SARS-CoV-2 genomic surveillance identifies naturally occurring truncation of ORF7a that limits immune suppression

Graphical abstract

Highlights

- ORF7a mutations are found in SARS-CoV-2 genomes isolated from around the globe
- The ORF7aΔ115 isolate displays a replication defect
- An ORF7a mutation limits viral suppression of the interferon response

Authors

Artem Nemudryi, Anna Nemudraia, Tanner Wiegand, ..., Diane Bimczok, Mark A. Jutila, Blake Wiedenheft

Correspondence

bwiedenheft@gmail.com

In brief

Nemudryi et al. use global SARS-CoV-2 phylogenomics to identify mutations that frequently occur in the C-terminal end of ORF7a accessory protein. They isolate one of these mutant viruses from a patient sample and demonstrate that ORF7a truncation affects viral mechanisms responsible for suppressing host immune response.
SARS-CoV-2 genomic surveillance identifies naturally occurring truncation of ORF7a that limits immune suppression

Artem Nemudryi,1,3,5 Anna Nemudraia,1,5 Tanner Wiegand,1,6 Joseph Nichols,1,6 Deann T. Snyder,1 Jodi F. Hedges,1 Calvin Cicha,1 Helen Lee,1 Karl K. Vanderwood,2 Diane Bimczok,1 Mark A. Jutila,1 and Blake Wiedenheft1,4,7,*
1Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717, USA
2Gallatin City-County Health Department, Bozeman, MT 59715, USA
3Twitter: @artemnemudryi
4Twitter: @WiedenheftLab
5These authors contributed equally
6These authors contributed equally
7Lead contact
*Correspondence: bwiedenheft@gmail.com
https://doi.org/10.1016/j.celrep.2021.109197

SUMMARY

Over 950,000 whole-genome sequences of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have been determined for viruses isolated from around the world. These sequences are 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 isolate one of these mutant viruses from a patient sample and use viral challenge experiments to link this isolate (ORF7a\textsuperscript{D115}) to a growth defect. ORF7a is implicated in immune modulation, and we show that the C-terminal truncation negates anti-immune activities of the protein, which results in elevated type I interferon response to the viral infection. Collectively, this work indicates that ORF7a mutations occur frequently, and that these changes affect viral mechanisms responsible for suppressing the immune 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 2.8 million deaths worldwide (https://coronavirus.jhu.edu/) (Andersen et al., 2020). Although 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 the proposed function of ORF8 is downregulating antigen presentation (Konno et al., 2020; Lei et al., 2020; Morin et al., 2020; Tan et al., 2020; Xia 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). Although 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\textsuperscript{+} monocytes and triggers expression of pro-inflammatory cytokines, including interleukin-6 (IL-6) and tumor necrosis factor alpha (TNF-\alpha) (Zhou et al., 2021). 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.

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 suggest 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 truncations of ORF7a

Genome sequencing has been used to track the rise and spread of new SARS-CoV-2 lineages over the course of the pandemic...
As part of this effort, we sequenced SARS-CoV-2 ge nomes isolated from patients in Bozeman, Montana (Figure 1A; Table S1). In total, we determined 55 whole-genome sequences of SARS-CoV-2 viruses isolated from patients between April and July 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 55 whole-genome sequences isolated from patients in Bozeman to 4,871 genomes subsampled from GISAID. Subsampling was performed using Augur utility (https://nextstrain.org) by selecting 10 genomes per country per month since the start of the pandemic. Outer ring shows SARS-CoV-2 lineages assigned to genome sequences (Rambaut et al., 2020a). Major lineages include A (pink) that is associated with initial outbreak in China and B (blue) that emerged later in Europe. Minor lineages (i.e., C–N) are offshoots of lineage B. Red branches identify truncated ORF7a variants (n = 205) detected in the global data and merged into the alignment. The red dot highlights 7 of the 55 ORF7a variants that were isolated in Bozeman between April and July (2020). White dots highlight 48 viral genomes isolated in Bozeman that have wild-type ORF7a sequences.

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 55 Bozeman samples and 4,871 genomes subsampled from GISAID. Subsampling was performed using Augur utility (https://nextstrain.org) by selecting 10 genomes per country per month since the start of the pandemic. Outer ring shows SARS-CoV-2 lineages assigned to genome sequences (Rambaut et al., 2020a). Major lineages include A (pink) that is associated with initial outbreak in China and B (blue) that emerged later in Europe. Minor lineages (i.e., C–N) are offshoots of lineage B. Red branches identify truncated ORF7a variants (n = 205) detected in the global data and merged into the alignment. The red dot highlights 7 of the 55 ORF7a variants that were isolated in Bozeman between April and July (2020). White dots highlight 48 viral genomes isolated in Bozeman that have wild-type ORF7a sequences.

(C) Distribution of different mutations that occur along the ORF7a coding sequence.

(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 55 whole-genome sequences of SARS-CoV-2 viruses isolated from patients between April and July 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 55 whole-genome sequences isolated from patients in Bozeman to 4,235 genomes subsampled from 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 55 SARS-CoV-2 genomes determined from patients in Bozeman, only 1 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 54 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 et al., 2020a) (Figure 1B). The genetic variability in SARS-CoV-2 circulating in Bozeman (April to July 2020) represents 14 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 recurring (7 out of 55 genomes) 115-nt deletion in the gene-encoding accessory protein ORF7a (27,549–27,644 nt). The ORF7a<sup>Δ115</sup>
mutation was found in patient swabs collected over a period of \(\sim1.5\) months (Table S1). RT-PCR and Sanger sequencing verified that all these seven ORF7a\(^{1115}\) variants are bona fide mutations and not a sequencing artifact (Figures S1B and S1C). In addition to the ORF7a\(^{1115}\) deletion, these genomes also share 10 single-nucleotide variants (SNVs) compared with the Wuhan-Hu-1 reference genome. Seven of the 10 SNVs are found frequently worldwide and are a signature of the B.1.1 lineage of SARS-CoV-2 (Rambaut et al., 2020a). The remaining three 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\(^{1115}\) 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.

To determine if the ORF7a\(^{1115}\) genotype is unique to Boezeman, 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 genome. 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 sequences 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 (Figures 1A and 1B). A portion (36.5%) of these truncated ORF7a variants appeared in two or more patient samples (116 maximum [max]). Although 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 S1B).

**ORF7a truncation has a loss-of-function effect**

ORF7a is a type I transmembrane (TM) protein with an N-terminal immunoglobulin-like (Ig-like) ectodomain, stalk, TM domain, and short cytosolic tail (Figure 2A). The two \(\beta\) 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 diislysin (KRKTE) endoplasmic reticulum (ER) retrieval signal (ERRS) that mediates protein trafficking to the ER-Golgi intermediate compartment (ERGIC) (Nelson et al., 2005).

The \(\Delta115\) nt mutation in ORF7a introduces a premature stop codon that eliminates \(\delta5, \delta6, \delta7\), two of the cystines that form disulfides (Cys58 and Cys67), the TM, and the cytosolic tail (Figures 2A and 2B). This truncation is expected to destabilize the protein structure and significantly impact 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 wild-type and truncated ORF7a proteins in HEK293T-HACE2 cells. The wild-type ORF7a accumulated in the perinuclear region of the cell, which is consistent with previously reported ERGIC localization (Figure 2D) (Martin-Sancho et al., 2020; Nelson et al., 2003). 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).

Extent of the truncation and the defect in intracellular targeting suggests possible loss of ORF7a function. Along with several SARS-CoV-2 accessory proteins, ORF7a has been implicated in suppression of host IFN response to the infection (Xia et al., 2020). To test the effect of the truncated ORF7a protein on IFN signaling, we generated reporter HEK293T cells with integrated luciferase reporter gene controlled by the interferon-stimulated response element (ISRE). In these cells, activation of interferon response drives reporter expression, which is quantified using a luciferase assay. To test immune suppression, we over-expressed the wild-type and the deletion mutant of ORF7a (ORF7a\(^{1115}\)) from plasmids and the stimulated cells with IFN-\(\alpha2b\). The ORF7a WT protein suppressed IFN response by 34% (\(p = 0.0074\)) compared with cells transfected with control plasmid, which is consistent with recent results by Xia et al. (2020). In contrast, the ORF7a\(^{1115}\) protein failed to suppress IFN response and showed no significant difference to a control (mCherry) plasmid (Figure 2F). These results indicate that the \(\Delta115\) mutation in ORF7a has a loss-of-function effect that negates antagonism of IFN signaling.

**ORF7a mutation has no collateral effect on ORF7b**

The ORF7a gene overlaps with the downstream ORF7b gene (Figure S2). To determine if the \(\Delta115\) 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 (RNA-seq) of SARS-CoV-2 overlap with the \(\Delta115\) 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 2G). In this assay, we infected 293T-HACE2 cells (MOI = 0.05) with ORF7a\(^{WT}\) or ORF7a\(^{1115}\) viral strains and extracted total RNA from cells at 24 h postinfection (hpi). Both viruses produced specific RT-PCR products corresponding to ORF7a and ORF7b sgRNAs. Finally, to verify that the ORF7a\(^{1115}\) variant does not impact ORF7b translation, we probed cell lysates 24 hpi with an anti-ORF7b antibody (Figure 2H). We detected ORF7b protein in both viral strains with band intensities that suggest similar expression levels. Collectively, these data indicate that the ORF7a\(^{1115}\) mutation has no collateral effect on the adjacent gene.

**ORF7a\(^{1115}\) isolate displays a replication defect**

To investigate the phenotype of SARS-CoV-2 \(\Delta115\) variant, we infected 293T-HACE2 and Vero E6 cells with ORF7a\(^{WT}\) and ORF7a\(^{1115}\) viruses at an MOI of 0.05 and measured viral RNA...
replication using qRT-PCR (Figures 3A and 3B). 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 et al., 2020a). Over the course of 120 h, ORF7aWT 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\(^{Δ115}\) 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 (\(p = 0.0135\)) and 24 hpi in 293T-hACE2 cells (\(p = 6.2 \times 10^{-5}\)).

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 qRT-PCR. Compared with wild type, ORF7a\(^{Δ115}\) RNA level was reduced 2.1-fold inside the host at 6 hpi (\(p = 0.002\)) and 4.3-fold at 24 hpi (\(p = 0.02\)) (Figure 3C).
RNA in the host cell differs from the corresponding decrease in the supernatant RNA at 24 hpi (4.3-fold versus 16.8-fold, respectively; Figures 3B and 3C). This disparity suggests that observed growth defect in the SARS-CoV-2 Δ115 variant might result from defects in genomic RNA replication and transcription, as well as in viral egress.

Figure 3. The ORF7aΔ115 mutation results in replication and immune suppression defects
(A and B) VeroE6 (A) or HEK293T-hACE2 (B) cells were infected with ORF7aWT or ORF7aΔ115 SARS-CoV-2 strain at MOI = 0.05. Viral RNA was measured in supernatant using qRT-PCR at different time points after infection. Measured RNA levels were normalized (using ΔCt method) to RNA levels at 0.5 hpi.
(C) Total RNA was extracted from infected HEK293T-hACE2 cells, and viral RNA was measured. The ΔΔCt method was used to normalize viral RNA levels to host RNA (ACTB) and 0.5 hpi time point. Data in (A)–(C) are presented as mean ± SD of three biological replicates. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, or nsp > 0.05 (no significant difference).
(D) Volcano plot showing IFN-I response in 293T-hACE2 cells infected with ORF7aWT SARS-CoV-2 strain at MOI = 0.05. Expression of IFN-I response genes was studied 24 hpi using qRT-PCR array targeting 91 human transcripts (88 targets and 3 references). Experiment was performed in three biological replicates. Dashed lines show regulation (≥2-fold) and statistical significance (p < 0.05) thresholds. Each dot represents mean (n = 3) normalized expression of a single gene relative to non-infected host. Genes that passed the threshold are labeled.
(E) Volcano plot showing IFN-I response in ORF7aΔ115 versus ORF7aWT infection.
(F) Data shown in (D) and Figure S3A were plotted as a heatmap. Genes were classified into three groups. Group #1 genes are oppositely regulated between the two viral strains. Group #2 genes are upregulated by both ORF7aWT and ORF7aΔ115. Group #3 genes are downregulated in both. Genes that have statistically significant difference in expression between two viral strains are marked with asterisk. Gray shows genes with no detectable expression over 40 cycles of qRT-PCR.
ORF7a\(^{\Delta115}\) virus induces elevated IFN-I response in vitro

Suppressed or delayed IFN responses characterize SARS-CoV-2 infections (Lei et al., 2020). The C-terminal truncation of ORF7a limits suppression of the type I IFN signaling (Figure 2F). Therefore, we hypothesized that ORF7a\(^{\Delta115}\) virus isolated from the community is deficient in antagonizing IFN response.

To test this hypothesis, we measured expression of type I IFN response genes upon infection with ORF7a\(^{WT}\) and ORF7a\(^{\Delta115}\) strains using a qRT-PCR array targeting 91 human transcripts (Table S2). The array included three 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 HEK293T-hACE2 with SARS-CoV-2 isolates (Banerjee et al., 2020; Blanco-Melo et al., 2020). Out of 88 type I IFN response genes included in the qRT-PCR array, 42 were differentially expressed (≥2-fold change, p < 0.05) in ORF7a\(^{WT}\)-infected cells and 55 in ORF7a\(^{\Delta115}\) infection, as compared with the non-infected control (Figure 3D; Figure S3B). Both viruses stimulated expression of multiple type I interferon genes, antiviral ISGs, ILs, and proinflammatory cytokines (Table S2).

To understand the effect of ORF7a\(^{\Delta115}\) on the host type I IFN response, we compared expression of genes targeted with qRT-PCR array between the two viruses. In general, the ORF7a\(^{\Delta115}\) variant induces elevated type I IFN response to the infection (Figures S3C and S3D). The median expression fold change of 88 genes targeted in the qRT-PCR array was 1.72 in ORF7a\(^{WT}\) and 2.51 in ORF7a\(^{\Delta115}\) virus (p = 1.4E–12; Figure S3C). Analysis of genes differentially expressed between the two viral infections results in an asymmetric volcano plot (Figure 3E). A subset of ISGs had expression levels that are significantly different (p < 0.05) between the two viral infections (Figures 3E and 3F; 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., 2021; Zang et al., 2020). Additionally, polymorphisms in IFITM3, MX1, and TLR7 are linked to the severity of COVID-19 (Andolfi et al., 2020; Zhang et al., 2020a; Pati et al., 2021). Upregulation of ISGs indicates that the ORF7a\(^{\Delta115}\) virus is inept in antagonizing the IFN-I response to the infection, which is consistent with results from our ISRE reporter assay (Figure 2F).

**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 (Addetta 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 isolated a virus with a deletion mutation in ORF7a (ORF7a\(^{\Delta115}\)) and show that this genotype results in an in vitro growth defect. This growth defect is associated with elevated IFN response to SARS-CoV-2.

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 ORF7b) 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). Using the ISRE reporter system, we show that C-terminal truncation negates the ability of ORF7a to suppress IFN signaling (Figure 2). The ORF7a\(^{\Delta115}\) variant shows limited replication in HEK293T-hACE2, as well as in Vero E6, the latter of which does not express type I IFN genes (Emeny and Morgan, 1979; Osada et al., 2014). Although 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 Klingström, 2010; Wang et al., 2020). ISGs exert anti-viral activities that combat the infection (Schoggins, 2019). We hypothesize that the replication defect of the ORF7a\(^{\Delta115}\) strain in Vero E6 and 293T-hACE2 cells results from an IFN-dependent response that the virus fails to suppress.

Several groups have replaced ORF7a with a reporter gene (Hou et al., 2020; Thi Nhu Thao et al., 2020; Xie et al., 2020; Thi Nhu Thao et al., 2020) show that complete deletion of ORF7a reduces replication of the synthetic virus, which agrees well with growth defect in the ORF7a\(^{\Delta115}\) strain. However, two other studies have not observed this effect. We believe that our results with a naturally occurring ORF7a deletion provide valuable insight to this discussion. Finally, it should be noted that different mutations (i.e., complete deletion versus partial) might have a different outcome for the viral phenotype.

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 (Zhang et al., 2020b; Tay et al., 2020; Zhou et al., 2020). Here, we detected enhanced IFN signaling in cells infected with a naturally occurring ORF7a\(^{\Delta115}\) variant of SARS-CoV-2.

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.

Deletions in ORF7a, ORF7b, and ORF8 have previously been reported (Addetta et al., 2020; Holland et al., 2020; Rosenthal et al., 2020; Su et al., 2020). Similar deletions emerged in Middle East Respiratory Syndrome (MERS)- and SARS-CoV, which were linked to viral attenuation (Chinese SARS Molecular Epidemiology Consortium, 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). Although this mutation affects clinical...
features of COVID-19, it has no evident effect on viral replication or host transcriptional responses in human cell culture (Gamage et al., 2020).

Several studies illustrate how persistent SARS-CoV-2 infections in immunocompromised hosts associate with accelerated viral evolution (Avanzato et al., 2020; Choi et al., 2020). In these hosts there is less selective pressure on viral anti-immune proteins, which might permit loss-of-function mutations in accessory genes. The ORF7aΔ115 variant that we isolated was present in only seven genomes over a course of ~1.5 months in a local community and then disappeared. We hypothesize that deletions in accessory genes, such as ORF7aΔ115, might emerge in immunocompromised patients but do not persist in the immunocompetent population. Similarly, a deletion mutation in ORF8 (i.e., Δ382), which was associated with milder COVID-19 disease, disappeared soon after it was discovered (Young et al., 2020). However, founder effect or mutations that co-occur with other mutations that increase transmissibility (e.g., spike mutations) can result in fixation of an otherwise attenuating mutation. A new more transmissible lineage of SARS-CoV-2 (i.e., B.1.1.7) has recently emerged (Davies et al., 2021; Rambaut et al., 2020b). Aside from changes in the Spike protein, this lineage also includes a premature stop codon in the accessory ORF8 that truncates most of the protein (Volz et al., 2021). This lineage might illustrate the second scenario for fixing mutations in accessory genes.

**Limitations of study**

The viral replication defect and cellular immune response assay presented here were performed using HEK293T-hACE2 and Vero E6 cells that are not the natural targets of SARS-CoV-2. However, both cell lines are permissive for SARS-CoV-2 infection and are used extensively to study efficacy of antiviral compounds (Riva et al., 2020), virion assembly (Klein et al., 2020), SARS-CoV-2 entry (Hoffmann et al., 2020), replication (Thi Nhu Thao et al., 2020), and anti-immune activities of SARS-CoV-2 proteins (Xia et al., 2020). Although we acknowledge that our conclusions are limited to Vero E6 and 293T-hACE2 cells, we anticipate that the results presented here will provide important context for clinical comparisons, similar to those recently published for ORF8 (Young et al., 2020).

Studies of naturally occurring SARS-CoV-2 variants have limitations, including unavailability of isogenic controls. Here, we compared two viral isolates, ORF7aWT and ORF7aΔ115, that in addition to Δ115 differ at several other nucleotide positions (Figure S3A). Although ISRE reporter assay unambiguously links the immunosuppression defect to the truncation of ORF7a, it is not possible to rule out a role for these mutations on the viral life cycle.

**STAR★METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Cell cultures
  - Plasmids
  - Antibodies
  - Human clinical sample collection and preparation
  - Production and titration of coronavirus stocks
  - Symptom onset data and clinical test results

- **METHOD DETAILS**
  - Quantitative reverse transcription PCR (qRT-PCR)
  - RT-PCR and SARS-CoV-2 genome sequencing
  - SARS-CoV-2 genome assembly
  - RT-PCR and Sanger sequencing
  - Phylogenetic and ORF7a mutational analysis
  - Western blot
  - Immunocytochemistry
  - IFN inhibition assay
  - SARS-CoV-2 replication assays
  - Type I interferon response assay

- **QUANTIFICATION AND STATISTICAL ANALYSIS**

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109197.

**ACKNOWLEDGMENTS**

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. Fienniken, K. Daughenbaugh, and other members of the COVID task force at MSU for assistance establishing the COVID testing center. We thank the Lefcot 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 Database and contributing laboratories (Table S4). The phylogenetic analysis in this paper would not have been possible without their willingness to share data. The graphical abstract was created with BioRender.

**AUTHOR CONTRIBUTIONS**

Conceptualization, B.W., A. Nemudryi, and A. Nemudraia; methodology, B.W., A. Nemudraia, A. Nemudryi, D.T.S., and J.F.H.; sample acquisition, D.B. and B.W.; investigation & data collection, A. Nemudraia, A. Nemudryi, D.T.S., J.F.H., J.N., H.L., and K.K.V.; genomics and bioinformatics analysis, T.W., A. Nemudryi, and 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., and M.A.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.

Received: February 3, 2021
Revised: April 4, 2021
Accepted: May 10, 2021
Published: June 1, 2021
REFERENCES

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 isolates. J. Clin. Virol. 129, 104523.

Andersen, K.G., Rambaut, A., Lipkin, W.J., Holmes, E.C., and Garry, R.F. (2020). The proximal origin of SARS-CoV-2. Nat. Med. 26, 450–452.

Andolfi, I., Russo, R., Lasorsa, A.V., Cantalupo, S., Rosato, B.E., Bonfiglio, F., Frisso, G., Pasquale, A., Cassese, G.M., Serville, G., et al. (2020). Common variants at 21q22.3 locus influence MX1 gene expression and susceptibility to severe COVID-19. medRxiv. https://doi.org/10.1101/2020.12.18.20248470.

Arunachalam, P.S., Wimmers, F., Mok, C.K.P., Perera, R.A.P.M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., et al. (2020). Systems biological assayment of immunity to mild versus severe COVID-19 infection in humans. Science 369, 1210–1220.

Avanzato, V., 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 178, 1901–1912.e9.

Banerjee, A., El-Sayes, N., Budovsky, P., Richard, D., Aegius, 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. https://doi.org/10.1101.2020.06.18.158154.

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.; Seattle Flu Study Investigators (2020). Cryptic transmission of SARS-CoV-2 in Washing-Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. Cell 178, 1901–1912.e9.

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. https://doi.org/10.1101.2020.06.18.158154.

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.; Seattle Flu Study Investigators (2020). Cryptic transmission of SARS-CoV-2 in Washing-Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. Cell 178, 1901–1912.e9.

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. https://doi.org/10.1101.2020.06.18.158154.

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.; Seattle Flu Study Investigators (2020). Cryptic transmission of SARS-CoV-2 in Washing-Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. Cell 178, 1901–1912.e9.

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. https://doi.org/10.1101.2020.06.18.158154.

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.; Seattle Flu Study Investigators (2020). Cryptic transmission of SARS-CoV-2 in Washing-Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. Cell 178, 1901–1912.e9.

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. https://doi.org/10.1101.2020.06.18.158154.
## STAR METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Biological samples** | | |
| Nasopharyngeal swabs | Bozeman Health Deaconess Hospital, MT, USA | N/A |
| **Bacterial and virus strains** | | |
| SARS-CoV-2 Isolate USA-WA1/2020 | BEI Resources | NR-52281 |
| **Experimental models: cell lines** | | |
| HEK293T-hACE2 cells | BEI Resources | NR-52511 |
| Vero E6 | ATCC | Cat#CRL-1586 |
| **Recombinant DNA** | | |
| pEGFP-N1 | Clontech | 6085-1 |
| pLV-mCherry | Addgene | 36084 |
| pORF7a-CoV2 | This study | N/A |
| pORF7a \(\Delta 115\)-Cov2 | This study | N/A |
| pGreenFire1-ISRE | SBI | TR016PA-1 |
| **Antibodies** | | |
| Rabbit anti-ACTB | ABclonal | Cat. AC026, RRID:AB_2768234 |
| Mouse anti-FLAG | ThermoFisher Scientific | Cat: MA1-91878-1MG, RRID:AB_2537619 |
| Sheep anti-ORF7b | MRC I PPU, College of Life Sciences, University of Dundee | N/A |
| Goat anti-mouse IgG peroxidase-conjugated | Jackson ImmunoResearch | Cat: 115-035-003, RRID:AB_10015289 |
| Goat anti-rabbit IgG peroxidase-conjugated | Jackson ImmunoResearch | Cat: 111-035-003, RRID:AB_2313567 |
| Donkey Anti-Sheep IgG peroxidase-conjugated | Jackson ImmunoResearch | Cat: 713-035-147, RRID:AB_2340710 |
| Mouse anti-FLAG | Invitrogen | Cat: MA1-91878-1MG, RRID:AB_2537619 |
| Goat anti-mouse-AlexaFluor-488 | Invitrogen | Cat: A11001, RRID:AB_2534069 |
| **Chemicals, peptides, and recombinant proteins** | | |
| QIAamp viral RNA mini kit | QIAGEN | 52904 |
| 2019-nCoV RUO kit | IDT | 10006713 |
| Positive template control (PTC) plasmid | IDT | 10006625 |
| TaqPath 1-Step RT-qPCR master mix | Thermo Fisher Scientific | A15300 |
| SuperScript IV reverse transcriptase | Thermo Fisher Scientific | 180900010 |
| R9.4.1 flow cells | Nanopore Technologies | FLO-MIN106 |
| AMX, LNB, SFB, EB and SQB | Nanopore Technologies | SQK-LSK109 |
| Flow cell priming kit | Nanopore Technologies | EXP-FLP002 |
| NEBNext Ultra II end-prep | New England Biolabs | E7546S |
| NEBNext quick ligation module | New England Biolabs | E6056S |
| Native barcoding expansion kits | Nanopore Technologies | EXP-NBD104, EXP-NBD114 |
| Q5 high-fidelity DNA polymerase | New England Biolabs | M0491S |
| DNA clean & concentrator kit | Zymo Research | D4005 |
| Lipofectamine 3000 | Invitrogen | L3000-015 |
| Pierce ECL western blotting substrate | ThermoFisher Scientific | #32106 |
| X-Ray film | Santa Cruz Biotech | sc-201696 |
| IFN-\(\alpha\)2b | Invivogen | rcyc-hifna2b |
| Luciferase assay system | Promega | E1500 |

(Continued on next page)
**RESOURCE AVAILABILITY**

**Lead contact**
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Blake Wiedenheft (bwiedenheft@gmail.com).

**Materials availability**
Plasmids generated in this study are available upon request from the lead contact.

**Data and code availability**
The complete SARS-CoV-2 genome sequences are deposited to GISAID EpiCoV and GenBank databases. GISAID accession IDs and URLs to GenBank records are provided in Table S1. Raw data for IFN response assay presented in Figures 3 and S3 is available in Table S2.

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Type I interferon response (SAB Target List) | Bio-Rad | N/A |
| H96 qPCR array | Bio-Rad | 1725270 |
| SsoAdvanced universal SYBR green supermix | Bio-Rad | 1725270 |
| Deposited data | | |
| URLs for SARS-CoV-2 genome sequences are listed in the Table S1 | GenBank | N/A |
| Deposited data | | |
| Oligonucleotides | IDT | N/A |
| Software and algorithms | | |
| SDS software v1.4 | Applied Biosystems | 4379633 |
| RStudio v1.2.1335 | The R project | https://www.r-project.org/ |
| artic-ncov2019 | ARTIC network | https://artic.network/ncov-2019 |
| MinKNOW software | Oxford Nanopore Technologies | https://community.nanoporetech.com/protocols/experiment-companion-minknow/ v/mke_1013_v1_revbm_11apr2016 |
| PrimePCR analysis software | Bio-Rad | https://www.bio-rad.com/en-us/category/qpcr-analysis-software?id=42a656e3-3ad7-43e9-bb8d-602737f1de67a |
| CD-HIT | GitHub | https://github.com/weizhongli/cdhit |
| BioStrings | Bioconductor | https://bioconductor.org/packages/release/bioc/html/Biostrings.html |
| ggplot2 | CRAN | https://cran.r-project.org/web/packages/ggplot2/index.html |
| Augur | Nextstrain | https://github.com/nextstrain/augur |
| Nextclade | Nextstrain | https://clades.nextstrain.org/ |
| pangolin | Github | https://github.com/cov-lineages/pangolin |
| Trimal | Github | https://github.com/scapella/trimal |
| IQ-Tree | IQ-Tree | http://www.iqtree.org/ |
| TreeTime | Github | https://github.com/jkanev/treetime |
| APE v5.3 | CRAN | https://cran.r-project.org/web/packages/ape/index.html |
| tidyquant | CRAN | https://cran.r-project.org/web/packages/tidyquant/index.html |
| stats | CRAN | https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html |
| UGENE | Unipro | http://ugene.net/ |
EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell cultures
HEK293T-hACE2 cells (BEI, NR-52511) were generously provided by Dr. Seth Pincus (Montana State University) and maintained at 37°C 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 37°C 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 ORF7aD115 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. pLV-mCherry was a gift from Pantelis Tsoulfas (Addgene plasmid # 36084; http://www.addgene.org/36084/; RRID:Addgene_36084).

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. Patients tested for SARS-CoV-2 included both in-patients and out-patients. The latter included individuals who developed symptoms and sought medical care and asymptomatic (at the time of testing) individuals who were exposed to known COVID-19 case and therefore were tested. 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. Information about age and sex of the patients is provided in Table S1.

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., 2021). 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 stock’s 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 0000165) 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**

**Quantitative reverse transcription PCR (qRT-PCR)**

qRT-PCR 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 qRT-PCR 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 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: 25°C for 2 min, 50°C for 15 min, 95°C for 2 min followed by 45 cycles of 95°C for 3 s and 55°C 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). 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., 2017; Wu et al., 2020). 20X coverage was used as a threshold both for consensus and variant calling. Nucleotide positions with less than 20X coverage were masked with Ns in the final consensus sequences. Consensus sequences were uploaded to GISAID (https://www.gisaid.org/), accession IDs are provided in Table S1.

**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., 2017; Wu et al., 2020). 20X coverage was used as a threshold both for consensus and variant calling. Nucleotide positions with less than 20X coverage were masked with Ns in the final consensus sequences. Consensus sequences were uploaded to GISAID (https://www.gisaid.org/), accession IDs are provided in Table S1.

**RT-PCR and Sanger sequencing**

To verify the ORF7a a115 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: 98°C for 2 min, 35 cycles of 98°C for 30 cycles of 98°C for 3 s and 55°C 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.

**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) (Bedford et al., 2020; Lemieux et al., 2021). 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. 55 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. Clades were assigned to SARS-CoV-2 genomes using Nextclade and pangolin (https://cov-lineages.org/) utilities.

### Western blot

293T-hACE2 cells were transfected using Lipofectamine 3000 (Invitrogen, L3000-015) with plasmids encoding for ORF7aWT, ORF7aΔ115 or control pEGFP-N1 plasmid. To detect ORF7b expression cells were infected with ORF7aWT or ORF7aΔ115 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 4°C for 30 min. The lysates were clarified by centrifugation (10'000 g, 20 min) and stored at −80°C. For western blot lysates were mixed with 6xLaemmli SDS-PAGE buffer and heated at 98°C 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

HEK293T-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 ORF7aΔ115, 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. Blocking, cells were stained with mouse monoclonal anti-Flag antibody (Invitrogen, MA1-91878) (1:200 in 1% BSA in PBST) overnight at 4°C. 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.

### IFN inhibition assay

To generate ISRE reporter system, HEK293T cells were transduced with lentiviruses carrying pGreenFire1-ISRE (Cat. TR016PA-1, on Superfrost Plus microscope slides (#22-037-246)). Immunostained cells were imaged using a Leica SP8 confocal microscope.

### SARS-CoV-2 replication assays

HEK293T-hACE2 and Vero E6 cells were seeded in the 48-well plate and infected with ORF7aWT or ORF7aΔ115 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 μl) 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 qRT-PCR with CDC N1 and N2 primers and probes as described above. Relative quantification was used (ΔΔCT method) to calculate viral replication versus 0.5 hpi time point as 2⁻ΔΔCT. 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: 74104) with on-column DNase digestion step (QIAGEN, 79254) according to the manufacturers protocol. Viral RNA was quantified using qRT-PCR 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

HEK293T-hACE2 were seeded in the 48-well plate and infected with ORF7aWT or ORF7aΔ115 SARS-CoV-2 strains at MOI = 0.05 in 50 ul 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 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: 95°C for 2 min, 40 cycles of 95°C for 5 s and 60°C 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. qRT-PCR data is shown as mean of three biological replicates (each with three technical replicates) ± standard deviation (SD). Data in ISRE reporter assay was analyzed with one-way ANOVA with Tukey’s post hoc pairwise comparisons. 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 in figures: * p < 0.05, ** p < 0.01, *** p < 0.001 or ns (no significant difference, p > 0.05).