medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

SARS-CoV-2 genomic surveillance identifies naturally occurring truncations of ORF7a that limit immune suppression.

Artem Nemudryi[1,3,5,*], Anna Nemudraia[1,*], Tanner Wiegand[1,†], Joseph Nichols[1,†], Deann T Snyder[1], Jodi F Hedges[1], Calvin Cicha[1], Helen Lee[1], Karl K Vanderwood[2], Diane Bimczok[1], Mark Jutila[1] and Blake Wiedenheft[1,4,‡]

Abstract

Over 200,000 whole genome sequences of SARS-CoV-2 have been determined for viruses isolated from around the world. These sequences have been critical for understanding the spread and evolution of SARS-CoV-2. Using global phylogenomics, we show that mutations frequently occur in the C-terminal end of ORF7a. We have isolated one of these mutant viruses from a patient sample and used viral challenge experiments to demonstrate that Δ115 mutation results in a growth defect. ORF7a has been implicated in immune modulation, and we show that the C-terminal truncation results in distinct changes in interferon stimulated gene expression. Collectively, this work indicates that ORF7a mutations occur frequently and that these changes affect viral mechanisms responsible for suppressing the immune response.

Keywords

SARS-CoV-2, ORF7a, IFN response

[1] *Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717, USA*
[2] *Gallatin City-County Health Department, Bozeman, MT, 59715, USA*
[3] *Twitter: @artemnemudryi*
[4] *Twitter: @WiedenheftLab*
[5] *Lead contact*
[*] *These authors contributed equally.*
[†] *These authors contributed equally.*
[‡] **Correspondence: bwiedenheft@gmail.com (B.W.)**

Highlights

- **ORF7a mutations are found in SARS-CoV-2 genomes isolated from around the globe.**
- **ORF7a mutation results in a replication defect.**
- **An ORF7a mutation limits viral suppression of the interferon response.**

Introduction

The spillover of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) into the human population has resulted in a global pandemic of Coronavirus Disease 2019 (COVID-19) with more than 2.1 million deaths worldwide (https://coronavirus.jhu.edu/) (Andersen et al., 2020). While much of the research on this virus is focused on the Spike protein, recent reports demonstrate that accessory proteins of SARS-CoV-2 might be involved in COVID-19 pathogenesis by modulating antiviral host responses (Young et al., 2020; Zhang et al., 2020). SARS-CoV-2 uses multiple strategies to evade host immunity. Accessory proteins ORF3b, ORF6 and ORF7a antagonize various steps of type I interferon (IFN-I) production and signaling, while ORF8 disrupts antigen presentation by downregulating MHC-I (Konno et al., 2020; Lei et al., 2020; Miorin et al., 2020; Xia et al., 2020; Zhang et al., 2020). To occlude signal transmission from IFN receptors, ORF7a subverts phosphorylation of STAT2, suppressing transcriptional activation of antiviral IFN-stimulated genes (ISGs) that can otherwise restrict viral replication (Martin-Sancho et al., 2020; Xia et al., 2020). While the proposed function for ORF7a is intracellular, antibodies against ORF7a are elevated in the serum of COVID-19 patients (Hachim et al., 2020). Work that is currently under review indicates that recombinant ORF7a protein interacts with $CD14^+$ monocytes and triggers expression of pro-inflammatory cytokines, including interleukin-6 (IL-6) and tumor necrosis factor alpha (TNF-α) (Zhou, Z. et al., 2020). However, it is unclear if these interactions happen in COVID-19 patients. Together, these data suggest that ORF7a plays a dual role in SARS-CoV-2 infection by modulating both the IFN and inflammatory responses.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Here, we show that C-terminal mutations in ORF7a occur frequently in samples isolated from patients around the globe. These mutations are not derived from a single lineage and they do not persist over time. Using samples collected from infected patients, we have isolated a virus containing a deletion mutation in ORF7a that truncates the C-terminal half of the protein. In vitro viral challenge experiments demonstrate that this mutation results in a replication defect and obviates viral suppression of the immune response. Collectively, these data suggests that ORF7a truncations are defective in suppressing the host immune response, which may explain why these mutations quickly disappear in the immunocompetent population.

Results

SARS-CoV-2 genomic surveillance identifies truncation of ORF7a accessory protein

Genome sequencing has been used to track the rise and spread of new SARS-CoV-2 lineages over the course of the pandemic (https://www.gisaid.org/) (Bedford et al., 2020; Korber et al., 2020). As part of this effort, we sequenced SARS-CoV-2 genomes isolated from patients in Bozeman, Montana (Figure 1A, Table S1). In total, we determined 56 whole genome sequences of SARS-CoV-2 viruses isolated from patients between April and July of 2020 using an amplicon-based Nanopore protocol from the ARTIC network (https://artic.network/) (Quick et al., 2017; Tyson et al., 2020). To place the local outbreak in the context of the global SARS-CoV-2 evolutionary trends, we aligned 56 whole genome sequences isolated from patients in Bozeman to 4,235 genomes subsampled (10 genomes per month per country) from 181,003 SARS-CoV-2 genomes downloaded from the GISAID. The subsampling protocol removes redundancy and bias introduced by uneven distribution of the global SARS-CoV-2 sequencing effort. The resulting alignment was used to build a phylogenetic tree (Figure 1B). Of the 56 SARS-CoV-2 genomes determined from patients in Bozeman, only one was derived from the WA1 lineage (lineage A in Figure 1B) that was introduced to Washington state from Wuhan, China in January 2020 (Bedford et al., 2020; Holshue et al., 2020). The remaining 55 genomes associate with clade B.1 and its offshoots, which are characterized by the D614G mutation in the spike and have prevailed globally since the spring of 2020 (Rambaut, Andrew et al., 2020)(Figure 1B). The genetic variability in SARS-CoV-2 circulating in Bozeman (April to July 2020) represents fourteen independent viral lineages, reflecting multiple introductions of the virus to the community from many different sources (Figure S1A).

During the annotation of these genomes, we noticed a reoccurring (7 out of 56 genomes) 115 nt deletion in the gene encoding accessory protein ORF7a (27,549 - 27,644 nt). The ORF7a$^{\Delta 115}$ mutation was found in patient swabs collected over a period of 1.5 months (Table S1). RT-PCR and Sanger sequencing verified that all these seven ORF7a$^{\Delta 115}$ variants are bona fide mutations and not a sequencing artifact (Figure S1B, C). In addition to the ORF7a$^{\Delta 115}$ deletion, these genomes also share 10 single nucleotide variants (SNVs) compared to the Wuhan-Hu-1 reference genome. Seven of the 10 SNVs are found frequently worldwide and are a signature of the B1.1 lineage of SARS-CoV-2 (Rambaut, Andrew et al., 2020). The remaining 3 mutations are rare, do not co-occur in any other genomes on GISAID, and lead to amino acid changes in ORF3a (Q38P, L95F) and N (R195I) proteins. Interestingly, one of the genomes in the ORF7a$^{\Delta 115}$ cluster has two additional SNVs that likely were acquired during circulation in the community. This virus was sampled from a patient 41 days after sampling the first ORF7a$^{\Delta 115}$ variant, which agrees well with estimated SARS-CoV-2 mutation rates (2 nucleotides / month) (van Dorp et al., 2020).

To determine if the ORF7a$^{\Delta 115}$ genotype is unique to Bozeman we downloaded an alignment of 180,971 SARS-CoV-2 genomes from GISAID, extracted ORF7a sequences and determined their mutational profiles. In total, we identified 845 unique ORF7a gene variants that are different from the Wuhan-Hu-1 reference sequence. Next, we looked for mutations with major effect on the ORF7a amino acid sequence. We identified 189 unique ORF7a variants (825 total sequences) with frameshifts, deletions or missense mutations causing premature stop codons (Figure 1C). To understand the evolutionary history of these mutations, we integrated these genome sequence into our phylogenetic analysis. Detected ORF7a variants appear to have emerged independently on every continent and were not confined to any single lineage (Figure 1B; Table S1). Next, we aligned translated ORF7a sequences to look for patterns. Most of the identified ORF7a mutations (126 out of 189) truncate the C terminus but preserve the N-terminal half of the protein (Figure 1C and D). A portion (36.5%) of these truncated ORF7a variants appeared in two or more patient samples (116 max). While some of these ORF7a mutants appear to have arisen on multiple independent occasions, others come from genomes that form monophyletic clades, suggesting that viruses with ORF7a truncations are capable of transmission within a host population (Figure 1C).

ORF7a truncation changes protein localization

ORF7a is a type I transmembrane protein with an N-terminal immunoglobulin-like (Ig-like) ectodomain, stalk, transmembrane (TM) domain and a short cytosolic tail (Figure 2A). The two β sheets of the ectodomain are held together by two disulfide bonds (i.e., Cys23-Cys58 and Cys35-Cys67) (Figure 2B). The cytosolic tail of ORF7a contains a dilysine (KRKTE) ER retrieval signal (ERRS) that mediates protein trafficking to the ER-Golgi intermediate compartment (ERGIC) (Nelson et al., 2005).

The Δ115 nt mutation in ORF7a introduces a premature stop codon that eliminates β5, β6 and β7, two of the cystines that form disulfides (Cys58 and Cys67), the TM and the cytosolic tail (Figures 2A and B). This truncation is expected to destabilize the protein structure and significantly impact

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

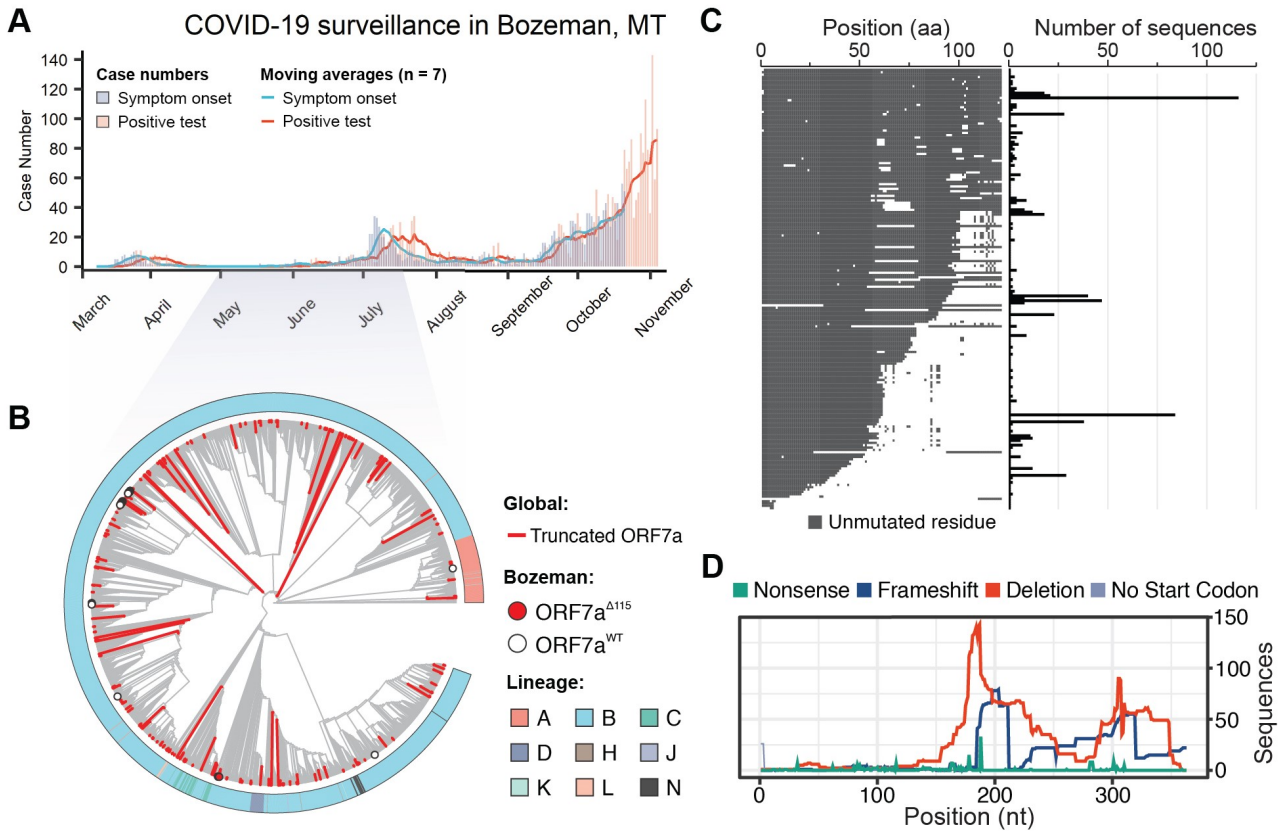

Figure 1. SARS-CoV-2 genomic surveillance identifies global reoccurrence of ORF7a truncations. A) Symptom onset (purple) and PCR-based SARS-CoV-2 test results (coral) for patients in Bozeman, Montana are shown with vertical bars. Seven day moving averages, shown with lines, were used to indicate epidemiological trends. **B)** Phylogenetic analysis of SARS-CoV-2 genomes sampled in Bozeman and globally. The tree was constructed from an alignment of 56 Bozeman samples and 4,871 genomes subsampled from GISAID. Subsampling was performed using Augur utility (Nextstrain) by selecting 10 genomes per country per month since the start of the pandemic. Red branches identify truncated ORF7a variants (n= 205) detected in the global data and merged into the alignment. The red dot highlights seven of the 56 ORF7a variants that were isolated in Bozeman between April and July (2020). White dots highlight 49 viral genomes isolated in Bozeman that have wild-type ORF7a sequences. **C)** Alignment of unique ORF7a protein sequences with five or more mutations as compared to the SARS-CoV-2 reference strain (Wuhan-Hu-1). Conserved residues are depicted in gray, while mutations and deletions are shown in white. The frequency of each genotype is shown on the right. **D)** Distribution of different mutations that occur along the ORF7a coding sequence.

protein function and trafficking in the host cell. To determine how loss of TM and sorting signal affects protein localization, we cloned and expressed Flag-tagged wildtype and truncated ORF7a proteins in HEK 293T-hACE2 cells. The wildtype ORF7a accumulated in the perinuclear region of the cell, which is consistent with the previously reported ERGIC localization (Figure 2D) (Martin-Sancho et al., 2020; Nelson et al., 2005). In contrast, the truncated ORF7a is distributed throughout the cytoplasm and does not associate with specific subcellular compartments, which is consistent with the loss of the TM domain and the ERRS signals required for protein targeting (Figure 2E). Overall, we predict that the extent of the truncation and the loss of intracellular targeting is likely to affect ORF7a function (Gordon et al., 2020; Taylor et al., 2015).

ORF7a mutation has no collateral effect on ORF7b

The ORF7a gene overlaps with the downstream ORF7b gene (Figure S2). To determine if the Δ115 mutation in ORF7a impacts ORF7b, we first examined whether it eliminates a transcription-regulatory sequence (TRS) that is required for subgenomic RNA (sgRNA) synthesis. None of the TRSs identified via direct RNA sequencing of SARS-CoV-2 overlap with the Δ115 mutation site (Kim et al., 2020), suggesting that ORF7b transcription is not affected (Figure S2). To confirm this prediction, we used RT-PCR with primers that detect ORF7a and ORF7b sgRNAs (Figure 2F). In this assay, we infected 293T-hACE2 cells (MOI = 0.05) with ORF7a^{WT} or ORF7a$^{\Delta 115}$ viral strains and extracted total RNA from cells at 24 hours post infection (hpi). Both viruses produced specific RT-PCR products corresponding to ORF7a and ORF7b sgR-

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

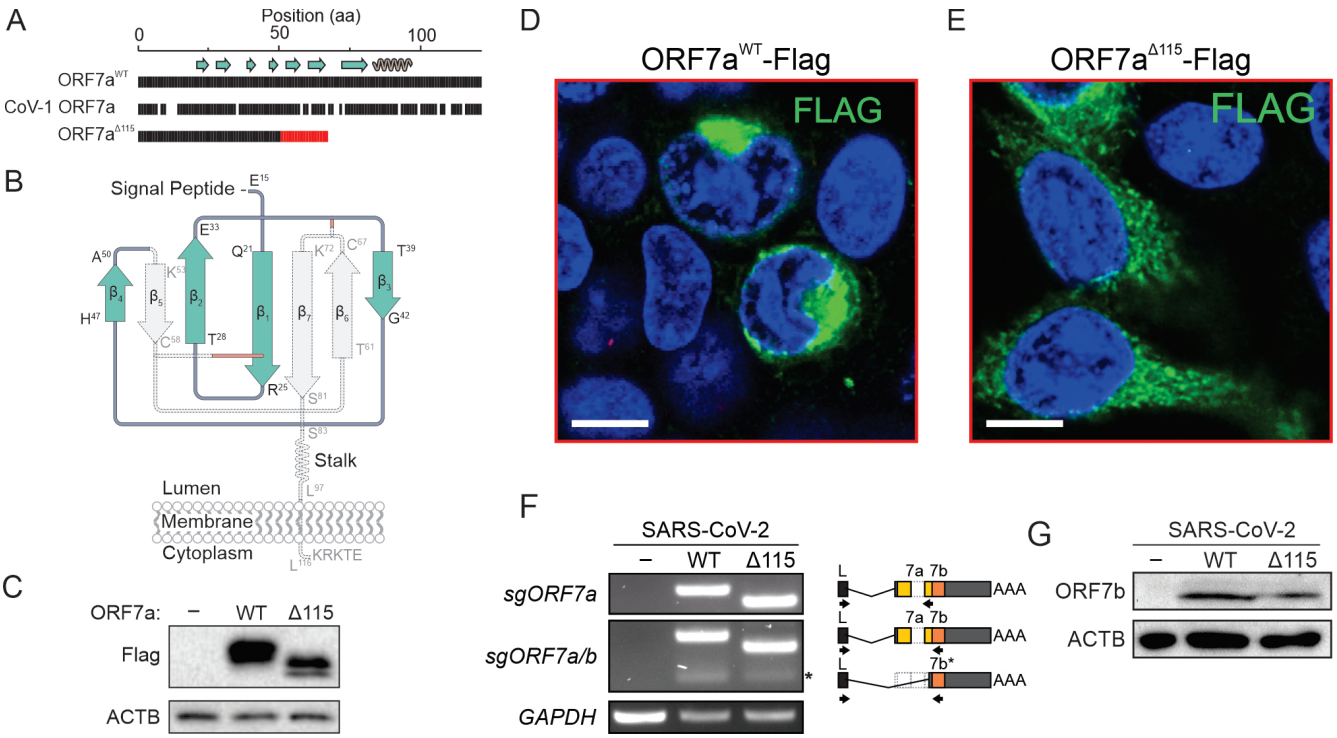

Figure 2. ORF7a truncation results in loss of cellular localization. A) Amino-acid (aa) sequence alignment of SARS-CoV-2 ORF7a^{WT}, ORF7a$^{\Delta 115}$ and SARS-CoV-1 ORF7a. Gaps show non-matching positions, red shows 17 aa sequence resulting from a frameshift in the ORF7a mutant. Beta strands (arrows) and alpha helices (coil) are shown above the alignment. **B)** Diagram of SARS-CoV-2 ORF7a Ig-like fold. Disulfide bonds that stabilize the β-sandwich structure are shown with red lines. The portion of the protein eliminated by the deletion is shown in gray. **C)** C-terminally Flag-tagged ORF7a^{WT} and ORF7a$^{\Delta 115}$ were cloned and overexpressed in HEK 293T-hACE2. Protein expression was confirmed with Western Blot using anti-Flag antibody. β-actin (ACTB) was used as a loading control. **D)** Flag-tagged ORF7a^{WT} or **E)** ORF7a$^{\Delta 115}$ expressed in HEK 293T-hACE2 cells. Immunostaining preformed using an anti-Flag antibody (green). Cell nuclei were stained with Hoechst 33342 (blue). White scale bar is 10 μm. **F)** ORF7a and ORF7b sgRNAs were identified by RT-PCR. Lower MW band corresponding to ORF7b sgRNA is indicated with asterisk (*). Diagram on the right shows primer (arrows) positions. Specificity of PCR products was confirmed with sanger sequencing. GAPDH was used as a control for cDNA synthesis. **G)** Lysates from SARS-CoV-2 infected cells were probed with antibodies raised against ORF7b protein. ACTB was used as a loading control.

NAs. Finally, to verify that the ORF7a$^{\Delta 115}$ variant does not impact ORF7b translation, we probed cell lysates 24 hpi with an anti-ORF7b antibody (Figure 2G). We detected ORF7b protein in both viral strains with band intensities that suggest similar expression levels. Collectively, these data indicate that the ORF7a$^{\Delta 115}$ mutation has no collateral effect on the adjacent gene.

ORF7a truncation results in a replication defect

To determine if ORF7a truncations impact viral replication, we infected 293T-hACE2 and Vero E6 cells with ORF7a^{WT} and ORF7a$^{\Delta 115}$ viruses at an MOI 0.05 and measured viral RNA levels using RT-qPCR (Figure 3A and B). In this assay, we used viral strains sharing the same haplotype (C241T, C3037T, C14408T, A23403G) that defines the SARS-CoV-2 B.1 lineage and its derivatives (Figure S3A) (Rambaut, Andrew et al., 2020). Over the course of 120 hours, ORF7a^{WT} viral RNA level increased 960-fold in supernatants from Vero E6 cells (Figure 3A) and 270-fold in supernatants from 293T-hACE2 cells (Figure 3B). The ORF7a$^{\Delta 115}$ virus displayed limited replication that resulted in only 18.6-fold and 7-fold RNA level increase at 120 hpi in VeroE6 and 293T-hACE2, respectively. This growth defect reaches statistical significance starting at 6 hpi in Vero E6 cells (p-value = 0.0135) and 24 hpi in 293T-hACE2 cells (p-value = $6.2*10^{-5}$).

To determine if this phenotype results from a defect in viral egress only, or if viral RNA replication in the host is also affected, we examined early steps in the infection. We extracted total RNA from the infected 293T-hACE2 cells at 0.5, 6 and 24 hpi, and measured SARS-CoV-2 RNA with RT-qPCR. Compared to wildtype, ORF7a$^{\Delta 115}$ RNA level was reduced 2.1-fold inside the host at 6 hpi (p-value = 0.002) and 4.3-fold at 24 hpi (p-value = 0.02) (Figure 3C). This decrease in viral RNA in the host cell differs from the corresponding decrease in the supernatant RNA at 24 hpi (4.3-fold vs 16.8-fold,

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

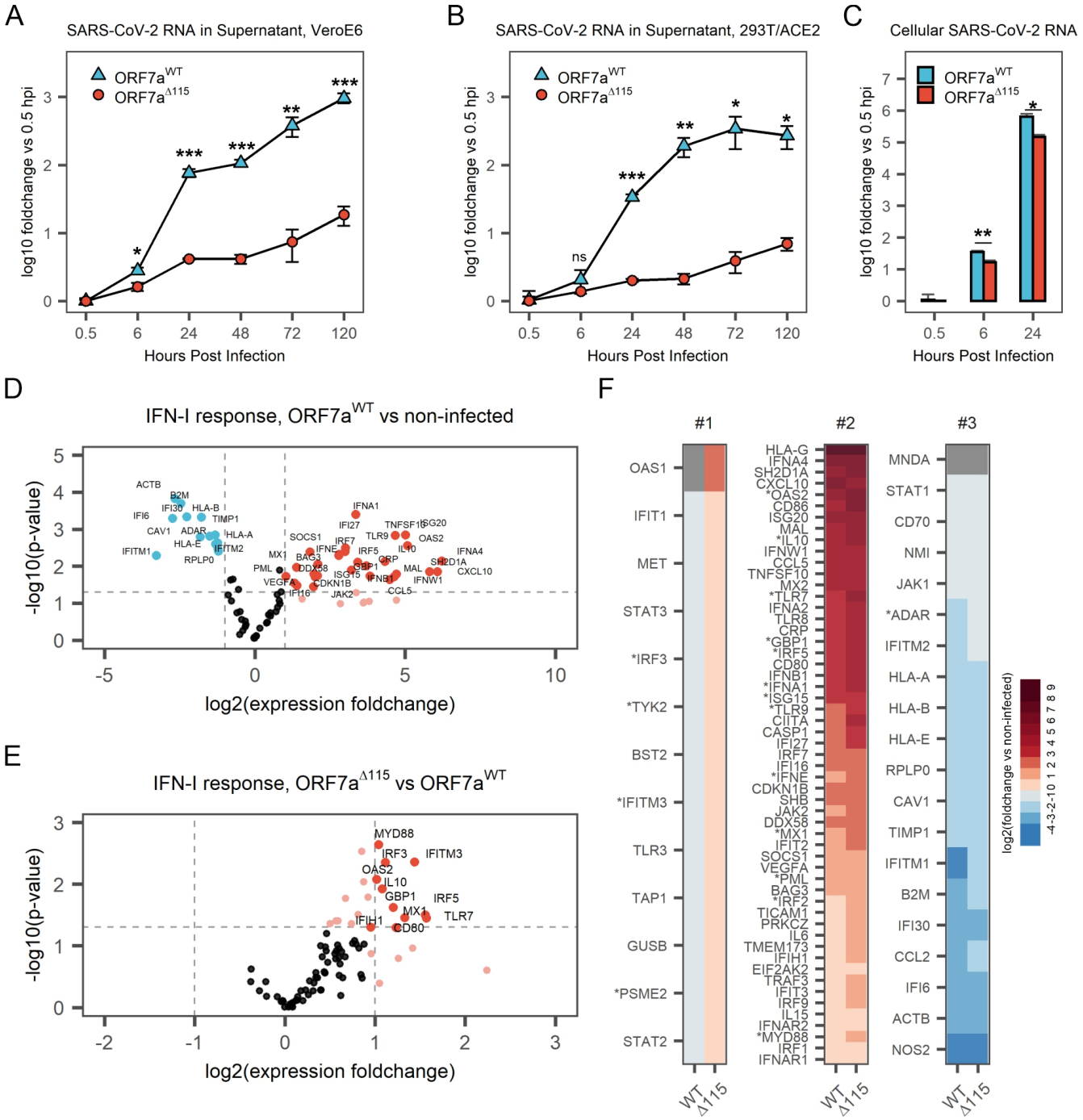

Figure 3. ORF7a truncation results in loss of cellular localization. A) Amino-acid (aa) sequence alignment of SARS-CoV-2 ORF7a^{WT}, ORF7a$^{\Delta 115}$ and SARS-CoV-1 ORF7a. Gaps show non-matching positions, red shows 17 aa sequence resulting from a frameshift in the ORF7a mutant. Beta strands (arrows) and alpha helices (coil) are shown above the alignment. **B)** Diagram of SARS-CoV-2 ORF7a Ig-like fold. Disulfide bonds that stabilize the β-sandwich structure are shown with red lines. The portion of the protein eliminated by the deletion is shown in gray. **C)** C-terminally Flag-tagged ORF7a^{WT} and ORF7a$^{\Delta 115}$ were cloned and overexpressed in HEK 293T-hACE2. Protein expression was confirmed with Western Blot using anti-Flag antibody. β-actin (ACTB) was used as a loading control. **D)** Flag-tagged ORF7a^{WT} or **E)** ORF7a$^{\Delta 115}$ expressed in HEK 293T-hACE2 cells. Immunostaining preformed using an anti-Flag antibody (green). Cell nuclei were stained with Hoechst 33342 (blue). White scale bar is 10 μm. **F)** ORF7a and ORF7b sgRNAs were identified by RT-PCR. Lower MW band corresponding to ORF7b sgRNA is indicated with asterisk (*). Diagram on the right shows primer (arrows) positions. Specificity of PCR products was confirmed with sanger sequencing. GAPDH was used as a control for cDNA synthesis. **G)**) Lysates from SARS-CoV-2 infected cells were probed with antibodies raised against ORF7b protein. ACTB was used as a loading control.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

respectively; Figure 3B, C). This disparity suggests that the ORF7a$^{\Delta 115}$ mutation might affect genomic RNA replication and transcription, as well as viral egress.

ORF7a$^{\Delta 115}$ mutation alters host Type I interferon response to SARS-CoV-2.

The ORF7a protein is not essential for SARS-CoV-2 replication (Xia et al., 2020). When overexpressed in vitro, ORF7a suppresses type I IFN (IFN-I) signaling (Xia et al., 2020). Interferon stimulated genes (ISGs) exert anti-viral activities and thus we hypothesized that the replication defect in ORF7a$^{\Delta 115}$ results from impaired suppression of host anti-viral immune responses. To test this hypothesis, we measured expression of type I IFN response genes upon infection with ORF7a^{WT} and ORF7a$^{\Delta 115}$ strains using an RT-qPCR array targeting 91 human transcripts (Table S2). The array included 3 reference genes that we used to normalize expression of 88 target genes. In agreement with recent transcriptomic studies, we detected a type I IFN-mediated antiviral response after infecting HEK 293T-hACE2 with SARS-CoV-2 isolates (Figure 3D, Figure S3B) (Banerjee et al., 2020; Blanco-Melo et al., 2020). Out of 88 type I IFN response genes included in the RT-qPCR array, 42 were differentially expressed ($\geq$2-fold change, p-value $<$0.05) in ORF7a^{WT} infected cells and 55 in ORF7a$^{\Delta 115}$ infection, as compared to the non-infected control (Figure 3D, Figure S3B). Both viruses stimulated expression of multiple type-I interferon genes, antiviral ISGs, interleukins and proinflammatory cytokines (Table S2).

To understand the effect of ORF7a$^{\Delta 115}$ on the host type I IFN response, we compared expression of genes targeted with RT-qPCR array between the two viruses. In general, the ORF7a$^{\Delta 115}$ variant induces elevated type I IFN response to the infection (Figure S3C, D). Analysis of genes differentially expressed between the two viral infections produced an asymmetric volcano plot gravitating towards upregulation (Figure 3E). A subset of ISGs had expression levels that are significantly different (p-value $<$0.05) between the two viral infections (Figure 3E, F, Table S2). This subset includes sensors (TLR7), signal transducers (MYD88, OAS2), transcriptional regulators (IRF3, IRF5) and restriction factors (GBP1, IFITM3, MX1) known to combat infections by RNA viruses (Schoggins, 2019). One of these differentially expressed ISGs (IFITM3) restricts SARS-CoV-2 entry (Shi et al., 2020; Zang et al., 2020). Additionally, polymorphisms in IFITM3, MX1 and TLR7 are linked to the severity of COVID-19 (Andolfo et al., 2020; Jin, R. et al., 2020; Pati et al., 2021). This upregulation of ISGs indicates that the ORF7a$^{\Delta 115}$ mutation limits viral suppression of the IFN-I response, which is consistent with previous work suggesting that ORF7a prevents transcriptional activation by STAT1/STAT2 complex (Xia et al., 2020). Collectively, our data show that a naturally occurring truncation variant of the ORF7a protein alters the host innate immune responses to the in vitro infection.

Discussion

In this study, we use genomic surveillance and phylogenetics to identify ORF7a variants that have emerged globally throughout the SARS-CoV-2 pandemic. Local occurrence of ORF7a mutations has been reported by others (Addetia et al., 2020; Holland et al., 2020; Rosenthal et al., 2020). However, the effect of these mutations on viral fitness and host immune responses has not been investigated. Here we have isolated an ORF7a$^{\Delta 115}$ variant and show that this mutation results in an in vitro growth defect. This defect is coupled with elevated IFN response to SARS-CoV-2 and specifically ISGs.

Interferon systems are a frontline of host defense that signals infection and elicits an antiviral state (McNab et al., 2015). Pretreatment with type I, type II and type III IFNs restricts SARS-CoV-2 infection and replication in cell culture (Mantlo et al., 2020; Miorin et al., 2020). SARS-CoV-2 deploys multiple proteins (e.g., nsp6, nsp13, ORF6, ORF7a, ORF7b and ORF9b) to shut down IFN signaling and dampen innate immune responses (Jiang et al., 2020; Lei et al., 2020; Xia et al., 2020). In particular, when overexpressed ORF7a subverts STAT2 phosphorylation, blocking IFN-dependent transcriptional activation of antiviral ISGs (Xia et al., 2020). Both type I and III IFNs signal through the JAK-STAT pathway and overlap substantially in the transcriptional responses they induce (Kotenko and Durbin, 2017). Therefore, ORF7a likely suppresses both IFN types, however its role in suppressing the type III IFN response has yet to be verified (Xia et al., 2020). The ORF7a$^{\Delta 115}$ variant impairs viral replication in HEK 293T-hACE2 as well as in Vero E6, the latter of which do not express type I IFN genes (Emeny and Morgan, 1979; Osada et al., 2014). While type I IFN genes are deleted in Vero E6 cells, the type III IFNs are intact and are activated upon infection with RNA viruses, which results in ISG upregulation (Stoltz and Klingstrom, 2010; Wang et al., 2020). Taken together, these data suggest that the growth defect of the ORF7a$^{\Delta 115}$ strain in Vero E6 cells might result from a type III IFN-dependent response that the virus fails to suppress.

Cytokine profiling and RNA-seq of clinical samples reveal suppressed or delayed IFN responses in COVID-19 patients (Arunachalam et al., 2020; Hadjadj et al., 2020; Lucas et al., 2020). Dysregulated immune responses coupled with exacerbated inflammatory cytokines drive pathogenesis of the disease and correlate with its severity (Jin, X. et al., 2020; Tay et al., 2020; Zhou, F. et al., 2020). Here, we detected enhanced IFN signaling in cells infected with a naturally occurring ORF7a$^{\Delta 115}$ variant of SARS-CoV-2. We anticipate that major truncation and loss of protein targeting (Figure 2) limit the ability of ORF7a to suppress IFN signaling. Given the importance of the IFN response in COVID-19, we anticipate that ORF7a mutations might affect SARS-CoV-2 pathogenicity. Clinical comparisons will be required to examine this potential effect on COVID-19 features.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Deletions in ORF7a, 7b and ORF8 have previously been reported (Addetia et al., 2020; Holland et al., 2020; Rosenthal et al., 2020; Su et al., 2020). Similar deletions emerged in MERS- and SARS-CoV, which were linked to viral attenuation (Chinese, 2004; Muth et al., 2018). A 382-nt deletion in ORF8 is associated with milder SARS-CoV-2 infection in patients (Young et al., 2020). While this mutation affects clinical features of COVID-19, it has no evident effect on viral replication or host transcriptional responses in human cell culture (Pekosz et al., 2020). Overexpression of ORF8 results in post-translational downregulation of MHC-I, therefore the protein might modulate the interaction between infected cells and immune system (Zhang et al., 2020).

A new lineage of SARS-CoV-2 (i.e., B.1.1.7) has recently emerged (Rambaut, Andrew et al., 2020). Aside from changes in the Spike protein, this lineage also includes a premature stop codon in ORF8 that truncates most of the protein (Volz et al., 2021). It has been hypothesized that B.1.1.7 might originate from an immunocompromised patient with chronic SARS-CoV-2 infection treated with convalescent plasma (Baang et al., 2021; Bazykin et al., 2021; Rambaut, Andrew et al., 2020). Several studies illustrate how persistent SARS-CoV-2 infections associate with accelerated viral evolution (Avanzato et al., 2020; Choi et al., 2020). In immunocompromised hosts the selective pressure on anti-immune accessory genes is lifted, which might permit loss-of-function mutations. However, these mutations will not persist in the immunocompetent population (Young et al., 2020), unless they co-occur with mutations that increase virus transmissibility or become fixed due to the founder effect (Lauring and Hodcroft, 2021; Rambaut, Andrew et al., 2020).

Limitations of Study

The viral replication defect and cellular immune responses presented here were performed using in vitro cell culture. While HEK 293T-hACE2 and Vero E6 cells support virus entry and replication (Cantuti-Castelvetri et al., 2020; Chu et al., 2020), they may not recapitulate particular aspects of an in vivo infection (Takayama, 2020). We anticipate that the in vitro results presented here will provide important context for clinical comparisons, similar to those recently published for ORF8 (Young et al., 2020).

Materials and Methods

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell cultures

HEK 293T-hACE2 cells (BEI, NR-52511) were generously provided by Dr. Seth Pincus (Montana State University) and maintained at 37C and 5% CO2 in Dulbecco's Modified Eagle Medium (DMEM, Gibco, cat. #12100-061) supplemented with 10% fetal bovine serum (FBS, ATLAS Biologicals, Lot. #F31E18D1), sodium bicarbonate (3.7 g/L), 50 I.U./mL penicillin and 50 μg/mL streptomycin. Vero E6 cells (ATCC® cat. #CRL-1586) were maintained at 37C and 5% CO2 in Eagle's Modified Eagle Medium (EMEM, ATCC®, cat. #30-2003) supplemented with 10% fetal bovine serum (FBS, ATLAS Biologicals, Lot. #F31E18D1) and 50 I.U./mL penicillin and 50 μg/mL streptomycin. All cell lines tested negative for mycoplasma.

Plasmids

Human-codon optimized dsDNA gene fragments (gBlocks) encoding for SARS-CoV-2 C-terminally FLAG-tagged ORF7aWT and ORF7a$^{\Delta 115}$ proteins were synthesized by Integrated DNA Technologies (IDT). Sequence of ORF7a from Wuhan-Hu-1 reference (genbank MN908947.3) was used for ORF7aWT gBlock. The gene fragments were cloned into the pEGFP-N1 (Clontech cat# 6085-1) backbone to replace the eGFP gene.

Antibodies

Western blotting: Rabbit anti-ACTB (Cat: AC026, 1:20,000) antibodies were from ABclonal, Mouse anti-FLAG antibodies (Cat: MA1-91878-1MG, 1:1000) were from ThermoFisher Scientific, sheep anti-ORF7b antibodies (1:120) were from MRC I PPU, College of Life Sciences, University of Dundee, Scotland; Goat anti-mouse IgG peroxidase-conjugated (Cat: 115-035-003, 1:10,000), Goat anti-rabbit IgG peroxidase-conjugated (Cat: 111-035-003, 1:10,000) and Donkey Anti-Sheep IgG peroxidase-conjugated (Cat: 713-035-147, 1:10,000) antibodies were from Jackson ImmunoResearch. Immunocytochemistry: Mouse anti-FLAG antibodies (Cat: MA1-91878-1MG, 1:200), Goat anti-mouse-AlexaFluor-488 (Cat: A11001, 1:2,000) antibodies were from Invitrogen.

Human clinical sample collection and preparation

Clinical samples were obtained with local IRB approval (protocol #DB033020) and informed consent from patients undergoing testing for SARS-CoV-2 at Bozeman Health Deaconess Hospital. Nasopharyngeal swabs from patients that tested positive for SARS-CoV-2 were collected in viral transport media. RNA was extracted from all patient samples using QIAamp Viral RNA Mini Kit (QIAGEN) a biosafety level 3 (BSL3) laboratory. All samples were heat-inactivated before removing from BSL3.

Production and titration of coronavirus stocks

The nasopharyngeal swabs in the viral transport media were used to generate viral stocks as previously described (Harcourt et al., 2020). Briefly, 100 ul of the viral transport media was two-fold serially diluted and applied to Vero E6 cells When infected cells showed extensive cytopathic effect, the media was collected and utilized to generate a greater volume second passage in Vero E6 cells. When CPE was apparent, supernatants were centrifuged 1,000 RCF for 5 minutes to remove cellular debris. Virus titer in clarified supernatants was determined with plaque assay in Vero E6 cells as described(Loveday et al., 2020). For plaque assay, Vero E6 cells were incubated with viral inoculum at limiting dilutions. Inoculated cells were overlayed with either 0.75% methylcellulose, DMEM supplemented with 2% FBS and 1% pen-strep and incubated for 4 days. Cells were fixed and stained with 0.5% methylene blue in 70% ethanol. Plaques were counted and the overall titer was calculated. Viral stocks identity was confirmed via whole genome sequencing on Oxford Nanopore. All SARS-CoV-2 experiments were performed in a BSL3 laboratory.

Symptom onset data and clinical test results

Suspect cases of COVID-19 were tested in a CLIA lab and instructed to self-quarantine until notified of the RT-qPCR test results. All laboratory confirmed positive cases of COVID-19 were contacted via telephone by local public health nurses to complete contact tracing. During this interview, the nurses collected recorded symptoms, symptom onset date, travel history, contact with other known laboratory confirmed cases, close contacts and activities on the two days before symptom onset up until notification of a positive test. Data collection was conducted as part of a public health response. The study was reviewed by the Montana State University Institutional Review Board (IRB) For the Protection of Human Subjects (FWA 00000165) and was exempt from IRB oversight in accordance with Code of Federal regulations, Part 46, section 101. All necessary patient/participant consent has been obtained and the appropriate institutional forms have been archived.

METHOD DETAILS

Reverse Transcription quantitative PCR (RT-qPCR)

RT-qPCR was performed using CDC primers (N1 and N2) and probes from the 2019-nCoV RUO Kit (IDT# 10006713). SARS-CoV-2 RNA was quantified using one-step RT-qPCR in ABI 7500 Fast Real-Time PCR System according to CDC protocol (https://www.fda.gov/media/ 134922/download). In brief, 20 μL reactions included 8.5 μL of Nuclease-free Water, 1.5 μL of

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Primer and Probe mix (IDT, 10006713), 5 μL of TaqPath 1-Step RT-qPCR Master Mix (ThermoFisher, A15299) and 5 μL of the template. Nuclease-free water was used as negative template control (NTC). Amplification was performed using following program: 25C for 2 min, 50C for 15 min, 95C for 2 min followed by 45 cycles of 95C for 3 s and 55C for 30 s. Run data was analyzed in SDS software v1.4 (Applied Biosystems). The NTC showed no amplification throughout the 40 cycles of qPCR.

RT-PCR and SARS-CoV-2 genome sequencing

For sequencing 10 μL of RNA from SARS-CoV-2 positive patient sample was reverse transcribed using SuperScript IV (Thermo Fisher Scientific) according to the supplier's protocol. The protocol developed by ARTIC Network was used to generate and sequence amplicon library that covers whole SARS-CoV-2 genome on Oxford Nanopore using ligation sequencing kit (SQK-LSK109) (https://artic.network/ncov-2019) (Grubaugh et al., 2019; Tyson et al., 2020). Briefly, two multiplex PCR reactions were performed with primer pools from ARTIC nCoV-2019 V3 Panel (IDT, Table S3) using Q5 High-Fidelity DNA Polymerase (New England Biolabs). Reactions were performed with the following thermocycling conditions: 98C for 2min, 30 cycles of 98C for 15 s and 65C for 5 min. Two resulting amplicon pools for each patient sample were combined and used for library preparation. After end repair (NEB E7546) samples were barcoded using Native Barcoding Expansion Kits EXP-NBD104 and EXP-NBD114 from Oxford Nanopore. A total of 24 barcoded samples were pooled together and Nanopore adaptors were ligated. After clean-up 20 ng of multiplexed library DNA was loaded onto the MinION flowcell for sequencing. A total of 0.3 Gb of raw sequencing data was collected per patient sample.

SARS-CoV-2 genome assembly

MinKNOW software was used to basecall raw Nanopore reads in high-accuracy mode. ARTIC bioinformatic pipeline for COVID-19 was used to analyze reads (https://artic.network/ncov-2019). Pipeline included demultiplexing with guppy barcoder and generating consensus sequences with minimap2 and calling single nucleotide variants with nanopolish relative to Wuhan-Hu-1/2019 reference genome (Li, 2018; Quick et al., 2016; Wu et al., 2020). Consensus sequences were uploaded to GISAID (https://www.gisaid.org/), accession IDs are provided in Supplemental Table S1.

RT-PCR and Sanger sequencing

To verify the ORF7a$^{\Delta 115}$ mutation we used RT-PCR with primers that flank the deletion region (Table S3). Reactions were performed with Q5 High-Fidelity DNA Polymerase in 25 μL volume (New England Biolabs) using following program: 98C for 2 min, 35 cycles of 98C for 15 s and 65C for 5 min, 35 cycles. PCR products were analyzed on 1% agarose gels stained with SYBR Safe (Thermo Fisher Scientific). The remaining volume of the reaction was purified using DNA Clean & Concentrator kit (Zymo Research) and sent to Psomagen for Sanger sequencing. Each PCR product was sequenced with both forward and reverse primers used for PCR. To detect sgRNAs of ORF7a and ORF7b primers annealing to Leader region of SARS-CoV-2 and inside the ORF were used (Table S3). PCR products were cut out from 1% agarose gel and purified with Zymoclean Gel DNA Recovery Kit. Purified products were sequenced with forward and reverse primers. Sequence was verified by aligning to SARS-CoV-2 genome in UGENE software (Unipro).

Phylogenetic and ORF7a mutational analysis

An alignment of 181,003 SARS-CoV-2 genomes was downloaded from the GISAID database at 9:11 AM on 2020-11-11, and sequences without corresponding metadata entries were removed. The nucleotide positions encoding ORF7a in the reference sequence (Wuhan-Hu-1; EP_ISL_402125) were extracted for the remaining 180,971 sequences, and ORF7a sequences with 100% sequence identity were clustered using CD-HIT (with settings: -c 1 -aL 1). One representative was randomly selected from each of the 846 ORF7a sequence clusters and mutations in each nucleotide and translated sequence were determined using the Biostrings package in R. Graphs of these mutations were rendered with the ggplot2 package. To construct a phylogenetic tree of global and Bozeman SARS-CoV-2 sequences, up to ten sequences from each country in the world were selected for each month of the pandemic using the Filter utility in the Augur pipeline (–group-by country year month – sequences-per-group 10). This resulted in an alignment of 4,235 GISAID sequences. Then, the NextClade online tool was used to determine the quality of 73 SARS-CoV-2 genomes (% genome covered) sequenced from Bozeman patients. 56 Bozeman sequences that were determined to be of "Good" quality were merged with the alignment of GISAID sequences, as well as an alignment of 669 representative SARS-CoV-2 genomes that had non-synonymous mutations in ORF7a. The bat coronavirus RaTG13 (MN996532.2) and the Wuhan- Hu-1 SARS-CoV-2 sequences were also merged, resulting in an alignment of 4,959 full genome sequences, and positions that are hypervariable or prone to sequencing artifacts were masked with a VCF file downloaded from EMBL at 12:45PM on 2020-05- 29. The alignment was trimmed of columns composed of 90% or more gaps using Trimal (-gt 0.9), and IqTree was used to construct a maximum likelihood phylogeny (-m GTR -ninit 2 -n 2 -me 0.05). Branch lengths and internal nodes were rescaled on the resulting tree using TreeTime, and tree was re-rooted to the RaTG13 sequence using the APE package before being visualized with the ggTree package in R.

Western Blot

293T-hACE2 cells were transfected using Lipofectamine 3000 (Invitrogen, L3000-015) with plasmids encoding for ORF7aWT, ORF7a$^{\Delta 115}$ or control pEGFP-N1 plasmid. To detect ORF7b expression cells were infected with ORF7aWT or ORF7a$^{\Delta 115}$ SARS-CoV-2 strains at MOI = 0.05. After 24 hours, cells were washed two times with PBS and lysed in RIPA buffer (150 mM sodium chloride, 1% NP-40, 0.5% sodium deoxycholate, 0.2% SDS, 50 mM Tris, pH 8.0) at 4C for 30 min. The lysates were clarified by centrifugation (10,000 g, 20 min) and stored at -80C. For Western blot lysates were mixed with 6xLaemmli SDS-PAGE buffer and heated at 98C for 5 min, resolved in 12% SDS-PAGE gel and transferred onto a PVDF (ORF7b) or nitrocellulose (ORF7a) membrane using Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad, #1703930). Membranes were blocked and probed with indicated antibodies. The proteins were visualized using Pierce ECL Western Blotting Substrate (ThermoFisher Scientific, #32106) and exposed to X-Ray film (sc-201696, Santa Cruz Biotech).

Immunocytochemistry

HEK 293T-hACE2 were transfected with plasmid DNA using Lipofectamine 3000 (Invitrogen REF: L3000-015) according to manufacturer's instructions. Cells were either transfected with expression vectors for SARS-CoV-2 ORF7aWT, SARS-CoV-2 ORF7a$^{\Delta 115}$, or mock transfected. 24 hours after transfection, cells were detached and seeded at 50% confluency on glass coverslips (Fisherbrand, #12-545-80) pre-coated with poly-D-lysine (Cultrex, #3439-100-01) in a 24 well plate. The next day, media was removed, and cells were washed with PBS and fixed with ice-cold 100% methanol on ice for 10 minutes. Half of the methanol was replaced with PBS three times before the methanol solution was discarded and cells were washed three times in PBS (5 minutes each wash). Cells were blocked with 1% BSA (Fisher Scientific BP9703-100) in PBS + 0.1% Tween-20 (PBST) for 30 minutes at room temperature. After blocking, cells were stained with mouse monoclonal anti-Flag antibody (Invitrogen, MA1-91878) (1:200 in 1% BSA in PBST) overnight at 4C. The next day, cells were washed 3X in PBS (5 min each) and stained with secondary antibody diluted 1:2000 in 1% BSA in PBST for 1 h at RT in the dark (Goat anti-mouse AlexaFluor488; Invitrogen, A11001). After 3x PBS washes nuclei were stained with Hoechst 33342 in PBS for 5 minutes at RT. Coverslips were then mounted with ProLong Gold antifade mounting media (Invitrogen, P36934) on Superfrost Plus microscope slides (#22-037-246). Immunostained cells were imaged using a Leica SP8 confocal microscope.

SARS-CoV-2 replication assays

HEK 293T-hACE2 and Vero E6 cells were seeded in the 48-well plate and infected with ORF7aWT or ORF7a$^{\Delta 115}$ SARS-CoV-2 strains at MOI = 0.05 in 50 μL of the cell culture media. Infections were performed in triplicates. Next, 0.5 h post infection (hpi) 500 μL of the fresh cell culture media was added to each well. The supernatants (140 ul) from the infected cells were harvested at the following time points 0.5, 6, 24, 48, 72 and 120 hpi. The RNAs from the supernatants were extracted using QIAamp Viral RNA Mini Kit (QIAGEN) and used as an input in RT-qPCR with CDC N1 and N2 primers and probes as described above. Relative quantification was used (ΔCt method) to calculate viral replication versus 0.5 hpi timepoint as 2-ΔCt. The infected cells were harvested 0.5, 6 and 24 hpi and washed with PBS. Total RNA from cells was extracted using RNeasy kit (QIAGEN, Cat No./ID:

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

74104) with on-column DNase digestion step (QIAGEN, 79254) according to the manufacturers protocol. Viral RNA was quantified using RT-qPCR CDC N1 primers and probe. Human ACTB endogenous control was quantified with TaqMan assay (ThermoFisher, 4333762T). ΔΔCt method was used to normalize viral RNA level to host RNA and 0.5 hpi. Viral replication was calculated for each replicate as 2-ΔΔCt.

Type I interferon response assay

HEK 293T-hACE2 were seeded in the 48-well plate and infected with ORF7aWT or ORF7a$^{\Delta 115}$ SARS-CoV-2 strains at MOI = 0.05 in 50 μl of the media. After 30 min on infection the 500 μL of the cell culture media was added to each well. At 24 hpi the cells were washed once with PBS, detached using Trypsin and the total RNA was extracted using RNeasy kit (QIAGEN, Cat No./ID: 74104) with on-column DNase digestion (QIAGEN, Cat No./ID: 79254) according to the manufacturer's protocol. Non-infected cells were processed using the same protocol with mock infection. After extraction, 0,5-1 μg of total RNA was reverse transcribed Reverse using SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific) according to the supplier's protocol. The cDNA was diluted 10-fold in water and analyzed with Type I interferon response (SAB Target List) H96 qPCR array. Each reaction contained 2 μL of cDNA, 10 μL of 2X SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) and 8 μL of water. Reactions were performed on QuantStudio 3 Real-Time PCR System instrument with following thermocycling conditions: 95C for 2min, 40 cycles of 95C for 5 s and 60C for 30 s. After amplification, melt curve analysis was performed to examine product specificity. Assay was performed in triplicates for non-infected cells and both viruses. Run data was analyzed in PrimePCR Analysis software (Bio-Rad). Expression of IFN-I response genes was normalized to geometric mean of 3 reference transcripts (TBP, GAPDH, HPRT1) and non-infected control (ΔΔCt method).

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed in RStudio v1.2.1335. Moving averages (n = 7) for symptom onset and positive COVID-19 tests data were calculated using geom_ma function from tidyquant package and plotted with ggplot2 in RStudio. RT-qPCR data is shown as mean of three biological replicates (each with three technical replicates) $\pm$ standard deviation (SD). Replication of viruses and expression of IFN-I response genes was compared using Student's t-test. Medians of IFN-I response between two viruses were compared with Wilcoxon signed-rank test. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ or ns (no significant difference, $p > 0.05$).

Acknowledgements

We are grateful to members of Bozeman Health that provided de-identified patient samples. Specifically, C. Nero, D. Smoot, W. Wallace, C. Faurot-Daniels, V. Lawrence and M. Blauvelt. We are also grateful to M. Flenniken, K. Daughenbaugh, and other members of the COVID task force at MSU for assistance establishing the COVID testing center. We thank the Lefcort lab for generous use of their fluorescent confocal microscope and Dr. Pincus for providing the HEK 293T-hACE2 cells. Research in the Wiedenheft lab is supported by the National Institutes of Health (1R35GM134867), the Montana State University Agricultural Experimental Station, the MJ Murdock Charitable Trust, the Gianforte Foundation and the MSU Office of the Vice President for Research. We thank the GISAID's EpiFlu^{TM} Database and contributing laboratories (Table S4). The phylogenetic analysis in this paper would not have been possible without their willingness to share data.

Author contributions

Conceptualization, B.W., A. Nemudryi, and A. Nemudraia; Methodology, B.W., A. Nemudraia, A. Nemudryi, D.S and J.H.; Sample acquisition, D.B. and B.W.; Investigation & Data Collection, A. Nemudraia., A. Nemudryi, D.S., J.H., J.N., H.L., K.V.; Genomics and bioinformatic analysis, T.W., A. Nemudryi, C.C.; Writing – Original Draft, B.W., A. Nemudryi, A. Nemudraia. and T.W.; Writing – Review & Editing, B.W, A. Nemudryi, A. Nemudraia, T.W., M.J.

Declaration of interests

B.W. is the founder of SurGene LLC, and VIRIS Detection Systems Inc. B.W., A. Nemudryi, and A. Nemudraia are inventors on patents related to CRISPR-Cas systems and applications thereof.

References

1. Addetia, A., Xie, H., Roychoudhury, P., Shrestha, L., Loprieno, M., Huang, M.-L., Jerome, K.R., and Greninger, A.L. (2020). Identification of multiple large deletions in ORF7a resulting in in-frame gene fusions in clinical SARS-CoV-2 isolates. Journal of Clinical Virology. 129. DOI: 10.1016/j.jcv.2020.104523.
2. Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., and Garry, R.F. (2020). The proximal origin of SARS-CoV-2. Nat Med. 26(4), 450-452. Published online 2020/04/15 DOI: 10.1038/ s41591-020-0820-9.
3. Andolfo, I., Russo, R., Lasorsa, A.V., Cantalupo, S., Rosato, B.E., Bonfiglio, F., Frisso, G., Pasquale, A., Cassese, G.M., Servillo, G., et al. (2020). Common variants at 21q22.3 locus influence MX1 gene expression and susceptibility to severe COVID-19. bioRxiv. DOI: 10.1101/ 2020.12.18.20248470.
4. Arunachalam, P.S., Wimmers, F., Mok, C.K.P., Perera, R., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., et al. (2020). Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. Science. 369(6508), 1210-1220. Published online 2020/08/14 DOI: 10.1126/ science.abc6261.
5. Avanzato, V.A., Matson, M.J., Seifert, S.N., Pryce, R., Williamson, B.N., Anzick, S.L., Barbian, K., Judson, S.D., Fischer, E.R., Martens, C., et al. (2020). Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. Cell. 183(7), 1901-1912.e1909. DOI: 10.1016/j.cell.2020.10.049.
6. Baang, J.H., Smith, C., Mirabelli, C., Valesano, A.L., Manthei, D.M., Bachman, M.A., Wobus, C.E., Adams, M., Washer, L., Martin, E.T., et al. (2021). Prolonged Severe Acute Respiratory Syndrome Coronavirus 2 Replication in an Immunocompromised Patient. The Journal of Infectious Diseases. 223(1), 23-27. DOI: 10.1093/ infdis/jiaa666.
7. Banerjee, A., El-Sayes, N., Budylowski, P., Richard, D., Maan, H., Aguiar, J.A., Baid, K., D'Agostino, M.R., Ang, J.C., Tremblay, B.J.M., et al. (2020). Experimental and natural evidence of SARS-CoV-2 infection-induced activation of type I interferon responses. bioRxiv. DOI: 10.1101/ 2020.06.18.158154.
8. Bazykin, G., Stanevich, O., Danilenko, D., Fadeev, A., Komissarova, K., Ivanova, A., Sergeeva, M., Safina, K., Nabieva, E., Klink, G., et al. (2021). Emergence of Y453F and Δ69-70HV mutations in a lymphoma patient with long-term COVID-19. https://virological.org/.
9. Bedford, T., Greninger, A.L., Roychoudhury, P., Starita, L.M., Famulare, M., Huang, M.-L., Nalla, A., Pepper, G., Reinhardt, A., Xie, H., et al. (2020). Cryptic transmission of SARS-CoV-2 in Washington state. Science. 370(6516), 571-575. DOI: 10.1126/ science.abc0523.
10. Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.-C., Uhl, S., Hoagland, D., Møller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., et al. (2020). Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. Cell. 181(5), 1036-1045.e1039. DOI: 10.1016/ j.cell.2020.04.026.
11. Cantuti-Castelvetri, L., Ojha, R., Pedro, L.D., Djannatian, M., Franz, J., Kuivanen, S., van der Meer, F., Kallio, K., Kaya, T., Anastasina, M., et al. (2020). Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. Science. 370(6518), 856-860. DOI: 10.1126/science.abd2985.
12. Chinese, S.M.E.C. (2004). Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science. 303(5664), 1666-1669. Published online 2004/01/31 DOI: 10.1126/science.1092002.
13. Choi, B., Choudhary, M.C., Regan, J., Sparks, J.A., Padera, R.F., Qiu, X., Solomon, I.H., Kuo, H.-H., Boucau, J., Bowman, K., et al. (2020). Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. New England Journal of Medicine. 383(23), 2291-2293. DOI: 10.1056/NEJMc2031364.
14. Chu, H., Chan, J.F.-W., Yuen, T.T.-T., Shuai, H., Yuan, S., Wang, Y., Hu, B., Yip, C.C.-Y., Tsang, J.O.-L., Huang, X., et al. (2020). Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. The Lancet Microbe. 1(1), e14-e23. DOI: 10.1016/s2666-5247(20)30004-5.
15. Emeny, J.M., and Morgan, M.J. (1979). Regulation of the Interferon System: Evidence that Vero Cells have a Genetic Defect in Interferon Production. Journal of General Virology. 43(1), 247-252. DOI: 10.1099/0022-1317-43-1-247.
16. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 583(7816), 459-468. Published online 2020/05/01 DOI: 10.1038/s41586-020-2286-9.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

17. Hachim, A., Kavian, N., Cohen, C.A., Chin, A.W.H., Chu, D.K.W., Mok, C.K.P., Tsang, O.T.Y., Yeung, Y.C., Perera, R.A.P.M., Poon, L.L.M., et al. (2020). ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. Nature Immunology. 21(10), 1293-1301. DOI: 10.1038/s41590-020-0773-7.

18. Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Pere, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al. (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science. 369(6504), 718-724. Published online 2020/07/15 DOI: 10.1126/science.abc6027.

19. Harcourt, J., Tamin, A., Lu, X., Kamili, S., Sakthivel, S.K., Murray, J., Queen, K., Tao, Y., Paden, C.R., Zhang, J., et al. (2020). Severe Acute Respiratory Syndrome Coronavirus 2 from Patient with Coronavirus Disease, United States. Emerging Infectious Diseases. 26(6), 1266-1273. DOI: 10.3201/eid2606.200516.

20. Holland, L.A., Kaelin, E.A., Maqsood, R., Estifanos, B., Wu, L.I., Varsani, A., Halden, R.U., Hogue, B.G., Scotch, M., Lim, E.S., et al. (2020). An 81-Nucleotide Deletion in SARS-CoV-2 ORF7a Identified from Sentinel Surveillance in Arizona (January to March 2020). Journal of Virology. 94(14). DOI: 10.1128/jvi.00711-20.

21. Holshue, M.L., DeBolt, C., Lindquist, S., Lofy, K.H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., et al. (2020). First Case of 2019 Novel Coronavirus in the United States. N Engl J Med. 382(10), 929-936. Published online 2020/02/01 DOI: 10.1056/NEJMoa2001191.

22. Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., et al. (2019). The reactome pathway knowledgebase. Nucleic Acids Research. DOI: 10.1093/nar/gkz1031.

23. Jiang, H.-w., Zhang, H.-n., Meng, Q.-f., Xie, J., Li, Y., Chen, H., Zheng, Y.-x., Wang, X.-n., Qi, H., Zhang, J., et al. (2020). SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70. Cellular Molecular Immunology. 17(9), 998-1000. DOI: 10.1038/s41423-020-0514-8.

24. Jin, R., Dong, T., Knight, J.C., Xiang, H., Hu, Z., Sun, J., Feng, Y., Dai, Y., Zhang, C., Liang, L., et al. (2020). Interferon-Induced Transmembrane Protein 3 Genetic Variant rs12252-C Associated With Disease Severity in Coronavirus Disease 2019. The Journal of Infectious Diseases. 222(1), 34-37. DOI: 10.1093/infdis/jiaa224.

25. Jin, X., Zhang, B., Zhou, X., Qiu, Y., Song, Y., Feng, F., Feng, J., Song, Q., Jia, Q., and Wang, J. (2020). Clinical characteristics of 82 cases of death from COVID-19. Plos One. 15(7). DOI: 10.1371/journal.pone.0235458.

26. Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J.W., Kim, V.N., and Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. Cell. 181(4), 914-921.e910. DOI: 10.1016/j.cell.2020.04.011.

27. Konno, Y., Kimura, I., Uriu, K., Fukushi, M., Irie, T., Koyanagi, Y., Sauter, D., Gifford, R.J., Nakagawa, S., and Sato, K. (2020). SARS-CoV-2 ORF3b Is a Potent Interferon Antagonist Whose Activity Is Increased by a Naturally Occurring Elongation Variant. Cell Reports. 32(12). DOI: 10.1016/j.celrep.2020.108185.

28. Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B., et al. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell. 182(4), 812-827.e819. DOI: 10.1016/j.cell.2020.06.043.

29. Kotenko, S.V., and Durbin, J.E. (2017). Contribution of type III interferons to antiviral immunity: location, location, location. Journal of Biological Chemistry. 292(18), 7295-7303. DOI: 10.1074/jbc.R117.777102.

30. Lauring, A.S., and Hodcroft, E.B. (2021). Genetic Variants of SARS-CoV-2—What Do They Mean? Jama. DOI: 10.1001/jama.2020.27124.

31. Lei, X., Dong, X., Ma, R., Wang, W., Xiao, X., Tian, Z., Wang, C., Wang, Y., Li, L., Ren, L., et al. (2020). Activation and evasion of type I interferon responses by SARS-CoV-2. Nat Commun. 11(1), 3810. Published online 2020/08/01 DOI: 10.1038/s41467-020-17665-9.

32. Loveday, E.K., Hain, K.S., Kochetkova, I., Hedges, J.F., Robison, A., Snyder, D.T., Brumfield, S.K., Young, M.J., Jutila, M.A., Chang, C.B., et al. (2020). Effects of inactivation method on SARS-CoV-2 virion protein and structure. medRxiv. DOI: 10.1101/2020.11.14.383026.

33. Lucas, C., Wong, P., Klein, J., Castro, T.B.R., Silva, J., Sundaram, M., Ellingson, M.K., Mao, T., Oh, J.E., Israelow, B., et al. (2020). Longitudinal analyses reveal immunological misfiring in severe COVID-19. Nature. 584(7821), 463-469. Published online 2020/07/28 DOI: 10.1038/s41586-020-2588-y.

34. Mantlo, E., Bukreyeva, N., Maruyama, J., Paessler, S., and Huang, C. (2020). Antiviral activities of type I interferons to SARS-CoV-2 infection. Antiviral Research. 179. DOI: 10.1016/j.antiviral.2020.104811.

35. Martin-Sancho, L., Lewinski, M.K., Pache, L., Stoneham, C.A., Yin, X., Pratt, D., Churas, C., Rosenthal, S.B., Liu, S., De Jesus, P.D., et al. (2020). Functional Landscape of SARS-CoV-2 Cellular Restriction. bioRxiv. Published online 2020/10/08 DOI: 10.1101/2020.09.29.319566.

36. McNab, F., Mayer-Barber, K., Sher, A., Wack, A., and O'Garra, A. (2015). Type I interferons in infectious disease. Nature Reviews Immunology. 15(2), 87-103. DOI: 10.1038/nri3787.

37. Miorin, L., Kehrer, T., Sanchez-Aparicio, M.T., Zhang, K., Cohen, P., Patel, R.S., Cupic, A., Makio, T., Mei, M., Moreno, E., et al. (2020). SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. Proceedings of the National Academy of Sciences. 117(45), 28344-28354. DOI: 10.1073/pnas.2016650117.

38. Muth, D., Corman, V.M., Roth, H., Binger, T., Dijkman, R., Gottula, L.T., Gloza-Rausch, F., Balboni, A., Battilani, M., Rihtaric, D., et al. (2018). Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. Sci Rep. 8(1), 15177. Published online 2018/10/13 DOI: 10.1038/s41598-018-33487-8.

39. Nelson, C.A., Pekosz, A., Lee, C.A., Diamond, M.S., and Fremont, D.H. (2005). Structure and Intracellular Targeting of the SARS-Coronavirus Orf7a Accessory Protein. Structure. 13(1), 75-85. DOI: 10.1016/j.str.2004.10.010.

40. Osada, N., Kohara, A., Yamaji, T., Hirayama, N., Kasai, F., Sekizuka, T., Kuroda, M., and Hanada, K. (2014). The Genome Landscape of the African Green Monkey Kidney-Derived Vero Cell Line. DNA Research. 21(6), 673-683. DOI: 10.1093/dnares/dsu029.

41. Pati, A., Padhi, S., Suvankar, S., and Panda, A.K. (2021). Minor Allele of Interferon-Induced Transmembrane Protein 3 Polymorphism (rs12252) Is Covered Against Severe Acute Respiratory Syndrome Coronavirus 2 Infection and Mortality: A Worldwide Epidemiological Investigation. The Journal of Infectious Diseases. 223(1), 175-178. DOI: 10.1093/infdis/jiaa630.

42. Pekosz, A., Gamage, A.M., Tan, K.S., Chan, W.O.Y., Liu, J., Tan, C.W., Ong, Y.K., Thong, M., Andiappan, A.K., Anderson, D.E., et al. (2020). Infection of human Nasal Epithelial Cells with SARS-CoV-2 and a 382-nt deletion isolate lacking ORF8 reveals similar viral kinetics and host transcriptional profiles. PLOS Pathogens. 16(12). DOI: 10.1371/journal.ppat.1009130.

43. Quick, J., Grubaugh, N.D., Pullan, S.T., Claro, I.M., Smith, A.D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T.F., Beutler, N.A., et al. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc. 12(6), 1261-1276. Published online 2017/05/26 DOI: 10.1038/nprot.2017.066.

44. Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., and Pybus, O.G. (2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nature Microbiology. 5(11), 1403-1407. DOI: 10.1038/s41564-020-0770-5.

45. Rambaut, A., Loman, N., Pybus, O., Barclay, W., Barrett, J., Carabelli, A., Connor, T., Peacock, T., Robertson, D.L., and Volz, E. (2020). Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. https://virological.org/t/563.

46. Rosenthal, S.H., Kagan, R.M., Gerasimova, A., Anderson, B., Grover, D., Hua, M., Liu, Y., Owen, R., and Lacbawan, F. (2020). Identification of eight SARS-CoV-2 ORF7a deletion variants in 2,726 clinical specimens. bioRxiv. DOI: 10.1101/2020.12.10.418855.

47. Schoggins, J.W. (2019). Interferon-Stimulated Genes: What Do They All Do? Annu Rev Virol. 6(1), 567-584. Published online 2019/07/10 DOI: 10.1146/annurev-virology-092818-015756.

48. Shi, G., Kenney, A.D., Kudryashova, E., Zani, A., Zhang, L., Lai, K.K., Hall-Stoodley, L., Robinson, R.T., Kudryashov, D.S., Compton, A.A., et al. (2020). Opposing activities of IFITM proteins in SARS-CoV-2 infection. The EMBO Journal. DOI: 10.15252/embj.2020106501.

49. Stoltz, M., and Klingstrom, J. (2010). Alpha/Beta Interferon (IFN-/)-Independent Induction of IFN-1 (Interleukin-29) in Response to Hantaan Virus Infection. Journal of Virology. 84(18), 9140-9148. DOI: 10.1128/ jvi.00717-10.

50. Su, Y.C.F., Anderson, D.E., Young, B.E., Linster, M., Zhu, F., Jayakumar, J., Zhuang, Y., Kalimuddin, S., Low, J.G.H., Tan, C.W., et al. (2020). Discovery and Genomic Characterization of a 382-Nucleotide Deletion in ORF7b and ORF8 during the Early Evolution of SARS-CoV-2. mBio. 11(4). Published online 2020/07/23 DOI: 10.1128/mBio.01610-20.

51. Takayama, K. (2020). In Vitro and Animal Models for SARS-CoV-2 research. Trends Pharmacol Sci. 41(8), 513-517. Published online 2020/06/20 DOI: 10.1016/j.tips.2020.05.005.

52. Tay, M.Z., Poh, C.M., Renia, L., MacAry, P.A., and Ng, L.F.P. (2020). The trinity of COVID-19: immunity, inflammation and intervention. Nat Rev Immunol. 20(6), 363-374. Published online 2020/04/30 DOI: 10.1038/s41577-020-0311-8.

53. Taylor, J.K., Coleman, C.M., Postel, S., Sisk, J.M., Bernbaum, J.G., Venkataraman, T., Sundberg, E.J., Frieman, M.B., and García-Sastre, A. (2015). Severe Acute Respiratory Syndrome Coronavirus ORF7a Inhibits Bone Marrow Stromal Antigen 2 Virion Tethering through a Novel Mechanism of Glycosylation Interference. Journal of Virology. 89(23), 11820-11833. DOI: 10.1128/jvi.02274-15.

54. Tyson, J.R., James, P., Stoddart, D., Sparks, N., Wickenhagen, A., Hall, G., Choi, J.H., Lapointe, H., Kamelian, K., Smith, A.D., et al. (2020). Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. bioRxiv. Published online 2020/09/11 DOI: 10.1101/2020.09.04.283077.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.22.21252253; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

55. van Dorp, L., Richard, D., Tan, C.C.S., Shaw, L.P., Acman, M., and Balloux, F. (2020). No evidence for increased transmissibility from recurrent mutations in SARS-CoV-2. Nature Communications. 11(1). DOI: 10.1038/s41467-020-19818-2.

56. Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R., Laydon, D.J., Dabrera, G., O'Toole, Á., et al. (2021). Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. medRxiv. DOI: 10.1101/2020.12.30.20249034.

57. Wang, C., Shan, L., Qu, S., Xue, M., Wang, K., Fu, F., Wang, L., Wang, Z., Feng, L., Xu, W., et al. (2020). The Coronavirus PEDV Evades Type III Interferon Response Through the miR-30c-5p/SOCS1 Axis. Frontiers in Microbiology. 11. DOI: 10.3389/fmicb.2020.01180.

58. Xia, H., Cao, Z., Xie, X., Zhang, X., Chen, J.Y.-C., Wang, H., Menachery, V.D., Rajsbaum, R., and Shi, P.-Y. (2020). Evasion of Type I Interferon by SARS-CoV-2. Cell Reports. 33(1). DOI: 10.1016/j.celrep.2020.108234.

59. Young, B.E., Fong, S.-W., Chan, Y.-H., Mak, T.-M., Ang, L.W., Anderson, D.E., Lee, C.Y.-P., Amrun, S.N., Lee, B., Goh, Y.S., et al. (2020). Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. The Lancet. 396(10251), 603-611. DOI: 10.1016/s0140-6736(20)31757-8.

60. Yu, G., and He, Q.Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst. 12(2), 477-479. Published online 2015/12/15 DOI: 10.1039/c5mb00663e.

61. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology. 16(5), 284-287. DOI: 10.1089/omi.2011.0118.

62. Zang, R., Case, J.B., Yutuc, E., Ma, X., Shen, S., Gomez Castro, M.F., Liu, Z., Zeng, Q., Zhao, H., Son, J., et al. (2020). Cholesterol 25-hydroxylase suppresses SARS-CoV-2 replication by blocking membrane fusion. Proceedings of the National Academy of Sciences. 117(50), 32105-32113. DOI: 10.1073/pnas.2012197117.

63. Zhang, Y., Zhang, J., Chen, Y., Luo, B., Yuan, Y., Huang, F., Yang, T., Yu, F., Liu, J., Liu, B., et al. (2020). The ORF8 Protein of SARS-CoV-2 Mediates Immune Evasion through Potently Downregulating MHC-I bioRxiv. DOI: 10.1101/2020.05.24.111823.

64. Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet. 395(10229), 1054-1062. DOI: 10.1016/s0140-6736(20)30566-3.

65. Zhou, Z., Huang, C., Zhou, Z., Huang, Z., Su, L., Kang, S., Chen, X., Chen, Q., He, S., Rong, X., et al. (2020). Structural Insight Reveals SARS-CoV-2 Orf7a as an Immunomodulating Factor for Human CD14+ Monocytes. SSRN Electronic Journal. DOI: 10.2139/ssrn.3699795.