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.

Gene Reports 23 (2021) 101100

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Comparative analysis of mutational hotspots in the spike protein of SARS-CoV-2 isolates from different geographic origins

Sanghoo Lee [a,*], Mi-Kyeong Lee [b], Hyeongkyun Na [a], Jinwoo Ahn [a], Gayeon Hong [a], Youngkee Lee [a], Jimyeong Park [a], Yejin Kim [a], Yun-Tae Kim [c], Chang-Ki Kim [d], Hwan-Sub Lim [b], Kyoung-Ryul Lee [a,b,c,d,**]

[a] Center for Companion Biomarker, Seoul Clinical Laboratories (SCL) Healthcare Inc., Republic of Korea
[b] Department of Molecular Diagnostics, Seoul Clinical Laboratories, Republic of Korea
[c] Center for Technology Innovation, Seoul Clinical Laboratories, Republic of Korea
[d] Center for Clinical Trial, Seoul Clinical Laboratories, Republic of Korea

ARTICLE INFO

Keywords:
SARS-CoV-2
Spike protein
Spike gene
Mutational hotspots
D614G mutation
Different geographic origins

ABSTRACT

The spike (S) protein mutations of SARS-CoV-2 are of major concern in terms of viral transmission and pathogenesis. Hence, we developed a PCR-based method to rapidly detect the 6 mutational hotspots (H49Y, G476S, V483A, H519Q, A520S, and D614G) in the S protein and applied this method to analyze the hotspots in the viral isolates from different geographical origins. Here, we identified that there was only the D614G mutation in the viral isolates. As of September 30, 2020, the analysis of 113,381 sequences available from the public repositories revealed that the SARS-CoV-2 variant carrying G614 has become the most prevalent form globally. Our results support recent epidemiological and genomic data demonstrating that the viral infectivity and transmission are enhanced by the S protein D614G mutation.

1. Introduction

A coronavirus disease (COVID-19) referred to as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a pandemic with high transmission and mortality rate worldwide (Zhu et al., 2020). SARS-CoV-2 has four structural proteins consisting of the spike (S), envelope (E), membrane (M), and nucleocapsid (N) (Kim et al., 2020). Among them, the S protein mediates the virus's entry into host cells (Shang et al., 2020). Especially, the D614G mutation in the S protein, an A-to-G base change at the position 23,403 in the Wuhan reference strain, is closely related to the emergence of a more transmissible form of SARS-CoV-2 (Zhang et al., 2020; Korber et al., 2020). In addition to the D614G, other mutations in the S gene region have recently been identified in isolates from different geographical origins (Shah et al., 2020; Chand et al., 2020). Therefore, the sequence analysis of the genomic region encoding the viral S protein is needed to understand the viral transmission and to design the viral vaccines (Amanat and Krammer, 2020).

Here, we analyzed the 6 hotspots in the S protein of SARS-CoV-2 isolates from different geographical origins and then compared our PCR-based sequencing results with the frequencies of the S mutations in publicly available SARS-CoV-2 genomes.

2. Materials and methods

For this study, nasopharyngeal samples were isolated from 16 Koreans, 12 Finns, and 8 North Americans who tested positive by reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR) at

Abbreviations: COVID-19, Coronavirus disease; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; S, Spike; E, envelope; M, membrane; N, nucleocapsid; RdRp, RNA-dependent RNA polymerase; RT-qPCR, reverse transcriptase-quantitative polymerase chain reaction; CT, cycle threshold; ACE2, angiotensin-converting enzyme-2; TMPRSS2, transmembrane serine protease2; Nsp3, nonstructural protein; Orf, open reading frame; RDB, receptor-binding domain; NGS, next-generation sequencing.

* Correspondence to: S. Lee, Center for Companion Biomarker, Seoul Clinical Laboratories (SCL) Healthcare Inc., 23F Heung-deok IT Valley, Heungdeok 1-ro, Giheung-gu, Yongin, Gyeonggi-do 16954, Republic of Korea.

** Correspondence to: K.-R. Lee, Center for Companion Biomarker, SCL & SCL Healthcare Co., Ltd., 25F Heung-deok IT Valley, Heungdeok 1-ro, Giheung-gu, Yongin, Gyeonggi-do 16954, Republic of Korea.

E-mail addresses: sprout30@scllab.co.kr (S. Lee), dkrlee@scllab.co.kr (K.-R. Lee).

https://doi.org/10.1016/j.genrep.2021.101100
Received 12 December 2020; Received in revised form 1 March 2021; Accepted 11 March 2021
Available online 23 March 2021

Seoul Clinical Laboratories (SCL) in Korea, and collected from March to April 2020. The twelve Finnish samples originate from those sent directly from Finland to SCL via Finnair flight for the SARS-CoV-2 diagnostic test, and the 8 North American samples originate from those sent to SCL from a US military hospital located in Korea for the viral testing.

This study was approved by the Institutional Review Board of Seoul Clinical Laboratories (SCL-IRB-20-036). Written consents to participate in this study were waived by not applicable because all personal identifiers were removed in the whole study process and only samples tested positive in RT-PCR diagnostic test performed at the SCL were used.

The extraction of SARS-CoV-2 RNA from nasopharyngeal samples was performed using MagNA Pure 96 DNA and Viral NA Small Volume Kit (Roche Diagnostics, Germany).

Based on the distribution of the frequent S protein mutations in SARS-CoV-2 isolates from North America and Europe (Zhang et al., 2020; Banerjee et al., 2020), we focused on the 6 major mutational hot spots, H49Y, G476S, V483A, H519Q, A520S, and D614G.

The RT-qPCR kit (Allplex™ 2019-nCoV Assay Kit, Seegene, Seoul, Korea) targets all E, RNA-dependent RNA polymerase (RdRp), and N genes. The ranges of cycle threshold (CT) value of all the samples tested positive in this study were $9.34 \leq CT \leq 30.18$ for E, $11.44 \leq CT \leq 32.21$ for RdRp, and $12.45 \leq CT \leq 32.77$ for N.

cDNA synthesis was performed with a commercially available kit (Invitrogen, USA) according to the manufacturer's instructions. Briefly, 50 ng/μl random hexamer and 10 mM dNTP mix were added to each sample (20 pg to 100 pg), reacted at 65 °C for 5 min. Then, 10× RT buffer, 25 mM $MgCl_2$, 0.1 M DTT, RNAaseOUT™ (40 U/μl), and Superscript® RT (200 U/μl) were added to each product and reacted as follows: 50 °C, 50 min; 25 °C, 10 min; 50 °C, 50 min; 80 °C, 10 min. After adding RNase H treatment (37 °C, 20 min reaction), the products were stored at −20 °C or immediately subjected to PCR experiments.

Target sequences of interest were amplified using the designed primers by a multiplex PCR method (Sol™ 2× multiplex PCR Smart Mix, Solgent, Korea). PCR programs were run as follows: 95 °C 10 min; 95 °C 30 s, 58 °C 30 s, 72 °C 1 min, 40 cycles; 72 °C 5 min.

All PCR products were confirmed by electrophoresis analysis and sequenced by Sanger's method (ABI 3500xL Genetic Analyzer, USA).

The primers covering the 6 hotspots in the S gene were designed to obtain amplicons of 100–200 bp for sequencing, according to the reference sequence (GenBank accession number: MN908947.3) (Table 1).

As of September 30, 2020, SARS-CoV-2 full genome sequences were downloaded from GISAID (http://gisaid.org), Nextstrain (www.nextstrain.org), and NCBI virus repository (http://www.ncbi.nlm.nih.gov/labs/virus), and then the 113,381 high-quality sequences were collected after removing duplicate samples, partially sequenced genomes, and samples with unclear collection data. These data were used to analyze the frequency of D614G mutations in the S protein of SARS-CoV-2 isolated from different geographical origins. An alignment tool of the NCBI virus server and MEGA-X software was used to perform multiple sequence alignments of the S protein.

The workflow of the complete work is shown in Fig. 1.

Table 1
Primers used for sequencing of the target hotspots in the S gene of SARS-CoV-2.

Target	Nucleotide region	Primer sequence (5′-3′)	Size (bp)
H49Y	21617–21643	Forward: AGAACTCAATTACCCCCTGCA	158
	21759–21780	Reverse: GTCCCAGAGACATGTATAGCAT	
G476S/ V483A	22931–22956	Forward: AGGAAGTCTAATCTCAAACCTTTTGA	197
	23107–23127	Reverse: GCTGGTGCATGTAGAAGTTCA	
H519Q/ A520S	23054–23075	Forward: CAACCCACTAATGGTGTTGGTT	164
	23196–23217	Reverse: AGAACACCTGTGCCTGTTAAAC	
D614G	23326–23347	Forward: ACCATGTTCTTTTGGTGGTGTC	157
	23460–23482	Reverse: AGAACCTGTAGAATAAACACGCC	

3. Results

The PCR-based sequencing method using the primer sets was first applied to rapidly detect the H49Y, G476S, V483A, H519Q, A520S, and D614G mutations in the S protein of the SARS-CoV-2 isolates from Korean, Finnish, and North American. Here, we identified that only the D614G mutation was present in the tested samples, where the D614G was identified in 1 of 16 Korean isolates (6.25%), 10 of 12 Finnish isolates (83.33%), and 8 of 8 North American isolates (100%) (Fig. 2), as compared to the reference sequence (NC_045512.2). No other mutations were detected except the D614G.

Based on these PCR-based sequencing results, we analyzed the frequency of the mutation in the 113,381 high-quality S protein sequences available from the GISAID, Nextstrain, and NCBI virus repositories as of September 30, 2020. Globally, the frequency of the G614 has increased constantly from January to September 2020 and the cumulative frequency of the G614 was about 84% of all the deposited sequences (Fig. 3A). In North America (n = 29,198), the D614G mutation was first reported at the end of February 2020 (Fig. 3B). On the other hand, the mutation was first reported at the end of January in Europe (n = 62,006) (Fig. 3C) and Asia (n = 6665) (Fig. 3D). The mutation first reported in China and Europe originated from the viral genomes identified on January 24, 2020, and January 28, 2020, in China (hCoV-19/Zhejiang/HZ103/2020, GISAID accession ID: EPI_ISL_422425) and Germany (hCoV-19/Germany/BavPat1-ChVir929/2020, GISAID accession ID: EPI_ISL_406862), respectively.

The frequencies of the D614G mutation between March and April 2020, when we collected SARS-CoV-2 clinical isolates for this study showed different patterns with geographic origins. During that period, the frequencies of G614 mutation were most pronounced in Europe (62–92%), followed by North America (19–95%). Especially, the G614 emerged at a higher frequency in February 2020 in Europe (25–83%) (Fig. 3C), compared to Asia (3–8%) (Fig. 3D) and Korea (0%) (Fig. 3E). In this study, the frequencies of G614 of the viral samples from each region, which were identified by our PCR-based sequencing method, showed a similar pattern with the global distribution of the G614 between March and April 2020.

In Korea (n = 834), the G614 was first reported in a SARS-CoV-2 sample collected on March 14, 2020, as of the date of collection (Fig. 3E). Although the number of samples collected from Koreans between March and April 2020 was significantly smaller than that collected from the other regions during the same period, the G614 frequency was between 8 and 83% (Fig. 3E).

4. Discussion

In this study, we used a PCR-based sequencing method to rapidly detect the 6 mutational hotspots in the S protein of SARS-CoV-2 isolates from different geographic origins, showing that there was only the D614G mutation in the samples.

Our sequencing results showed a trend similar to the frequency of the rapidly increasing D614G mutation until September 30, 2020. Together with the sequencing results, geographical analysis of 113,381 sequences available from the public repositories as of September 30, 2020, showed that the D614G is the most prevalent form worldwide including North America, Europe, Asia, and South Korea, supporting previously reported studies (Korber et al., 2020; Isabel et al., 2020; Yurkovetskiy et al., 2020). Our analysis of the publicly available sequences also supports recent studies showing that the D614G began in Europe and has spread to Asia through North America and Oceania (Korber et al., 2020; Isabel et al., 2020).

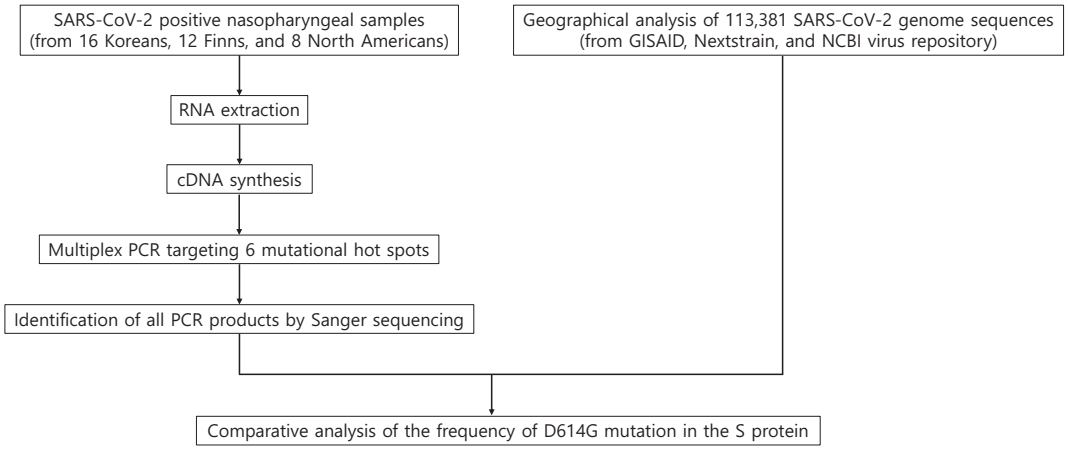

Fig. 1. The workflow of the complete work.

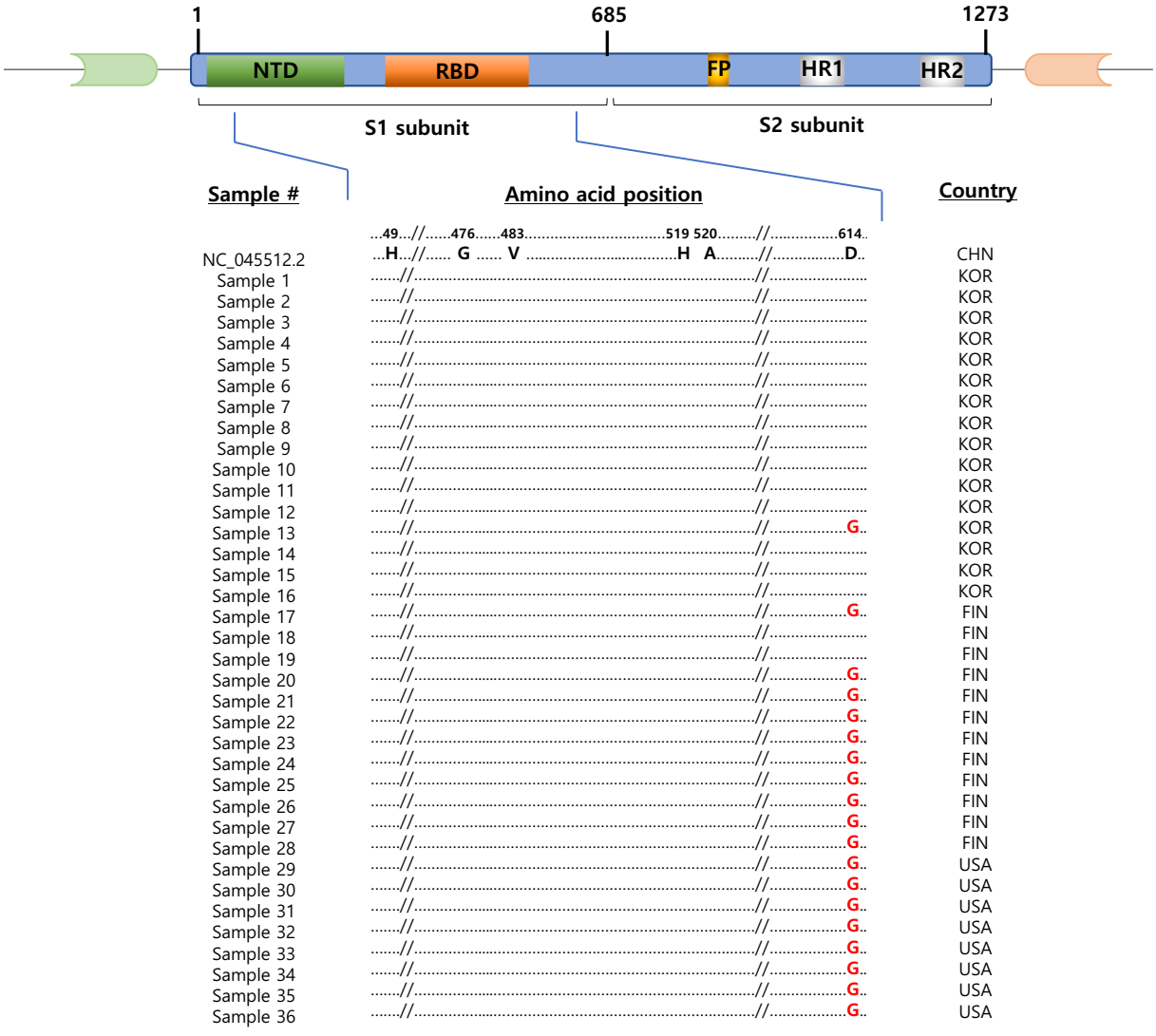

Fig. 2. Sequencing results and sequence alignment of the 6 target hotspots in the S protein of SARS-CoV-2 by Sanger sequencing. The dot indicates the same amino acid. The double slash indicates that a region composed of identical amino acids is omitted. The letters in red indicate that the amino acid D is mutated to G at position 614 as compared to the reference sequence (NC_045512.2). CHN, KOR, FIN, and USA indicate China, Korea, Finland, and North America, respectively.

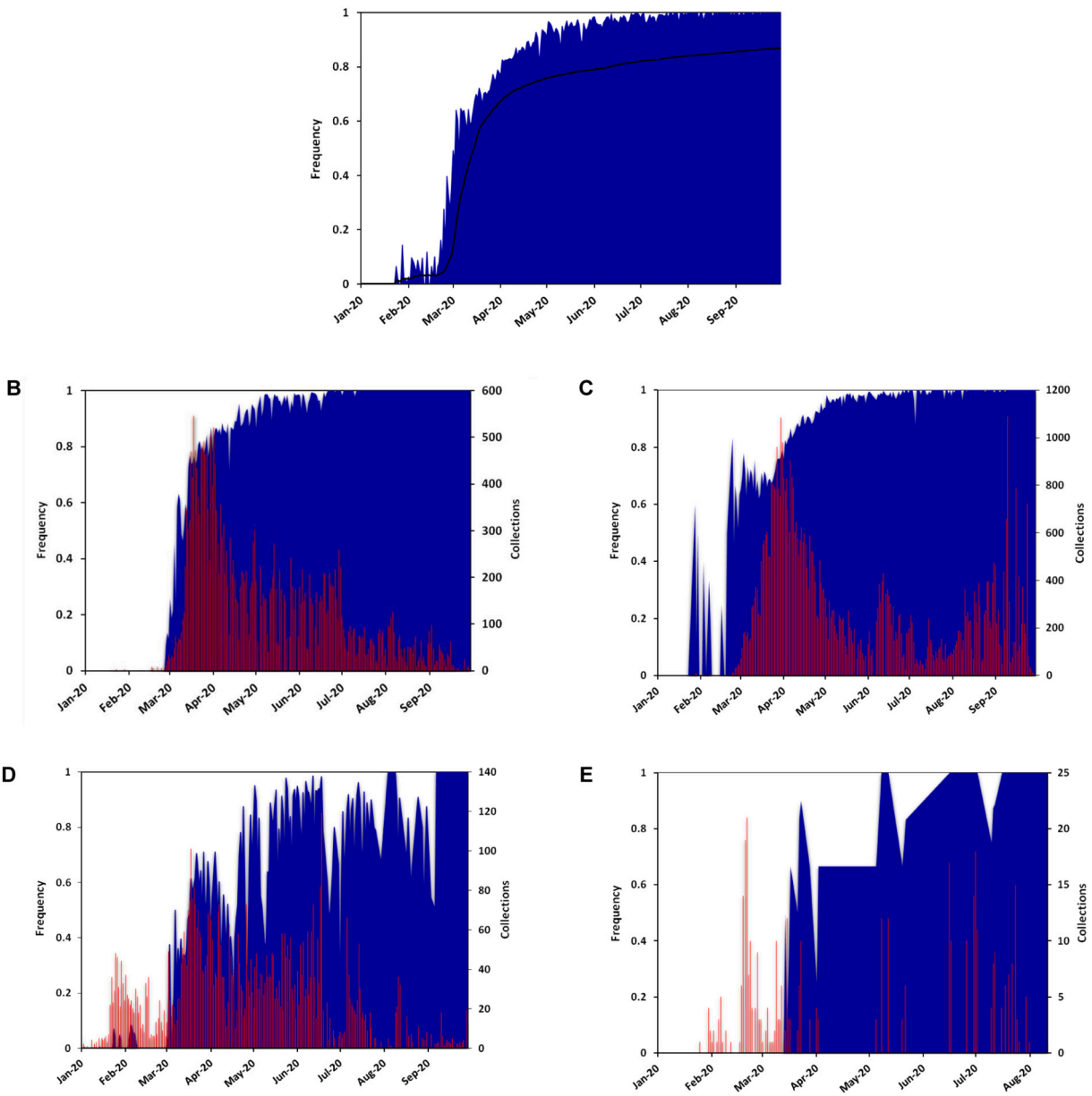

Fig. 3. The frequency of the S protein D614G (blue) over time in sequences collected from worldwide (n = 113,381) (A), North America (n = 29,198) (B), Europe (n = 62,006) (C), Asia (6665) (D), and Korea (n = 834) (E) in the public databases as of September 30, 2020. In the global frequency (A), the overlaid black line indicates the cumulative frequency of D614G in sequences collected up to and including each date. Red bars represent the number of sequences collected for each date.

The predominance of the D614G across the globe suggests that the mutation enhances human-to-human transmission by conferring a replication advantage to SARS-CoV-2 (Yurkovetskiy et al., 2020; Volz et al., 2021). The D614G increases cell entry with enhanced human receptor angiotensin-converting enzyme-2 (ACE2)-binding affinity, and the D614G-mediated cell entry is also highly dependent on transmembrane serine protease2 (TMPRSS2) (Ozono et al., 2021). A recent study has suggested that differences in ACE2 expression across different geographic regions may be a driving force for the positive selection of the D614G in the viral S protein. Interestingly, the study has also revealed that ACE2 expression is higher in Asians than in North Americans and Europeans, suggesting an association of the D614G with increased transmission of SARS-CoV-2 in specific populations with lower ACE2 expression (Huang et al., 2021). Moreover, the D614G showed the highest host entry activity among mutations such as V367F, G476S, V483A, and H49Y that frequently occur in the S protein (Ozono et al., 2021). A recent study showed that the G614 variant may be closely associated with lower C_T value in the RT-qPCR test, suggesting higher viral loads in the upper respiratory tract of patients infected with the virus (Korber et al., 2020). Also, our results support previous studies showing that the D614G variant can significantly enhance the viral infectivity and transmission compared to its D614 form (Zhang et al., 2020; Korber et al., 2020; Yurkovetskiy et al., 2020; Bhattacharyya et al., 2020; Hou et al., 2020).

Interestingly, a recent study has revealed that the D614G makes a neutrophil elastase cleavage site increasing the viral spread in high α1-antitrypsin-deficient individuals (Bhattacharyya et al., 2021). The α1-antitrypsin, which prevents lung tissue damage by neutrophil elastase, acts as an inhibitor of the elastase. Deficiency of the inhibitor facilitates entry of virus carrying the D614G into host cells. However, genotype-dependent deficiency of the inhibitor varies by human populations. Importantly, it is suggested that deficiency of the inhibitor is more

prevalent in North Americans and Europeans than in East Asians. Therefore, the virus carrying the D614G is more likely to infect and spread in North America and Europe than in East Asia, supporting the results in our study. Although the D614G confers the viral infectivity and transmissibility, its effect on disease severity remains unclear because other mutations often co-occur. Therefore, further studies are needed to elucidate the association of COVID-19 severity with the D614G mutation among human populations.

Positive selection resulting from the D614G in the S protein is closely related to the adaptive evolution of SARS-CoV-2. However, it has been shown that mutations in other viral genes of SARS-CoV-2 evolve along with the D614G. The D614G is frequently accompanied by a silent mutation 3037C>T and a missense mutation 14408C>T that causes P323L in viral RdRp in mild and severely affected patients (Korber et al., 2020; Biswas and Mudi, 2020). Specific mutations in the RdRp and S genes may alter the viral replication capability, which may cause increased infectivity and transmissibility. Furthermore, a recent study has suggested that the prevalent D614G and combined with other mutations such as L5F, V341I, K458R, I472V, D936Y, and S943T are more infectious (Li et al., 2020). Among frequent mutations observed in the S protein, a non-synonymous mutation, V483A, which increases the binding with ACE2 receptor, was mainly identified in the viral genomes isolated from North Americans (Laamarti et al., 2020). More recently, analysis of prevalent clades by continents revealed that the clade GH (carrying 241C>T in 5′ untranslated region, 3037C>T in nonstructural protein 3 (Nsp3), P314L in Nsp12b, Q57H in open reading frame 3a (Orf3a), and D614G in S protein) was most predominant in North America, followed by G (carrying 241C>T, 3037C>T, P314L, and D614G) and GR (carrying 241C>T, 3037C>T, P314L, RG203KR in N, and D614G). Notably, the clade GH was exclusively predominant compared to other Clades in South Korea (Park et al., 2021). Although clade G is now most prevalent in Europe, viral strains from Finland exhibited the highest proportion of clade GH (Alm et al., 2020). These results suggest that co-occurrence of D614G with other mutations may contribute to the adaptive evolution of SARS-CoV-2 under the current COVID-19 pandemic.

In addition to the D614G mutation occurring in the S protein, structural and multi-omics studies revealed that D7611G in S protein, T265I in Nsp2, S1920P in Nsp3, L3605F in Nsp6, co-occurring P5731L and Y5768C in Nsp13 may play crucial roles in modulating the efficiency of the viral entry into host cells and its pathogenesis. Also, these studies suggest that the D7611G mutation in the S protein reduces the binding of the S protein with ACE2, which leads to interaction with the TMPRSS2 receptor that may contribute to its high infectivity and transmissibility in North American strains (Gupta et al., 2020). P1327L, Y1364C, and S2540F mutations in Orf1b protein involved in the viral RNA replication and processing have been also identified in strains from North America, suggesting that the viral replication processes have favorably evolved in North American populations (Kumar et al., 2020). Notable, the structure of the receptor-binding domain (RDB) in the S protein of SARS-CoV-2 carrying the D614G has more open conformations such as a two-open conformation and an all-open state, which may increase the efficiency of ACE2 binding and fusion (Yurkovetskiy et al., 2020).

In Korea, SARS-CoV-2 strains with D614G have also become the most dominant form as of September 2020. Combining the results of our PCR-based sequencing and publicly available sequences results, it is presumed that patient infected with the virus carrying the G614 has already entered Korea from abroad between January and March 2020 before Korean health authorities tightened virus controls on overseas arrivals using methods such as extensive diagnostic testing, contact tracing and two weeks of quarantine.

Although next-generation sequencing (NGS) techniques have several advantages in analyzing the complete SARS-CoV-2 genomes, Sanger sequencing is still considered a gold standard and also used to analyze complete viral genomes (Moniruzzaman et al., 2020) or specific target genes (Tabibzadeh et al., 2020). For these reasons, our sequencing method may be useful in rapidly detecting the hotspot mutations in the viral S protein in general laboratories without an expensive NGS system.

The effect of SARS-CoV-2 carrying the D614G on the viral transmission and vaccine effectiveness targeting the variant clinically remains unclear. In this regard, further research is needed to clarify our understanding of the evolutionary mechanism of the SARS-CoV-2 D614G variant and the vaccine efficacy against the variant, which is rapidly spreading worldwide.

CRediT authorship contribution statement

S Lee & K-R Lee conceived and coordinated the projects. S Lee, Y-T Kim, C-K Kim, H-S Lim, and K-R Lee shaped up the study design. H Na, G Hong, J Park, and Y Kim performed the experiments. H. Na, M-K Lee, J Ahn, and Y Lee analyzed the sequences downloaded from the public databases. S Lee and K-R Lee wrote the manuscript with help of M-K Lee. S Lee and K-R Lee edited the final manuscript.

Declaration of competing interest

The authors declare that there is no conflict of interest.

Acknowledgment

We thank molecular diagnostics department at SCL for fully supporting the SARS-CoV-2 diagnostic test.

References

Alm, E., Broberg, E.K., Connor, T., Hodcroft, E.B., Komissarov, A.B., Maurer-Stroh, S., Melidou, A., Neher, R.A., O'Toole, Á., Pereyaslov, D., 2020. Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020, Euro. Surveill. 25(32):pii=2001410.

Amanat, F., Krammer, F., 2020. SARS-CoV-2 vaccines: status report. Immunity 52 (4), 583–589.

Banerjee, A.K., Begum, F., Ray, U., 2020. Mutation hot spots in spike protein of COVID-19. Preprints, 2020040281. https://doi.org/10.20944/preprints202004.0281.v1.

Bhattacharyya, C., Das, C., Ghosh, A., Singh, A.K., Mukherjee, S., Majumder, P.P., Basu, A., Biswas, N.K., 2020. 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. https://doi.org/10.1101/2020.05.04.075911.

Bhattacharyya, C., Das, C., Ghosh, A., Singh, A.K., Mukherjee, S., Majumder, P.P., Basu, A., Biswas, N.K., 2021. SARS-CoV-2 614G creates an elastase cleavage site enhancing its spread in high AAT-deficient regions. Infect. Genet. Evol. 90, 104760.

Biswas, S.K., Mudi, S.R., 2020. Spike protein D614G and RdRp P323L: the SARS-CoV-2 mutations associated with severity of COVID-19. Genomics Inform. 18 (4), e44.

Chand, G.B., Banerjee, A., Azad, G.K., 2020. Identification of twenty-five mutations in surface glycoprotein (Spike) of SARS-CoV-2 among Indian isolates and their impact on protein dynamics. Gene Rep. 21, 100891.

Gupta, V., Haider, S., Verma, M., Ponnusamy, K., Malik, Z.M., Singhvi, N., Verma, H., Kumar, R., Sood, U., Hira, P., Satija, S., Lal, R., 2020. Multi-omics and integrated network approach to unveil evolutionary patterns, mutational hotspots, functional crosstalk and regulatory interactions in SARS-CoV-2. bioRxiv. https://doi.org/10.1101/2020.06.20.162560 preprint.

Hou, Y.J., Chiba, S., Halfmann, P., Here, C., Kuroda, M., Dinnon III, K.H., Leist, S.R., Schäfer, A., Nakajima, N., Takahashi, K., Lee, R.E., Mascenik, T.M., Graham, R., Edwards, C.E., Tse, L.V., Okuda, K., Markmann, A.J., Bartelt, L., de Silva, A., Margolis, D.M., Boucher, R.C., Randell, S., Suzuki, T., Gralinski, L.E., Kawaoka, Y., Baric, R.S., 2020. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. Science, eabe8499. https://doi.org/10.1126/science.abe8499.

Huang, S.-W., Miller, S.O., Yen, C.-H., Wang, S.-F., 2021. Impact of genetic variability in ACE2 expression on the evolutionary dynamics of SARS-CoV-2 spike D614G mutation. Genes 12, 16.

Isabel, S., Graña-Miraglia, L., Gutierrez, J.M., Bundalovic-Torma, C., Groves, H.E., Isabel, M.R., Eshaghi, A., Patel, S.N., Gubbay, J.B., Poutanen, T., Guttman, D.S., Poutanen, S.M., 2020. Evolutionary and structural analyses of SARS-CoV-2 D614G spike protein mutation now documented worldwide. Sci. Rep. 10, 14031.

Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J.W., Kim, V.N., Chang, H., 2020. The architecture of SARS-CoV-2 transcriptome. Cell 181, 914–921.

Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B., Hastie, K.M., Parker, M.D., Partridge, D.G., Evans, C.M., Freeman, T.M., de Silva, T.I., Sheffield COVID-19 Genomics Group, McDanal, C., Perez, L.G., Tang, H., Moon-Walker, A., Whelan, S.P., LaBranche, C.C., Saphire, E.O., Montefiori, D.C., 2020. Tracking changes in SARS-

CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 182, 812–827.
Kumar, R., Verma, H., Singhvi, N., Sood, U., Gupta, V., Singh, M., Kumari, R., Hira, P., Nagar, S., Talwar, C., Nayyar, N., Anand, S., Rawat, C.D., Verma, M., Negi, R.K., Singh, Y., Lal, R., 2020. Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogeographical distribution. mSystems 5, e00505-20.
Laamarti, M., Alouane, T., Kartti, S., Chemao-Elfihri, M.W., Hakmi, M., Essabbar, A., Laamarti, M., Hlali, H., Bendani, H., Boumajdi, N., Benhrif, O., Allam, L., Hafidi, N. E., Jaoudi, E.R., Allali, I., Marchoudi, N., Fekkak, J., Benrahma, H., Nejjari, C., Amzazi, S., Belyamani, L., Ibrahimi, A., 2020. Large scale genomic analysis of 3067 SARS-CoV-2 genomes reveals a clonal geo-distribution and a rich genetic variations of hotspots mutation. PLoS One 15 (11), e0240345.
Li, Q., Wu, J., Nie, J., Zhang, L., Hao, H., Liu, S., Zhao, C., Zhang, Q., Liu, H., Nie, L., Qin, H., Wang, M., Lu, Q., Li, X., Sun, Q., Liu, J., Zhang, L., Li, X., Huang, W., Wang, Y., 2020. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell 182, 1284–1294.
Moniruzzaman, M., Hossain, M.U., Islam, M.N., Rahman, M.H., Ahmed, I., Rahman, T.A., Bhattacharjee, A., Amin, M.R., Rashed, A., Keya, C.A., Das, K.C., Salimullah, M., 2020. Coding-complete genome sequence of SARS-CoV-2 isolate from Bangladesh by Sanger sequencing. Microbiol. Resour. Announc. 9 (28), e00626-20.
Ozono, S., Zhang, Y., Ode, H., Sano, K., Tan, T.S., Imai, K., Miyoshi, K., Kishigami, S., Ueno, T., Iwatani, Y., Suzuki, T., Tokunaga, K., 2021. SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. Nat. Commun. 12, 848.
Park, A.K., Kim, I.-H., Kim, J., Kim, J.-M., Kim, H.M., Lee, C.Y., Han, M.-K., Rhie, G.-E., Kwon, D., Nam, J.-G., Park, Y.-J., Gwack, J., Lee, N.-J., Woo, S.H., No, J.S., Lee, J., Ha, J., Rhee, J., Yoo, C.-K., Kim, E.-J., 2021. Genomic surveillance of SARS-CoV-2: distribution of clades in the Republic of Korea in 2020. Osong Public Health Res. Perspect. 12, 37–43.
Shah, A., Rashid, F., Aziz, A., Jan, A.U., Suleman, M., 2020. Genetic characterization of structural and open reading Fram-8 proteins of SARS-CoV-2 isolates from different countries. Gene Rep. 21, 100886.
Shang, J., Wan, Y., Luo, C., Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., Li, F., 2020. Cell entry mechanisms of SARS-CoV-2. Proc. Natl. Acad. Sci. 117, 11727–11734.
Tabibzadeh, A., Zamani, F., Laali, A., Esghaei, M., Tameshkel, F.S., Keyvani, H., Makiani, M.J., Panahi, M., Motamed, N., Perumal, D., Khoonsari, M., Ajdarkosh, H., Sohrabi, M., Ghanbari, B., Savaj, S., Mosavi-Jarrahi, A., Niya, M.H.K., 2020. SARS-CoV-2 molecular and phylogenetic analysis in COVID-19 patients: a preliminary report from Iran. Infect. Genet. Evol. 84, 104387.
Volz, E.M., Hill, V., McCrone, J.T., Price, A., Jorgensen, D., O'Toole, Á., Southgate, J., Johnson, R., Jackson, B., Nascimento, F.F., Rey, S.M., Nicholls, S.M., Colquhoun, R. M., Filipe, A.S., Shepherd, J., Pascall, D.J., Shah, R., Jesudason, N., Li, K., Jarrett, R., Pacchiarini, N., Bull, M., Geidelberg, L., Siveroni, I., COG-UK Consortium, Goodfellow, I., Loman, N.J., Pybus, O.G., Robertson, D.L., Thomson, E.C., Rambaut, A., Connor, T.R., 2021. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. Cell 184, 64–75.e11.
Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A., Diehl, W.E., Dauphin, A., Carbone, C., Veinotte, K., Egri, S.B., Schaffner, S. F., Lemieux, J.E., Munro, J.B., Rafique, A., Barve, A., Sabeti, P.C., Kyratsous, C.A., Dudkina, N.V., Shen, K., Luban, J., 2020. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. Cell 183, 1–13.
Zhang, L., Jackson, C.B., Mou, H., Ojha, A., Peng, H., Quinlan, B.D., Rangarajan, E.S., Pan, A., Vanderheiden, A., Suthar, M.S., Li, W., Izard, T., Rader, C., Farzan, M., Choe, H., 2020. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. Nat. Commun. 11, 6013.
Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., Tan, W., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 382, 727–733.