245 were local cases (72.1%).

For the 340 SARS-CoV-2 viruses sequenced, the amino acids of the S protein for 44 (12.9%) were identical to the reference sequence, WIV04 (GenBank accession MN996528). These 44 viruses were collected from January to April 2020. Reviewing the nonsynonymous substitutions for the remaining 296 viruses, majority of them harbored a maximum of four nonsynonymous substitutions, three viruses harbored  $\geq 10$  nonsynonymous substitutions (Table 1). Excluding the three viruses that harbored  $\geq 10$  nonsynonymous substitutions, an average of 1.60 nonsynonymous substitutions were found for the 293 viruses. For these 293 viruses, when compared with the number of nonsynonymous substitutions among imported cases ( $N = 70$ ) and local cases ( $N = 223$ ), statistical differences were not found for these two groups of viruses (unpaired  $t$  test,  $p$ -value  $> 0.05$ ).

The number of nonsynonymous substitutions observed was unusually high for those three viruses. These two thresholds accounted for 0.31% (4/1273) and 0.78% (10/1273) of amino acid changes, respectively. A total of 43 nonsynonymous substitution patterns were found, 33 (76.7%) of them were detected only once (Table S1). Although these 33 sequences were only sequenced once, the specimens were sequenced in both forward and reverse directions. In addition, all of the nonsynonymous substitutions were also checked by viewing the chromatogram trace files to make sure that the corresponding nucleotide substitutions were genuine.

Of the nonsynonymous substitutions found, three of them were of great interest. They were 8V (6.5%, 22/340), 12F (33.8%, 115/340), and 614G (79.1%, 269/340). The other nonsynonymous

**TABLE 1** The summary of SARS-CoV-2 viruses having nonsynonymous substitutions in the S protein

No. of nonsynonymous substitutions in the S protein	SARS-CoV-2 viruses	
	No.	%
0	44	12.9
1	135	39.7
2	142	41.8
3	14	4.1
4	2	0.6
10	1	0.3
11	2	0.6
Total	340	100%

Abbreviation: SARS-CoV-2 viruses, severe acute respiratory syndrome coronavirus 2.

substitutions were only found at a maximum of 1.5% (5/340). For 8V and 12F, both of them were only detected at specific time intervals and then they disappeared. They were only detected transiently from February–March 2020 to July–September 2020, respectively. On the other hand, the 8V and 12F were detected in 100% (22/22) and 94.8% (109/115) of the local cases, respectively. For the 614G, it first appeared in March 2020, then 614G was detected in each month subsequently. From August 2020 onwards, 614G was detected in all the viruses sequenced (Table S2).

The S gene nucleotide sequences generated in this study have been deposited in the Global Initiative for Sharing All Influenza Data (GISAID) database. The corresponding patient information is also available in Table S3.

## 4 | DISCUSSIONS

In Hong Kong, the first COVID-19 case was detected from respiratory samples collected from a man on January 22, 2020. A total of 8847 COVID-19 cases were detected in 2020. Regarding the number of cases confirmed daily, three peaks were observed, namely during March–April 2020, July–August 2020, and November–December 2020.<sup>11</sup> Each peak was characterized by consecutive daily reported cases of  $\geq 10$ . In this study, we analyzed 340 S protein sequences from 340 COVID-19 patients detected in Hong Kong in 2020. Although it only accounted for a small proportion of cases (3.8%, 340/8847) detected in 2020, we were capable of detecting dominating variants during different time intervals. An epidemiological study showed that the cases of 8V were involved in two clusters, “Hotpot dinner gathering at Kwun Tong,” and “Fook Wai Ching She in Maylun Apartments in North Point.”<sup>11</sup> A research group in Hong Kong also detected the 8V viruses for the 50 viruses collected before March 2020. The 8V viruses accounted for 54% of genomes sequenced (27/50).<sup>12</sup> The 12F was detected from July to September 2020, it coincided with the peak observed during July

2020. This finding was concordant with the genome study performed by another research group in Hong Kong.<sup>13</sup> The 614G was first detected in Mar 2020 and since then its proportion had increased from 50.0% (11/22) to 98.9% (85/86) by July 2020. Its proportion remained 100% from Aug 2020. Although 614G viruses might increase infectivity over 614D viruses,<sup>14</sup> it was not correlated with the three peaks observed in Hong Kong. However, the peaks were more likely related to the containment strategies implemented in Hong Kong.<sup>10,15,16</sup>

For the three viruses harboring  $\geq 10$  nonsynonymous substitutions in the S protein, two of them were 20I/501Y.V1 (also known as VOC 202012/01, lineage B.1.1.7), the remaining was 20H/501Y.V2 (B.1.351). These three cases were first identified by the intensive surveillance system targeting returning travelers to Hong Kong.<sup>17,18</sup> The two 20I/501Y.V1 cases shared a novel set of spike mutations (Table S1, 69–70 deletion, 144 deletion, 501Y, 570D, 681H, 716I, 982A, 1118H) and they were concordant as described.<sup>3</sup> The 20H/501Y.V2 case harbored another set of spike mutations (Table S1, 80A, 215G, 242–244 deletion, 417N, 484K, 501Y, 701V) and was also in line with the report by another study.<sup>4</sup> Among those mutations, 501Y and 484K were shown to have an impact on biological functions, such as receptor binding; however, the significance of other mutations remained to be determined.<sup>3–5</sup>

Our data show that, over the 1-year period of 2020, the mutations of S protein harbored by each virus were few overall. The low genetic diversity of the S protein might be due to the lack of immunological pressure. Although dominating variants were found in different time periods, they were only detected for few months. The genetic diversity of the S protein is still being determined after vaccination. Laboratory surveillance to monitor variants of SARS-CoV-2 is important. It is common that laboratories use specific PCR assays to track the signature mutations in variants. Coincidentally, variants identified in UK, South Africa and Brazil shared 501Y.<sup>2–4</sup> However, these assays should be used with cautions. It is not known whether new variants will harbor novel mutations in the S protein or other regions.<sup>19,20</sup> It is expected that the SARS-CoV-2 is likely to become an endemic infection. To track variants, a surveillance system should be implemented in addition to monitoring mutations in the S protein.

## ACKNOWLEDGMENT

We thank our colleagues Alan K.L. Tsang, Peter C.W. Yip, and Peter K.C. Cheng for technical assistance in identifying variants during the current SARS-CoV-2 pandemic.

## CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

## AUTHOR CONTRIBUTIONS

**Gannon C. K. Mak:** conceptualized the study design, led the data collection, data analysis, data interpretation, and manuscript writing. **Angela W. L. Lau and Andy M. Y. Chan:** coordinated and carried out

the experiments. **Edman T. K. Lam, Rickjason C. W. Chan, and Dominic N. C. Tsang:** supervised and oversaw all aspects of experiments. All authors participated in the drafting of the manuscript and approval of the final version.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article.

## ORCID

Gannon C. K. Mak  <http://orcid.org/0000-0003-0893-533X>

Andy M. Y. Chan  <http://orcid.org/0000-0001-6744-5870>

## REFERENCES

1. WHO. The effects of virus variants on COVID-19 vaccines. Available at: <https://www.who.int/news-room/feature-stories/detail/the-effects-of-virus-variants-on-covid-19-vaccines> (Accessed 23 April 2021).
2. Mallapaty S. COVID mink analysis shows mutations are not dangerous - yet. *Nature*. 2020; 587(7834):340–341. <https://doi.org/10.1038/d41586-020-03218-z> PMID: 33188367
3. Public Health England. Investigation of novel SARS-CoV-2 variant Variant of Concern 202012/01. Version 1, release date 21/12/2020. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/959438/Technical\\_Briefing\\_VOC\\_SH\\_NJL2\\_SH2.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959438/Technical_Briefing_VOC_SH_NJL2_SH2.pdf) (Accessed 5 March 2021).
4. Tegally H, Wilkinson E, Giovanetti M, et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *bioRxiv*. 2020. <https://doi.org/10.1101/2020.12.21.20248640>
5. Voloch CM, Silva FR Jr., De Almeida LGP, et al. Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. *bioRxiv*. 2020 <https://doi.org/10.1101/2020.12.23.20248598>
6. Zhang W, Davis BD, Chen SS, Sincuir Martinez JM, Plummer JT, Vail E. Emergence of a novel SARS-CoV-2 variant in Southern California. *JAMA*. 2021;325(11):1324–1326e211612. <https://doi.org/10.1001/jama.2021.1612>
7. 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(2):281–292.e6. <https://doi.org/10.1016/j.cell.2020.02.058>
8. The Government of the Hong Kong Special Administrative Region. COVID-19 Vaccination Programme. Available at: <https://www.covidvaccine.gov.hk/en/programme> (Accessed 23 April 2021)
9. Mak GCK, Lau AWL, Chan AMY, Chan DYW, Tsang DNC. The D614G substitution in the S gene and clinical information for patients with COVID-19 detected in Hong Kong. *J Clin Virol*. 2020;130: 104550. <https://doi.org/10.1016/j.jcv.2020.104550>
10. Lam HY, Lam TS, Wong CH, 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;98:51–58. <https://doi.org/10.1016/j.ijid.2020.06.057>
11. Centre for Health Protection. Latest situation of cases of COVID-19. Available at: [https://www.chp.gov.hk/files/pdf/local\\_situation\\_covid19\\_en.pdf](https://www.chp.gov.hk/files/pdf/local_situation_covid19_en.pdf) (Accessed 23 April 2021)
12. Leung KS, Ng TT, Wu AK, et al. Territorywide study of early coronavirus disease outbreak, Hong Kong, China. *Emerg Infect Dis*. 2021;27(1):196–204. <https://doi.org/10.3201/eid2701.201543>

13. To KK, Chan WM, Ip JD, 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>
14. Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812-827.e19. <https://doi.org/10.1016/j.cell.2020.06.043>
15. Cheng VC, Wong SC, Chuang VW, 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;81(1): 107-114. <https://doi.org/10.1016/j.jinf.2020.04.024>
16. Siu GK, Lee LK, Leung KS, et al. Will a new clade of SARS-CoV-2 imported into the community spark a fourth wave of the COVID-19 outbreak in Hong Kong? *Emerg Microbes Infect*. 2020;9(1): 2497-2500. <https://doi.org/10.1080/22221751.2020.1851146>
17. Hong Kong Free Press. Covid-19: Gov't barring Hongkongers returning from UK, despite earlier stance on returnees from China, 24 December 2020. Available at: <https://hongkongfp.com/2020/12/24/covid-19-govt-barring-hongkongers-returning-from-uk-despite-earlier-stance-on-returnees-from-china/> (Accessed 5 March 2021).
18. Radio Television Hong Kong. HK sees first mutant Covid strain from South Africa, 31 December 2020. Available at: <https://news.rthk.hk/rthk/en/component/k2/1568042-20201231.htm> (5 March 2021).
19. Deng X, Garcia-Knight MA, Khalid MM, et al. Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell*. 2021;20. <https://doi.org/10.1016/j.cell.2021.04.025>
20. Tse H, Lung DC, Wong SC, et al. Emergence of a severe acute respiratory syndrome coronavirus 2 virus variant with novel genomic architecture in Hong Kong. *Clin Infect Dis*. 2021:ciab198. <https://doi.org/10.1093/cid/ciab198>

#### SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Mak GCK, Lau AWL, Chan AMY, Lam ETK, Chan RCW, Tsang DNC. The surveillance of spike protein for patients with COVID-19 detected in Hong Kong in 2020. *J Med Virol*. 2021;93:5644–5647. <https://doi.org/10.1002/jmv.27063>