



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Letter to the editor

Identification of multiple large deletions in ORF7a resulting in in-frame gene fusions in clinical SARS-CoV-2 isolates



Peculiar among human RNA viruses, coronaviruses have large genomes containing accessory genes that are not required for replication [1]. Here, we report two large deletions in ORF7a, both of which produce new open reading frames (ORFs) through the fusion of the N-terminus of ORF7a and a downstream ORF.

We recovered genomes from the SARS-CoV-2 isolates from two separate patients as part of an ongoing University of Washington IRB-approved genomic surveillance project [2,3]. Sequencing libraries were generated from double-stranded cDNA [4] using Swift Biosciences' SARS-CoV-2 Normalase Amplicon Sequencing kit or the Nextera DNA Flex Pre-Enrichment kit (Illumina) followed by enrichment using a SARS-CoV-2 xGen enrichment panel (NC_045512; IDT). Sequence reads were trimmed using Trimmomatic v0.38 [5], aligned to the SARS-CoV-2 reference genome (NC_045512.2) using BMAP (<https://sourceforge.net/projects/bbmap/>), trimmed of synthetic PCR primers using Primerclip (<https://github.com/swiftbiosciences/primerclip>) if appropriate, and visualized in Geneious v11.1.4 [6]. Sequencing reads and consensus genomes are available under NCBI BioProject PRJNA610428.

In WA-UW-4570, we identified a 392-nucleotide deletion originating at nt 27,494 in ORF7a. This deletion results in the loss of the entirety of ORF7b and creates a new ORF through the fusion of the N-terminus of ORF7a with ORF8 (Fig. 1a). Notably, this is the largest deletion recorded to date in SARS-CoV-2. In WA-UW-5812, we identified a 227 nt deletion also resulting in a new ORF through the fusion of the N-terminus of ORF7a with ORF7b (Fig. 1b). Both deletions were confirmed by RT-PCR (Fig. 1c and d) and Sanger sequencing. In contrast to a previously sequenced SARS-CoV-2 isolate with a deletion in ORF7a [7], our isolates retained the presumptive signal peptide and part of the β sheet of the Ig superfamily fold of ORF7a (Fig. 1e and f) and lost the protein's transmembrane and cytoplasmic domains (PDB 6W37) [8].

We next assessed the genetic relatedness of the two strains. WA-UW-4570 and WA-UW-5812 differed by 12 single nucleotide variants and a 3-bp deletion in the spike protein of WA-UW-4570 and belonged to the SARS-CoV-2 phylogenetic lineages A.1 and B.1 respectively (<https://github.com/hCoV-2019/pangolin>) [9]. This suggests that the ORF7a deletions arose independently in the two strains. Both samples were recovered from the only SARS-CoV-2-positive nasopharyngeal swab available from each patient. The ORF1ab qRT-PCR cycle thresholds on the Hologic Panther Fusion for these isolates were 27.6 (WA-UW-4570) and 28.5 (WA-UW-5812). WA-UW-4570 was collected from an individual prophylactically taking hydroxychloroquine and mycophenolate mofetil while the WA-UW-5812 individual was not taking any medications for SARS-CoV-2 infection.

ORF7a of SARS-CoV-2 interacts with the ribosomal transport proteins HEATDR3 and MDN1 [10], and has been demonstrated to inhibit cellular translation in SARS-CoV [11]. Based on the size of the deletions recovered and structure of ORF7a, we hypothesize that most biochemical functions of ORF7a would be inactivated by these deletions. Interestingly, ORF6 of SARS-CoV-2 interacts with the mRNA export proteins NUP98 and RAE1, and may inhibit cellular translation [10]. This redundancy may partially explain the ORF7a deletions observed in our isolates. We predict global sequencing projects may yield additional clinical SARS-CoV-2 isolates with deletions in ORF6 or ORF7a, but not both.

Declaration of Competing Interest

None

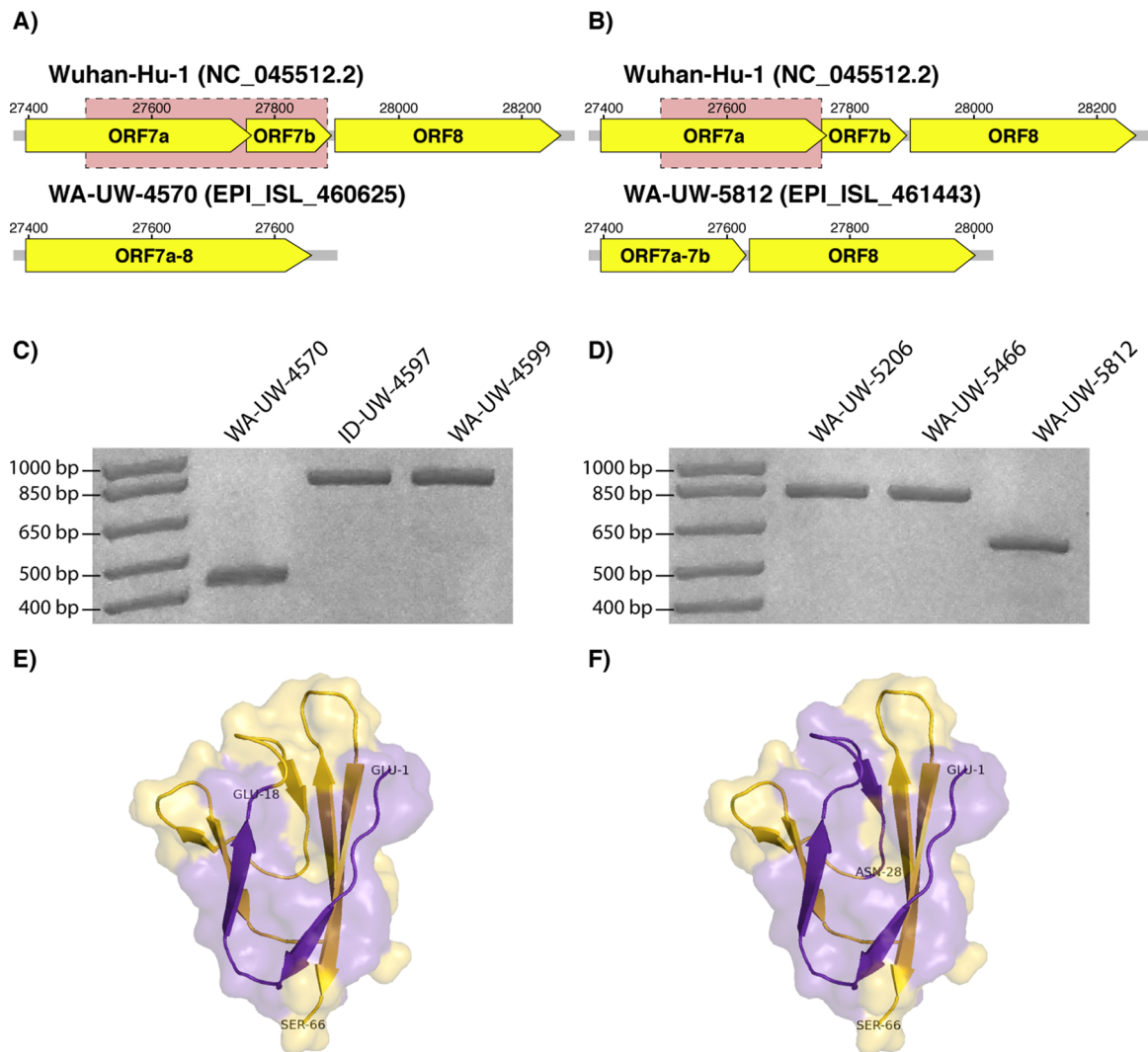


Fig. 1. A 392-nucleotide deletion starting at nt 29,424 was identified in **a)** WA-UW-4570 and resulted in the fusion of ORF7a and ORF8. A 227-nucleotide deletion beginning at nt 27,524 was identified in **b)** WA-UW-5812 and resulted in the fusion of ORF7a and ORF7b. The deletions in **c)** WA-UW-4570 and **d)** WA-UW-5812 were confirmed by RT-PCR. An 826-bp product is expected for strains with an intact ORF7a. 434-bp and 599-bp amplicons were obtained for WA-UW-4570 and WA-UW-5812, respectively. The deletions in **e)** WA-UW-4570 and **f)** WA-UW-5812 severely truncate the β sheet and result in the loss of the transmembrane and cytoplasmic domains of ORF7a (PDB 6W37) [8]. Residues retained are highlighted in purple, while those lost are highlighted in gold (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

References

- [1] K. Narayanan, C. Huang, S. Makino, SARS coronavirus accessory proteins, *Virus Res.* 133 (2008) 113–121.
- [2] 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.M. Ott, J. Quick, N.J. Loman, K.M. Neugebauer, A.L. Greninger, K.R. Jerome, P. Roychoudhury, H. Xie, L. Shrestha, M.-L. Huang, V.E. Pitzer, A. Iwasaki, S.B. Omer, K. Khan, I.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 during the early epidemic in the United States, *Cell* 181 (2020) 990–996 e5.
- [3] V. Peddu, R.C. Shean, H. Xie, L. Shrestha, G.A. Perchetti, S.S. Minot, P. Roychoudhury, M.-L. Huang, A. Nalla, S.B. Reddy, Q. Phung, A. Reinhardt, K.R. Jerome, A.L. Greninger, Metagenomic analysis reveals clinical SARS-CoV-2 infection and bacterial or viral superinfection and colonization, *Clin. Chem.* (2020).
- [4] 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 (2017) 177–182.
- [5] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* 30 (2014) 2114–2120.
- [6] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A. Drummond, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data, *Bioinformatics* 28 (2012) 1647–1649.
- [7] L.A. Holland, E.A. Kaelin, R. Maqsood, B. Estifanos, L.I. Wu, A. Varsani, R.U. Halden, B.G. Hogue, M. Scotch, E.S. Lim, An 81 nucleotide deletion in SARS-CoV-2 ORF7a identified from sentinel surveillance in Arizona (Jan-Mar 2020), *J. Virol.* (2020) JVI.00711-20, jvi;JVI.00711-20v1.
- [8] C.A. Nelson, A. Pekosz, C.A. Lee, M.S. Diamond, D.H. Fremont, Structure and intracellular targeting of the SARS-coronavirus ORF7a accessory protein, *Structure* 13 (2005) 75–85.
- [9] A. Rambaut, E.C. Holmes, V. Hill, Á O'Toole, J. McCrone, C. Ruis, L. du Plessis, O.G. Pybus, A dynamic nomenclature proposal for SARS-CoV-2 to assist genomic epidemiology, Preprint, *Microbiology* (2020).
- [10] D.E. Gordon, G.M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K.M. White, M.J. O'Meara, V.V. Rezelj, J.Z. Guo, D.L. Swaney, T.A. Tummino, R. Huettnerlein, R.M. Kaake, A.L. Richards, B. Tutuncuoglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B.J. Polacco, H. Braberg, J.M. Fabius, M. Eckhardt, M. Soucheray, M.J. Bennett, M. Kadir, M.J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z.Z.C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I.T. Kirby, J.E. Melnyk, J.S. Chiorba, K. Lou, S.A. Dai, I. Barrio-Hernandez, D. Memon, C. Hernandez-Armenta, J. Lyu, C.J.P. Mathy, T. Perica, K.B. Pilla, S.J. Ganesan, D.J. Saltzberg, R. Rakesh, X. Liu, S.B. Rosenthal, L. Calviello, S. Venkataramanan, J. Liboy-Lugo, Y. Lin, X.-P. Huang, Y. Liu, S.A. Wankowicz, M. Bohn, M. Safari, F.S. Ugur, C. Koh, N.S. Savar, Q.D. Tran, D. Shengjuler, S.J. Fletcher, M.C. O'Neal, Y. Cai, J.C.J. Chang, D.J. Broadhurst, S. Klippenstein, P.P. Sharp, N.A. Wenzell, D. Kuzuoglu, H.-Y. Wang, R. Trenker, J.M. Young, D.A. Cavero, J. Hiatt, T.L. Roth, U. Rathore, A. Subramanian, J. Noack, M. Hubert, R.M. Stroud, A.D. Frankel, O.S. Rosenberg, K.A. Verba, D.A. Agard,

M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdin, A. Ashworth, O. Schwartz, C. d'Enfert, S. Mukherjee, M. Jacobson, H.S. Malik, D.G. Fujimori, T. Ideker, C.S. Craik, S.N. Floor, J.S. Fraser, J.D. Gross, A. Sali, B.L. Roth, D. Ruggero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. García-Sastre, K.M. Shokat, B.K. Shoichet, N.J. Krogan, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing, *Nature* (2020).

- [11] S.A. Kopecky-Bromberg, L. Martinez-Sobrido, P. Palese, 7a protein of severe acute respiratory syndrome coronavirus inhibits cellular protein synthesis and activates p38 mitogen-activated protein kinase, *JVI* 80 (2006) 785–793.

Amin Addetia, Hong Xie, Pavitra Roychoudhury, Lasata Shrestha,
Michelle Loprieno, Meei-Li Huang, Keith R. Jerome,
Alexander L. Greninger*
*Department of Laboratory Medicine, University of Washington, Seattle, WA,
USA*
E-mail address: agrening@uw.edu (A.L. Greninger).

* Corresponding author.