Clinical Infectious Diseases

BRIEF REPORT

Specific Redistribution of Severe Acute Respiratory Syndrome Coronavirus 2 Variants in the Respiratory System and Intestinal Tract

Pengcheng Du,[1,2,a] Chuan Song,[1,2,a] Rui Li,[1,2,a] Yangzi Song,[1,2,a] Jiarui Li,[1,2] Nan Ding,[1,2] Ju Zhang,[1,2] Rui Song,[3] Junyan Han,[1,2] Guiju Gao,[3] Jinglin Yue,[1,2] Ang Duan,[1,2] Yanyi Huang,[4] Jing An,[5] Jianbin Wang,[6,b] Fujie Zhang,[3,b] Chen Chen,[1,2,b] and Hui Zeng[1,2,b]

[1]Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, People's Republic of China, [2]Beijing Key Laboratory of Emerging Infectious Diseases, Beijing, People's Republic of China, [3]Clinical and Research Center of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, People's Republic of China, [4]Biomedical Pioneering Innovation Center, Beijing Advanced Innovation Center for Genomics, College of Chemistry and Molecular Engineering, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, People's Republic of China, [5]Department of Microbiology, School of Basic Medical Science, Capital Medical University, Beijing, People's Republic of China, and [6]School of Life Sciences, Beijing Advanced Innovation Center for Structural Biology, Chinese Institute for Brain Research, Tsinghua University, Beijing, People's Republic of China

Intrahost analysis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genomic sequences identified 2 viral haplotypes comprised of 3 genetically linked mutations from the respiratory and intestinal tracts of a patient with coronavirus disease 2019. Spatiotemporal data suggest that this patient initially had dual infection of 2 SARS-CoV-2 variants, which subsequently redistributed into the 2 systems.

Keywords. severe acute respiratory syndrome coronavirus 2; SARS-CoV-2; dual infection; organ-specific redistribution.

The coronavirus disease 2019 (COVID-19) pandemic has engendered a growing global health crisis with unpredictable long-term consequences for human health [1, 2]. The causative agent of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is transmitted by direct contact and exhaled respiratory droplets, with increasing evidence of potential fecal-oral transmission [3]. SARS-CoV-2 RNA has been detected in respiratory tract specimens and feces for weeks after acute infection. Live virus has been cultured from multiple sample types, including feces [4, 5]. In addition to airway epithelial cells, human intestinal epithelium has been shown to be a productive site of SARS-CoV-2 replication, suggesting that the intestine may be an alternative target organ for SARS-CoV-2 [6].

Genetic variation driven by SARS-CoV-2 adaptions occurring between and within human hosts may have important effects on the transmissibility and pathogenicity of the virus. To date, studies have documented >19 000 SARS-CoV-2 mutations in public databases [7]. Genomic heterogeneity of the virus also exists within individual patients [8], although it is still unclear whether these genetic variations represent specific adaptations to different organs. Here, by using deep sequencing, we identified 2 distinct SARS-CoV-2 variants in a COVID-19 patient distinguished by 3 genetically linked mutations and observed changes in the distribution of these variants in specimens collected from the respiratory and intestinal tracts of the patient during the course of their infection.

RESULTS

Clinical Course

A 30-year-old man tested positive for SARS-CoV-2 RNA by quantitative reverse-transcription polymerase chain reaction (qRT-PCR), and was admitted to Beijing Ditan Hospital on 29 February 2020. He had contact with a confirmed COVID-19 patient on 18 February. Prior to the admission, he exhibited mild symptoms (sore throat, dry cough, and runny nose), but no fever or diarrhea (Supplementary Figure 1). During hospitalization, chest computed tomographic scans taken on hospital days 1 and 14 were normal. He was treated with interferon-α, but no other antiviral medication. The patient's pharyngeal swabs, sputum, and feces were collected at 3-day intervals and tested by qRT-PCR to monitor viral load; the last positive test was carried out on hospital day 33 (Supplementary Figure 1). Additionally, serum specimens collected on hospital days 10, 18, and 26 were tested for SARS-CoV-2–specific antibodies. Immunoglobulin M antibodies were negative at all 3 time points and immunoglobulin G antibodies were positive on days 18 and 26 (Supplementary Figure 1). The patient was discharged on day 37 in accordance with the Chinese Diagnosis and Treatment Protocols (no clinical symptoms and 2 consecutive negative qRT-PCR tests).

Genomic Sequences and Single-Nucleotide Polymorphisms

SARS-CoV-2 genome sequencing was conducted using 2 strategies to confirm sequence data: (1) whole metatranscriptome sequencing using a TNBSEQ-T7 platform and (2) targeted SARS-CoV-2 sequencing on an Illumina NextSeq500

Received 6 September 2020; editorial decision 14 October 2020; published online 21 October 2020.

[a]P. D., C. S., R. L., and Y. S. contributed equally to this work.

[b]H. Z., C. C., F. Z., and J. W. contributed equally to this work.

Correspondence: H. Zeng, Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University and Beijing Key Laboratory of Emerging Infectious Diseases, No. 8 Jingshundongjie, Beijing 100015, China (zenghui@ccmu.edu.cn).

Clinical Infectious Diseases® 2021;73(9):e2814–7

DOI: 10.1093/cid/ciaa1617

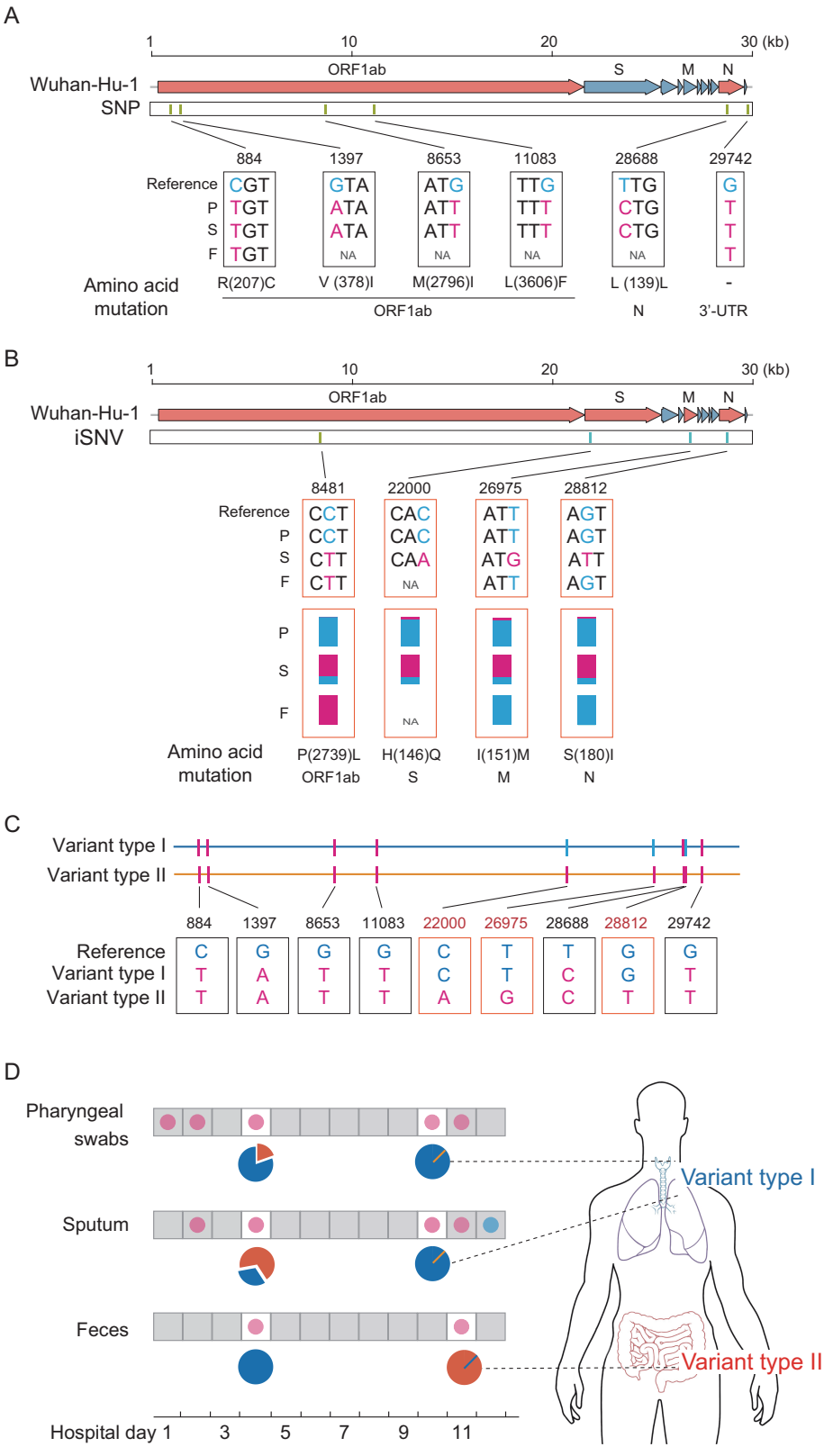

Figure 1. The distribution of single-nucleotide polymorphisms (SNPs) and intrahost single-nucleotide variations (iSNVs) and definition of the 2 variant types. *A*, The 6 SNPs and their genomic locations and genes. The scale and arrow map on the top represent the genetic positions and genes of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) reference strain Wuhan-Hu-1. The green lines below mark the positions of the 6 SNPs. The characters below represent the nucleotide substitutions and corresponding amino acid substitutions (P: pharyngeal swab; S: sputum; F: feces). *B*, The genomic position and composition of the 4 iSNVs identified. The characters display the nucleotide substitutions, and the histograms below show the allele frequencies. The characters at the bottom display the corresponding mutations of amino acid residues.

(Supplementary Methods). After patient admission and consenting, the parallel sequencing using the 2 strategies was performed on 3 samples (pharyngeal swab, sputum, and feces) collected on 3 March 2020 (hospital day 4, see Supplementary Tables 1 and 2). Sequence comparison to the reference genome (Wuhan-Hu-1, accession number NC_045512.2) revealed 6 high-quality single-nucleotide polymorphisms (SNPs) (C884T, G1397A, G8653T, G11083T, T28688C, and G29742T) in all 3 samples (Figure 1A). The first 4 SNPs were located in ORF1a and led to nonsynonymous mutations (R207C, V378I, M2796I, and L3606F); the other 2 synonymous SNPs were located in the N gene (T28688C) and the 3′-UTR region (G29742T), respectively (Figure 1A). The 6 unique SNPs were not detected in other patients in our hospital during the same period.

Haplotype Identification Based on Intrahost Single-Nucleotide Variations

We further analyzed the sequence data for the presence of intrahost single-nucleotide variations (iSNVs) in the 3 samples. We detected 4 iSNVs (C8481T, C22000A, T26975G, and G28812T; Figure 1B) in the pharyngeal swab and sputum samples, 3 of which (C8481T, C22000A, and G28812T) have been identified as SNP sites previously [7]. These 4 iSNVs were located in the ORF1a, S, M, and N genes, respectively, and led to nonsynonymous mutations (Figure 1B). In the pharyngeal swab sample, the mutation allele frequencies (MuAFs) at these 4 sites (C8481T, C22000A, T26975G, and G28812T) were 0%, 6.25%, 8.2%, and 4.69%. In marked contrast, the MuAFs at these 4 sites in sputum were 74.29%, 75.35%, 76.49%, and 79.30%. Frequencies of all 4 mutant iSNV alleles were significantly correlated across the 2 samples ($P < .001$) (Supplementary Figures 2–4). Haplotype analysis of these iSNVs identified 3 positions that were closely linked and located near the 3′-end of the viral genome at positions 22 000, 26 975, and 28 812 bp. Using these positions, 2 haplotypes were defined corresponding to "C-T-G" (variant I) and "A-G-T" (variant II) (Figure 1C). Variant I made up approximately 90% of the viral sequences in the pharyngeal swab, whereas variant II was predominant (~70%) in the sputum sample (Supplementary Table 3). In the fecal sample, we detected 1088 T nucleotides at position 26 975, and 289 G nucleotides at position 28 812 of the genome. We did not find any alternative allele at either of these 2 sites, which suggested that variant II was extremely low in feces.

Dynamic Alteration of 2 Variants in the Respiratory and Digestive Systems

To investigate any spatiotemporal changes in the 2 genetic variants over the course of infection, we performed a follow-up sequencing analysis on pharyngeal swab, sputum, and fecal samples collected on hospital days 10 or 11. We confirmed the presence of the same 6 SNPs originally identified on hospital day 4. However, the distribution of iSNV MuAFs was significantly different (Supplementary Figure 5; Supplementary Table 3). Variant I became overwhelmingly dominant (>99%) in the pharyngeal swab and sputum, though we still observed a few reads supporting continued presence of variant II in pharyngeal swab (0.07%) and sputum (0.04%; Figure 1D). In feces, the variant distribution was profoundly reversed, with variant II predominating and only 0.04% of variant I, despite the fact that this was the only variant detectable in the fecal sample on hospital day 4. This might be caused by stochastic sampling of the extremely low proportion minor variants in feces at the 2 time points (Supplementary Figure 6). These results based on MuAFs at the 3 iSNV sites revealed dynamic changes in distribution of the variant types in respiratory and gastrointestinal systems.

DISCUSSION

Here, we present spatiotemporal dynamic of molecular data demonstrating that 2 SARS-CoV-2 variants might initially coinfect the respiratory tract of a single patient. The case presented here is a young adult with mild symptoms, and without being subjected to specific antiviral therapy, except for atomized interferon-α. Our results were supported by sequencing data from both whole metatranscriptome sequencing and targeted SARS-CoV-2 sequencing. The ultra-deep sequencing provides high-quality unbiased data, and the targeted approach extended the coverage and sequencing depth of the viral genome (Supplementary Table 2; Supplementary Figure 2). Spatiotemporal analysis of the iSNVs provided concrete evidence to support the notion that a dual SARS-CoV-2 infection occurred in this patient. For instance, 3 iSNVs were highly genetically linked in both pharyngeal swabs and sputum in the early stage of infection (hospital day 4), presented as 2 haplotypes. Moreover, only haplotype variant I was identified in samples from the patient's respiratory tract, whereas only variant II was found in intestinal specimens approximately 1 week later. We provide concrete evidence to support the notion of presence of 2 distinct variants in a patient. First, the 3 iSNVs in variant II are highly correlated in both pharyngeal swabs and sputum in the early stage of infection. Second, the MuAF of the 3 iSNVs vary in a coordinated pattern across different samples. Finally, near-homogenous populations were identified in the samples from respiratory system (variant I) and digestive system (variant II, see Supplementary Materials). The dual infection in this patient might be ascribable to 2 possibilities: (1) The 2 distinct variants evolved previously, then were transmitted to this

C, The definition of the 2 variant types based on the 3 iSNVs. On the top are the genomic locations of the 6 SNPs and the 3 iSNVs used to define the 2 variant types, and the alleles are displayed at the bottom. *D*, The distribution and dynamics of the 2 variant types in the respiratory and digestive systems. The proportion of the 2 variant types in pharynx, sputa, and feces collected on hospital days 4 and 10/11, respectively. The pie charts present the proportion of the 2 variants (blue: variant type I; orange: variant type II). The purple and blue dots represent the positive or negative results of real-time reverse-transcription polymerase chain reaction around the 2 sequencing time points.

patient; or (2) a single variant infected this patient and then evolved into 2 distinct variants due to compensatory genetic changes at these sites.

It is known that SARS-CoV-2 infection is not confined to the lung, but can involve other parts of the respiratory tract and digestive system [9]. Multiplicity of RNA virus infection has been widely studied. The dynamic distribution of SARS-CoV-2 variants within hosts has been observed in the course of infection, as well as at different anatomical sites [10]. Moreover, the composition of different variants could also alter dynamically: Minor haplotypes could become dominant during interhost transmission [11]. Here, within an individual infected with SARS-CoV-2, we observed the spatiotemporal alteration of the dominant haplotypes in the organs, which possibly hinted at distinct tropism or adaptive advantages of the 2 variants under immune selective pressures in different organs and might allow immunological escape of the virus [12]. Alternatively, this phenomenon could simply be the random and variable outcome of the heterologous replication of virus within the interconnected respiratory and intestinal tracts. Given that 3 mutations causing amino acid alternations were identified in 3 important genes (S, M, and N), more investigations on organ-specific distribution of the variants and functional impacts of the mutations are urgently needed, particularly by the intrahost studies of polymorphic viral variants or quasispecies. Finally, the dual infection of SARS-CoV-2 variants augments the viral complexity in each individual. As the existence of multiple variants in vivo and associated pathogenic changes have implications for disease progression and therapeutics, more work is necessary to resolve whether increasing complexity in patients will increase difficulties of vaccine and drug development, and make the virus even harder to be controlled at the population level.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. The authors thank all healthcare workers involved in the diagnosis and treatment of COVID-19 in Beijing Ditan Hospital, as well as Professor Joseph D. Tucker and Dr Jonathan Julino from the University of North Carolina and Professor Mark Edwards from the University of Oxford for their kind assistance in polishing the English writing.

Potential conflicts of interest. The authors: No potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

References

1. Zhu N, Zhang D, Wang W, et al; China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med **2020**; 382:727–33.
2. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science **2020**; 368:860–8.
3. Zhang H, Kang Z, Gong H, et al. Digestive system is a potential route of COVID-19: an analysis of single-cell coexpression pattern of key proteins in viral entry process. Gut **2020**; 69:1010.
4. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA **2020**; 323:1843–4.
5. Sehmi P, Cheruiyot I. Presence of live SARS-CoV-2 virus in feces of coronavirus disease 2019 (COVID-19) patients: a rapid review. medRxiv [Preprint]. June 29, **2020** [cited 2020 Oct 12]. Available from: https://doi.org/10.1101/2020.06.27.20105429. Accessed 12 October 2020.
6. Lamers MM, Beumer J, van der Vaart J, et al. SARS-CoV-2 productively infects human gut enterocytes. Science **2020**; 369:50–4.
7. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data—from vision to reality. Euro Surveillance **2017**; 22:30494.
8. Shen Z, Xiao Y, Kang L, et al. Genomic diversity of severe acute respiratory syndrome-coronavirus 2 in patients with coronavirus disease 2019. Clin Infect Dis **2020**; 71:713–20.
9. Guan WJ, Ni ZY, Hu Y, et al; China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med **2020**; 382:1708–20.
10. Jary A, Leducq V, Malet I, et al. Evolution of viral quasispecies during SARS-CoV-2 infection. Clin Microbiol Infect **2020**; doi:10.1016/j.cmi.2020.07.032.
11. Sun F, Wang X, Tan S, et al. SARS-CoV-2 quasispecies provides insight into its genetic dynamics during infection. BioRxiv [Preprint]. August 20, **2020** [cited 2020 Oct 12]. Available from: https://doi.org/10.1101/2020.08.20.258376. Accessed 12 October 2020.
12. Karamitros T, Papadopoulou G, Bousali M, Mexias A, Tsiodras S, Mentis A. SARS-CoV-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies. J Clin Virol **2020**; 131:104585.