Phylogenomics of SARS-CoV-2 in Emergency Shelters for People Experiencing Homelessness

Amanda M. Casto[1,2], Julia H. Rogers[1], Amy C. Link[1], Michael Boeckh[1,2,3], Michael L. Jackson[4], Timothy M. Uyeki[5], Janet A. Englund[3,6], Lea M. Starita[3], Helen Y. Chu[1]

[1]Division of Allergy and Infectious Diseases, Department of Medicine, University of Washington, Seattle, WA, USA

[2] Vaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[3]Brotman Baty Institute for Precision Medicine, Seattle, WA, USA

[4]Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA

[5]Influenza Division, Centers for Disease Control and Prevention, Atlanta, GA, USA

[6]Seattle Children's Research Institute, Seattle, WA, USA

Summary: Using genomic data, we estimate that 28 cases of SARS-CoV-2 infection in 8 shelters in King County, Washington, were the result of at least 12 introduction events and that at least 16 of these cases resulted from intra-shelter viral transmission.

Corresponding Author: Amanda M. Casto, MD, PhD**,** UW Medicine**,** 750 Republican St.**,** Chu Lab, Box 358061, Seattle, WA 98109, USA, amcasto@uw.edu**,** 864.221.0919

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Corresponding Author

Amanda M. Casto, MD, PhD

UW Medicine

750 Republican St.

Chu Lab,

Box 358061,

Seattle, WA 98109

amcasto@uw.edu

864.221.0919

Alternate Corresponding Author

Helen Y. Chu, MD

UW Medicine

750 Republican St.

Chu Lab,

Box 358061,

Seattle, WA 98109

helenchu@uw.edu

617.935.9509

Abstract

Residents and staff of emergency shelters for people experiencing homelessness (PEH) are at high risk of infection with SARS-CoV-2. The importance of shelter-related transmission of SARS-CoV-2 in this population remains unclear. It is also unknown whether there is significant spread of shelter-related viruses into surrounding communities. We analyzed genome sequence data for 28 SARS-CoV-2-positive specimens collected from 8 shelters in King County, Washington between March and October, 2020. We identified at least 12 separate SARS-CoV-2 introduction events into these 8 shelters and estimated that 57% (16 out of 28) of the examined cases of SARS-CoV-2 infection were the result of intra-shelter transmission. However, we identified just a few SARS-CoV-2 specimens from Washington that were possible descendants of shelter viruses. Our data suggest that SARS-CoV-2 spread in shelters is common, but we did not observe evidence of wide-spread transmission of shelter-related viruses into the general population.

Key Words: SARS-CoV-2, Homeless, Shelters, Residential Facilities, Transmission, Genome, Phylogenetic Tree

Introduction

Persons living and working in residential or communal housing facilities are at high risk of COVID-19, the disease caused by infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Outbreaks of SARS-CoV-2 in skilled nursing facilities [1–5], hospitals [6], prisons [7–9], and dormitories [10,11] have been well documented. These outbreaks have frequently resulted in significant morbidity and mortality and may be an important contributor to SARS-CoV-2 transmission in the communities where these facilities are located [12]. Residents and staff of emergency shelters that serve persons experiencing homelessness (PEH) may be especially vulnerable to outbreaks of SARS-CoV-2 [13–15]. Shelters are frequently crowded and sleeping spaces, restrooms, and bathing facilities are often shared among many residents. Also, unlike residents of skilled nursing facilities, hospitalized patients, and incarcerated persons, sheltered PEH may regularly leave the shelter [15,16], putting them at risk for acquiring SARS-CoV-2 outside the shelter environment.

Recent studies of SARS-CoV-2 outbreaks in emergency shelters have provided insights into the impact of the pandemic on residents of these facilities. A point prevalence of infection of 30% or greater among shelter residents has been observed in multiple studies during outbreak investigations [13,16–18]. In at least one case, a high shelter prevalence of SARS-CoV-2 infection was observed, even though the prevalence of infection in the surrounding community was low [6]. Studies have also documented simultaneous SARS-CoV-2 infections among shelter residents and staff [13,17–20] with a high percentage of asymptomatic infections among those testing positive [17,18,21,22]. Higher case numbers have been observed in shelters with more transient residents [16] and in shelters with shared or communal sleeping facilities [21]. SARS-CoV-2 spread has been reported among persons residing at different shelters who shared day-use facilities [19]. Finally, one study found higher rates of SARS-CoV-2 infection in emergency shelter residents than among unsheltered PEH [20], an observation perhaps attributable to crowded living conditions in shelters.

Important questions remain about SARS-CoV-2 transmission within emergency shelters and the possible role these facilities play in viral spread in surrounding communities. Specifically, it is unknown what proportion of SARS-CoV-2 infections in shelter residents and staff represent viruses acquired in the surrounding community versus within the shelter environment. It is also unknown how likely it is for viruses transmitted in shelters to spread into outside communities. In this study, we used whole genome sequencing to assess the relationship among SARS-CoV-2 specimens collected from emergency shelter residents and staff in King County, Washington from March through October of 2020 as part of an on-going surveillance program. We also used publicly available SARS-CoV-2 sequences to evaluate the relationship between these specimens and those collected from the community at large during the same time period.

Methods

Origin of Specimens

We conducted SARS-CoV-2 testing of both residents and staff of 13 emergency shelters located throughout King County, Washington from March to October, 2020, using two different approaches: routine surveillance and surge testing (Supplementary Tables S1, S2). These approaches and the lettering system used to denote shelters have been described previously [22]. Briefly, routine surveillance consisted of testing shelter residents and staff who expressed interest in being tested to members of our study team who were present on-site [22]. During March 2020, residents and staff were also required to have self-reported symptoms (such as cough, fever, sinus congestion, rhinorrhea, or sore throat) consistent with an acute respiratory illness within the previous week to be eligible for testing. After March 2020, testing through routine surveillance was open to all interested shelter residents and staff, regardless of symptoms. Surge testing consisted of one-day testing events where all residents and staff at a shelter were offered testing in response to a previously detected case or cases of SARS-CoV-2 infection at the facility. Informed consent was obtained from all study participants and specimens were obtained via participant self-collected, mid-

turbinate nasal swabs [23]. Demographic data was collected from residents and staff who underwent SARS-CoV-2 testing via a standardized questionnaire that was administered by study staff [22]. The study protocol was approved by the University of Washington Institutional Review Board. The Centers for Disease Control and Prevention determined that the data and specimen collection was not conducted or sponsored by the Federal government and the study protocol did not require Human Research Protection Office review under 45 CFR 46.102(l).

SARS-CoV-2 Testing

Specimens collected through March 18, 2020 were tested for SARS-CoV-2 using a University of Washington laboratory-developed, multiplexed RT-PCR assay targeting the SARS-CoV-2 Orf1b gene and the human RNase P gene. This assay was run in duplicate for each specimen. Specimens collected after March 18 were tested using a multiplexed RT-PCR assay that targeted both the viral Orf1b and S genes labeled with FAM fluor and the human RNaseP gene labeled with VIC or HEX fluor. For positive specimens, a cycle threshold (Ct) value was calculated by averaging the individual Ct values for each viral target as described previously [24].

Genome Sequencing and Assembly

Sequencing was attempted on all positive specimens with Ct values ≤30. RNA from specimens meeting this criterion was converted to cDNA and libraries were constructed with the Illumina TruSeq RNA Library Prep for Enrichment kit (Illumina). Library enrichment for SARS-CoV-2 sequences was performed with the Twist Respiratory Virus Research Panel (Twist Bioscience, South San Francisco, CA) or the Illumina COVID-seq Kit (Illumina, San Diego, CA). Sequencing was performed on a MiSeq or NextSeq machine. Raw reads were assembled into consensus sequences using the Seattle Flu Assembly Pipeline (https://github.com/seattleflu/assembly, Supplementary Methods). GISAID accession numbers for all consensus sequences are provided in Supplementary Table S3.

Sequence Analysis

Sequence masking, sequence alignment, and phylogenetic tree construction were performed using the Nextstrain augur computational toolset (Supplementary Methods) [25]. Trees were visualized using the Nextstrain auspice software [25]. Genomes not collected as part of this study were downloaded from the GISAID database for use in our analyses (Supplementary Methods). Quality of these publicly available genomes was assessed using the Nextclade online tool (https://clades.nextstrain.org/). Nextstrain clade assignment of genomes was also performed using Nextclade and Pangolin lineage assignment was done using the Pangolin COVID-19 lineage assigner online tool (https://pangolin.cog-uk.io) [26].

Results

Clade/lineage membership of sequenced specimens from shelters

A total of 72 specimens from 11 shelters in King County were positive for SARS-CoV-2 from March through October, 2020. Full genome sequences (>90% complete) were generated for 28 specimens from 25 unique residents and 3 unique staff members from 8 different shelters (Table 1, Supplementary Figure S1). The remaining 44 specimens could not be sequenced due to low viral RNA quantity (high Ct values) and/or quality (Supplementary Table S4, S5; Supplementary Note S1). The Pangolin lineages represented by the 28 shelter genomes included A.1 (1), B.1 (14), B.1.1.291 (1), B.1.234 (8), and B.1.240 (4). The lineage classification of sequenced genomes grouped by collection month is shown in Figure 1A with classification data for all King County (N = 575) and Washington (N = 3167) SARS-CoV-2 genomes collected from March to October, 2020, and deposited in GISAID shown in Figures 1B and 1C, respectively. Most of the 28 shelter genomes were classified as Nextstrain Clade 20A (13) or Clade 20C (13) with one genome classified as 19B and one as 20B (Supplementary Figure S2). A phylogenetic tree of the 28 shelter genomes is shown in Figure 2A while the tree in Figure 2B includes all shelter genomes plus a random sample of publicly available

SARS-CoV-2 genomes from Washington (N = 445) and from the United States (N = 757) collected during the same time period.

Genetic relationships among specimens from the same shelter

We next assessed the genetic relationship among viruses from the same shelter. We sequenced more than one specimen from 6 shelters – H, M, D (n = 2 each), I (n = 4), L (n = 7), and E (n = 9). All sequenced specimens from Shelters H and M were Clade 20C and Lineage B.1. Only one nucleotide change separated H1 from H2, collected 24 days apart, and M1 from M2, collected on the same day. In contrast, the two specimens from Shelter D, collected 11 days apart, were from different clades and lineages. The four specimens from Shelter I belonged to 3 different clades and lineages, 19B/A.1 (I1), 20A/B.1.240 (I4), and 20C/B.1 (I2, I3). Specimens I2 and I3, collected two days apart, differed at just two positions and were the only viruses from Washington with the alternative alleles C10138T, C10645T, and C23673T (Figure 2C, Supplementary Note S2, Supplementary Figure S1).

Shelters L and E both had a large number of sequenced specimens. All sequenced specimens from Shelter L were Clade 20C/Lineage B.1. L1 and L2 were genetically identical and differed by just a one nucleotide change from L3 and L6. We also noted that specimens L4 – L7 formed a monophyletic clade with just three other Washington specimens. These 7 genomes all shared the A27253G alternative allele, not seen in any other sequences from Washington (Supplementary Note S2, Supplementary Figure S1). As a group, specimens L4 – L7, all collected on April 23, were more diverse than specimens L1 – L3, all collected on April 1.

All Shelter E specimens were Clade 20A, but E1 – E8 were Lineage B.1.234 and E9 was Lineage B.1.240. E1 - E8 also all share the C3411T and T9875C alternative alleles seen in only 3 other specimens from Washington (Supplementary Note S2, Supplementary Figure S1). Like the Shelter L specimens, the specimens in this group that were the last to be collected (E7 and E8) were the most

divergent from the other specimens in the group. Specimen E9 was not a member of the monophyletic clade formed by specimens E1 – E8 (Figure 2D) and it differed from those specimens by 13 – 16 single nucleotide changes (for comparison, pairwise genetic distances among specimens E1 – E8 ranged from 0 to 3). E9 is also a chronologic outlier among the Shelter E specimens as E1 – E8 were collected over a 9 day period in September (September 9 – 18) while E9 was not collected until October 9.

Genetic relationships among specimens from different shelters

We next examined the relationships among viruses collected at different shelters. We first assessed the specimens collected at shelters H, L, and M in late March and April. All of these specimens were Clade 20C and Lineage B.1. Specimens L1, L2, H1, and M1 were all genetically identical and all specimens from shelters H, L, and M were closely related (pairwise genetic distances among specimens from these 3 shelters ranged from 0 – 6). However, we also noted that the diversity of Clade 20C viruses circulating in Washington during this time period was very limited and that the genotype represented by L1, L2, H1, and M1 was observed in 108 additional sequenced specimens collected in the state between March 10 and May 20.

We were also interested in the relationship among shelter viruses from late July through October. These were all Clade 20A viruses, despite the co-circulation of viruses from other clades in King County during that time. We were particularly interested in the relationship among specimens I4, C1, J1, and E9, which were all Lineage B.1.240 and were collected from August to October, 2020. Of the 276 genomes in GISAID that were collected in King County during those months, only 12 (4.3%), including the 4 genomes from this study, were Lineage B.1.240. To assess the relationship among these viruses in a different way, we calculated the genetic distance between all pairs of sequenced viruses collected in King County ≤30 days apart. We selected one virus to represent each putative introduction of virus into a shelter (that is, E1 was used as a representative for E1 - E8). We found that the pairwise distances between specimens C1, J1, and E9 were all less than half the

average distance between random viral pairs collected in King County during the same time period (Supplementary Figure S3).

Genetic relationship of shelter viruses to those collected outside the study

We examined 5943 genomic sequences for SARS-CoV-2 specimens collected outside this study in Washington from March 2020 to January 2021 for evidence of transmission of shelter viruses into the wider community. We searched for viruses in the GISAID database collected in Washington that carried the same set of alternative alleles (with or without additional alternative alleles) as at least one shelter virus (see Supplementary Note S2, Supplementary Table S6). We excluded specimens L1, L2, H1, and M1 from this analysis as the genotype shared by these specimens was commonly observed. We found a total of 21 GISAID viruses that were possible descendants of shelter viruses. The last of these was collected on November 14, 2020. Six possible descendant viruses were observed for specimen I4, the most identified for any shelter virus.

Discussion

Our analysis of genome sequences for 28 SARS-CoV-2 specimens collected in emergency shelters in King County, Washington from March to October, 2020, resulted in several notable findings. First, for all pairs of specimens collected 10 days apart or less from the same shelter (specimens L1 – L3, L4 – L7, M1 – M2, I2 – I3, and E1 – E8), we found genetic evidence suggesting that they resulted from a single viral introduction event into the shelter followed by intra-shelter transmission. We found no evidence for the co-circulation of more than one viral genotype in any shelter at any one time. Second, there were two instances (Shelters H and L) where genetically similar viruses were obtained from a shelter several weeks apart. This suggests prolonged intra-shelter transmission, prolonged shedding by shelter residents/staff, and/or multiple introductions of genetically similar viruses from an outside source. Third, the temporal and genetic relationship among sequenced specimens from Shelters H, L, and M are consistent with transmission of virus

among these shelters. Further, we noted that all shelter viruses collected after late July were from Clade 20A and that four viruses collected from August to October from four different shelters were Lineage B.1.240, despite the co-circulation of viruses from other clades and lineages in King County during this time. Finally, the number of viruses collected outside the study that were potentially descended from a shelter virus was fairly small (21 total out of 5943).

Our findings suggest that intra-shelter transmission was the primary source of the SARS-CoV-2 cases we examined, despite the potential for residents and staff of shelters to be exposed to viruses outside the shelter environment. We estimate that these cases were the result of 12 different viral introductions into shelters (one introduction into Shelters H, M, L, J, and C; two into D and E; and three into I) and that at least 57% of these cases were the result of intra-shelter transmission. The latter estimate makes the assumption that for each of the 12 viral introduction events, the index case was among the specimens we sequenced. As this is unlikely, 57% is a conservative estimate of the percentage of the 28 examined cases that were due to intra-shelter transmission. This finding supports the prioritization of control measures to limit intra-shelter spread of SARS-CoV-2 such as mask mandates, promotion/facilitation of vaccination, shelter-wide testing in response to infection in residents or staff, and isolation of infected residents.

Our results also provide insight into the relationship among the viruses from different shelters. The detection of SARS-CoV-2 infections at shelters H, L, and M in late March/early April and the public health response to these cases has previously been described [19]. As residents of these shelters shared day-use facilities, it was hypothesized prior to the availability of SARS-CoV-2 genomic data that these cases were all related via direct spread of virus among shelter residents and staff. Genetic data from some of these cases was consistent with this hypothesis. Because four of the cases from these 3 shelters were due to a genotype of virus that was extremely common in King County from March to May 2020, it remains possible that this cluster of cases was due to more than one introduction of virus from the broader community into a population consisting of shelter

residents and staff. However, the chronologic overlap in cases at these 3 shelters, the known opportunity for interaction among residents and staff of these shelters, and the genetic data taken together are very suggestive of inter-shelter viral transmission. We also found that four viruses collected in four different shelters from August to October 2020 were all members of a lineage that was uncommon in King County at that time. Given that these shelters are geographically dispersed throughout the county, these findings may suggest that the connections between residents of different shelters in the same region (e.g. social, shared use of community resources, etc) are more important conduits for viral transmission than the geographic connections between shelters and the communities where they are located.

Given that there is mounting evidence that SARS-CoV-2 can spread rapidly and extensively within shelters and even between shelters, there may be concern about shelter outbreaks fueling outbreaks in surrounding communities. To our knowledge, there is little data to inform this concern. Most studies of SARS-CoV-2 transmission within residential facilities have focused on individual outbreaks or on the determination of viral point prevalence with data collection occurring during a short time window. This prevents assessment of the genetic relationship between the viruses causing the outbreak and viruses collected later in communities that may be connected to these facilities. In contrast, our sequenced specimens were collected over 8 months and we included publicly available sequences collected over 11 months in our analyses. We observed very little evidence of transmission of SARS-CoV-2 from emergency shelter residents and staff into the broader community. Though there are a handful of viruses collected outside the study which may be descendants of shelter-related viruses, these are relatively few in number and none have led to on-going chains of transmission in Washington that have been detected as of yet by sequencing. Generating any conclusions from this observation is hampered by both the modest number of specimens we sequenced and by the fact that genomic data are available for only a small fraction of circulating viruses in Washington from March 2020 to January 2021. However, we can say that

among the specimens we analyzed we did not see evidence that SARS-CoV-2 transmission in shelters led to large number of cases in the broader community.

There were several limitations of our dataset that may affect the generalizability of our results. This study analyzed a relatively small number of SARS-CoV-2 genomes. Though study staff collected over 5,700 specimens from shelter residents and staff, the positivity rate was just 1.3% and fewer than half of the positive specimens contained enough viral RNA to generate high quality genomes. Additionally, as testing was voluntary, it is likely that there were also undetected SARS-CoV-2 cases in the study population during the study period. These limitations hindered our ability to "re-construct" complete chains of viral transmission within shelters. Next, the non-random nature of the specimens we were able to sequence (typically those with the lowest Ct values) and of the shelter residents and staff who volunteered for testing likely introduced some biases into our dataset. Particularly relevant to our conclusions, residents who left their shelters frequently during the day may have been less likely to be present when SARS-CoV-2 testing was being conducted. These residents may have a higher risk of acquiring and transmitting SARS-CoV-2 outside their shelters. Similarly, the publicly available SARS-CoV-2 sequences from King County and Washington are not an unbiased sample of all viruses circulating in these areas and instead represent a conglomerate of sample sets collected for various reasons by various organizations (Supplementary Methods). Finally, we were limited by the lack of demographic data that exist for many publicly available SARS-CoV-2 sequences. Though some specimens collected outside our study could be descendants of study specimens, we cannot be completely sure that these putative descendant viruses are not from shelter residents or staff.

The objective of this study was to better understand SARS-CoV-2 transmission within emergency shelters and between emergency shelter residents and staff and the broader community in King County, Washington. Our results can be used to inform inferences about the relationship among SARS-CoV-2 cases in similar populations where genetic data are not available and to guide

public health decision-making regarding SARS-CoV-2 infection control measures. Our findings that intra-shelter viral spread is common support aggressive responses to the detection of SARS-CoV-2 in a shelter, such as shelter-wide testing and quarantine of infected residents, to prevent viral spread among residents and staff of the affected shelter. However, more molecular epidemiology studies analyzing genetic data are needed to determine the generalizability of our findings in this population. More work will also be needed to understand how SARS-CoV-2 transmission patterns in shelters evolve over time and are affected by specific interventions aimed at reducing viral spread.

Conflicts of Interest

MB consults for Moderna, VirBio, and Merck and has received research support from Regeneron, Ridgeback, Merck, VirBio, Glaxo-Smith-Kline, and Gates Ventures outside the submitted work. MLJ has received research funding from Sanofi Pasteur. JAE serves as a consultant for Sanofi Pasteur, Teva Pharmaceuticals, AstraZeneca, and Meissa Vaccines and receives research support from Novavax, Glaxo-Smith-Kline, Merck, and Pfizer. HYC has served as a consultant with Ellume, Pfizer, the Bill and Melinda Gates Foundation, Glaxo-Smith-Kline, and Merck and has received research funding from Gates Ventures, the Bill and Melinda Gates Foundation, and Emergent Ventures and support and reagents from Ellume and Cepheid. LMS has received research support from Gates Ventures. All other authors have no conflicts of interest to declare.

Funding

This work was supported by grants from the Centers for Disease Control (75D30120C09322) and from Gates Ventures.

Meetings where the information has previously been presented

None

References

1. Taylor J, Carter RJ, Lehnertz N, et al. Serial Testing for SARS-CoV-2 and Virus Whole Genome Sequencing Inform Infection Risk at Two Skilled Nursing Facilities with COVID-19 Outbreaks - Minnesota, April-June 2020. MMWR Morb Mortal Wkly Rep. 2020; 69(37):1288–1295.

2. Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. N Engl J Med. 2020; 382(22):2081–2090.

3. McMichael TM, Currie DW, Clark S, et al. Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. N Engl J Med. 2020; 382(21):2005–2011.

4. Ladhani SN, Chow JY, Janarthanan R, et al. Increased risk of SARS-CoV-2 infection in staff working across different care homes: enhanced CoVID-19 outbreak investigations in London care Homes. J Infect. 2020; 81(4):621–624.

5. Voeten HACM, Sikkema RS, Damen M, et al. Unravelling the modes of transmission of SARS-CoV-2 during a nursing home outbreak: looking beyond the church super-spread event. Clin Infect Dis. 2020; .

6. Lucey M, Macori G, Mullane N, et al. Whole-genome sequencing to track SARS-CoV-2 transmission in nosocomial outbreaks. Clin Infect Dis. 2020; .

7. Wallace M, Marlow M, Simonson S, et al. Public Health Response to COVID-19 Cases in Correctional and Detention Facilities - Louisiana, March-April 2020. MMWR Morb Mortal Wkly Rep. 2020; 69(19):594–598.

8. Gouvea-Reis FA, Oliveira PD, Silva DCS, et al. COVID-19 Outbreak in a Large Penitentiary Complex, April-June 2020, Brazil. Emerg Infect Dis. 2021; 27(3).

9. Maxmen A. California's San Quentin prison declined free coronavirus tests and urgent advice - now it has a massive outbreak. Nature. 2020; 583(7816):339–340.

10. Gorny AW, Bagdasarian N, Koh AHK, et al. SARS-CoV-2 in migrant worker dormitories: Geospatial epidemiology supporting outbreak management. Int J Infect Dis. 2020; 103:389–394.

11. Ngiam JN, Chew N, Tham SM, et al. Demographic shift in COVID-19 patients in Singapore from an aged, at-risk population to young migrant workers with reduced risk of severe disease. Int J Infect Dis. 2020; 103:329–335.

12. Lemieux JE, Siddle KJ, Shaw BM, et al. Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. Science. 2020; :eabe3261.

13. Mosites E, Parker EM, Clarke KEN, et al. Assessment of SARS-CoV-2 Infection Prevalence in Homeless Shelters - Four U.S. Cities, March 27-April 15, 2020. MMWR Morb Mortal Wkly Rep. 2020; 69(17):521–522.

14. Perri M, Dosani N, Hwang SW. COVID-19 and people experiencing homelessness: challenges and mitigation strategies. CMAJ. 2020; 192(26):E716–E719.

15. Tan LF, Chua JW. Protecting the Homeless During the COVID-19 Pandemic. Chest. 2020; 158(4):1341–1342.

16. Karb R, Samuels E, Vanjani R, Trimbur C, Napoli A. Homeless Shelter Characteristics and Prevalence of SARS-CoV-2. West J Emerg Med. 2020; 21(5):1048–1053.

17. Baggett TP, Keyes H, Sporn N, Gaeta JM. Prevalence of SARS-CoV-2 Infection in Residents of a Large Homeless Shelter in Boston. JAMA. 2020; 323(21):2191–2192.

18. Imbert E, Kinley PM, Scarborough A, et al. Coronavirus Disease 2019 (COVID-19) Outbreak in a San Francisco Homeless Shelter. Clin Infect Dis. 2020; .

19. Tobolowsky FA, Gonzales E, Self JL, et al. COVID-19 Outbreak Among Three Affiliated Homeless Service Sites - King County, Washington, 2020. MMWR Morb Mortal Wkly Rep. 2020; 69(17):523–526.

20. Yoon JC, Montgomery MP, Buff AM, et al. COVID-19 Prevalence among People Experiencing Homelessness and Homelessness Service Staff during Early Community Transmission in Atlanta, Georgia, April-May 2020. Clin Infect Dis. 2020; .

21. Ghinai I, Davis ES, Mayer S, et al. Risk Factors for Severe Acute Respiratory Syndrome Coronavirus 2 Infection in Homeless Shelters in Chicago, Illinois-March-May, 2020. Open Forum Infect Dis. 2020; 7(11):ofaa477.

22. Rogers JH, Link AC, McCulloch D, et al. Characteristics of COVID-19 in Homeless Shelters : A Community-Based Surveillance Study. Ann Intern Med. 2021; 174(1):42–49.

23. McCulloch DJ, Kim AE, Wilcox NC, et al. Comparison of Unsupervised Home Self-collected Midnasal Swabs With Clinician-Collected Nasopharyngeal Swabs for Detection of SARS-CoV-2 Infection. JAMA Netw Open. 2020; 3(7):e2016382.

24. Srivatsan S, Han PD, Raay K van, et al. Preliminary support for a “dry swab, extraction free” protocol for SARS-CoV-2 testing via RT-qPCR [Internet]. Molecular Biology; 2020 Apr. Available from: http://biorxiv.org/lookup/doi/10.1101/2020.04.22.056283

25. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018; 34(23):4121–4123.

26. Rambaut A, Holmes EC, O’Toole Á, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol. 2020; 5(11):1403–1407.

Figure Legends

Figure 1: Plots showing number of SARS-CoV-2 full genome sequences by Pangolin lineage. Legend includes all pangolin lineages represented among shelter specimens as well as all pangolin lineages representing ≥2% of all King County or Washington specimens from March to October, 2020. Other lineages are shown collectively as "Other". A) Count of sequenced specimens from study shelter residents (N = 28) by month and lineage. B) Count of SARS-CoV-2 genomes from King County in GISAID (N = 575) by month and lineage. C) Count of SARS-CoV-2 genomes from Washington in GISAID (N = 3167) by month and lineage.

Figure 2: Phylogenetic trees showing the relationship among 28 SARS-CoV-2 specimens collected from residents and staff in 8 emergency shelters housing PEH from March to October, 2020. Black dots are used to facilitate the differentiation of similar colors representing different shelters. A) Tree showing the relationship among the genomes from 28 SARS-CoV-2 specimens collected from residents and staff. The x-axis values represent the number of single nucleotide changes among genomes. B) Tree showing the 28 shelter genomes plus a random sample of publicly available SARS-CoV-2 genomes from Washington (darker gray, N = 445) and from the United States (lighter gray, N = 757) collected during the same time period for context. The x-axis values represent date of specimen collection. C) Tree showing Clade 20C shelter genomes plus all publicly available genomes from specimens collected in King County (gray nodes) from March to June, 2020. Inset shows detail of branches containing shelter genomes. X-axis values represent number of single nucleotide changes relative to the Wuhan/Hu-1 reference genome. D) Tree showing Clade 20A shelter genomes plus all publicly available genomes from specimens collected in King County (gray nodes) from July to

October 2020. Inset shows detail of branches containing shelter genomes. X-axis values represent number of single nucleotide changes relative to the Wuhan/Hu-1 reference genome.

Specimen Name	# of Sequenced Specimens	Source of Specimen(s)	Date of Collection	Nextstrain Clade	Pangolin Lineage
I1	1	Resident	11 March	19B	A.1
I2	1	Resident	18 June	20C	B.1
I3	1	Resident	20 June	20C	B.1
I4	1	Resident	6 August	20A	B.1.240
H1	1	Resident	30 March	20C	B.1
H2	1	Resident	23 April	20C	B.1
L1 – L3	3	Residents	1 April	20C	B.1
L4 – L6	3	Residents	23 April	20C	B.1
L7	1	Staff	23 April	20C	B.1
M1 – M2	2	Staff	13 April	20C	B.1
D1	1	Resident	16 July	20B	B.1.1.291
D2	1	Resident	27 July	20A	B.1
E1 – E5	5	Residents	9 September	20A	B.1.234
E6	1	Resident	11 September	20A	B.1.234
E7	1	Resident	16 September	20A	B.1.234
E8	1	Resident	18 September	20A	B.1.234
E9	1	Resident	9 October	20A	B.1.240
C1	1	Resident	14 September	20A	B.1.240
J1	1	Resident	15 September	20A	B.1.240
Total	28				

Table 1: SARS-CoV-2-positive specimens with full genome sequences collected from emergency shelter residents and staff - King County, Washington - March to October, 2020

Figure 1

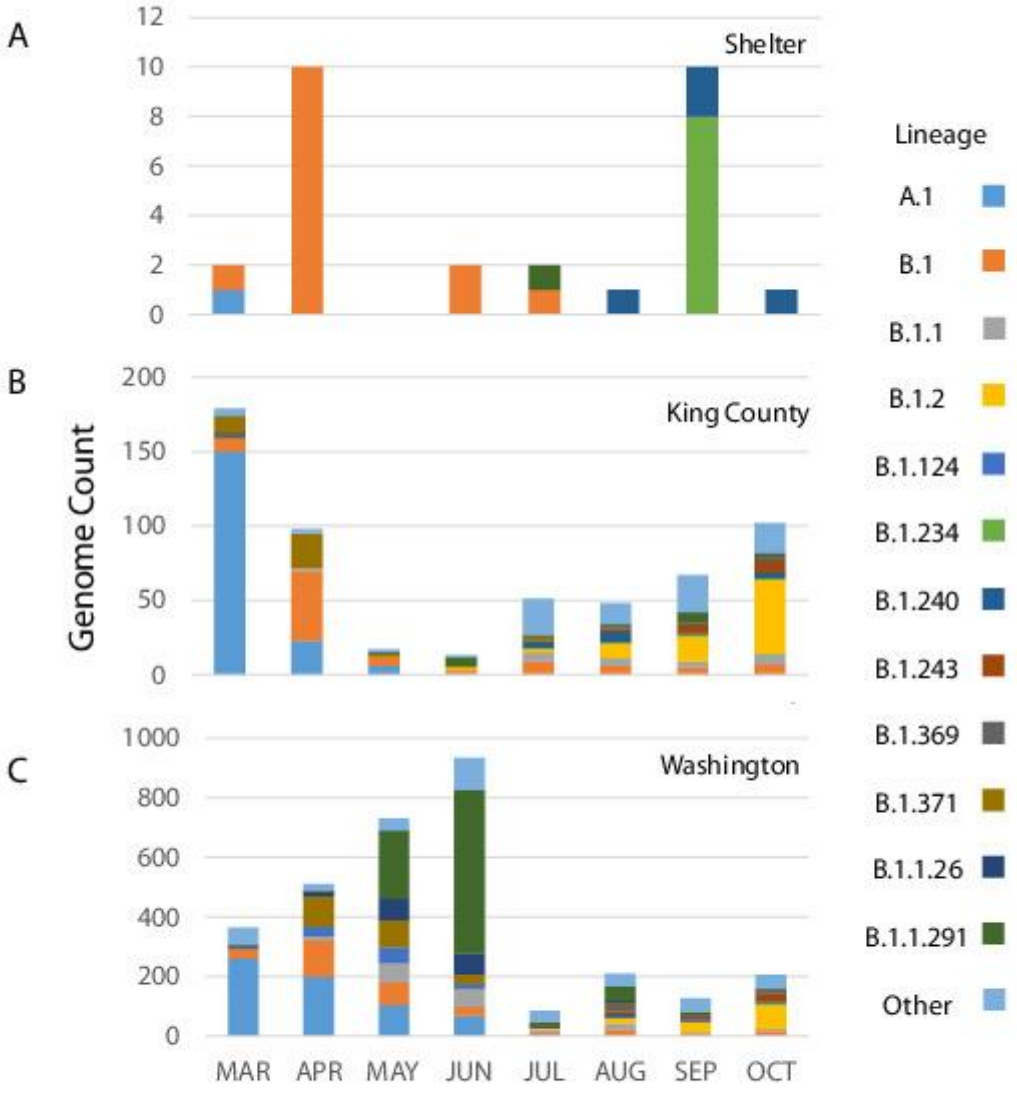
A
Shelter
12
10
8
6
4
2
0
B
King County
200
150
100
50
0
Genome Count
C
Washington
1000
800
600
400
200
0
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
Lineage
A.1
B.1
B.1.1
B.1.2
B.1.124
B.1.234
B.1.240
B.1.243
B.1.369
B.1.371
B.1.1.26
B.1.1.291
Other

Figure 2