



Host- and Species-Dependent Quasispecies Divergence of Severe Acute Respiratory Syndrome Coronavirus-2 in Non-human Primate Models

Eun-Ha Hwang^{1,2}, Hoyin Chung^{1,3}, Green Kim^{1,2}, Hanseul Oh¹, You Jung An¹, Philyong Kang⁴, Choong-min Ryu⁵, Jong-Hwan Park², Jungjoo Hong^{1*} and Bon-Sang Koo^{1*}

OPEN ACCESS

Edited by:

Eric Daniel Becraft,
University of North Alabama,
United States

Reviewed by:

Hadi M. Yassine,
Qatar University, Qatar
Juan C. De La Torre,
The Scripps Research Institute,
United States

*Correspondence:

Jungjoo Hong
hong75@kribb.re.kr
Bon-Sang Koo
porco9@kribb.re.kr

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 14 April 2021

Accepted: 02 June 2021

Published: 09 July 2021

Citation:

Hwang E-H, Chung H, Kim G,
Oh H, An YJ, Kang P, Ryu C-m,
Park J-H, Hong J and Koo B-S (2021)
Host- and Species-Dependent
Quasispecies Divergence of Severe
Acute Respiratory Syndrome
Coronavirus-2 in Non-human Primate
Models. *Front. Microbiol.* 12:694897.
doi: 10.3389/fmicb.2021.694897

¹ National Primate Research Center, Korea Research Institute of Bioscience and Biotechnology, Cheongju, South Korea,

² Laboratory of Animal Medicine, College of Veterinary Medicine, Chonnam National University, Gwangju, South Korea,

³ Department of Microbiology, College of Natural Sciences, Chungbuk National University, Cheongju, South Korea,

⁴ Futuristic Animal Resource Center, Korea Research Institute of Bioscience and Biotechnology, Cheongju, South Korea,

⁵ Infectious Disease Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea

Recently, newly emerging variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have been continuously reported worldwide. However, the precise evaluation of SARS-CoV-2 microevolution in host is very limited because the exact genetic information of infected virus could not be acquired in human researches. In this report, we performed deep sequencing for seed virus and SARS-CoV-2 isolated in eight cynomolgus and rhesus macaques at 3 days postinoculation and evaluated single-nucleotide polymorphisms (SNPs) in SARS-CoV-2 by variant analysis. A total of 69 single-nucleotide variants (SNVs) were present in the 5'-untranslated region (UTR), 3'-UTR, *ORF1ab*, *S*, *ORF3a*, *ORF8*, and *N* genes of the seed virus passaged in VERO cells. Between those present on the seed virus and those on each SARS-CoV-2 isolated from the lungs of the macaques, a total of 29 variants was identified in 4 coding proteins (*ORF1ab*, *S*, *ORF3a*, and *N*) and non-coding regions (5'- and 3'-UTR). Variant number was significantly different according to individuals and ranged from 2 to 11. Moreover, the average major frequency variation was identified in six sites between the cynomolgus monkeys and rhesus macaques. As with diverse SNPs in SARS-CoV-2, the values of viral titers in lungs were significantly different according to individuals and species. Our study first revealed that the genomes of SARS-CoV-2 differ according to individuals and species despite infection of the identical virus in non-human primates (NHPs). These results are important for the interpretation of longitudinal studies evaluating the evolution of the SARS-CoV-2 in human beings and development of new diagnostics, vaccine, and therapeutics targeting SARS-CoV-2.

Keywords: SARS-CoV-2, non-human primate, single nucleotide variant, genetic variants, single nucleotide polymorphism

INTRODUCTION

Following the first outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in a Wuhan seafood market in China in December 2019, the virus has spread rapidly worldwide, resulting in significant severe respiratory disease in humans (Huang et al., 2020; Zhu et al., 2020). Associated clinical features range from asymptomatic infection to mild or severe disease, including fever, fatigue, dry cough, olfactory and taste disorder, acute respiratory distress syndrome, and multiorgan failure (Wu and McGoogan, 2020). As of March 31, 2021, a total of 127,877,462 cases and 2,796,561 deaths were confirmed in 216 countries on the World Health Organization COVID-19 dashboard. Recently, the broad vaccination was initiated in many countries. However, the rapid evolution of this virus causes concern for the efficacy of vaccine.

RNA viruses generally have higher nucleotide substitution rates than other viruses due to the low fidelity of RNA polymerase and absence of postreplication repair systems (Duffy et al., 2008). SARS-CoV-2, a single-stranded, positive-sense, enveloped RNA virus, belongs to the genus *Betacoronavirus* within the family *Coronaviridae* that contains four other human coronaviruses of SARS-CoV, Middle East respiratory syndrome (MERS) virus, human coronavirus (HCoV)-OC43, and HCoV-HKU1 (Wu and McGoogan, 2020). The SARS-CoV-2 gene is 29,891 nucleotides in length and includes a 5'-untranslated region (UTR), open reading frame (ORF) 1a and 1b, spike (S), envelope (E), membrane (M), ORF5, ORF7a, ORF8, nucleocapsid (N), ORF10, and 3'-UTR (Wu et al., 2020). Mutations in SARS-CoV-2 have been continuously reported based on observed changes in the nucleotide or amino acid sequences of the analyzed macromolecules (Korber et al., 2020; Mercatelli and Giorgi, 2020). At present, eight distinct subgroups of SARS-CoV-2 have been reported worldwide in the Global Initiative for Sharing All Influenza Data (GISAID) database. The heterogeneity in nucleotide sequences of recent isolates is 15% greater than those of the first isolates (Benvenuto et al., 2020).

Recently, deep sequencing techniques have revealed that viruses consist of mixed bases at each nucleotide position rather than identical consensus nucleotide sequences in their host. This intrahost ensemble of virus variants has been called viral quasispecies, mutant clouds, and mutant spectra (Domingo and Perales, 2019). These groups of viruses are not simply a collection of various mutants but rather synergize to contribute to the overall characteristics of the population through cooperative interactions (Vignuzzi et al., 2006). Among SARS-CoV-2 patients, the presence of intrahost quasispecies has been identified in each structural and non-structural viral protein (Al Khatib et al., 2020; Karamitros et al., 2020; Shen et al., 2020). This microevolutionary event could affect the infectivity, virulence, and transmissibility of the virus. Minor alleles of the mutant spectrum not only confer resistance to neutralizing antibodies, cytotoxic T cells, and therapeutics but also change the immune response and virion stability (Domingo and Perales, 2019). Therefore, viral quasispecies can change the tissue tropism, virulence, infectivity, and transmissibility of

viruses, challenging the development of vaccines, therapeutics, and diagnostics.

A non-human primate (NHP) model is one of the preclinical animal models closest to humans based on phylogenetic, physiological, anatomical, and immunological aspects. For this reason, NHP models have been continuously used for the study of human infectious disease, especially SARS-CoV-2 (Gardner and Luciw, 2008; Koo et al., 2020). Viral quasispecies have been reported previously in humans (Karamitros et al., 2020; Shen et al., 2020). However, research on viral quasispecies in humans has limitations with regard to single-nucleotide polymorphism (SNP) analysis due to a lack of knowledge of the exact genetic information of the seed virus. In this report, we performed SNP and variant analyses using two NHP models infected with identical strains of SARS-CoV-2 with the goal of furthering the comprehensive understanding of evolution of this deadly virus.

MATERIALS AND METHODS

Virus

The SARS-CoV-2 strain (registration number 43326) used in this study was isolated from a Korean patient traveling to China (the National Culture Collection for Pathogens) (Cheongju, Korea). Following plaque purification of SARS-CoV-2, this virus was cultivated twice in VERO cells with Dulbecco's minimum essential medium (Welgene Inc., Daegu, Korea) containing 2% fetal bovine serum and 1% penicillin (10,000 IU/ml)/streptomycin (10,000 IU/ml) (Gibco, NY, United States) at 37°C in a 5% CO₂ incubator. At 3 days postinoculation (dpi), the cell supernatant was collected, centrifuged at 3,000 g for 10 min, and stored at -80°C. Virus titration was performed using VERO cells, and the 50% tissue culture infectious dose 50/ml (TCID₅₀/ml) was calculated according to the Reed-Muench method (Reed and Muench, 1938).

Animal Experiments

A total of eight, 3- to 6-year-old, healthy, Cambodian-origin cynomolgus (*Macaca fascicularis*) and Chinese-origin rhesus (*Macaca mulatta*) macaques were equally selected and reared in indoor cages at the National Primate Research Center (Korea Research Institute of Bioscience and Biotechnology) (Table 1). All animals were fed fresh fruits and commercial monkey feed during the experimental procedures (Teklad, Global 20% Protein Primate Diet). They were transferred and reared in negative-pressure, NHP specific isolators in an Animal Biosafety Level 3 (ABSL-3) facility (Three-Shine Inc., Seoul, Korea). After 1 week, these animals were inoculated under anesthesia with 12.5 ml (2.1 × 10⁶ TCID₅₀/ml) of the same batch of virus via oral (5 ml), intratracheal (4 ml), nasal (1 ml), conjunctival (0.5 ml), and intravenous (2 ml) routes, as previously reported (Koo et al., 2020). At 3 dpi, all six lobes of the lungs of all animals were aseptically collected, homogenized with sterilized 10-fold (w/v) phosphate buffered solution (pH 7.4), and filtered with a 0.2-μm pore size syringe filter for virus isolation. The presence of viable viruses from lung tissues was examined

TABLE 1 | Information on the animals and SARS-CoV-2 used in this report.

Animal	Species	Birth date	Sex	Body weight (kg)	Tissues	Viral loads	
						TCID ₅₀ /ml	Copies/ml
CM1	<i>Macaca fascicularis</i>	July 30, 2014	Male	5.05	Lung	9.3×10^2	2.7×10^8
CM2	<i>Macaca fascicularis</i>	September 7, 2013	Male	4.57	Lung	4.9×10^2	1.3×10^8
CM3	<i>Macaca fascicularis</i>	December 29, 2014	Female	3.2	Lung	4.3×10^4	1.7×10^9
CM4	<i>Macaca fascicularis</i>	September 29, 2013	Female	3.68	Lung	6.3×10^4	1.3×10^9
RM1	<i>Macaca mulatta</i>	March 11, 2015	Female	4.89	Lung	2.6×10^3	1.6×10^8
RM2	<i>Macaca mulatta</i>	January 16, 2015	Female	3.95	Lung	1.6×10^4	1.3×10^8
RM3	<i>Macaca mulatta</i>	October 6, 2014	Male	4.5	Lung	5.4×10^2	3.0×10^7
RM4	<i>Macaca mulatta</i>	June 6, 2015	Male	4.954	Lung	4.3×10^2	5.9×10^7

by virus isolation using VERO cells, and the TCID₅₀/ml was calculated according to the Reed–Muench method (Reed and Muench, 1938). Hematological analysis was performed from ethylenediaminetetraacetic acid (EDTA) blood samples with an auto-hematology analyzer (Mindray Inc., Shenzhen, China).

Deep Sequencing and Variant Analysis

The whole genome sequence of SARS-CoV-2 was sequenced using the BTSeq™ SARS-CoV-2 WGS Service (Celemics Inc., Seoul, Korea). Reverse transcriptase was used to synthesize complementary DNA (cDNA). The viral cDNA library was amplified with primers designed based on a published SARS-CoV-2 sequence database (data not shown). Using the amplified DNA, a DNA library was constructed through DNA fragmentation, end repair, adaptor ligation, and PCR amplification for next-generation sequencing (NGS). The DNA library was purified with CeleMag DNA Clean-up Beads (Celemics Inc.). Quantification was conducted using an Agilent 2200 TapeStation (Agilent Technologies, Sta. Clara, CA, United States). After quantification, NGS was performed using a MiSeq next-generation sequencer with 150PE (Illumina, San Diego, CA, United States). Sequence data were imported as fastq files and edited using the CLC Genomics Workbench 11.0.1. All low quality (<0.05), ambiguous (0), short (under 64 nucleotides), and adapter sequences were removed in the trimming procedure. Trimmed reads were mapped to a reference strain (GenBank number MN908947.3). Intrahost variant analysis of SARS-CoV-2 was performed against the genome of seed virus using the basic variant analysis in the CLC Genomics Workbench 11.0.1. Minor alleles were identified based on criteria of >50 read depth, >5% minor allele frequency, and >5 minor allele count. The frequency difference of the major and minor alleles was statistically verified using Fisher's exact test, with a $p \leq 0.05$ considered statistically significant.

Ethical Approval

All experimental procedures were performed in an ABL3 facility and approved by the Institutional Animal Care and Use Committee of the Korea Research Institute of Bioscience and Biotechnology (permit number KRIBB-AEC-20064).

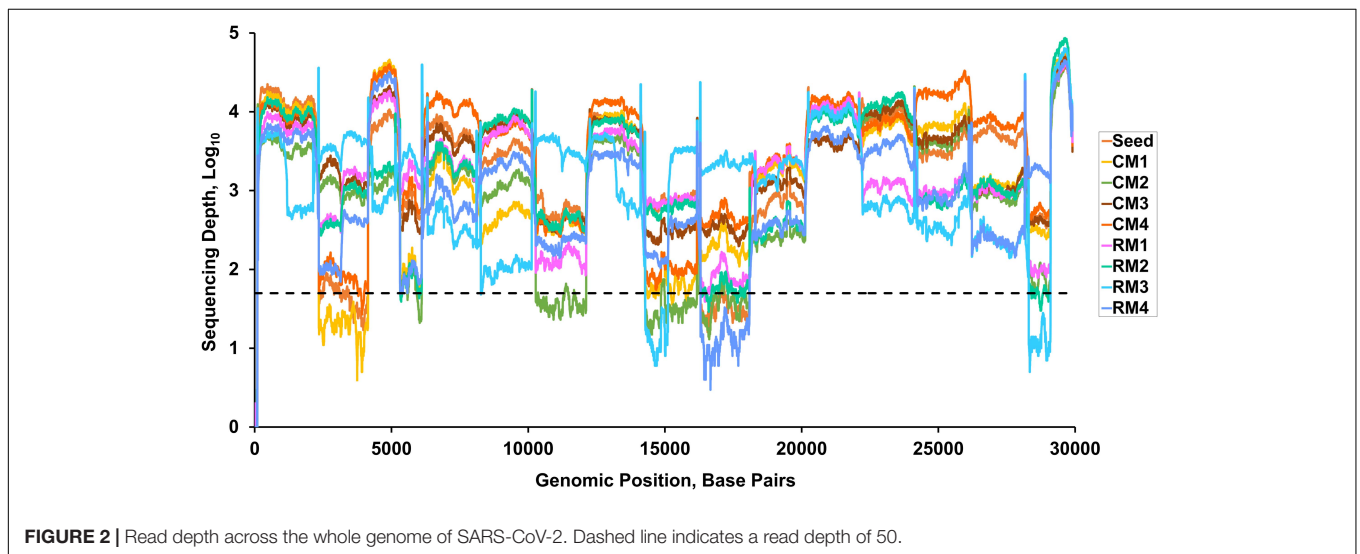
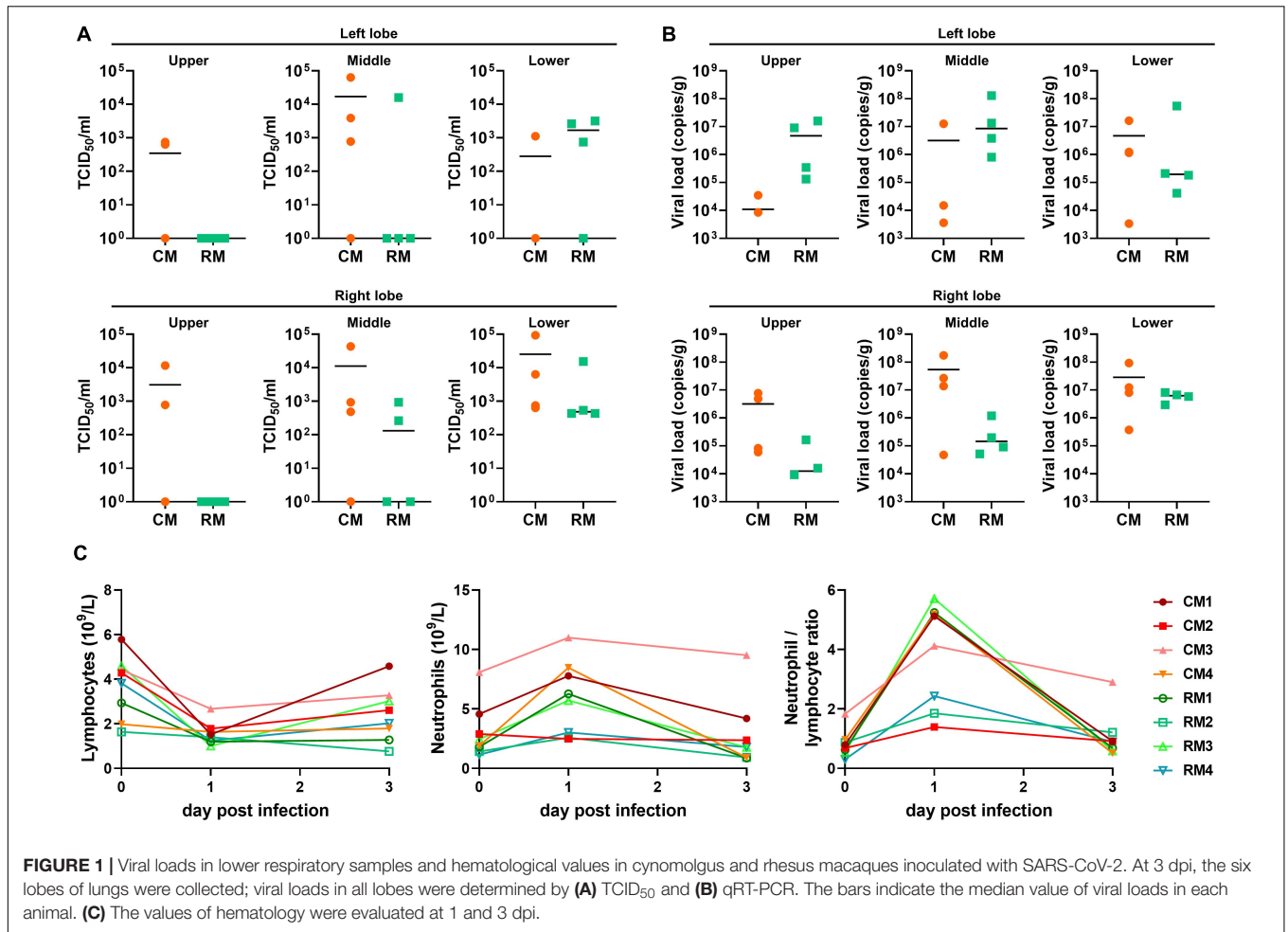
RESULTS

Sequencing Data

Viable SARS-CoV-2 was successfully isolated in the lungs of all monkeys at 3 dpi. Virus titers in each animal were successfully determined expressed as TCID₅₀ and viral copy load (Figure 1). One of the six lobes with the highest viral titers in each animal was selected for further analysis based on TCID₅₀ and viral copy number (Table 1). In addition, hematological values, body weight, rectal temperature, and respiratory rates of animals were described in Figure 1 and Supplementary Table 1. Using deep sequencing, an average of 2,595,447 reads per sample were obtained after quality control. After assembly mapping these data to the reference genome of SARS-CoV-2 (NCBI accession number MN908947), the average number of mapped reads per sample was 1,294,312. The mean read depth was 4,915×, and the breadth of coverage rate was 99.74% of the 29,903 nucleotides in the SARS-CoV-2 genome (Supplementary Table 2). Similar fluctuations in the pattern of read depth at each nucleotide position were identified among samples (Figure 2). All sequence data were submitted and deposited in the GenBank Sequence Read Archive (BioProject number PRJNA688005).

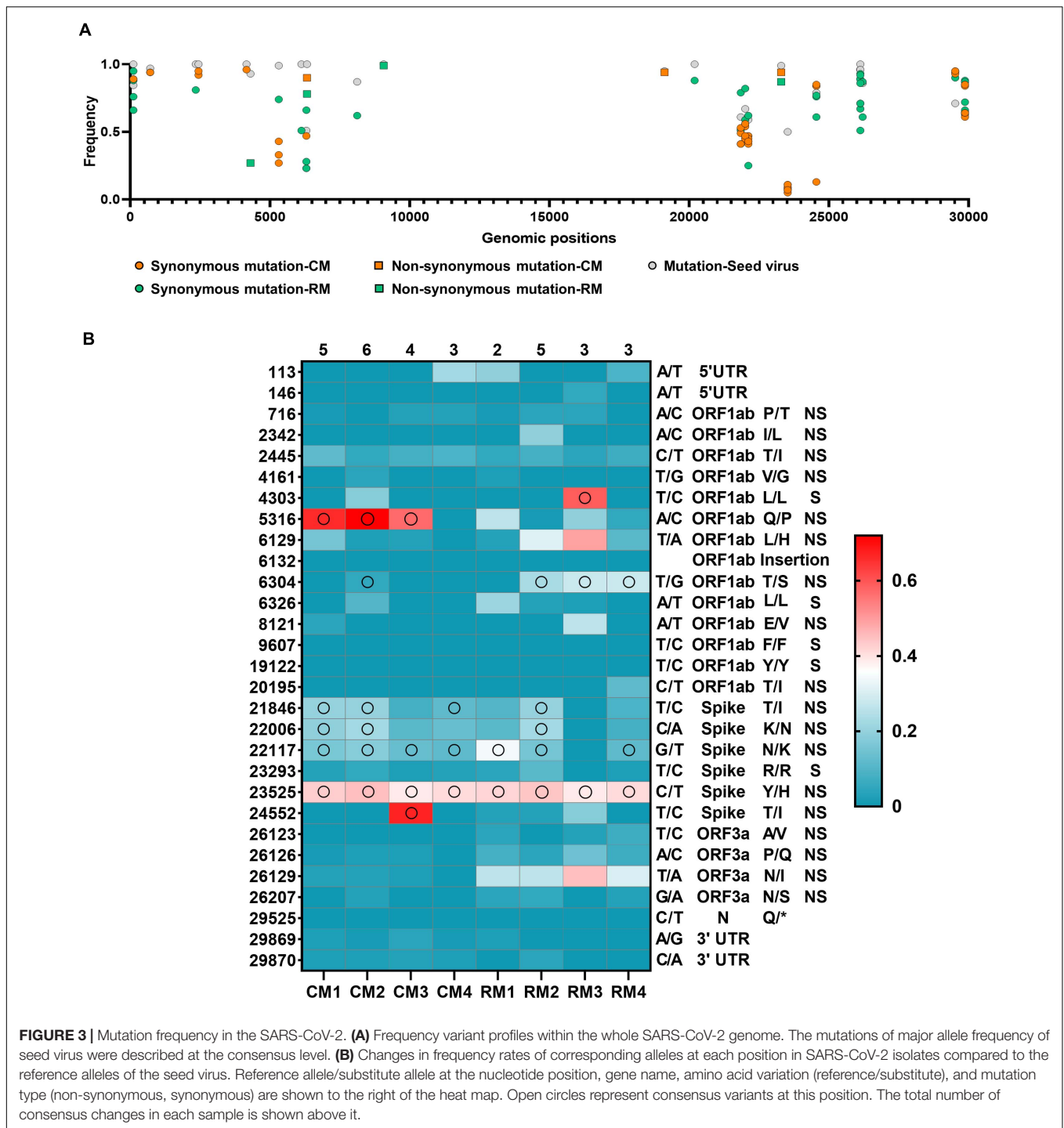
Intrahost Variant Analysis

A total of 69 SNVs were identified in the 5'- and 3'-UTRs and the *ORF1ab*, *S*, *ORF3a*, *ORF8*, and *N* genes of the seed virus (Supplementary Table 3). Frequency mutation of major alleles in each isolate was evaluated in the different sites between isolates and seed virus at the consensus level (Figure 3A). Non-synonymous mutations were mainly observed in the *S* gene of the seed virus. Different consensus sequences were identified in three sites of the *ORF1ab* and four sites of the *S* genes between the seed and isolated SARS-CoV-2 viruses (Figure 3B). All four mutation sites in the *S* gene were non-synonymous substitutions, whereas two of three substitutions in *non-structural protein* (*NSP*) 3 within *ORF1ab* corresponded to non-synonymous substitutions (Figure 3). On variant analysis based on >5% minor allele frequency and >5 minor allele count, a total of 29 variants was identified in 4 coding genes (*ORF1ab*, *S*, *ORF3a*, and *N*) and non-coding regions (5'- and 3'-UTR). The most common variant type (28/29) was SNV, with one insertion site (Figure 4A). Among the analyzed genes, the number of identified SNVs was highest



in *ORF1ab* (14/28), followed by *S* (6/28), *ORF3a* (4/28), 5'-UTR (2/28), 3'-UTR (2/28), and *N* (1/28) (Figure 3B). These SNVs in viral coding genes mainly corresponded to non-synonymous mutations (18/24, 75.0%) rather than synonymous (5/24, 20.8%)

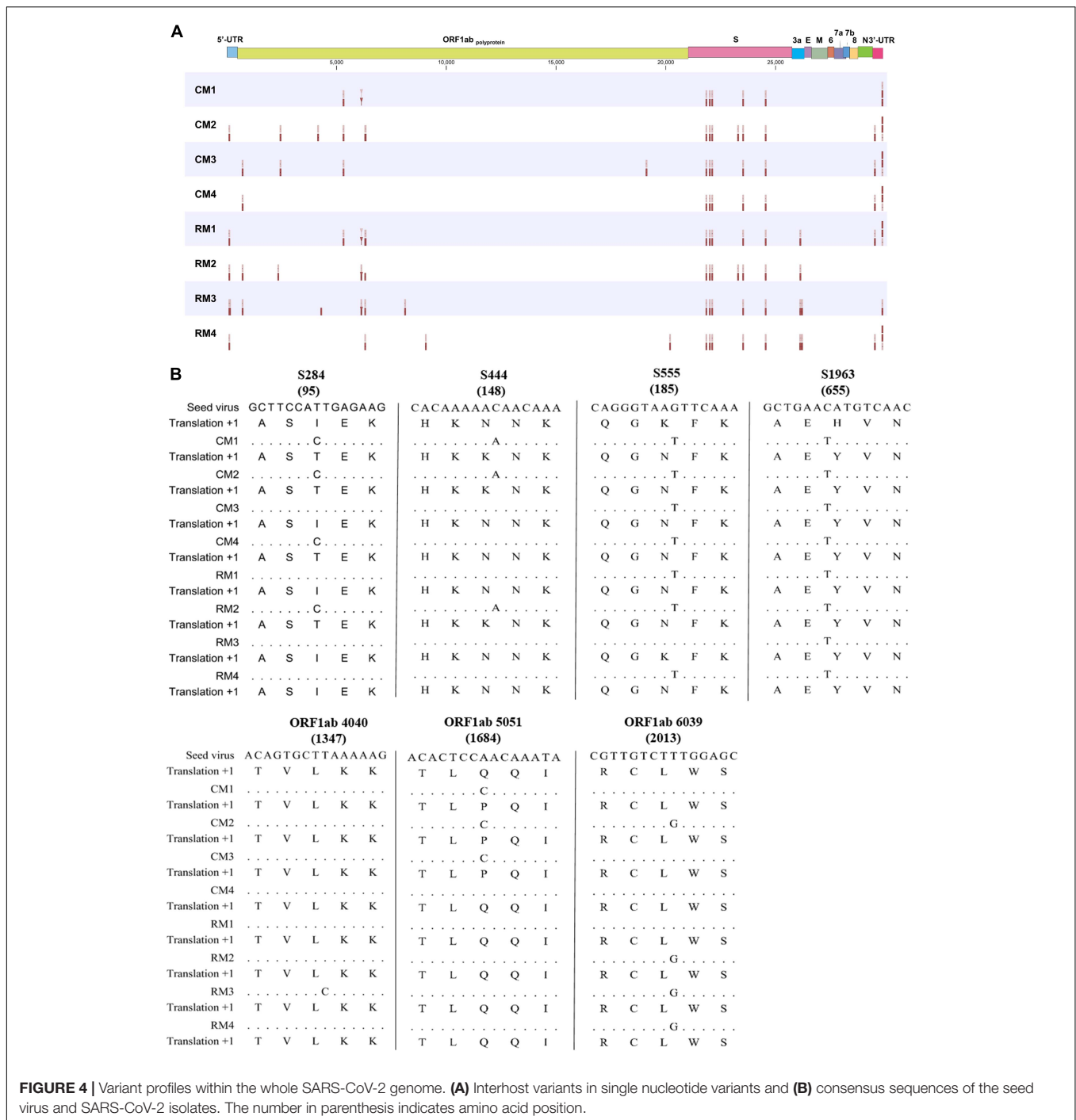
or stop mutations (1/24, 4.2%). The rate of non-synonymous mutations was highest in *ORF3a* (4/4, 100%), followed by *S* (5/6, 83.3%) and *ORF1ab* (9/13, 69.2%). The stop codon was only observed in the *N* gene.



Interhost Variant Analysis

Interhost distinct SNVs and consensus sequences were identified among the genomes of SARS-CoV-2 isolated in individual macaques, even though the same seed virus replicated in the lungs of these monkeys for only 3 days. Different consensus sequences were identified in three sites of the *ORF1ab* and four sites of the *S* genes between the seed and isolated SARS-CoV-2 viruses (Figure 4B). All four mutation sites in the

S gene were non-synonymous substitutions, whereas two of three substitutions in *NSP 3* within *ORF1ab* corresponded to non-synonymous substitutions. A total of seven different consensus sequences at six sites were identified among some, but not all, monkeys. Variant analysis comparing isolates to the seed virus revealed that the number of SNVs ranged from 9 to 16 among the individual monkeys (Figure 4). The number of statistically different variant sites ranged from 2



to 11 by monkey. These differences in SNVs were mainly identified in the *ORF1ab* gene. The differences at six sites (nucleotide positions 5316, 6129, 6304, 26123, 26129, and 26207) were statistically verified based on average values of major allele frequency variations between the two macaque species (**Supplementary Table 4**). Moreover, the site numbers of frequency differences in SARS-CoV-2 in rhesus monkeys were more frequent compared to those in cynomolgus monkeys in variant analysis (**Supplementary Table 5**).

DISCUSSION

SARS-CoV-2 infection has been the cause of the most prevalent and devastating respiratory disease worldwide after the first outbreak in late 2019. Recently, newly emerging variants have been continuously reported to cause significantly higher morbidity and mortality than previous viruses despite enforcement of strict biosecurity policies in many countries. Therefore, there is an increasingly urgent need for vaccines

and therapeutics against SARS-CoV-2. The precise evaluation of its viral genomic evolution should be conducted for the development of effective vaccines and therapies against SARS-CoV-2. Macaque models of cynomolgus and rhesus monkeys were successfully constructed via the combined routes, which have been reported as possible natural routes of transmission in human (Li H. et al., 2020; Lu et al., 2020). In this report, we characterized the genomic mutations of SARS-CoV-2 in two macaque species infected with the same seed virus by deep analysis.

Seventy-six SNVs were identified in the 5' and 3'-UTRs and genes that encode for four viral proteins (ORF1ab, S, ORF3a, and N) of the seed virus, which was serially passaged in VERO cells. The highest number of SNVs was observed in S gene. The S protein of SARS-CoV-2 is a well-known surface glycoprotein and is closely associated with its antigenicity, pathogenicity, and tissue tropism. Most SNVs in the S gene corresponded to non-synonymous mutations, indicating the possibility of a changed phenotype. The second highest number of SNVs was found in *ORF1ab*, a gene known to encode for a replicase/transcriptase involved in the virulence, virus-cell interaction, virus-host response, as well as replication, of SARS-CoV-2 (Graham et al., 2008). An *ORF3a* with high mutation rate could induce changes in virulence, infectivity, ion channel formation, and virus release; however, the SNVs identified in this report were not found in the functional domains of ORF3a (Hassan et al., 2020; Issa et al., 2020). Overall, the mutation sites we identified within the *ORF1ab*, S, and *ORF3a* genes seem to be related to increased fitness in VERO cells by serial passage and inversely induced attenuated pathogenicity and changed antigenicity for an original host, human (Hanley, 2011). Therefore, these sites should be evaluated for their exact role in pathogenicity and antigenicity of SARS-CoV-2 in humans, as they may represent a new target for development of novel vaccines and therapeutics.

After identical seed SARS-CoV-2 were replicated in the lungs of monkeys for 3 days, a total of 29 variant sites were identified between the seed virus and SARS-CoV-2 isolates. Based on the consensus sequence, only seven sites of allele frequency changes in the genome were determined between the seed virus and the isolates from each individual macaque. Distinct results obtained with two evaluation methods indicated that precise SNV analysis using deep sequencing should be conducted for exact evaluation of viral evolution. The nucleotide substitution rates of SARS-CoV-2 have proven to be relatively fast compared to other viruses (Duffy et al., 2008; Li X. et al., 2020). A total of eight genotypes with distinct phenotypes have been reported worldwide in the GISAID database within approximately 1 year of the first outbreak (Elbe and Buckland-Merrett, 2017). The presence and evolution of these genetic mutations in SARS-CoV-2 was evaluated using consensus sequences. Our results indicate that more complicated and faster viral evolution could proceed in host monkeys, as expected based on previous reports. Importantly, the allele frequency changes with non-synonymous mutations were most frequently observed in the S gene, which could lead to significant possible changes in antigenicity, virulence, and transmissibility of the virus. Almost all changes in *ORF1ab* occurred in the Nsp3 region, which

codes for the papain-like protease PLpro that regulates SARS-CoV-2 viral spread and innate immunity (Shin et al., 2020). In this report, the grades of lymphopenia as a cardinal marker for clinical severity were significant differentiated between cynomolgus and rhesus macaques at 3 dpi ($p = 0.02$, data not shown). Distinct frequency allele change between two species was identified in only ORF 3a (nucleotide position 26129). However, viral loads and other hematological pathological indicators (thrombocytopenia, neutrophilia, and neutrophil/lymphocyte ratio) were identified variable according to each animal, and the link between specific genetic mutation and phenotypic changes could not be determined (**Supplementary Table 1**). However, this intrahost ensemble of virus variants could contribute to the overall phenotypes of the population through cooperative interactions (Vignuzzi et al., 2006). In conclusion, non-synonymous mutations in the regions of genomes associated with changes in the antigenicity and virulence of SARS-CoV-2 occurred in hosts after a very short replication period. This phenomenon could help explain the pattern of reinfection with newly emerging genotypes that has been reported in human patients (Tillett et al., 2020; To et al., 2020).

A key finding of this report is that the same seed virus differentially evolved in eight individual monkeys despite the short replication time of 3 days (**Figure 3**). Different frequencies of variation were identified in the genomes of SARS-CoV-2 among the individual macaques; the number and position of the variation sites were different among individuals. Hypervariable regions were identified within the partial regions of genome of ORF1ab and S proteins. The 11 variations sites in RM3 were statistically identified, whereas only 2 sites were determined in CM4. These variable SARS-CoV-2 microevolution could affect distinct virus replication pattern in each macaque (**Figure 1**). The absolute number of SNV sites and pathogenicity have been reported to decrease due to bottleneck transmission among hosts (Gutiérrez et al., 2012). However, distinct viral microevolution that occurs within individual hosts could lead to new genotypes after interhost transmission. Distinct simplified SNVs were observed in a contact group after interhost bottleneck transmission of influenza virus (Varble et al., 2014). Additionally, differences in average allele frequency between the two macaque species was verified in six sites of SARS-CoV-2. In humans, differences in virulence according to race and ethnicity have been reported in the United States (Vahidy et al., 2020). Our results suggest that host genetics could induce differences in viral evolution of SARS-CoV-2, which could in turn lead to phenotypic differences, including virulence, in humans.

Recent analyses of SNVs in SARS-CoV-2 isolated from human patients have been reported (Karamitros et al., 2020; Shen et al., 2020). However, given the lack of information on the seed virus when SARS-CoV-2 is isolated from human patients, there remains difficulty in evaluating its exact viral evolution in humans. In this report, the high intrahost, interhost, and interspecies genomic variability of SARS-CoV-2 following a short replication period were verified using two species of macaques. Further researches should be conducted to determine the impact of each viral variant on clinical outcomes in NHPs. This information should be valuable not only for understanding

the viral evolution of SARS-CoV-2 in hosts but also for the development of new diagnostics, vaccines, and therapeutics for this fatal disease.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee of the Korea Research Institute of Bioscience and Biotechnology.

AUTHOR CONTRIBUTIONS

E-HH and B-SK performed the experiments and wrote the manuscript. HC, GK, HO, YA, PK, C-MR, and J-HP performed

the data analysis. JH and B-SK designed the study and supervised the experiments. All authors read and approved the final version of the manuscript.

FUNDING

This study was supported by grants (PRM1752011) from the Ministry of Science and ICT and the Korea Research Institute of Bioscience and Biotechnology Research Initiative Program (KGM4572121), South Korea.

ACKNOWLEDGMENTS

We thank the breeders and veterinarians (National Primate Research Center) who contributed to the NHPs care.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.694897/full#supplementary-material>

REFERENCES

- Al Khatib, H. A., Benslimane, F. M., Elbashir, I. E., Coyle, P. V., Al Maslamani, M. A., Al-Khal, A., et al. (2020). Within-Host Diversity of SARS-CoV-2 in COVID-19 Patients with Variable Disease Severities. *Front. Cell. Infect. Microbiol.* 10:575613. doi: 10.3389/fcimb.2020.575613
- Benvenuto, D., Giovanetti, M., Salemi, M., Prosperi, M., De Flora, C., Junior Alcantara, L. C., et al. (2020). The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathog. Glob. Health* 114, 64–67. doi: 10.1080/20477724.2020.1725339
- Domingo, E., and Perales, C. (2019). Viral quasispecies. *PLoS Genet.* 15:e1008271. doi: 10.1371/journal.pgen.1008271
- Duffy, S., Shackelton, L. A., and Holmes, E. C. (2008). Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* 9, 267–276. doi: 10.1038/nrg2323
- Elbe, S., and Buckland-Merrett, G. (2017). Data, disease and diplomacy: gISAID's innovative contribution to global health. *Glob. Chall.* 1, 33–46. doi: 10.1002/gch2.1018
- Gardner, M. B., and Luciw, P. A. (2008). Macaque models of human infectious disease. *ILAR J.* 49, 220–255. doi: 10.1093/ilar.49.2.220
- Graham, R. L., Sparks, J. S., Eckerle, L. D., Sims, A. C., and Denison, M. R. (2008). SARS coronavirus replicase proteins in pathogenesis. *Virus Res.* 133, 88–100. doi: 10.1016/j.virusres.2007.02.017
- Gutiérrez, S., Michalakis, Y., and Blanc, S. (2012). Virus population bottlenecks during within-host progression and host-to-host transmission. *Curr. Opin. Virol.* 2, 546–555. doi: 10.1016/j.coviro.2012.08.001
- Hanley, K. A. (2011). The double-edged sword: how evolution can make or break a live-attenuated virus vaccine. *Evolution* 4, 635–643. doi: 10.1007/s12052-011-0365-y
- Hassan, S. S., Choudhury, P. P., Basu, P., and Jana, S. S. (2020). Molecular conservation and Differential mutation on ORF3a gene in Indian SARS-CoV2 genomes. *Genomics* 112, 3226–3237. doi: 10.1016/j.ygeno.2020.06.016
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506.
- Issa, E., Merhi, G., Panossian, B., Salloum, T., and Tokajian, S. (2020). SARS-CoV-2 and ORF3a: nonsynonymous Mutations, Functional Domains, and Viral Pathogenesis. *mSystems* 5, e00266–20.
- Karamitros, T., Papadopoulou, G., Bousali, M., Mexias, A., Tsiodras, S., and Mentis, A. (2020). SARS-CoV-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies. *J. Clin. Virol.* 131:104585. doi: 10.1016/j.jcv.2020.104585
- Koo, B.-S., Oh, H., Kim, G., Hwang, E.-H., Jung, H., Lee, Y., et al. (2020). Transient Lymphopenia and Interstitial Pneumonia with Endotheliitis in SARS-CoV-2-Infected Macaques. *J. Infect. Dis.* 222, 1596–1600. doi: 10.1093/infdis/ji aa486
- Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., et al. (2020). Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* 182, 812–827.e19.
- Li, H., Wang, Y., Ji, M., Pei, F., Zhao, Q., Zhou, Y., et al. (2020). Transmission Routes Analysis of SARS-CoV-2: a Systematic Review and Case Report. *Front. Cell. Dev. Biol.* 8:618. doi: 10.3389/fcell.2020.00618
- Li, X., Zai, J., Zhao, Q., Nie, Q., Li, Y., Foley, B. T., et al. (2020). Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J. Med. Virol.* 92, 602–611. doi: 10.1002/jmv.25731
- Lu, S., Zhao, Y., Yu, W., Yang, Y., Gao, J., Wang, J., et al. (2020). Comparison of nonhuman primates identified the suitable model for COVID-19. *Signal. Transduct. Target Ther.* 5:157.
- Mercatelli, D., and Giorgi, F. M. (2020). Geographic and Genomic Distribution of SARS-CoV-2 Mutations. *Front. Microbiol.* 11:1800. doi: 10.3389/fmicb.2020.01800
- Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* 27, 493–497. doi: 10.1093/oxfordjournals.aje.a118408
- Shen, Z., Xiao, Y., Kang, L., Ma, W., Shi, L., Zhang, L., et al. (2020). Genomic diversity of SARS-CoV-2 in Coronavirus Disease 2019 patients. *Clin. Infect. Dis.* 71, 713–720.
- Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A., et al. (2020). Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* 587, 657–662. doi: 10.1038/s41586-020-2601-5

- Tillett, R. L., Sevinsky, J. R., Hartley, P. D., Kerwin, H., Crawford, N., Gorzalski, A., et al. (2020). Genomic evidence for reinfection with SARS-CoV-2: a case study. *Lancet Infect. Dis.* 21, 52–58. doi: 10.1016/s1473-3099(20)30764-7
- To, K. K.-W., Hung, I. F.-N., Ip, J. D., Chu, A. W.-H., Chan, W.-M., Tam, A. R., et al. (2020). Coronavirus disease 2019 (COVID-19) re-infection by a phylogenetically distinct severe acute respiratory syndrome coronavirus 2 strain confirmed by whole genome sequencing. *Clin. Infect. Dis.* 2020:ciaa1275. doi: 10.1093/cid/ciaa1275
- Vahidy, F. S., Nicolas, J. C., Meeks, J. R., Khan, O., Jones, S. L., Masud, F., et al. (2020). Racial and Ethnic Disparities in SARS-CoV-2 Pandemic: analysis of a COVID-19 Observational Registry for a Diverse US Metropolitan Population. *BMJ Open* 10:e039849. doi: 10.1136/bmjopen-2020-039849
- Varble, A., Albrecht, R. A., Backes, S., Crumiller, M., Bouvier, N. M., Sachs, D., et al. (2014). Influenza A virus transmission bottlenecks are defined by infection route and recipient host. *Cell. Host. Microbe*. 16, 691–700. doi: 10.1016/j.chom.2014.09.020
- Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E., and Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439, 344–348. doi: 10.1038/nature04388
- Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., et al. (2020). Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell. Host. Microbe*. 9, 267–276.
- Wu, Z., and McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 323, 1239–1242. doi: 10.1001/jama.2020.2648
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–733.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hwang, Chung, Kim, Oh, An, Kang, Ryu, Park, Hong and Koo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.