Rapid, widespread, and preferential increase of SARS-CoV-2 B.1.1.7 variant in Houston, TX, revealed by 8,857 genome sequences

**James M. Musser, MD, PhD¹; Randall J. Olsen, MD, PhD¹; Paul A. Christensen,
MD¹; S. Wesley Long, MD, PhD¹; Sishir Subedi, MS¹; James J. Davis, PhD²;
Jimmy Gollihar PhD³**

¹Houston Methodist Research Institute and Houston Methodist Hospital Department of
Pathology and Genomic Medicine, Houston, Texas

²Computing, Environment and Life Sciences, Argonne National Laboratory, Lemont,
Illinois

³CCDC Army Research Laboratory-South, Austin, Texas

Corresponding Author: James M. Musser, MD, PhD, Department of Pathology and
Genomic Medicine, Houston Methodist Hospital and Houston Methodist Research
Institute, Houston, TX 77030 (jmmusser@houstonmethodist.org).

Manuscript Word Count: 592

Abstract

Genetic variants of the SARS-CoV-2 virus have become of great interest worldwide because they have the potential to detrimentally alter the course of the SARS-CoV-2 pandemic, and disease in individual patients. We recently sequenced 20,453 SARS-CoV-2 genomes from patients with COVID-19 disease in metropolitan Houston (population 7 million), dating from March 2020 to early February 2021. We discovered that all major variants of concern or interest are circulating in the region. To follow up on this discovery, we analyzed 8,857 genome sequences from patients in eight Houston Methodist hospitals dispersed throughout the metroplex diagnosed from January 1, 2021 to March 7, 2021. This sample represents 94% of Houston Methodist cases and 4.8% of all reported cases in metropolitan Houston during this period. We discovered rapid, widespread, and preferential increase of the SARS-CoV-2 UK B.1.1.7 throughout the region. The estimated case doubling time in the Houston area is 6.9 days. None of the 648 UK B.1.1.7 samples identified had the E484K change in spike protein that can cause decreased recognition by antibodies.

[Introduction]

Genetic variants of the SARS-CoV-2 virus are now of considerable concern internationally because some have been reported to spread rapidly, cause unusually severe infections, and be less controlled by naturally acquired or vaccine-induced immunity.¹⁻⁶ By sequencing 20,453 SARS-CoV-2 genomes dating from March 2020 to early February 2021, we recently demonstrated all major variants of concern (VOCs) or interest (VOIs) are circulating in metropolitan Houston (population 7 million).⁴ These variants emerged in Houston in late December or early January. In the aggregate, the variants constituted only 1.1% of strains causing COVID-19 disease among patients in the eight Houston Methodist hospitals geographically dispersed throughout the metropolitan region. We identified 230 patients infected with a VOC or VOI, including 23 cases caused by B.1.1.7, a VOC reportedly increasing in frequency in several regions of the United States.⁵

Given their potential to detrimentally alter the course of the SARS-CoV-2 pandemic, and disease in individual patients, we tested the hypothesis that VOC and VOI are increasing in disease frequency in metropolitan Houston. We also sought to determine if any of the VOC/VOI caused an unusually severe clinical course.

Methods. We sequenced the genomes of 8,857 SARS-CoV-2 from patient samples acquired January 1, 2021 through March 7, 2021 (**Figure 1A**) using Illumina instruments.⁴ This sample represents 94% of the 9,445 cases of COVID-19 diagnosed in the Houston Methodist system during this period. VOCs were identified based on genome sequence data and designated by Pangolin (<https://cov->

lineages.org/pangolin.html). Patient metadata were acquired from the electronic medical record by standard informatics methods. Patient home address zip codes were used to visualize the geospatial distribution of the B.1.1.7 VOC. Figures were generated with Tableau version 2020.3.4 (<https://www.tableau.com/>).

Results. We identified 648 cases caused by VOC B.1.1.7 and discovered that the variant increased in frequency during the study period (**Figure 1A**); in contrast, the other VOCs/VOIs did not (data not shown). Patients infected with B.1.1.7 were distributed throughout the entire metropolitan region (**Figure 1B**). Together, these data indicate rapid, widespread, and preferential dissemination of B.1.1.7 in the Houston area, with an estimated case doubling time of 6.9 days (**Figure 1C**). None of the 648 B.1.1.7 VOC had the E484K change in spike protein that can cause decreased recognition by antibodies.

To gain additional insight into these B.1.1.7 cases, we examined available patient metadata. Compared to non-B.1.1.7 cases, there was no significant difference in patient age, gender, median length of stay, or mortality (**Table 1**). However, significantly more patients with B.1.1.7 were hospitalized. Nasopharyngeal specimens from patients with the B.1.1.7 VOC had significantly lower Ct values (frequently used as a proxy for virus load) compared to non-B.1.1.7 cases, suggesting higher virus loads on initial diagnosis (**Table 1**). This finding is consistent with increased transmissibility of B.1.1.7.

Discussion. Our genome sequence data document a very rapid upward trajectory of VOC B.1.1.7 cases in metropolitan Houston. Importantly, there was no analogous

increase in other variants. Our study has several limitations. During the study period, 197,817 cases of COVID-19 were reported in Harris County and eight contiguous counties (https://usafactsstatic.blob.core.windows.net/public/data/covid-19/covid_confirmed_usafacts.csv; last accessed March 12, 2021). Thus, our genome sample represents 4.8% of all cases in the metropolitan region and 94% of Houston Methodist cases. Our eight hospitals are widely distributed across the metropolitan region and serve patients who are demographically and geographically highly diverse, but it is likely our sample underrepresents SARS-CoV-2 genotypes circulating in difficult-to-reach populations, such as homeless and otherwise disenfranchised individuals. Finally, our hospitals care mainly for adult patients, which means that SARS-CoV-2 variants infecting pediatric patients are underrepresented in our study, although these are rare.

Author Contributions: Dr. Musser had full access to all study data and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Musser, Olsen, Christensen, Long

Data acquisition, analysis, or interpretation of data: All authors

Drafting of the manuscript: All authors

Statistical analysis: Christensen

Obtained funding: Musser and Davis

Supervision: Musser

Authorship: Musser, Olsen, Christensen, and Long contributed equally and are co-first authors

Conflict of Interest Disclosures: The authors report no conflict of interest.

Institutional Review Board statement: This work was approved by the Houston Methodist Research institutional review board (IRB1010-0199).

Funding/Support: The research was supported by the Houston Methodist Academic Institute Infectious Diseases Fund and many generous Houston philanthropists. James Davis was funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. 75N93019C00076.

Role of Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The findings and conclusions in this article are those of the authors and do not necessarily reflect the views of the U.S. Army.

Additional Contributions: We are indebted to Drs. Marc Boom and Dirk Sostman for their ongoing support. We thank Matthew Ojeda Saavedra, Prasanti Yerramilli, Layne Pruitt, Kristina Reppond, Madison N. Shyer, Jessica Cambric. Dr. Ilya J. Finkelstein, Dr. Jessica W. Podnar, Hung-Che Kuo and G. Nguyen for genome sequencing support. We acknowledge the assistance of Robert Olson, Maulik Shukla and Marcus Nguyen at the Bacterial and Viral Bioinformatics Resource Center for their assistance with sequence analysis. Dr. Parsa Hodjat contributed to formulating the real-time Tableau dashboard analysis. We thank Drs. Kathryn Stockbauer and Sasha Pejerrey for editorial contributions.

Additional Information: Genome data used in this study have been deposited to GISAID.

References

1. Horby P, Huntley C, Davies N, et al. *Paper from the New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) on new coronavirus (COVID-19) variant B.1.1.7*. Posted January 21, 2021. Accessed February 8, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955239/NERVTAG_paper_on_variant_of_concern__VOC__B.1.1.7.pdf2021.
2. Luring AS, Hodcroft EB. Genetic Variants of SARS-CoV-2-What Do They Mean? *JAMA*. 2021;325(6):529-531.
3. Li Q, Wu J, Nie J, et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell*. 2020;182(5):1284-1294 e1289.
4. Long SW, Olsen RJ, Christensen PA, et al. Sequence Analysis of 20,453 SARS-CoV-2 Genomes from the Houston Metropolitan Area Identifies the Emergence and Widespread Distribution of Multiple Isolates of All Major Variants of Concern. *medRxiv*. 2021 doi: 10.1101/2021.02.26.21252227.
5. Washington NL, Gangavarapu K, Zeller M, et al. Genomic epidemiology identifies emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *medRxiv*. 2021. doi: 10.1101/2021.02.06.21251159.
6. Zhang W, Davis BD, Chen SS, Sincuir Martinez JM, Plummer JT, Vail E. Emergence of a Novel SARS-CoV-2 Variant in Southern California. *JAMA*. 2021. doi: 10.1001/jama.2021.1612.

[Figure legend]

Figure 1. Epidemiologic curve, geospatial distribution, and growth model for the UK B.1.1.7 VOC in metropolitan Houston, Texas. **A**, Epidemiologic curve of COVID-19 cases in the Houston Methodist hospital system. The curve shows the increased frequency of cases caused by B.1.1.7 from January 1, 2021 to March 7, 2021. **B**, Geospatial distribution of B.1.1.7 cases identified in the study. The home address zip code for each patient was used and figures were generated using Tableau version 2020.3.4. **C**, Growth model for the B.1.1.7 variant in metropolitan Houston. The relative growth of the B.1.1.7 variant was fitted to logistic growth model using the curve fit procedure with nonlinear least-squares estimation in SciPy (1.1.0) (<https://docs.scipy.org/doc/scipy/reference/release.1.1.0.html>). Growth rate: 0.1012 [95% confidence interval (CI): 0.0631 - 0.1393]; doubling time: 6.91 days (95% CI: 5.02 - 11.08).

Table 1. Summary of pertinent patient metadata for the 8,857 unique patients

	B.1.1.7 Variant	Other Variants	Total	Statistical Analysis
No. (%) with data	648 (7.3)	8209 (92.7)	8857	
Patient Characteristics				
Median Age	54.0	54.5		P=0.5329 Mann-Whitney
Female	336 (51.9)	4366 (53.2)	4702	P=0.5134 Fisher's exact test
Male	312 (48.1)	3843 (46.8)	4155	
Admission Data				
Admitted	333 (51.4)	3747 (45.6)	4080	P=0.0002 Fisher's exact test
Not Admitted	315 (48.6)	4462 (54.4)	4777	Odds Ratio: 1.358 (95% CI 1.156 – 1.595)
Median LOS (Days)	5.5	5.3	5.3	P=0.3159 Mann-Whitney
Max Respiratory Support				
Room Air	60 (18.0)	755 (20.1)	815	P=0.5519 Chi-square
Supplemental Oxygen	238 (71.5)	2609 (69.6)	2847	
Mechanical Ventilation	32 (9.6)	366 (9.8)	398	
ECMO	3 (0.9)	17 (0.5)	20	
Mortality				
Alive	618 (95.4)	7859 (95.7)	8477	P=0.6154 Fisher's exact test
Deceased	30 (4.6)	350 (4.3)	380	Odds Ratio: 1.090 (95% CI 0.7478 – 1.587)
Median PCR Cycle Threshold				
Abbott Alinity	22.9 (n=228)	26.5 (n=2975)	(n=3202)	P=0.0006 Mann-Whitney
Hologic Panther	24.4 (n=62)	26.1 (n=1420)	(n=1482)	P=0.0435 Mann-Whitney

Data include 8,857 unique patients with high-quality sequence results between January 1, 2021, and March 7, 2021.

CI: confidence interval; ECMO: extracorporeal membrane oxygenation; LOS: length of stay