science.sciencemag.org/cgi/content/full/science.abc0523/DC1

Supplementary Materials for

Cryptic transmission of SARS-CoV-2 in Washington state

Trevor Bedford*†, Alexander L. Greninger†, Pavitra Roychoudhury†, Lea M. Starita†, Michael Famulare†, Meei-Li Huang, Arun Nalla, Gregory Pepper, Adam Reinhardt, Hong Xie, Lasata Shrestha, Truong N. Nguyen, Amanda Adler, Elisabeth Brandstetter, Shari Cho, Danielle Giroux, Peter D. Han, Kairsten Fay, Chris D. Frazar, Misja Ilcisin, Kirsten Lacombe, Jover Lee, Anahita Kiavand, Matthew Richardson, Thomas R. Sibley, Melissa Truong, Caitlin R. Wolf, Deborah A. Nickerson, Mark J. Rieder, Janet A. Englund, The Seattle Flu Study Investigators‡, James Hadfield, Emma B. Hodcroft, John Huddleston, Louise H. Moncla, Nicola F. Müller, Richard A. Neher, Xianding Deng, Wei Gu, Scot Federman, Charles Chiu, Jeff S. Duchin, Romesh Gautom, Geoff Melly, Brian Hiatt, Philip Dykema, Scott Lindquist, Krista Queen, Ying Tao, Anna Uehara, Suxiang Tong, Duncan MacCannell, Gregory L. Armstrong, Geoffrey S. Baird, Helen Y. Chu‡, Jay Shendure‡, Keith R. Jerome‡

*Corresponding author. Email: tbedford@fredhutch.org
†These authors contributed equally to this work.
‡These authors contributed equally to this work.

Published 10 September 2020 on *Science* First Release
DOI: 10.1126/science.abc0523

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.abc0523/DC1)

MDAR Reproducibility Checklist (.pdf)
Data S1 (.pdf)

Materials and Methods

Specimen collection

Specimens analyzed in this manuscript were obtained in a collaboration between the Washington State Department of Health, the Seattle Flu Study and the University of Washington Laboratory Medicine Department of Virology (UW Virology).

Specimens were collected by the Washington State Department of Health following CDC criteria. Specimens from UW Virology were obtained as part of clinical testing for SARS-CoV-2. Nasopharyngeal/oropharyngeal swabs were received from local healthcare providers and from the Washington State Department of Health to perform qualitative detection of SARS-CoV-2 RNA by a one-step real-time RT-PCR assay. Sequencing was performed on all samples with a positive or inconclusive RT-PCR assay result. The Seattle Flu Study (SFS) was established in the 2018-2019 flu season to combine clinical and innovative community sampling methods to measure how influenza, RSV, and other respiratory pathogens enter and circulate with the Seattle metropolitan region (*27*). Samples screened for COVID-19 were collected as part of routine clinical testing, and residual samples were utilized for this study. Samples were additionally collected as part of prospective community enrollment of individuals with acute respiratory illness. Detailed protocols for sample collection are described in Chu et al. (*28*).

Diagnostics and sequencing

Extracted nucleic acids (Magna Pure, 96 Roche) were screened for SARS-CoV-2 in a multiplexed Taqman assay with primer/probe sets targeting SARS-CoV-2 Orf1B (FAM) and human RNaseP (VIC) in duplicate (Life Technologies assay ID APGZJKF and A30064) in 384 well plates. Samples that are positive for SARS-CoV-2 are retested using the CDC-designed rtRT-PCR assays acquired directly from IDT (2019 nCoV Kit lot# 0000500389, 2019-nCoV_N positive control lot# 0000500326) according to CDC instructions with the exception that a 384-well plate and ViiA7 thermocycler was used.

SARS-CoV-2 genome sequencing was conducted using a targeted enrichment approach. RNA from positive specimens was converted to cDNA using random hexamers and reverse transcriptase (Superscript IV, Thermo) and a sequencing library was constructed using the Illumina TruSeq RNA Library Prep for Enrichment kit (Illumina). The sequencing library was enriched for SARS-CoV-2 using the Twist Respiratory Virus Research Panel (Twist). Libraries

were sequenced on a MiSeq or NextSeq instruments. The resulting reads were assembled against the SARS-CoV-2 reference genome Wuhan-Hu-1/2019 (Genbank accession MN908947) using the bioinformatics pipeline https://github.com/seattleflu/assembly. Consensus sequences were deposited to Genbank and GISAID.

For UW Virology samples, sequencing was performed as described previously (*29*). Libraries were sequenced on Illumina MiSeq or NextSeq instruments using 1x185 or 1x75 runs respectively. Consensus sequences were assembled using a custom bioinformatics pipeline (https://github.com/proychou/hCoV19) adapted for SARS-CoV-2 from previous work (*30*, *31*). Briefly raw reads were trimmed to remove adapters and low quality regions using BBDuk and a k-mer based filter was used to pull out reads matching the reference sequence NC_045512. Filtered reads were *de novo* assembled using SPAdes (*32*) and contigs were ordered against the reference using BWA-MEM (*33*). Gaps were filled by remapping reads against the assembled scaffold and a consensus sequence was called from this alignment using a custom script in R/Bioconductor. Consensus sequences were annotated using Prokka (*34*) and deposited to Genbank (MT152824, MT419827-MT419859, MT598633-MT598642, MT605815-MT605817, MT627212-MT627318, MT627608-MT627638, MT632500, MT632502, MT632506, MT632518, MT632541-MT632624, MT632837, MT632951, MT632990), GISAID (accessions in Data Table S1), and NCBI SRA (Bioproject PRJNA610428).

Strains USA/WA-S31/2020 and USA/WA-S242/2020 had site 17747 masked to 'N' in our standard pipeline due to 99% and 96% respectively of reads coming from the forward strand. However, USA/WA-S31/2020 had 1483 and USA/WA-S242/2020 had 1497 reads all with 'T'. Given the importance of this site to determining whether a virus falls into the Washington State outbreak clade, we updated both strains to call 'T' at site 17747.

We examined the date of sample collection and other available metadata to remove possible duplicates from sequencing. For the present analysis, we restricted the dataset to samples collected on or before 15 March 2020. The final dataset of 455 SARS-CoV-2 genomes from Washington State represent a convenience sample from the underlying outbreak. All Washington State sequences used in the paper are available here https://github.com/blab/ncov-cryptic-transmission.

Phylogenetics

SARS-CoV-2 genomes from the global COVID-19 pandemic were downloaded from GISAID (*9*, *10*) and processed using the Nextstrain (*25*) bioinformatics pipeline *Augur* to align genomes via MAFFT v7.4 (*35*), build maximum likelihood phylogeny via IQTREE v1.6 (*36*) and reconstruct nucleotide and amino acid changes on the ML tree. Branch locations were estimated using the maximum-likelihood discrete traits model in *Augur*. The resulting tree was visualized in the Nextstrain web application *Auspice* to view resulting inferences. Workflows to reproduce phylogenetic trees shown in Figure 1 and Figure 2 are available from https://github.com/blab/ncov-cryptic-transmission.

The 455 focal genomes from Washington State were placed on a background of 494 publicly available genomes from GISAID collected between December 2019 and 15 March 2020. These 494 genomes were subsampled from an available 10,484 global genomes to attempt to equitably subsample through space and through time while enriching for genomes that are similar to Washington State focal samples. This was done by subsampling to 60 genomes per continent-level geographic region per month. Within each of these bins, genomes that are closer in genetic distance to focal genomes are preferentially selected. This approach ensures spatiotemporal diversity of background samples while preferentially including genomes that are likely to be phylogenetically informative for the focal sample. Subsampling algorithms are included in the linked workflow. We only analyzed genomes with at least 27000 bases (90% coverage) and only analyzed genomes with completely specified date information.

Additionally, 384 SARS-CoV-2 aligned genomes from the Washington State outbreak clade were analyzed in BEAST (*37*) to estimate time of common ancestor and rate of epidemic growth. This analysis used an exponential growth coalescent model in which effective population size and rate of exponential growth are estimated. We assumed a HKY85 nucleotide substitution model (*38*) with gamma distributed rate variation and a strict molecular clock with an informed lognormal prior with mean 6×10^{-4} and standard deviation of 5×10^{-5}. This prior was chosen based on the global clock analysis described below. Full analysis details, including BEAST XML, are available at https://github.com/blab/ncov-cryptic-transmission.

To establish an informed prior on the rate of SARS-CoV-2 molecular evolution, we performed an additional BEAST (*37*) analysis in which we subsampled 309 SARS-CoV-2 viruses sampled globally and collected between December 2019 and July 2020. This sampling gives enough temporal range to more accurate assess rate of evolution. We assumed a HKY85 nucleotide

substitution model (*38*) with gamma distributed rate variation and a strict molecular clock with an CTMC rate reference prior. This analysis produced a posterior estimate of the rate of SARS-CoV-2 molecular evolution of 6.0×10^{-4} subs per site per year (95% HPD 5.3 to 6.7×10^{-4} subs per site per year).

Sequential Monte Carlo

Proportion of positive specimens was estimated following a sequential Monte Carlo procedure. We recorded the vector $(n_0, \ldots, n_{75})$ = **n** of daily totals for specimens collected between 1 January and 15 March (shown in **Fig. 3A**). We also recorded the vector **k** of daily positive specimens (shown in **Fig. 3B**). We sought to estimate the vector **p** of daily proportion positive. We chose a process model of daily change to **p** as a standard diffusion process with $p_t = N(p_{t-1}, \sigma)$, where the volatility parameter $\sigma = 0.01$. We chose an observation model as a standard binomial probability $k_t = B(n_t, p_t)$. We estimated **p** using a sequential Monte Caro procedure via a bootstrap filter with 2000 particles and a reweighting (*39*, *40*)

Supplementary Text

Seattle Flu Study Investigators

Principal Investigators: Helen Y. Chu[1,7], Michael Boeckh[1,2,7], Janet A. Englund[3,7], Michael Famulare[4], Barry R. Lutz[5,7], Deborah A. Nickerson[6,7], Mark J. Rieder[7], Lea M. Starita[6,7], Matthew Thompson[9], Jay Shendure[6,7,8], and Trevor Bedford[2,6,7]. Co-Investigators:

Amanda Adler[3], Elisabeth Brandstetter[1], Shari Cho[6,7], Chris D. Frazar[6], Danielle Giroux[6], Peter D. Han[6,7], James Hadfield[1], Shichu Huang[5], Michael L. Jackson[10], Anahita Kiavand[6,7], Louise E. Kimball[2], Kirsten Lacombe[3], Jennifer Logue[1], Victoria Lyon[1], Kira L. Newman[1] , Matthew Richardson[6,7], Thomas R. Sibley[2], Monica L. Zigman Suchsland[1], Melissa Truong[6,7] and Caitlin R. Wolf[1]. Affiliations:[1] Department of Medicine, University of Washington

[2] Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center

[3] Seattle Children's Research Institute

[4] Institute for Disease Modeling

[5] Department of Bioengineering, University of Washington

[6] Department of Genome Sciences, University of Washington

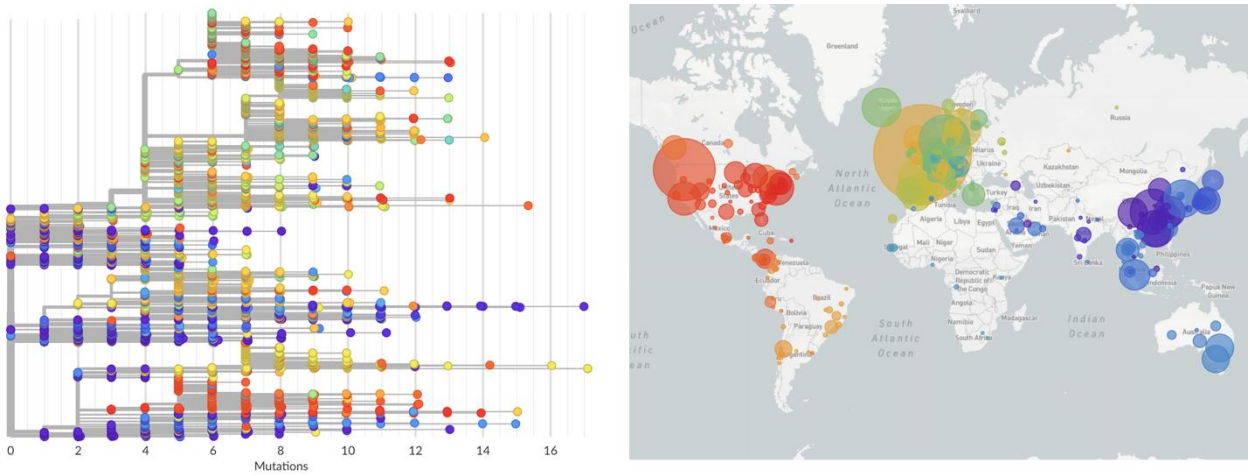
[7] Brotman Baty Institute For Precision Medicine

[8] Howard Hughes Medical Institute

[9] Department of Family Medicine, University of Washington

[10] Kaiser Permanente Washington Health Research Institute

Supplementary Figures



Supplementary Figure S1. Maximum-likelihood phylogeny of 6827 SARS-CoV-2 viruses sampled globally and collected between 24 December 2019 and 14 March 2020. Tips are colored based on location shown in the map on the right-hand side, branch lengths are proportional to number of mutations along a branch and the x-axis is labeled with the number of substitutions relative to the root of the phylogeny, here equivalent to basal Wuhan outbreak viruses. An interactive version of this figure is available at https://nextstrain.org/groups/blab/ncov/early-pandemic allowing zooming to specific clades of interest. This analysis represents a non-subsampled dataset of viruses present in GISAID as of 4 August 2020.

Data S1. GISAID EpiCoV acknowledgments table. This table lists consensus SARS-CoV-2 genomes used in this manuscript's analysis including strain name and GISAID EpiCoV accession. This includes SARS-CoV-2 genomes generated as part of the present study as well as background genomes shared by research groups from all over the world.

References and Notes

1. A. E. Gorbalenya, S. C. Baker, R. S. Baric, R. J. de Groot, C. Drosten, A. A. Gulyaeva, B. L. Haagmans, C. Lauber, A. M. Leontovich, B. W. Neuman, D. Penzar, S. Perlman, L. L. M. Poon, D. Samborskiy, I. A. Sidorov, I. Sola, J. Ziebuhr, Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group. bioRxiv 2020.02.07.937862 [Preprint]. 11 February 2020. https://doi.org/10.1101/2020.02.07.937862.

2. S. Jiang, Z. Shi, Y. Shu, J. Song, G. F. Gao, W. Tan, D. Guo, A distinct name is needed for the new coronavirus. *Lancet* **395**, 949 (2020). doi:10.1016/S0140-6736(20)30419-0 Medline

3. A. Rambaut, Phylogenetic analysis of nCoV-2019 genomes (Virological, 2020); https://virological.org/t/phylodynamic-analysis-176-genomes-6-mar-2020/356.

4. World Health Organization (WHO), "Coronavirus disease 2019 (COVID-19): Situation report – 55" (WHO, 2020); www.who.int/docs/default-source/coronaviruse/situation-reports/20200315-sitrep-55-covid-19.pdf?sfvrsn=33daa5cb_8.

5. F. Schlosser, B. F. Maier, O. Baranov, D. Brockmann, C. Jongen, A. Zachariae, A. Rose, Coronavirus COVID-19 Global Risk Assessment, Event Horizon - COVID-19 (2020); http://rocs.hu-berlin.de/corona/.

6. World Health Organization (WHO), "Coronavirus disease 2019 (COVID-19): Situation report – 38" (WHO, 2020); www.who.int/docs/default-source/coronaviruse/situation-reports/20200227-sitrep-38-covid-19.pdf?sfvrsn=2db7a09b_4.

7. K. Bray, "Coronavirus update: Addressing questions about a presumptive positive case in an adolescent" (Snohomish Health District Public Health Essentials, 2020); www.snohd.org/Blog.aspx?IID=13.

8. Washington State Department of Health, "2019 Novel Coronavirus Outbreak (COVID-19)" (2020); www.doh.wa.gov/Emergencies/Coronavirus.

9. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* **22**, 30494 (2017). doi:10.2807/1560-7917.ES.2017.22.13.30494 Medline

10. S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall* **1**, 33–46 (2017). doi:10.1002/gch2.1018 Medline

11. K. G. Andersen, A. Rambaut, W. I. Lipkin, E. C. Holmes, R. F. Garry, The proximal origin of SARS-CoV-2. *Nat. Med.* **26**, 450–452 (2020). doi:10.1038/s41591-020-0820-9 Medline

12. N. Imai, I. Dorigatti, A. Cori, C. Donnelly, S. Riley, N. M. Ferguson, "Report 2: Estimating the potential total number of novel Coronavirus cases in Wuhan City, China" (Imperial College London, 2020); http://hdl.handle.net/10044/1/77150.

13. A. Rambaut, E. C. Holmes, Á. O'Toole, V. Hill, J. T. McCrone, C. Ruis, L. du Plessis, O. G. Pybus, A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic

epidemiology. *Nat. Microbiol.* 10.1038/s41564-020-0770-5 (2020). doi:10.1038/s41564-020-0770-5 Medline

14. E. M. Volz, K. Koelle, T. Bedford, Viral phylodynamics. *PLOS Comput. Biol.* **9**, e1002947 (2013). doi:10.1371/journal.pcbi.1002947 Medline

15. M. L. Holshue, C. DeBolt, S. Lindquist, K. H. Lofy, J. Wiesman, H. Bruce, C. Spitters, K. Ericson, S. Wilkerson, A. Tural, G. Diaz, A. Cohn, L. Fox, A. Patel, S. I. Gerber, L. Kim, S. Tong, X. Lu, S. Lindstrom, M. A. Pallansch, W. C. Weldon, H. M. Biggs, T. M. Uyeki, S. K. Pillai, Washington State 2019-nCoV Case Investigation Team, First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med.* **382**, 929–936 (2020). doi:10.1056/NEJMoa2001191 Medline

16. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020). doi:10.1056/NEJMoa2001316 Medline

17. H. Nishiura, N. M. Linton, A. R. Akhmetzhanov, Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **93**, 284–286 (2020). doi:10.1016/j.ijid.2020.02.060 Medline

18. A. C. Miller, I. Singh, E. Koehler, P. M. Polgreen, A Smartphone-Driven Thermometer Application for Real-time Population- and Individual-Level Influenza Surveillance. *Clin. Infect. Dis.* **67**, 388–397 (2018). doi:10.1093/cid/ciy073 Medline

19. U.S. Centers for Disease Control and Prevention (CDC), "Update and Interim Guidance on Outbreak of 2019 Novel Coronavirus (2019-nCoV)" (CDC Health Alert Network, 2020); https://emergency.cdc.gov/han/han00427.asp.

20. U.S. Centers for Disease Control and Prevention (CDC), "Criteria to Guide Evaluation of Persons Under Investigation (PUI) for 2019-nCoV" (CDC, 2020); https://web.archive.org/web/20200222215422/https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-criteria.html.

21. A. S. Gonzalez-Reiche, M. M. Hernandez, M. Sullivan, B. Ciferri, H. Alshammary, A. Obla, S. Fabre, G. Kleiner, J. Polanco, Z. Khan, B. Alburquerque, A. van de Guchte, J. Dutta, N. Francoeur, B. S. Melo, I. Oussenko, G. Deikus, J. Soto, S. H. Sridhar, Y.-C. Wang, K. Twyman, A. Kasarskis, D. R. Altman, M. Smith, R. Sebra, J. Aberg, F. Krammer, A. Garcia-Sarstre, M. Luksza, G. Patel, A. Paniz-Mondolfi, M. Gitman, E. M. Sordillo, V. Simon, H. van Bakel, Introductions and early spread of SARS-CoV-2 in the New York City area. medRxiv 2020.04.08.20056929 [Preprint]. 16 April 2020. https://doi.org/10.1101/2020.04.08.20056929.

22. J. R. Fauver, M. E. Petrone, E. B. Hodcroft, K. Shioda, H. Y. Ehrlich, A. G. Watts, C. B. F. Vogels, A. F. Brito, T. Alpert, A. Muyombwe, J. Razeq, R. Downing, N. R. Cheemarla, A. L. Wyllie, C. C. Kalinich, I. Ott, J. Quick, N. J. Loman, K. M. Neugebauer, A. L. Greninger, K. R. Jerome, P. Roychoundhury, H. Xie, L. Shrestha, M.-L. Huang, V. E. Pitzer, A. Iwasaki, S. B. Omer, K. Khan, I. Bogoch, R. A. Martinello, E. F. Foxman, M.-

L. Landry, R. A. Neher, A. I. Ko, N. D. Grubaugh, Coast-to-coast spread of SARS-CoV-2 in the United States revealed by genomic epidemiology. medRxiv 2020.03.25.20043828 [Preprint]. 26 March 2020. https://doi.org/10.1101/2020.03.25.20043828.

23. W.-J. Guan, Z.-Y. Ni, Y. Hu, W.-H. Liang, C.-Q. Ou, J.-X. He, L. Liu, H. Shan, C.-L. Lei, D. S. C. Hui, B. Du, L.-J. Li, G. Zeng, K.-Y. Yuen, R.-C. Chen, C.-L. Tang, T. Wang, P.-Y. Chen, J. Xiang, S.-Y. Li, J.-L. Wang, Z.-J. Liang, Y.-X. Peng, L. Wei, Y. Liu, Y.-H. Hu, P. Peng, J.-M. Wang, J.-Y. Liu, Z. Chen, G. Li, Z.-J. Zheng, S.-Q. Qiu, J. Luo, C.-J. Ye, S.-Y. Zhu, N.-S. Zhong, China Medical Treatment Expert Group for Covid-19, Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020). doi:10.1056/NEJMoa2002032 Medline

24. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, J. Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020). doi:10.1126/science.abb3221 Medline

25. J. Hadfield, C. Megill, S. M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, R. A. Neher, Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018). doi:10.1093/bioinformatics/bty407 Medline

26. T. Bedford, blab/ncov-cryptic-transmission: Release 2020-05-31, version 2020-05-31, Zenodo (2020); http://doi.org/10.5281/zenodo.3871089.

27. H. Y. Chu, M. Boeckh, J. A. Englund, M. Famulare, B. R. Lutz, D. A. Nickerson, M. Rieder, L. Starita, M. Thompson, J. Shendure, T. Bedford, A. Adler, E. Brandstetter, J. Bosua, C. D. Frazar, P. D. Han, R. Gulati, J. Hadfield, S. C. Huang, M. L. Jackson, A. Kiavand, L. E. Kimball, K. Lacombe, J. Logue, V. Lyon, T. R. Sibley, M. L. Zigman Suchsland, C. R. Wolf, LB21. The Seattle Flu Study: A Community-Based Study of Influenza. *Open Forum Infect. Dis.* **6**, S1002 (2019). doi:10.1093/ofid/ofz415.2504

28. H. Y. Chu, M. Boeckh, J. A. Englund, M. Famulare, B. R. Lutz, D. A. Nickerson, M. J. Rieder, L. M. Starita, A. Adler, E. Brandstetter, C. D. Frazar, P. D. Han, R. K. Gularti, J. Hadfield, M. L. Jackson, A. Kiavand, L. E. Kimball, K. Lacombe, J. Logue, V. Lyon, K. L. Newman, T. R. Sibley, M. L. Zigman Suschsland, C. Wolf, J. Shendure, T. Bedford, The Seattle Flu Study: a multi-arm community-based prospective study protocol for assessing influenza prevalence, transmission, and genomic epidemiology. medRxiv 2020.03.02.20029595 [Preprint]. 6 March 2020. https://doi.org/10.1101/2020.03.02.20029595.

29. A. L. Greninger, D. M. Zerr, X. Qin, A. L. Adler, R. Sampoleo, J. M. Kuypers, J. A. Englund, K. R. Jerome, Rapid Metagenomic Next-Generation Sequencing during an Investigation of Hospital-Acquired Human Parainfluenza Virus 3 Infections. *J. Clin. Microbiol.* **55**, 177–182 (2016). doi:10.1128/JCM.01881-16 Medline

30. A. L. Greninger, G. M. Knudsen, P. Roychoudhury, D. J. Hanson, R. H. Sedlak, H. Xie, J. Guan, T. Nguyen, V. Peddu, M. Boeckh, M.-L. Huang, L. Cook, D. P. Depledge, D. M. Zerr, D. M. Koelle, S. Gantt, T. Yoshikawa, M. Caserta, J. A. Hill, K. R. Jerome, Comparative genomic, transcriptomic, and proteomic reannotation of human herpesvirus 6. *BMC Genomics* **19**, 204 (2018). doi:10.1186/s12864-018-4604-2 Medline

31. A. L. Greninger, P. Roychoudhury, H. Xie, A. Casto, A. Cent, G. Pepper, D. M. Koelle, M.-L. Huang, A. Wald, C. Johnston, K. R. Jerome, Ultrasensitive Capture of Human Herpes Simplex Virus Genomes Directly from Clinical Samples Reveals Extraordinarily Limited Evolution in Cell Culture. *MSphere* **3**, e00283-18 (2018). doi:10.1128/mSphereDirect.00283-18 Medline

32. A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, P. A. Pevzner, SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012). doi:10.1089/cmb.2012.0021 Medline

33. M. Vasimuddin, S. Misra, H. Li, S. Aluru, "Efficient architecture-aware acceleration of BWA-MEM for multicore systems," in *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* (IEEE, 2019), pp. 314–324.

34. T. Seemann, Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014). doi:10.1093/bioinformatics/btu153 Medline

35. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). doi:10.1093/molbev/mst010 Medline

36. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015). doi:10.1093/molbev/msu300 Medline

37. M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, A. Rambaut, Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018). doi:10.1093/ve/vey016 Medline

38. M. Hasegawa, H. Kishino, T. Yano, Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985). doi:10.1007/BF02101694 Medline

39. A. Doucet, N. de Freitas, N. Gordon, in *Sequential Monte Carlo Methods in Practice. Statistics for Engineering and Information Science*, A. Doucet, N. de Freitas, N. Gordon, Eds. (Springer, 2001), pp. 3–14.

40. N. J. Gordon, D. J. Salmond, A. F. M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F, Radar Signal Process.* **140**, 107–113 (1993). doi:10.1049/ip-f-2.1993.0015