

# Genome-wide analysis of SARS-CoV-2 strains circulating in Vietnam: Understanding the nature of the epidemic and role of the D614G mutation

Manh H. Dao<sup>1</sup>  | Lan T. Phan<sup>2</sup> | Thang M. Cao<sup>1</sup> | Quang C. Luong<sup>3</sup> |  
Hang T. T. Pham<sup>1</sup> | Nhung H. P. Vu<sup>1</sup>  | Nghia V. Khuu<sup>3</sup> | Thinh V. Nguyen<sup>3</sup> |  
Long T. Nguyen<sup>1</sup> | Hieu T. Nguyen<sup>1</sup> | Anh H. Nguyen<sup>1</sup> | Loan Kim Thi Huynh<sup>1</sup> |  
Thao P. Huynh<sup>1</sup> | Quan H. Nguyen<sup>1</sup> | Hieu C. Truong<sup>3</sup> | Hang Minh Nguyen<sup>4</sup> |  
Tung Xuan Trinh<sup>4</sup> | Dung T. Nguyen<sup>5</sup> | Trieu B. Nguyen<sup>6</sup> | Hung T. Do<sup>6</sup> |  
Quang D. Pham<sup>7,8</sup>  | Thuong V. Nguyen<sup>1</sup> 

<sup>1</sup>Microbiology and Immunology Department, Pasteur Institute of Ho Chi Minh City, Ho Chi Minh City, Vietnam

<sup>2</sup>Directorial Board, Pasteur Institute of Ho Chi Minh City, Ho Chi Minh City, Vietnam

<sup>3</sup>Department for Disease Control and Prevention, Pasteur Institute of Ho Chi Minh City, Ho Chi Minh City, Vietnam

<sup>4</sup>General Department of Preventive Medicine, Ha Noi, Vietnam

<sup>5</sup>Ho Chi Minh City Center for Diseases Control, Ho Chi Minh City, Vietnam

<sup>6</sup>Pasteur Institute of Nha Trang, Nha Trang, Vietnam

<sup>7</sup>Planning Division, Pasteur Institute of Ho Chi Minh City, Ho Chi Minh City, Vietnam

<sup>8</sup>Training Center, Pasteur Institute of Ho Chi Minh City, Ho Chi Minh City, Vietnam

## Correspondence

Thuong V. Nguyen, Pasteur Institute of Ho Chi Minh City, 167 Pasteur St, District 3, Ho Chi Minh City 700000, Vietnam.  
Email: [nguyenthuong@yahoo.com](mailto:nguyenthuong@yahoo.com)

## Funding information

This study was supported by Vietnam's Ministry of Science and Technology

## Abstract

Genome-wide analysis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains is essential to better understand infectivity and virulence and to track coronavirus disease 2019 (COVID-19) cases and outbreaks. We performed whole-genome sequencing of 27 SARS-CoV-2 strains isolated between January 2020 and April 2020. A total of 54 mutations in different genomic regions was found. The D614G mutation, first detected in March 2020, was identified in 18 strains and was more likely associated with a lower cycle threshold (<25) in real-time reverse-transcription polymerase chain reaction diagnostic tests than the original D614 (prevalence ratio = 2.75; 95% confidence interval, 1.19–6.38). The integration of sequencing and epidemiological data suggests that SARS-CoV-2 transmission in both quarantine areas and in the community in Vietnam occur at the beginning of the epidemic although the country implemented strict quarantine quite early, with strict contact tracing, and testing. These findings provide insights into the nature of the epidemic, as well as shape strategies for COVID-19 prevention and control in Vietnam.

## KEYWORDS

D614G, mutation, next-generation sequencing, SARS-CoV-2, Vietnam

## 1 | INTRODUCTION

First described as a pneumonia of unknown etiology in Wuhan, China, in December 2019, coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (family *Coronaviridae*, genus *Betacoronavirus*), a single-stranded, positive-sense RNA virus<sup>1,2</sup> which is presumed to have originated in horseshoe bats.<sup>3</sup> Efficient and widespread human-to-human transmission has subsequently triggered a global coronavirus pandemic, resulting in more than 100 million cases of COVID-19 with nearly 2.5 million deaths worldwide during the past year.<sup>4</sup>

Early in the pandemic, Vietnam closed its borders with China to prevent importation of COVID-19 cases after the first COVID-19 case was detected in Ho Chi Minh City on January 23, 2020.<sup>1</sup> Quarantine measures were enforced to travelers from China beginning in February 2020, and for travelers from South Korea, Japan, Europe, the United States, South East Asia, and all countries since March 21, 2020. Electronic health declaration forms were made mandatory for all passengers entering Vietnam to assist in detecting suspected cases for testing and quarantine.

As of February 24, 2021, there were 2412 confirmed cases, and more than half of these cases were identified among travelers entering Vietnam.<sup>5</sup> Several COVID-19 outbreaks which have occurred as a result of local person-to-person transmission of SARS-CoV-2 may give the viruses chances to develop mutations for a further efficient adaption and evolution and a wider circulation in humans.<sup>6</sup> Many mutations have been found amongst SARS-CoV-2 strains circulating around the world, particularly the D614G which has been shown to increase infectivity.<sup>7,8</sup> The relevance of the mutations to virulence or virus replication is still unknown.

Therefore, it is necessary to elucidate the nature of the COVID-19 epidemic and to explore the characteristics of circulating SARS-CoV-2 strains. In this study, we isolated and generated whole virus genome sequences from 27 SARS-CoV-2-infected patients in southern Vietnam. Findings from epidemiological and sequencing data of SARS-CoV-2 provide information for building better strategies of COVID-19 prevention, control, and research in Vietnam.

## 2 | MATERIALS AND METHODS

### 2.1 | Ethical approval

This study was approved by the Institutional Review Ethical Board of the Pasteur Institute in Ho Chi Minh City (PIHCM) (reference number: 05/GCN-PAS). All participants provided written informed consent.

### 2.2 | Samples

Between January and April 2020, swab specimens (both NP and OP specimens were collected and combined in a single virus transport

medium tube to maximize test sensitivity) were taken from more than 24,000 persons who traveled to Vietnam and were close contacts with confirmed cases. SARS-CoV-2 was detected in 71 samples by real-time reverse-transcription polymerase chain reaction (RT-PCR).<sup>9</sup> Viral isolation was successful in 33 SARS-CoV-2-positive samples. Whole-genome sequencing was performed for 27 SARS-CoV-2 strains from imported cases and three clusters in southern and central Vietnam.

### 2.3 | Virus isolation

Vero E6 cells (ATCC) were used for virus isolation, as previously described elsewhere.<sup>10,11</sup> Briefly, Vero E6 cells grown in 25 cm<sup>2</sup> culture flasks were inoculated with filtered eluates of real-time RT-PCR-positive NP and OP swabs. Cells were maintained in Dulbecco's modified Eagle's media supplemented with 0.2% bovine serum albumin and 16 µg/ml *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK)-trypsin, and incubated at 37°C, 5% CO<sub>2</sub>. Cultures were monitored daily for cytopathic effect (CPE). Once CPE was observed, culture supernatants were analyzed for SARS-CoV-2 RNA by real-time RT-PCR, and aliquots were stored at -70°C until sequencing was performed.

### 2.4 | Viral genome sequencing

SARS-CoV-2 isolates were sequenced using a Next-generation sequencing platform. Briefly, viral RNA was extracted from 140 µl culture supernatant using QiAmp viral RNA mini kit (Qiagen), then treated with 2 U/µl of Turbo DNase (Ambion) at 37°C for 30 min. cDNA was synthesized from 8 µl-treated RNA using a SuperScript III First-Strand Synthesis kit (Thermo Fisher Scientific) followed by RNase H digestion. dsDNA was generated using NEBNext® Ultra™ II Nondirectional RNA Second Strand Synthesis Module (BioLabs Inc.). The Nextera XT Library Prep Kit (Illumina) was used to prepare the viral library. The library was purified by using Agencourt® AMPure® XP beads (Beckman Coulter), and then qualified using Qubit High Sensitivity DNA Kit (Thermo Fisher Scientific) and a NEBNext® Library Quant Kit for Illumina®. The prepared library was sequenced using MiSeq reagent kit V2 on the MiSeq platform (Illumina).

### 2.5 | Sequence and phylogenetic analysis

The raw sequencing data for each sample were trimmed for quality and assembled by mapping to the reference genome from Wuhan (MT019529) using the CLC Genomics Workbench v10.1.1. Viral sequences were deposited in NCBI. We downloaded 233 full genome sequences of SARS-CoV-2 from the GISAID database based on two steps: (1) location and time that patients came from or went through; (2) randomly selecting 10 full genome sequences with high coverage from each location mentioned in step 1 at 1 month before patient's

onset day. Then, we aligned the obtained sequences with 206 reference sequences by the Mafft software (v7.452). Phylogenetic trees were constructed by the maximum-likelihood method with 1000 bootstrap replicates in MEGA (v10.0.5). Variant analysis was done using CLC (v10.1.1) and Nextclade (<https://clades.nextstrain.org>).

## 2.6 | Statistical analysis

To deal with a small sample size, we applied Poisson regression analysis, using a prevalence ratio (PR) as a conservative, consistent, and interpretable measure, to look for the predictors, including D614G, of a strong cycle threshold (<25) in real-time RT-PCR diagnosis of SARS-CoV-2 in the multivariate regression model. A backward elimination with a log-likelihood ratio test was used to identify the best-fitting model that described contributing variables. Moreover, we used exact Poisson bivariate regression to explore the correlation between the D614G mutation (dependent variable) and other mutations (independent variables) due to small samples. Data were entered using Microsoft Excel (2013), and all statistical analyses were carried out in Stata version 14.0 (StataCorp).

## 3 | RESULTS

### 3.1 | Sample characteristics

SAR-CoV-2 was detected in 71 swab specimens by real-time RT-PCR and virus isolation was successful in 33 SARS-CoV-2-positive samples. Whole-genome sequencing was performed for 27 SARS-CoV-2 strains, isolated from imported cases and three clusters in southern and central Vietnam: six cases in a bar outbreak in Ho Chi Minh City; two cases of a family cluster originating in a man from Wuhan, China,<sup>11</sup> two cases of a cluster sharing a room in a quarantine area (travelling back from United States and Australia, respectively) and 17 imported cases from European countries, the United States, and Australia (Figure S1). Of the 27 individuals from whom SARS-CoV-2 was isolated and sequenced, the median age was 28 years (range: 16–73 years), and 51.9% were males. Symptoms included fever (44.4%), cough (40.7%), sore throat (18.5%), and shortness of breath (11.1%) (Table 1).

### 3.2 | Viral sequencing and genomic epidemiology

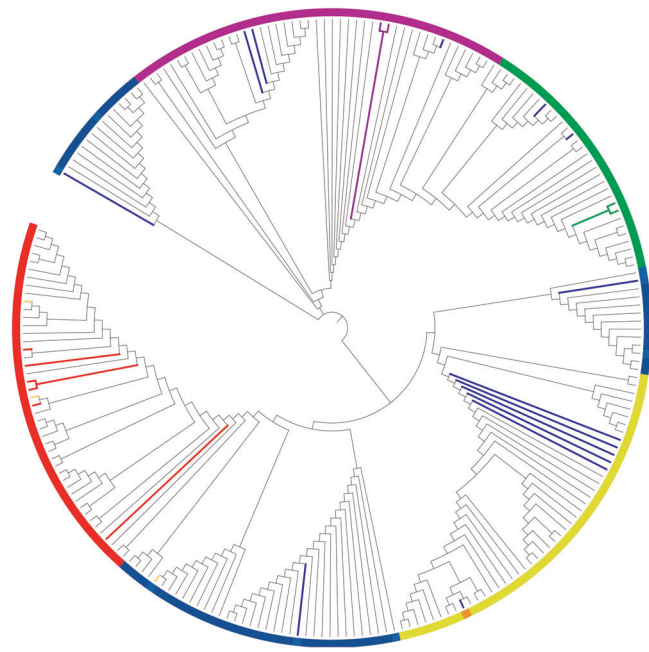
Full-length genomes of the 27 SARS-CoV-2 strains isolated from imported COVID-19 cases and three COVID-19 clusters occurring in southern Vietnam were analyzed. The SARS-CoV-2 strains from the first cluster, involving a Chinese man from Wuhan, China, and his son in January 2020, were identical and closely related to sequences found in China at the same time; both of the strains belong to 19A clade (Figure 1).

**TABLE 1** Selected characteristics among 27 SARS-CoV-2-infected patients

Characteristics	Frequency	Percent (%)
Sex		
Male	14	51.9
Female	13	48.1
Age (years)		
<20	2	7.4
20–39	15	55.6
40–59	6	22.2
60+	4	14.8
Median (range)	28 (16–73)	
Case category		
Imported cases	19	70.4
Locally acquired cases	8	29.6
Clinical symptoms		
Fever	12	44.4
Cough	11	40.7
Sore throat	5	18.5
Shortness of breath	3	11.1
No symptoms	6	22.2
Duration of treatment (in days)	17 (9–37)	
Median (range)		
Gen E cycle threshold (C <sub>t</sub> ) (n = 21)		
<25	11	52.4
≥25	10	47.6
Median (range)	24.4 (16.5–30.4)	
Mutation group		
D614	9	33.3
G614	18	66.7

The second cluster of two cases occurred at the end of March 2020 in a quarantine area, where a Vietnamese traveler returning from the United States shared a room with another Vietnamese travelling back from Australia. An NP specimen, taken from the first traveler on arrival, was positive by real-time RT-PCR, and she was isolated and hospitalized. Three days later, the one back from Australia became symptomatic for COVID-19, which was confirmed by RT-PCR. Sequences of these two SARS-CoV-2 strains were identical (>99%) and originated from the United States, classified as in 19B clade, but did not have D614G mutation (Figure 1).

The third cluster was linked to an outbreak with 19 confirmed cases occurring in a bar in District 2 of Ho Chi Minh City.<sup>12</sup> Six SARS-CoV-2 strains isolated from this outbreak showed high sequence identity, and they were closely related to strains from other two



**FIGURE 1** A maximum-likelihood phylogenetic tree was constructed based on 27 SARS-CoV-2 sequences from this study and 233 reference sequences from representative geographic regions. The first two strains discovered in Vietnam are shown in purple. The bar-related strains are shown in red, three strains from three patients who have not visited the bar are shown in orange. Two strains from two roommates in the quarantine area are shown in green. Strains from individuals traveling from other countries to Vietnam and immediately quarantined at the airport are shown in blue. Circle around the phylogenetics tree shows the clades of strains in the tree: Clade 19A (purple), clade 19B (green), clade 20A (blue), clade 20B (yellow), clade 20C (red), and clade 20D (orange)

patients who did not visit the bar: one traveler returning to Vietnam from the United States and one close contact of a case who also returned to Vietnam from the United States. All of the eight isolates had the D614G mutation and were classified as 20C clade (lineage B.1 according to Pangolin). These strains were close to another strain isolated from a traveler returning to Vietnam from the United States (clade 20A/lineage B.1) (Figure 1).

### 3.3 | Genetic variation of SARS-CoV-2

Complete genome sequences of 27 SARS-CoV-2 isolates were aligned with 206 SARS-CoV-2 sequences obtained from GenBank or GISAID. The directionality of alterations was inferred based on the SARS-CoV-2 strain from Wuhan (accession number: MT019529). The 54 different variants included 30 missense mutations, 19 synonymous mutations, and 5 mutations in the noncoding region (Tables S1, S2, and S3).

A total of 29 variants were found in ORF1ab, which roughly occupied two-thirds of the whole genome, including 19 missense and 12 synonymous mutations (Table S2). The largest numbers of

missense alterations were recognized in the NSP3, NSP4, and NSP13 regions. Particularly, G8388A in NSP3 and A8987T in NSP4 which lead to amino acid change including S2708D and I2907F were the most common (27/27 samples). We also found an alteration in the nonstructural protein RNA-dependent RNA polymerase (RdRp) that resulted in amino acid changes, including P314L (18/27 samples) (Table S3).

Moreover, four nonsynonymous and one synonymous mutation were recognized in the ORF3a region. ORF8 had one missense mutation at position T28144C that led to an amino acid change from Leucine to Serine (L84S). Nine mutations, including six missense mutations and three synonymous mutations, were found in the nucleocapsid (N) gene. Membrane (M) protein interacting with N protein to encapsulate viral RNA found two silent mutations. The spike (S) gene, encoding the S protein, mediates attachment of the virus to the host cell. There were three mutations in the Spike region including two missense mutations and one silent mutation. In particular, the A to G mutation at nucleotide position 23,403 caused Spike D614G amino acid change (Table S3). We also found that the appearance of the ORF1ab-14408 mutation was inversely correlated with D614G (Table S4).

Finally, there were 13 transversions and 40 transitions (Figure S2). As expected, transition mutations were generated at a higher frequency than transversions. The most popular base change transition was C→T ( $n = 28$ ). There were 19/40 transitions and 11 transversions that led to amino acid alterations. It is assumed that nonsynonymous transversions hardly conserve biochemical characteristics of the initial amino acid.<sup>13</sup>

### 3.4 | Association between D614G and lower $C_t$ of real-time PCR of SARS-CoV-2 E gene

A multivariate generalized Poisson regression model with the E gene ( $C_t < 25$ ) as the dependent variable showed that the presence of the D614G mutation in the SARS-CoV-2 spike protein was more likely to have a lower  $C_t$  ( $C_t < 25$ ) in real-time RT-PCR than the original (D614) (prevalence ratio = 2.75; 95% confidence interval, 1.19–6.38) (Table S5).

## 4 | DISCUSSION

Our study showed that the integration of sequencing and epidemiological data could elucidate the nature of the COVID-19 epidemic in Vietnam. Travelers from outside countries are quarantined for 14 days and must have a negative RT-PCR test before entering the community. This is useful to prevent imported cases, particularly in the early phase of the epidemic.<sup>14</sup> However, in a cluster of the two COVID-19 patients who shared the same room in a quarantine area, sequencing results showed high sequence identity with SARS-CoV-2 strains from the United States. This supports the hypothesis that the traveler returning from the United States transmitted the virus to

the traveler from Australia. This implicates that it is necessary to carry out strict regulations of prevention in the quarantine area to avoid person-to-person transmission.

From March 19 to April 4, 2020, most COVID-19 cases in Vietnam were imported. As having been previously described,<sup>15</sup> a British pilot was sick and visited a hospital where he was diagnosed with COVID-19. He had ever gone to a bar. Following this, rigorous contact tracing was done to detect further 18 COVID-19 cases; particularly, two British males who shared an apartment, coming back from Malaysia 2 weeks earlier before the outbreak occurred in the bar. These men were hypothesized to initiate the transmission of SARS-CoV-2 in the bar.<sup>15</sup> However, our sequencing data demonstrated all strains from patients linked to the bar were identical with strains from the three patients who had never gone to this bar, of which, two persons came back from the United States, and one was exposed to an infected one back from the United States. Hence, we recognized that the original case initiating the outbreak in the bar was unknown, suggesting a silent local transmission of COVID-19 in the community. In April 2020, the government issued a 3- to 4-week social distancing regulation, hidden Covid-19 cases, therefore, might not be able to spread. In Australia, SARS-CoV-2 genomics was demonstrated to be useful in rapidly identifying SARS-CoV-2 transmission chains and in significantly assisting public health responses.<sup>16</sup>

This study also delineated the circulating strain characteristics, including mutations in Vietnam. There were three alterations in the spike region: A23403G, C23731T, and G24794T. Eighteen samples had the signature mutation profile of A23403G which was related to the amino acid change, D614G. All 18 samples carrying D614G mutation were detected from March 14 onwards. These strains are classified into clades 20A, 20B, 20C, and 20D according to Next-strain. The D614G mutation had not been detected in the samples sequenced before March 14, 2020, which were classified mainly into clade 19A and 19B. The concurrent appearance of different viral clades (20A, 20B, 20C, and 20D) in Vietnam shortly after the clades 19A and 19B were detected suggest importation of viral strains carried by international travelers entering this country. The D614G was reportedly associated with increased infectivity of SARS-CoV-2.<sup>7,8</sup> However, current SARS-CoV-2 vaccine candidates are unlikely to be affected by D614G mutation.<sup>17-19</sup> We found that there was a mutation on ORF1ab-14408 that was inversely correlated with D614G. At present, the role of this mutation is unknown, so future research is needed.

D614G mutation has detected in many areas of the world.<sup>7</sup> This mutation located in the S1-S2 junction region near the furin recognition site (R667) for the cleavage of the S protein when the virions enter the host cells.<sup>20</sup> Moreover, D614G mutation was found in the six SARS-CoV-2 strains isolated from the outbreak occurring in a bar. In our study, those with D614G mutation in the SARS-CoV-2 spike protein were more likely to have a lower  $C_t$  (<25) in real-time RT-PCR tests than the original (D614). This might indirectly reflect a higher viral load in nasal and pharyngeal membranes in people infected with G614 than D614. The finding that the G614-mutation

strain is more contagious than D614 has been shown in a number of studies.<sup>7,8,21</sup> Therefore, it is possible that D614G facilitates the transmission of SARS-CoV-2 and may cause larger outbreaks.

Despite the limited number of samples sequenced, our findings are essential to inform on the circulation of variant strains of SARS-CoV-2 during the beginning of the COVID-19 epidemic in Vietnam. Evidence of SARS-CoV-2 transmission in the quarantine setting suggests the need for implementing extremely strict preventive measures. SARS-CoV-2 was also found to be silently transmitted in the community as early as March 2020, when the D614G strain was first detected and became dominant. Several mutations found in the current study need further exploration to understand their roles in viral infectivity, virulence, and even vaccine development. Until now, Vietnam has experienced and controlled a few waves of epidemics, the "silent" cases may appear through illegal border crossing. Therefore, the current measures of quarantine, case isolation, rigorous contact tracing, testing, and wearing face masks, which have been well applied in the country, need to be maintained and strengthened until an effective COVID-19 vaccine programme can be implemented.

## ACKNOWLEDGMENTS

The authors would like to thank Dr Richard Yanagihara with the Department of Pediatrics, John A. Burns School of Medicine, the University of Hawaii at Manoa for language proofreading before submission. We acknowledge the staff of the PIHCM and Southern Vietnam CDCs for their support in data collection, originating and submitting laboratories of the GISAID sequences. This study was supported by the Vietnam's Ministry of Science and Technology (Grant number: 32/20-ĐTĐL.CN-CNN).

## CONFLICT OF INTERESTS

The authors declare that they have no conflict of interest.

## AUTHOR CONTRIBUTIONS

Manh H. Dao, Lan T. Phan, Thuong V. Nguyen interpreted the data and wrote the first draft of the manuscript. Quang C. Luong, Thinh V. Nguyen, Hieu C. Truong, Hung T. Do, Trieu B. Nguyen, Dung T. Nguyen collected and interpreted epidemiological data. Manh H. Dao, Thang M. Cao, Nhung H. P. Vu, Hang T. T. Pham, Loan Kim Thi Huynh, Thao P. Huynh, Thao P. Huynh, Quan H. Nguyen, Long T. Nguyen, Anh H. Nguyen, Hieu T. Nguyen performed cell culture, sequencing and contributed to interpretation of virological findings. Nghia V. Khuu, Manh H. Dao, Hang T. T. Pham, Nhung H. P. Vu, Quang D. Pham and Thuong V. Nguyen clean data and data analysis. Thuong V. Nguyen, Hung T. Do, Quang D. Pham, Hang Minh Nguyen, Tung Xuan Trinh, and Lan T. Phan assisted in the interpretation of findings and overviewed the study. All authors have reviewed the manuscript, contributed critical revision of the manuscript, and approved the final version.

## DATA AVAILABILITY STATEMENT

All sequencing data are available in GISAID (<https://www.gisaid.org/>).



## ORCID

Manh H. Dao  <http://orcid.org/0000-0003-3874-5051>

Nhung H. P. Vu  <http://orcid.org/0000-0002-5854-7864>

Quang D. Pham  <http://orcid.org/0000-0003-2205-0327>

Thuong V. Nguyen  <http://orcid.org/0000-0003-0219-928X>

## REFERENCES

- Phan LT, Nguyen TV, Luong QC, et al. Importation and human-to-human transmission of a novel coronavirus in Vietnam. *N Engl J Med*. 2020;382:872-874. <https://doi.org/10.1056/NEJMc2001272>
- Riou J, Althaus CL. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill*. 2020;25:2000058. <https://doi.org/10.2807/1560-7917.ES.2020.25.4.2000058>
- Burki T. The origin of SARS-CoV-2. *Lancet Infect Dis*. 2020;20:1018-1019. [https://doi.org/10.1016/S1473-3099\(20\)30641-1](https://doi.org/10.1016/S1473-3099(20)30641-1)
- World Health Organization. COVID-19 Weekly Epidemiological Update 22. World Health Organization; 2020:1-3. [https://www.who.int/docs/default-source/coronaviruse/situation-reports/weekly-epidemiological\\_update\\_22.pdf](https://www.who.int/docs/default-source/coronaviruse/situation-reports/weekly-epidemiological_update_22.pdf)
- PAGE OF COVID-19 DISEASE. Ministry of Health—Newsletter about acute respiratory infections COVID-19. Accessed February 24, 2021. <https://ncov.moh.gov.vn/>. Accessed February 24, 2021.
- van Dorp L, Acman M, Richard D, et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect Genet Evol*. 2020;83:104351. <https://doi.org/10.1016/j.meegid.2020.104351>
- 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;182:812-827.e19. <https://doi.org/10.1016/j.cell.2020.06.043>
- Plante JA, Liu Y, Liu J, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*. 2020;592:116-121. <https://doi.org/10.1038/s41586-020-2895-3>
- Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill*. 2020;25:2000045. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>
- Wahba L, Jain N, Fire AZ, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270-273. <https://doi.org/10.1038/s41586-020-2012-7>
- Phan LT, Nguyen TV, Huynh LKT, et al. Clinical features, isolation, and complete genome sequence of severe acute respiratory syndrome coronavirus 2 from the first two patients in Vietnam. *J Med Virol*. 2020;92:2209-2215. <https://doi.org/10.1002/jmv.26075>
- Nguyen TV, Tran QD, Phan LT, et al. In the interest of public safety: rapid response to the COVID-19 epidemic in Vietnam. *BMJ Glob Heal*. 2021;6:e004100. <https://doi.org/10.1136/bmjgh-2020-004100>
- Lyons DM, Luring AS. Evidence for the selective basis of transition-to-transversion substitution bias in Two RNA viruses. *Mol Biol Evol*. 2017;34:3205-3215. <https://doi.org/10.1093/molbev/msx251>
- Ali M, Shah STH, Imran M, Khan A. The role of asymptomatic class, quarantine and isolation in the transmission of COVID-19. *J Biol Dyn*. 2020;14:389-408. <https://doi.org/10.1080/17513758.2020.1773000>
- Chau NVV, Hong NTT, Ngoc N, et al. Superspreading event of SARS-CoV-2 infection at a bar, Ho Chi Minh City, Vietnam. *Emerg Infect Dis J*. 2021;27. <https://doi.org/10.3201/eid2701.203480>
- Seemann T, Lane CR, Sherry NL, et al. Tracking the COVID-19 pandemic in Australia using genomics. *Nat Commun*. 2020;11:4376. <https://doi.org/10.1038/s41467-020-18314-x>
- Dearlove B, Lewitus E, Bai H, et al. A SARS-CoV-2 vaccine candidate would likely match all currently circulating variants. *Proc Natl Acad Sci*. 2020;117:23652 LP-23623662. <https://doi.org/10.1073/pnas.2008281117>
- McAuley AJ, Kuiper MJ, Durr PA, et al. Experimental and in silico evidence suggests vaccines are unlikely to be affected by D614G mutation in SARS-CoV-2 spike protein. *NPJ Vaccines*. 2020;5:96. <https://doi.org/10.1038/s41541-020-00246-8>
- Weissman D, Alameh M-G, de Silva T, et al. D614G spike mutation increases SARS CoV-2 susceptibility to neutralization. *Cell Host Microbe*. 2020;S1931-3128(20):30634-3063. <https://doi.org/10.1016/j.chom.2020.11.012>
- Follis KE, York J, Nunberg JH. Furin cleavage of the SARS coronavirus spike glycoprotein enhances cell-cell fusion but does not affect virion entry. *Virology*. 2006;350:358-369. <https://doi.org/10.1016/j.virol.2006.02.003>
- Bhattacharyya C, Das C, Ghosh A, et al. Global spread of SARS-CoV-2 subtype with spike protein mutation D614G is shaped by human genomic variations that regulate expression of TMPRSS2 and MX1 genes. *BioRxiv*. 2020. <https://doi.org/10.1101/2020.05.04.075911>

## SUPPORTING INFORMATION

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

**How to cite this article:** Dao MH, Phan LT, Cao TM, et al. Genome-wide analysis of SARS-CoV-2 strains circulating in Vietnam: Understanding the nature of the epidemic and role of the D614G mutation. *J Med Virol*. 2021;93:5660-5665. <https://doi.org/10.1002/jmv.27103>