De novo emergence of a remdesivir resistance mutation during treatment of persistent SARS-CoV-2 infection in an immunocompromised patient: A case report

Shiv Gandhi*[1†], Jonathan Klein*[2], Alexander Robertson[3]*, Mario A. Peña-Hernández*[2], Michelle J Lin[4], Pavitra Roychoudhury[4], Peiwen Lu[2], John Fournier[1], David Ferguson[5], Shah A. Mohamed Bakhash[4], M. Catherine Muenker[3], Ariktha Srivathsan[3], Elsio A. Wunder Jr[3], Nicholas Kerantzas[6], Wenshuai Wang[7], Anna Pyle[7,8,10], Craig B. Wilen[2,6], Onyema Ogbuagu[1], Alexander L. Greninger[4,9], Akiko Iwasaki[2,3,6,10], Wade L. Schulz[5,6], and Albert I. Ko[1,3†]

[1]Department of Medicine, Section of Infectious Diseases, Yale University School of Medicine, New Haven, CT, USA

[2]Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA

[3]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

[4]Department of Laboratory Medicine, University of Washington School of Medicine, Seattle, Washington, USA

[5]Center for Outcomes Research and Evaluation, Yale New Haven Hospital, New Haven, CT, USA

[6]Department of Laboratory Medicine, Yale School of Medicine, New Haven, CT, USA

[7]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT, USA

[8]Department of Chemistry, Yale University, New Haven, CT, USA

[9]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

[10]Howard Hughes Medical Institute, Chevy Chase, MD, USA.

*These authors contributed equally

Corresponding Authors[†]: S.G. (shiv.gandhi@yale.edu), A.I.K. (albert.ko@yale.edu)

Abstract

SARS-CoV-2 remdesivir resistance mutations have been generated *in vitro* but have not been reported in patients receiving treatment with the antiviral agent. We present a case of an immunocompromised patient with acquired B-cell deficiency who developed an indolent, protracted course of SARS-CoV-2 infection. Remdesivir therapy alleviated symptoms and produced a transient virologic response, but her course was complicated by recrudescence of high-grade viral shedding. Whole genome sequencing identified a mutation, E802D, in the nsp12 RNA-dependent RNA polymerase which was not present in pre-treatment specimens. *In vitro* experiments demonstrated that the mutation conferred a ~6-fold increase in remdesivir IC50 but resulted in a fitness cost in the absence of remdesivir. Sustained clinical and virologic response was achieved after treatment with casirivimab-imdevimab. Although the fitness cost observed *in vitro* may limit the risk posed by E802D, this case illustrates the importance of monitoring for remdesivir resistance and the potential benefit of combinatorial therapies in immunocompromised patients with SARS-CoV-2 infection.

Text

Remdesivir (RDV) is a nucleotide prodrug of an adenosine analog, which interferes with the SARS-CoV-2 RNA polymerase when converted to its active form[1]. Although there is a lack of consensus amongst guidelines on its use[2,3], RDV has been widely used in the treatment of patients hospitalized with COVID-19. Furthermore, RDV has been used in the treatment of protracted SARS-CoV-2 infection in immunocompromised patients, in which setting, case reports have found that the antiviral drug provided transient clinical improvement [4,5]. RDV resistance mutations have been generated *in vitro*[6] but as of yet, have not been identified in patients in association with treatment. Here we report a case of an immunocompromised patient

with persistent SARS-CoV-2 infection from whom a RDV resistance mutation was identified during recrudescence of viral shedding following treatment with the antiviral agent.

A woman in her 70s received a course of rituximab and bendamustine for the treatment of Stage IV Non-Hodgkin's lymphoma (NHL) which was completed in March 2019 and complicated by persistent lymphocytopenia and hypogammaglobinemia. The NHL was in remission when she developed a SARS-CoV-2 infection in May 2020, which presented as the acute onset of fever, anosmia, cough and rhinorrhea (Figure 1a) and was confirmed by reverse transcriptase-polymerase chain reaction (RT-PCR) on day 0 (Figure 1b). The patient was hospitalized twice in the following two months for persistent fever and new onset neutropenia and anemia (Figure 1c), during which she had low SARS-CoV-2 RT-PCR cycle threshold (Ct) values from longitudinal nasopharyngeal specimens (Ct 20.5 and 17.1 on days 17 and 36, respectively, Figure 1b), undetectable serum anti-SARS-CoV-2 IgG antibody (Abbott), and elevated serum inflammatory markers (Figure 1c). Chest computed tomography (CT) revealed bilateral ground glass opacities (day 38, Figure 1d), but she did not experience dyspnea, hypoxemia or recurrence of respiratory symptoms during her hospitalizations and subsequent course of illness. The workup for other sources of fever, which included a bone marrow biopsy, was unrevealing. $CD19^{+}$ B-cells were not identified in the bone marrow biopsy or by flow-cytometry of peripheral blood mononuclear cells (PBMC). The patient was initiated on treatment with filgrastim 1-2 times per week and monthly intravenous immunoglobulin (IVIG) to manage her neutropenia and hypogammaglobulinemia, respectively, and both were continued after hospital discharge on day 51.

The patient continued to have daily fevers, except for a 30-day period of defervescence (days 103-132), refractory neutropenia, anemia, and positive RT-PCR test results with Ct values <25

when she was hospitalized at our institution on day 145. Her chest CT on day 140 was notable for increased opacities compared with prior examinations (Figure 1d). The patient initiated a 10-day course of RDV (days 148-157), which precipitated a resolution of her fever (day 149, Figure 1a), normalization of CRP (day 156, Figure 1c) and improvement in the opacities on chest CT (day 162, Figure 1d). An initial virologic response to therapy was observed, with an increase in RT-PCR Ct values and decreasing numbers of plaque forming units (PFU) in viral cultures and lower proportions of SARS-CoV-2 subgenomic RNA (sgRNA), measured in amplicon sequencing data, in her respiratory tract specimens by day 152 (Figure 2a).

However, a recrudescence of viral shedding occurred during and after RDV therapy, as evidenced by decreasing RT-PCR Ct values which reached 18 on day 160 and increasing PFU and sgRNA in respiratory tract specimens on day 156 (Figure 2a). The patient received an 8g infusion of casirivimab-imdevimab[7,8] on day 163 after expanded access for the use of the monoclonal antibody (mAb) therapy was approved. A rapid and sustained virologic response was observed with undetectable cultured virus, low sgRNA and negative NP RNA in specimens by days 164, 166 and 217, respectively. The patient's anosmia resolved roughly 17 days after administration of casirivimab-imdevimab (day 180. Figure 1a). During the following five-month period (days 160-292), the patient did not have recurrence of COVID-19 related symptoms and had a chest CT which showed minimal opacities (day 170, Figure 1d), and had high anti-SARS-CoV-2 S1 protein IgG titers (Figure 2a). Peripheral blood neutrophil counts, hemoglobin levels and serum inflammatory markers normalized during convalescence (Figure 1c).

To identify SARS-CoV-2 mutations that arose during the course of illness, we longitudinally sampled multiple sources of the patient's tissues and secretions (Supplemental Table 1) and performed WGS on an Illumina NextSeq platform using the Swift SARS-CoV-2 multiplex

amplicon sequencing panel [9]. Viral N1 or N2 RNA was detected in 19 of 21 nasopharyngeal, 10 of 10 saliva, 10 of 10 stool and 0 of 10 whole blood specimens. Among the 39 samples with detectable N1 or N2 RNA, we assembled 27 whole SARS-CoV-2 genomes from nasopharyngeal (12), saliva (9) and stool (6) specimens. One of the nasopharyngeal specimens was obtained from the initial phase of illness (day 36, Figure 2b). Phylogenetic analysis found that sequenced genomes belonged to a single, non-intermixed, lineage (Pango B1) within Nextstrain clade 20C (Figure S1), indicating that viral genomes identified during the course of illness were derived from intra-host diversification following infection with a single strain.

Analysis of assembled viral genomes identified a mutation, E802D, in nsp12, whose detection in patient specimens was temporally associated with RDV therapy. E802D was not identified at an allele frequency above 1% in either the specimen obtained during the initial phase of illness (day 36) or in specimens collected during the first 5 days (days 148-152) of RDV therapy. The mutation was first detected 7 days after initiation of RDV therapy (day 155) and accounted for 22.6% and 96% of the allele frequency in nasopharyngeal and saliva specimens, respectively, by day 160. The E802 residue of nsp12 resides in the palm sub-domain which contains a portion of the residues that comprise the active site of the SARS-CoV-2 RNA polymerase[1,10]. The E802 residue participates in an electrostatic network with D804 and K807 which stabilizes the loop involved with binding to the nascent RNA (Figure 2c, S2a). E802D has been identified in an *in vitro* RDV resistance selection experiment and was found to confer a ~2.5-fold increase in IC_{50} to the drug[6]. Together, the temporality of E802D emergence in the patient, its location in nsp12 and the *in vitro* identification of the same mutation with a RDV resistance phenotype support the plausibility that that RDV treatment of the patient selected for variants with the E802D mutation, which in turn contributed to the observed rebound viral shedding.

To validate the resistance phenotype of E802D, we engineered this mutation and an E802A mutation, which had been also shown to confer RDV resistance[6], into an infectious molecular clone of SARS-CoV-2/WA01 (icSARS-CoV-2-mNG), which expresses the mNeon Green reporter and is ORF7a depleted[11]. Replication kinetics of the E802D and E802A mutants in Vero-E6 cells revealed decreased viral replication relative to parental icSARS-CoV-2-mNG (Figure 2d), suggesting that these mutations may impart a fitness cost. In the presence of high concentrations of RDV (5 μM), the E802D mutant replicated to higher titers than the parental virus (Figure S2b), confirming that this mutation confers resistance to RDV. RDV dose-response curves demonstrated that E802D and E802A mutants had significant increases in IC_{50} values (4.2 μM and 2.7 μM, respectively) relative to parental icSARS-CoV-2-mNG (0.7 μM) (Figure 2e). Assessment of RDV cytotoxicity demonstrated limited impact on cell death rates below 10 μM, indicating that the RDV resistance phenotype conferred by mutations at E802 may occur at physiologically relevant concentrations (Fig S2c)[6].

These findings suggest that substitutions at residue 802 in nsp12 are novel mediators of RDV resistance. We speculate that these mutations distort the active site in a way that either enables the enzyme to exclude RDV or alleviates the steric clash mediated by S861, thereby allowing the enzyme to escape RDV-mediated chain termination[1,10]. Further biochemical and structural characterization will be needed to delineate the mechanism by which substitutions at residue E802 confer resistance to RDV.

The therapeutic impact of RDV resistance mutations at E802 may be attenuated by a fitness cost imparted by these mutations which we observed *in vitro*. Although RDV therapy has been widely administered to patients during the pandemic, E802D and all substitutions at residue 802 have been found in 122 (0.003%) and 270 (0.007%), respectively, of the 4.1M genome

sequences obtained from patient isolates in the GISAID database (www.gisaid.org; accessed 10/14/21). Consistent with the observation of a fitness cost, the allele frequency of E802D in nasopharyngeal specimens from the patient decreased from 22.6% to <1% between the period (day 160-162) after completion of RDV therapy and prior to casirivimab-imdevimab administration (Figure 2b). However, E802 allele frequencies were significantly higher (50-96% at days 158-167) in saliva specimens during this period and did not decline until casirivimab-imdevimab was administered. An explanation for these discordant findings is unclear and may relate to the differential compartmentalization of variants within the respiratory tract or the stochastic nature of sampling the nasopharyngeal compartment as opposed to saliva where virus from a wider geographic distribution may pool. Given that saliva is infrequently sampled, RDV resistance mediated by E802 mutations may be under-detected. Continued surveillance will be needed to elaborate whether the occurrence of E802D, as well as other substitutions at this residue, will be limited because of an underlying fitness cost or alternatively, emerge to pose a broader risk for RDV resistance.

We identified two additional mutations, A504V in nsp14 (exonuclease) and I115L in nsp15 (endoRNAse), whose allele frequencies increased during and after RDV treatment (Figure S3). However, A504V in nsp14 was present (allele frequency 85.1%) in a specimen obtained during her initial phase of illness (day 36). Although I115L in nsp15 was not detected in the early-phase specimen, the plausibility of a mutation in the endoRNAse gene which confers RDV resistance is less clear.

The case highlights additional issues for the management of SARS-CoV-2 infection in immunocompromised patients. Firstly, the course of infection was notable for being indolent despite a six-month period of sustained high-grade viral shedding and underlying lung

involvement. We identified six mutations, including a frameshift mutation in ORF3a (N257fs), which were present throughout the course of illness prior to RDV treatment (Figure S4). Yet none provide a clear explanation for an attenuated phenotype. Longitudinal profiling of PBMC demonstrated profound depletion of circulating B-cell populations and a T-cell phenotype consistent with chronic antigen exposure as evidenced by the increased activation (CD38+, HLA-DR+) and exhaustion (PD-1, Tim-3) markers (Figure S5). Given reports that maladaptive immune responses contribute to the risk of severe COVID-19[13,14], we speculate that the compromised immune response was unable to control viral load yet conversely, protected the patient from developing severe complications.

Secondly, SARS-CoV-2 infection was associated with the development of severe refractory neutropenia and moderate anemia in the patient (Figure 1c). While we are unable to discern if the cytopenia was due to direct or indirect effects of infection, SARS-CoV-2 has been detected in the bone marrow of an immunocompromised patient who developed pancytopenia during the course of infection[15]. Our patient's cytopenia resolved after casirivimab-imdevimab therapy, suggesting a causal link between SARS-CoV-2 infection and cytopenia in this setting.

Finally, the case illustrates the potential benefit of mAb therapy for immunocompromised patients with persistent SARS-CoV-2 infection. Prior to initiation of mAb therapy and in the absence of detectable circulating anti-SARS-CoV-2 spike protein antibodies (Figure 2a), two non-synonymous mutations, H655Y (Figure S6a) and P812L (Figure S6b), emerged which were not present in the initial specimen (day 36), indicating that spike protein continues to evolve and adapt to the host in the absence of selective pressure from an adaptive immune response. One week after casirivimab-imdevimab was administered (day 170), we identified a new spike protein mutation, A348S in the RBD domain, which persisted to day 177 (Figure S6c). WGS did

not detect the mutation in specimens collected after day 177 because of low amounts or undetectable viral RNA. A348S is not proximal to the mAb binding sites (Figure S6d)[16] and was not identified in a comprehensive analysis of resistance mutations[7]. In the setting of a negligible cellular immune response, mAb therapy cleared viral shedding, abrogated the residual complaint of anosmia, and resolved the blood dyscrasias.

In summary, we identified the *de novo* emergence of a RDV-resistance mutation, E802D, following initiation of RDV in an immunocompromised patient with persistent SARS-CoV-2 infection. While the finding is limited to a single case and requires confirmation of its generalizability in larger patient populations, it suggests that RDV can impart selective pressure *in vivo* to drive evolution of the virus. E802D is associated with a fitness cost *in vitro* which may limit the broader impact of this mutation on the development of secondary resistance during treatment and the risk for primary resistance through transmission of resistant variants. Yet, our findings underscore the importance of immunocompromised hosts with uncontrolled viral replication as a source of genetic diversification[4,17] and selection of mutations that may potentially impart adverse consequences for antiviral therapy. Enhanced genomic surveillance of immunocompromised patients may thus be warranted. As observed in this case, initiation of anti-SARS-CoV-2 mAb may serve as a therapeutic option to achieve rapid and sustained virologic responses and improved clinical outcomes in immunocompromised patients.

Online Methods

Regulatory and ethical approvals

Regulatory approval to administer an 8g dose of casirivimab/imdevimab was obtained from the U.S. Food and Drug Administration under investigational new drug application no. 153390. This study was also approved by Yale Human Research Protection Program Institutional Review Boards (FWA00002571, protocol ID 2000027690, 2000029277). Informed consent was obtained from both the patient and healthcare workers.

RNA extraction and PCR

RNA was extracted on a Qiagen Bio Robot (Qiagen, Hilden, Germany) using the QIAamp Virus Bio Robot MDx Kit (Qiagen) following manufacturer's instructions. RT-PCR was performed using AgPath-ID One-Step RT-PCR kit (Life Technologies, Carlsbad, CA, USA) with CDC primers and probes for N1/N2 following protocol and concentrations defined previously[18]. Each 25 µL of reaction mix included 2X RT-PCR buffer, 25X enzyme mix, primers-probes, and 5 µL of the extracted RNA. Cycle parameters on the ABI 7500 real-time PCR system (Applied Biosystems) were 10 min at 48°C for reverse transcription, 10 min of inactivation at 95°C, 40 cycles of 15 sec at 95°C, and 45 sec at 60°C. Internal control (EXO), positive, and negative (water) control were included on every run.

Library preparation and sequencing

Using 11 µL of extracted RNA, single strand complementary DNA (sscDNA) was synthesized using SuperScript IV First-Strand Synthesis System (Thermofisher). Sequencing libraries were prepared using the Swift Normalase Amplicon Panel (SNAP) for SARS-CoV-2 (Swift

Biosciences, Ann Arbor, MI, USA) as described previously[9]. Libraries were quantified using fluorometric methods (Quant-iT dsDNA high sensitivity kit, Life Technologies, Carlsbad, CA, USA), normalized, and sequenced on the Illumina NextSeq 500 or 2000 instruments (Illumina, San Diego, CA, USA) using 2x150 reads.

Analysis of sequencing data

Reads were processed using a custom bioinformatic pipeline (https://github.com/greninger-lab/covid_swift_pipeline), as described[9]. Briefly, raw FASTQs were trimmed of adapter content and low-quality reads, then aligned to the NCBI reference Wuhan-1 (NC_045512). PCR primers were soft-clipped with Primerclip (https://github.com/swiftbiosciences/primerclip) and a consensus genome was generated from this alignment using bcftools v1.9[19]. Variants were examined longitudinally with a modified Python script adapted from LAVA (https://github.com/greninger-lab/lava).

A k-mer based approach was used to identify reads corresponding to subgenomic RNAs. After removing adapters and low-quality regions, junction-spanning reads corresponding to individual sgRNAs were filtered using BBDuk (https://sourceforge.net/projects/bbmap/) against a custom FASTA file composed of 30 nucleotides of the 5' leader sequence and 30 nucleotides of the respective 3' junctions of downstream ORFs[20], allowing an edit distance of 1. In order to avoid calling of genomic reads, we constrained matches to those that had 10 consecutive matches of 31-mers.

Isolation of patient plasma and PBMCs

As described previously, patient whole blood was collected in sodium heparin-coated vacutainers and kept gently agitating at room temperature until sample pick-up by IMPACT team members. All blood was processed on the day of collection. Plasma samples were collected after centrifugation of whole blood at 400*g* for 10 minutes at room temperature (RT) without brake. The undiluted serum was then transferred to 15-ml polypropylene conical tubes, and aliquoted and stored at −80 °C for subsequent analysis. PBMCs were isolated using Histopaque (Sigma-Aldrich, #10771-500ML) density gradient centrifugation in a biosafety level 2+ facility. After isolation of undiluted serum, blood was diluted 1:1 in room temperature PBS, layered over Histopaque in a SepMate tube (StemCell Technologies; #85460) and centrifuged for 10 minutes at 1,200*g*. The PBMC layer was isolated according to the manufacturer's instructions. Cells were washed twice with PBS before counting. Pelleted cells were briefly treated with an ACK lysis buffer for 2 minutes and then counted. Percentage viability was estimated using standard Trypan blue staining and an automated cell counter (Thermo-Fisher, #AMQAX1000).

Flow cytometry

As described previously, antibody clones and vendors used for flow cytometry analysis were as follows: BB515 anti-hHLA-DR (G46-6) (1:400) (BD Biosciences), , BV605 anti-hCD3 (UCHT1) (1:300) (BioLegend),, BV785 anti-hCD4 (SK3) (1:200) (BioLegend), APCFire750 or PE-Cy7 or BV711 anti-hCD8 (SK1) (1:200) (BioLegend), BV421 anti-hCCR7 (G043H7) (1:50) (BioLegend), AlexaFluor 700 anti-hCD45RA (HI100) (1:200) (BD Biosciences), PE anti-hPD1 (EH12.2H7) (1:200) (BioLegend), APC anti-hTIM3 (F38-2E2) (1:50) (BioLegend), BV711 anti-hCD38 (HIT2) (1:200) (BioLegend), BB700 anti-hCXCR5 (RF8B2) (1:50) (BD Biosciences), PE-Cy7 anti-hCD127 (HIL-7R-M21) (1:50) (BioLegend), PE-CF594 anti-hCD25 (BC96)

(1:200) (BD Biosciences), In brief, freshly isolated PBMCs were plated at 1–2 × 10^6 cells per well in a 96-well U-bottom plate. Cells were resuspended in Live/Dead Fixable Aqua (ThermoFisher) for 20 min at 4 °C. Following a wash, cells were blocked with Human TruStan FcX (BioLegend) for 10 min at RT. Cocktails of desired staining antibodies were added directly to this mixture for 30 min at RT. For reinfection stains, cells were first washed and supernatant aspirated; then to each cell pellet a cocktail of secondary markers was added for 30 min at 4 °C. Prior to analysis, cells were washed and resuspended in 100 µl 4% PFA for 30 min at 4 °C. Following this incubation, cells were washed and prepared for analysis on an Attune NXT (ThermoFisher). Data were analysed using FlowJo software version 10.6 software (Tree Star). The specific sets of markers used to identify each subset of cells are summarized in Extended Data Fig. 9[21].

SARS-CoV-2 specific-antibody measurements

ELISAs were performed as previously described[22]. Briefly, Triton X-100 and RNase A were added to serum samples at final concentrations of 0.5% and 0.5mg/ml respectively and incubated at room temperature (RT) for 30 minutes before use to reduce risk from any potential virus in serum. 96-well MaxiSorp plates (Thermo Scientific #442404) were coated with 50 µl/well of recombinant SARS Cov-2 S1 protein (ACROBiosystems #S1N-C52H3-100ug) at a concentration of 2 µg/ml in PBS and were incubated overnight at 4 °C. The coating buffer was removed, and plates were incubated for 1h at RT with 200 µl of blocking solution (PBS with 0.1% Tween-20, 3% milk powder). Serum was diluted 1:50 in dilution solution (PBS with 0.1% Tween-20, 1% milk powder) and 100 µl of diluted serum was added for two hours at RT. Plates were washed three times with PBS-T (PBS with 0.1% Tween-20) and 50 µl of HRP anti-Human

IgG Antibody (GenScript #A00166, 1:5,000) diluted in dilution solution added to each well. After 1 h of incubation at RT, plates were washed three times with PBS-T. Plates were developed with 100 µl of TMB Substrate Reagent Set (BD Biosciences #555214) and the reaction was stopped after 15 min by the addition of 2 N sulfuric acid. Plates were then read at a wavelength of 450 nm and 570nm.

Cell lines and virus

As previously described[21], Vero E6 kidney epithelial cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 1% sodium pyruvate (NEAA) and 5% fetal bovine serum (FBS) at 37°C and 5% CO2. The cell line was obtained from the ATCC and has been tested negative for contamination with mycoplasma. SARS-CoV-2, strain USA-WA1/2020, was obtained from BEI Resources (#NR-52281) and was amplified in Vero E6 cells. Cells were infected at a MOI 0.01 for four three days to generate a working stock and after incubation the supernatant was clarified by centrifugation (450g × 5min) and filtered through a 0.45-micron filter. The pelleted virus was then resuspended in PBS then aliquoted for storage at − 80°C. Viral titers were measured by standard plaque assay using Vero E6 cells. All experiments were performed in a biosafety level 3 with the Yale Environmental Health and Safety office approval.

Virus titration from nasopharyngeal swabs.

SARS-CoV-2 titers from longitudinal nasopharyngeal swabs were assessed by conventional plaque assay. Nasopharyngeal swabs were collected by trained clinical staff and swabs stored and frozen in universal viral transport media (VTM) prior to plaquing. Each nasopharyngeal sample was ten-fold serially diluted (dilution range from 1:10 to 1:1000000). Diluted NP

samples were added on a monolayer of Vero E6 cells. After 1 hour of infection, DMEM (2% FBS), and 1.2% of Avicel (Dupont, 9004346) were overlaid. 72 hours post inoculation, media was removed and 4% Paraformaldehyde added for 1 hour. Cells were stained with 0.1% of crystal violet to visualize SARS-CoV-2 plaques. Plaques were counted by single operator. Each sample was tested in duplicate.

Site-directed mutagenesis and full length icSARS-CoV-2 mNG E802D and E802A production.

The reverse genetics system used to generate E802D and E802A SARS-CoV-2 mutants is described at length elsewhere[11]. The icSARS-CoV-2 mNG (mNeonGreen) backbone previously reported was employed to construct both full-length E802D and E802A mutant viruses utilizing the following primer pairs: forward (E802D): *gttggactgaTactgaccttactaaaggac*, reverse (E802D): *aggtcagtAtcagtccaacattttgcttc*; and forward (E802A): *gttggactgCgactgaccttactaaaggac*, reverse (E802A): *aggtcagtcGcagtccaacattttgcttc*. Site-directed mutagenesis was performed using Pfu Turbo Polymerase (Agilent) according to manufacturer's instructions. E802D and E802A mutations were confirmed in icSARS-CoV-2 mNG backbone by Sanger sequencing using the primer: *gatgatactctctgacgatg*.

Steps to produce full-length infectious virus were performed as described previously[11], with the following minor modifications: Plasmid propagation and digestion were performed as described and excision of DNA bands from the gel was done by using a Blue-Light transilluminator (Accuris). Next, equal molar amounts of the seven plasmids were reduced by half relative to the prior reports to perform the full-length genomic ligation: F1 (.305 μg), F2 (.325 μg), F3 (.375

μg), F4 (.47 μg) were ligated using T4 DNA ligase with a final volume of 50 μl. Plasmids F5 (.375 μg), F6 (.36 μg) and F7 (.3 μg) were incubated with T4 DNase Ligase in a volume of 40 μl. The primary reactions were incubated at 4°C for 16 h. After this, the two reactions were mixed and volumed up to 100 μl by adding the corresponding additional volume of ligation reagents. The secondary ligation was incubated for 18 h at 4°C. The full-length genomes were purified by using the Genomic DNA Clean and Concentrator kit from Zymo Research according to manufacturer's instructions. Successful ligation products were visualized using a 0.8% agarose gel and standard gel imager.

Next, full-length *in vitro* RNA transcription was also performed with additional modifications. First, 500 ng of each ligation product was employed. Reactions were performed using the mMESSAGE mMACHINE T7 Transcription Kit (Thermo Scientific) with a cap analog-to-GTP ratio of 1:1. Incubation of the reaction was done at 32°C for 12 h. Following DNAse treatment was performed at 37°C for 15 min. Next, viral RNA was purified by using the RNA Clean and Concentrator kit from Zymo Research according to the manufacturer's instructions. Successful transcription products were visualized using a 0.8% agarose gel and standard gel imager.

RNA transcripts were next electroporated into BHK-N expressing cells (courtesy of Dr. Brett Lindenbach). 15 μg of each mutant transcript was electroporated into BHK cells using the Electro Square Porator Device (BTX). Electroporated cells were seeded into a T75 flask containing 1:2 ratio of BHK-N:Vero-E6 cells. Complete monolayer CPE was evident at 72 h post-electroporation and supernatant was centrifuged and .22 um filtered. Vero-E6 were infected

with 500 uL of each of the clarified supernatants. Supernatants were collected for each mutant virus at 48 hours post infection and tittered as described previously.

Sequencing and variant calling

RNA was extracted from SARS-CoV-2 mutants using the MagMAX viral/pathogen nucleic acid isolation kit, and tested with the multiplexed RT-qPCR variant assay[23]. Using a standard dilution series, Ct values were converted to RNA copies per µL. Extracted RNA was diluted to 1 million copies per µL and used as input into the COVIDSeq Test RUO version for library preparation, in duplicate. Pooled libraries were sequenced on the NovaSeq (paired-end 150) at the Yale Center for Genome Analysis. Single nucleotide variants were called at a minimum frequency threshold of 0.02 and minimum read depth of 400X using iVar (version 1.3.1)[24], and filtered between duplicate replicates. Introduction of the E802D and E802A was confirmed at a minimum frequency of 0.82-0.9 in initial stocks.

Replication kinetics of icSARS-CoV-2 mutants.

Isolates from initial stocks were expanded and used to infect Vero-E6 cells with 0.01 MOI of each E802D and E802A mutant virus to assess their replication kinetics. Supernatants were collected at 8, 24 and 48 hours post-infection. Titers for each time point were assessed by conventional plaque assay as described above. Each virus was assessed in triplicate.

Remdesivir resistance assay.

Each icSARS-CoV-2 mutant virus (E802D and E802A) was cultured in the presence of increasing concentrations of RDV, ranging from 0.39 to 40 µM. Infections were performed using

Vero-E6 in 384-well plates containing the corresponding RDV dilutions in phenol-red free DMEM media with 5% FBS and incubated at 37°C for 4 hours. Next, cells were infected with 0.01 MOI of each mutant virus and infected cell frequencies were measured at 48 hours post-infection by mNeonGreen expression by high content imaging (Cytation 5, BioTek)[12]. Cell culture supernatants were also collected for plaque assay. Cell viability in uninfected cells was assessed at 72 h post-infection using the CellTiter-Glo kit (Promega) according to the manufacturer's instructions.

Neutralization assay

As described previously[25], patient sera was isolated from whole blood and heat treated for 30 minutes at 56 °C. Sixfold serially diluted plasma, from 1:10 to 1:2430 were incubated with SARS-CoV-2 for 1 h at 37 °C. The mixture was subsequently incubated with Vero E6 cells in a 6-well plate for 1 hour for infection. Cells were overlayed with DMEM supplemented $NaHCO_3$, 2% FBS, and 0.6% Avicel mixture[12]. Plaques were resolved at 40h post infection by fixing in 4% formaldehyde for 1h followed by staining in 0.5% crystal violet. All experiments were performed in parallel with healthy, uninfected control sera to establish degree of protection as reflected in difference in plaque counts.

Phylogenetic analysis

Phylogenic trees were generated using the Nextstrain software package[26]. 326 background sequences from CT, USA collected April 1st-September 30th 2020 were downloaded from GISAID (Supplemental Table 3). The tree was rooted using the Wuhan reference strain.

Data availability

Sequencing data are available under NCBI BioProject no. PRJNA774781 (Supplementary Table 2).

Acknowledgements

The authors thank Dr. Michael Jacobs, Dr. Sanjay Bhagani and Dr. David Lowe at the Royal Free Hospital for their advice regarding the management of the patient, Dr. Jennifer Hamilton, David Stein, Dr. Wendy Kampman and Dr. David M. Weinreich at Regeneron Pharmaceuticals, Inc. for their guidance on the use of monoclonal antibody therapy and laboratory evaluations, Nathanial Price at Yale School of Medicine for his assistance in identifying data and samples for the analyses, and Dr. Chantal Vogels, Dr. Joseph Fauver, Mallery Breban and Dr. Nathan Grubaugh at the Yale School of Public Health for sequencing infectious clones and providing advice on the analysis and interpretation of the genomic data. Casirivimab and imdevimab was provided by Regeneron Pharmaceuticals, Inc. for the purpose of emergency compassionate use treatment. We also thank the health care staff at Yale New Haven Hospital, which include Dr. Jensa Morris, Dr. Anne Spichler, Dr. Frederick Altice and Dr. Nikhil Seval, who cared for the patient during her hospital course and provided key input, and most of all, the patient for her equanimity and contributions to this study.

This study was supported by the Beatrice Kleinberg Neuwirth Fund; the Sendas Family Fund, Yale School of Public Health; and Department of Internal Medicine at the Yale School of Medicine. SG was supported by a grant from the National Institutes of Health (2T32AI007517-21A1). C.B.W. was supported by Burroughs Welcome Fund, Ludwig Family Foundation, and Mathers Foundation.

Contributions

J.K., A.R., and M.P.H contributed to data collection, analysis, and writing. M.L., P.R., J.F., D. F., C.V., M.C.M., A.S., E.W, N.K., P.L, W.W., A.P. contributed to data collection and analysis. O.O., A.G, C.B.W, A.I., and W.L.S. conceived of the study and contributed to study design, interpretation, and editing of the manuscript. S.G. and A.I.K conceived of the study and contributed to study design, data interpretation, writing, and regulatory approvals.

Competing Interests

ALG reports institutional central testing contracts from Abbott and research grants from Merck and Gilead, outside of the proposed work. AIK received consulting fees from Tata Sons and is recipient of grants on COVID-19 from Merck, Regeneron and Serimmune all of which are outside the submitted work. OO received consulting fees from Gilead and ViiV, as well as research support and honoraria from Gilead, outside of the submitted work. WLS is a consultant for Hugo Health, founder of Refactor Health and is recipient of grants on COVID-19 from Merck and Regeneron, all of which are outside the submitted work.

References

1. Kokic, G. *et al.* Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat Commun* 12, 279 (2021).

2. Rochwerg, B. *et al.* A living WHO guideline on drugs for covid-19. *Bmj* 370, m3379 (2020).

3. IDSA Guidelines on the Treatment and Management of Patients with COVID-19. https://www.idsociety.org/practice-guideline/covid-19-guideline-treatment-and-management/.

4. Collaboration, T. C.-N. B. C.-19 *et al.* SARS-CoV-2 evolution during treatment of chronic infection. *Nature* 1–10 (2021) doi:10.1038/s41586-021-03291-y.

5. Choi, B. *et al.* Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *New Engl J Med* 383, 2291–2293 (2020).

6. Szemiel, A. M. *et al.* In vitro selection of Remdesivir resistance suggests evolutionary predictability of SARS-CoV-2. *Plos Pathog* 17, e1009929 (2021).

7. Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* 369, 1014–1018 (2020).

8. Baum, A. *et al.* REGN-COV2 antibodies prevent and treat SARS-CoV-2 infection in rhesus macaques and hamsters. *Science* eabe2402 (2020) doi:10.1126/science.abe2402.

9. Addetia, A. *et al.* Sensitive Recovery of Complete SARS-CoV-2 Genomes from Clinical Samples by Use of Swift Biosciences' SARS-CoV-2 Multiplex Amplicon Sequencing Panel. *J Clin Microbiol* 59, (2020).

10. Yin, W. *et al.* Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* 368, 1499–1504 (2020).

11. Xie, X. *et al.* Engineering SARS-CoV-2 using a reverse genetic system. *Nat Protoc* 16, 1761–1784 (2021).

12. Wei, J. *et al.* Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection. *Cell* 184, 76-91.e13 (2021).

13. Wang, E. Y. *et al.* Diverse Functional Autoantibodies in Patients with COVID-19. doi:10.1101/2020.12.10.20247205.

14. Wilk, A. J. *et al.* A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat Med* 26, 1070–1076 (2020).

15. Issa, N., Lacassin, F. & Camou, F. First case of persistent pancytopenia associated with SARS-CoV-2 bone marrow infiltration in an immunocompromised patient. *Ann Oncol* 31, 1418–1419 (2020).

16. Hansen, J. *et al.* Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* 369, 1010–1014 (2020).

17. Drouin, A. C. *et al.* Successful Clearance of 300 Day SARS-CoV-2 Infection in a Subject with B-Cell Depletion Associated Prolonged (B-DEAP) COVID by REGEN-COV Anti-Spike Monoclonal Antibody Cocktail. *Viruses* 13, 1202 (2021).

18. Corman, V. M. *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* 25, 2000045 (2020).

19. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993 (2011).

20. Alexandersen, S., Chamings, A. & Bhatta, T. R. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. *Nat Commun* 11, 6059 (2020).

21. Lucas, C. *et al.* Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* 584, 463–469 (2020).

22. Amanat, F. *et al.* A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med* 26, 1033–1036 (2020).

23. Vogels, C. B. F. *et al.* Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. *Plos Biol* 19, e3001236 (2021).

24. Grubaugh, N. D. *et al.* An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol* 20, 8 (2019).

25. Lucas, C. *et al.* Delayed production of neutralizing antibodies correlates with fatal COVID-19. *Nat Med* 27, 1178–1186 (2021).

26. Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121–4123 (2018).

Figures

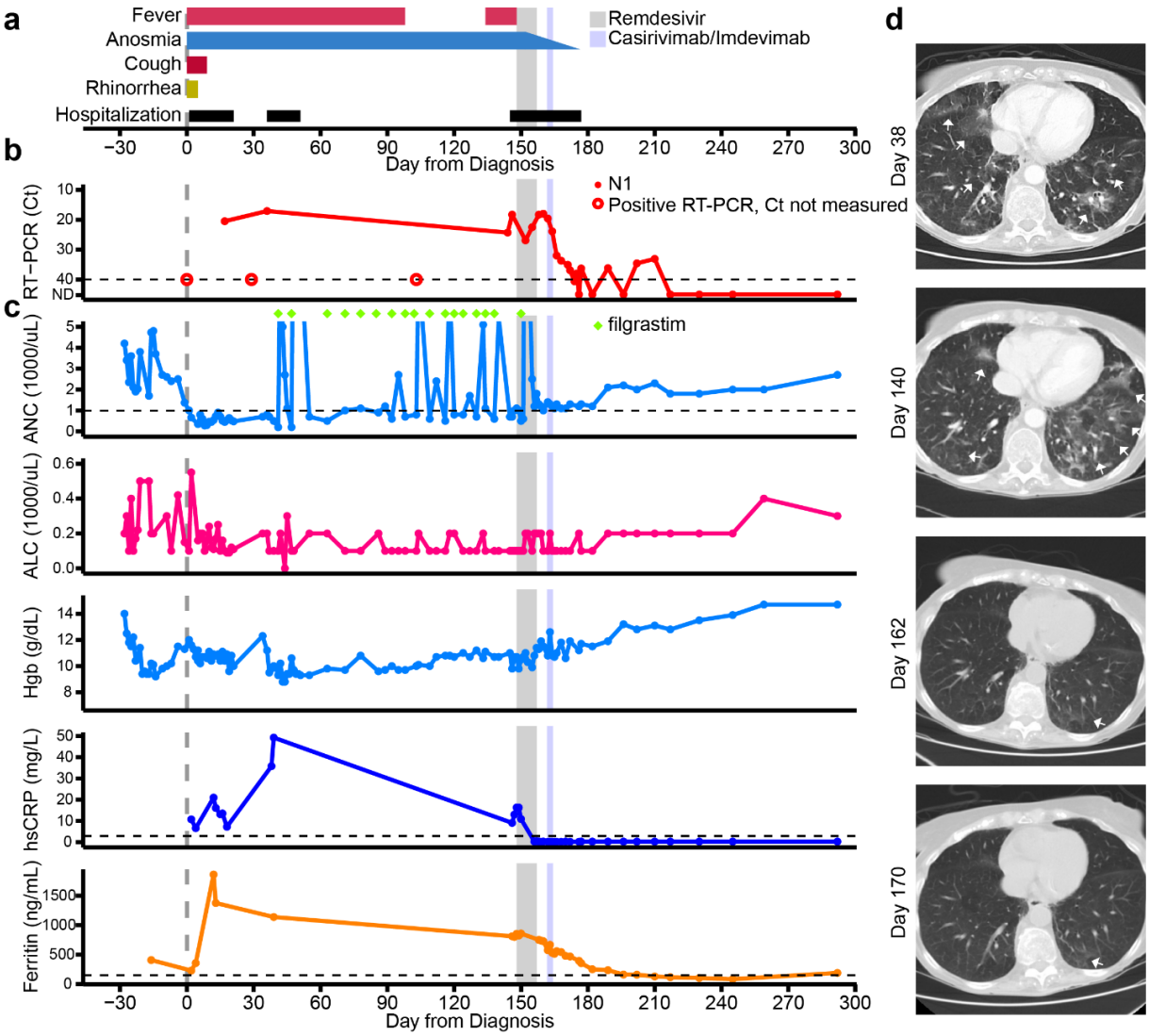


Fig. 1: Clinical course of the immunocompromised patient with persistent SARS-CoV-2 infection

Timeline of a) patient symptoms and hospitalizations, b) SARS-CoV-2 N1 RT-PCR Ct values and c) clinical laboratory parameters from time of laboratory diagnosis of SARS-CoV-2 infection (day 0) to end of the follow-up period (day 292) and d) computed tomography (CT) scans of chest at indicated days after time of initial diagnosis. The timing of remdesivir and

casirivimab-imdevimab are shown as grey and light blue shading, respectively. RT-PCR results that were positive but performed on assays that did not generate a Ct value are denoted by the open circle in panel b. The timing of filgrastim treatments are denoted by green diamonds in panel b. lsCRP values were converted to hsCRP values using a correction factor of 9.2. Ground-glass opacities marked by white arrows in panel d. Abbreviations: Real-time polymerase chain reaction (RT-PCR); Cycle threshold (Ct) Absolute neutrophil count (ANC); Absolute lymphocyte count (ALC); Hemoglobin (Hgb); high-sensitivity C-reactive protein (hsCRP).

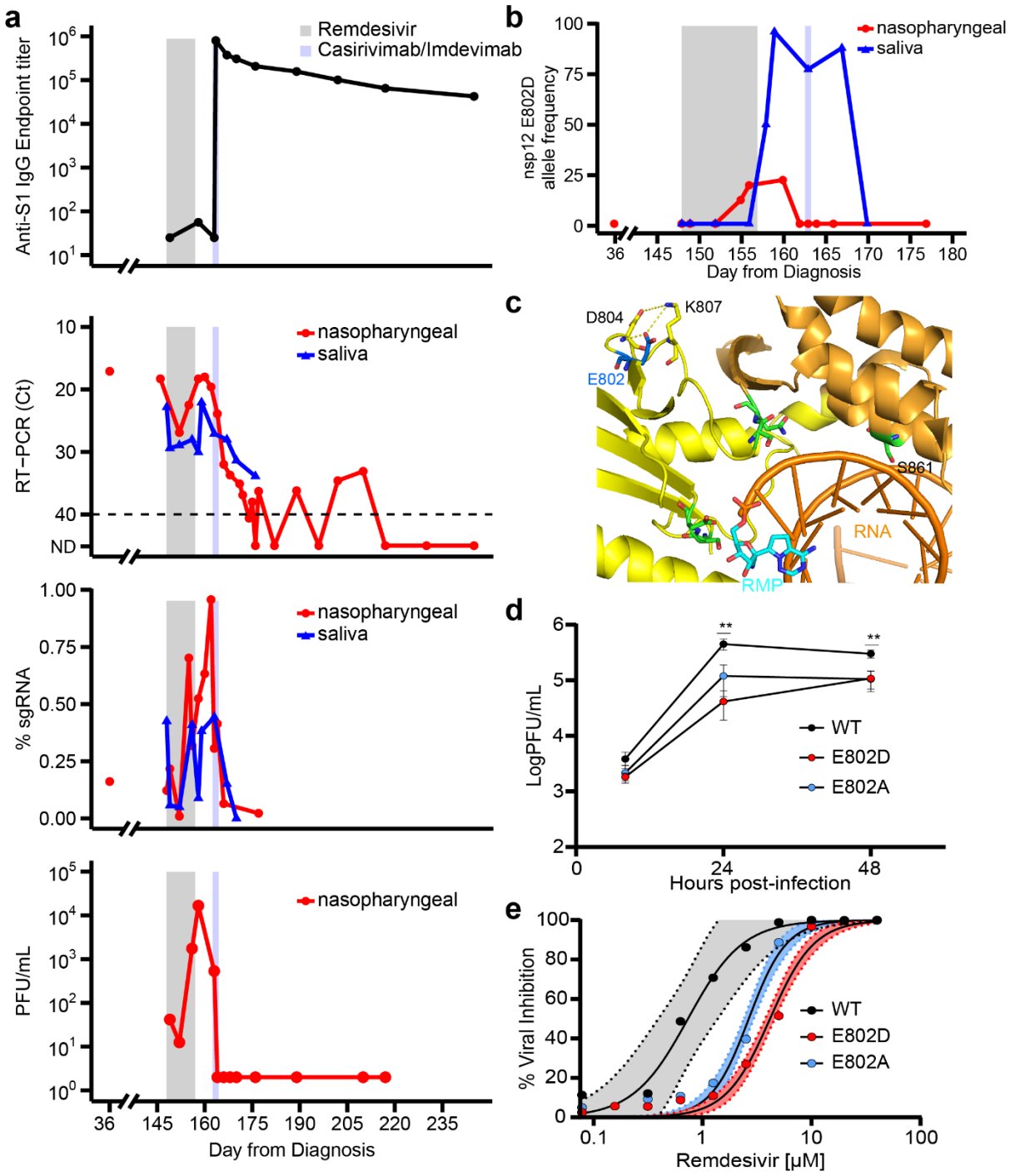


Fig. 2: *de novo* emergence of remdesivir resistance mutation during and following treatment with the antiviral agent

a) Anti-SARS-CoV-2 Spike Protein IgG ELISA endpoint titers (1st panel), RT-PCR Ct values (2nd panel), subgenomic RNA (sgRNA, 3rd panel), and plaque forming units (PFU) on viral culture (4th panel) during the course of illness which include the period when remdesivir (grey shading) and casirivimab/imdevimab (light blue shading) were administered. b) Allele frequencies of E802D in nsp12 as ascertained by whole genome sequencing. c) Crystal structure (PDB:7BV2) depicting nsp12 E802 (dark blue) and its putative hydrogen bonds (yellow dashed lines). In green are palm domain (yellow) residues that interact with the replicating RNA molecule (orange). Remdesivir monophosphate (RMP) is depicted in light blue. d) Viral growth curves from icSARS-CoV-2 mNG WT, E802D (patient), and E802A (control) nsp12 mutants on a ORF7a depleted backbone. Significance was assessed by unpaired two sample t-tests (** $p < 0.01$) with 3 biological replicates. e) SARS-CoV-2 inhibition by RDV assessed at 48 h post-infection (0.01 multiplicity of infection) as determined by quantitative image analysis of percentage of cells expressing mNG. Vertical lines indicate IC_{50} for WT (black) and E802D (red). Shading represents 95% confidence intervals of the IC_{50} estimate.

Supplemental Material

	Nasopharyngeal specimen				Saliva specimen			Stool specimen			Whole blood specimen		
Days after diagnosis	N1 RT-PCR Ct	N2 RT-PCR Ct	RP RT-PCR Ct	Viral Culture PFU/ml	N1 RT-PCR Ct	N2 RT-PCR Ct	RP RT-PCR Ct	N1 RT-PCR Ct	N2 RT-PCR Ct	RP RT-PCR Ct	N1 RT-PCR Ct	N2 RT-PCR Ct	RP RT-PCR Ct
36	**21.8**	**21.5**	**-**										
148	**26.1**	**25.4**	**32.9**		**22.9**	**22.2**	**22.3**						
149	**19.2**	**18.0**	**24.0**	**41.7**	**29.4**	**29.8**	**21.8**	**30.0**	**31.6**	**36.6**	ND	ND	36.0
152	**28.3**	**27.7**	**27.8**	**12.5**	**28.9**	**29.0**	**22.1**	**33.2**	**32.9**	**34.4**			
155	**26.1**	**25.6**	**27.6**										
156	**26.2**	**26.1**	**28.2**	**1741.7**	**28.0**	**27.6**	**23.3**	**32.8**	**ND**	**33.7**			
157											ND	ND	26.9
158	**24.2**	**23.7**	**34.5**	**16666.7**	**30.0**	**30.8**	**23.0**	**32.4**	**ND**	**33.1**	ND	ND	27.9
159					**22.1**	**21.2**	**23.5**	**31.7**	**ND**	**36.4**	ND	ND	28.2
160	**23.2**	**22.7**	**27.0**								ND	ND	29.2
162	**23.1**	**22.7**	**28.7**								ND	ND	27.1
163	**22.5**	**22.1**	**28.3**	**533.3**	**27.1**	**27.2**	**22.1**	**32.7**	**33.9**	**34.5**	ND	ND	30.6
164	**27.3**	**27.1**	**28.5**	**0.0**									
166	**35.5**	**36.1**	**29.3**	**0.0**									
167	36.2	36.9	30.7		**28.0**	**27.9**	**23.1**	34.1	41.1	31.5	ND	ND	30.1
168	39.5	38.7	28.2	0.0									
170	36.7	40.2	26.9	0.0	**31.4**	**32.4**	**22.4**	34.1	37.7	33.0	ND	ND	30.0
171	39.4	41.0	30.9										
176	39.7	ND	28.0	0.0	33.9	36.9	22.3	35.0	ND	32.5	ND	ND	29.8
177	**33.3**	**33.7**	**26.2**										
189											ND	ND	28.0
202	ND	ND	33.7								ND	ND	30.3
210	39.8	44.3	29.1	0.0									
217	35.9	37.1	24.4	0.0							ND	ND	28.1
245								34.5	ND	31.9	ND	ND	29.1

Supplementary Table 1: Serial patient samples from various tissues and secretions

Catalogue of patient samples collected for research during the course of illness. N1, N2, and RP cycle threshold values are reported for each sample. WGS was performed on the bolded samples (accession numbers in table S2). Days of RDV treatment are denoted by grey shading and casirivimab/imdevimab was administered on day 163 (blue shading). Note that NP Ct values depicted in Figures 1,2 were obtained from clinical platforms and are not shown here.

Abbreviations: ND=Not Detected

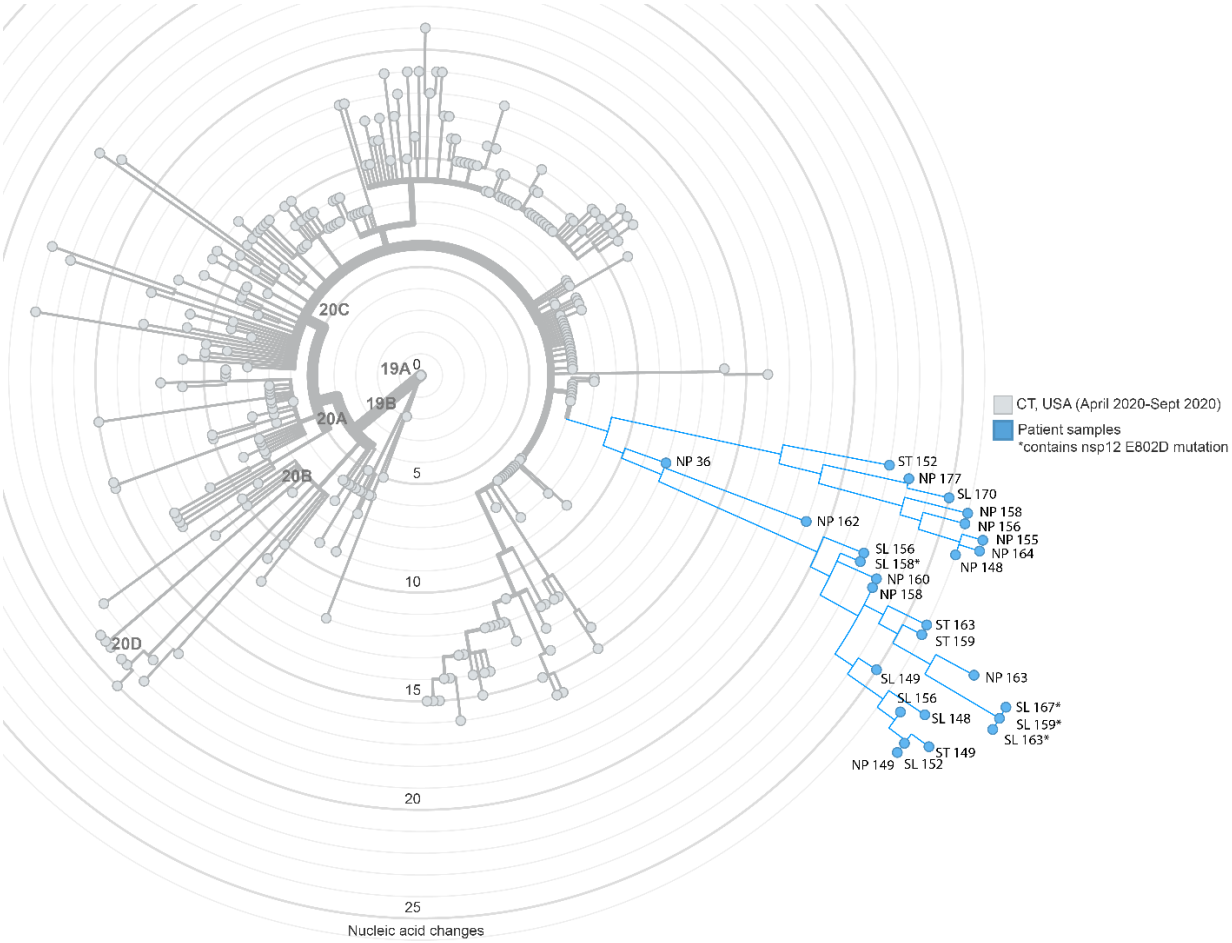


Fig. S1: Phylogenetic analysis of SARS-CoV-2 genomes from patient samples and contemporaneous samples from Connecticut, USA.

Nextstrain maximum likelihood analysis of patient whole genome samples (blue) in comparison to 283 contemporaneous local samples (grey) obtained from GISAID (10/30/21; Supplemental Table 3). Samples are annotated by type (nasopharyngeal (NP), saliva (SL), or stool (ST)) and day from diagnosis. Divergence from root reference genome (Wuhan-Hu-1) by nucleic acid changes.

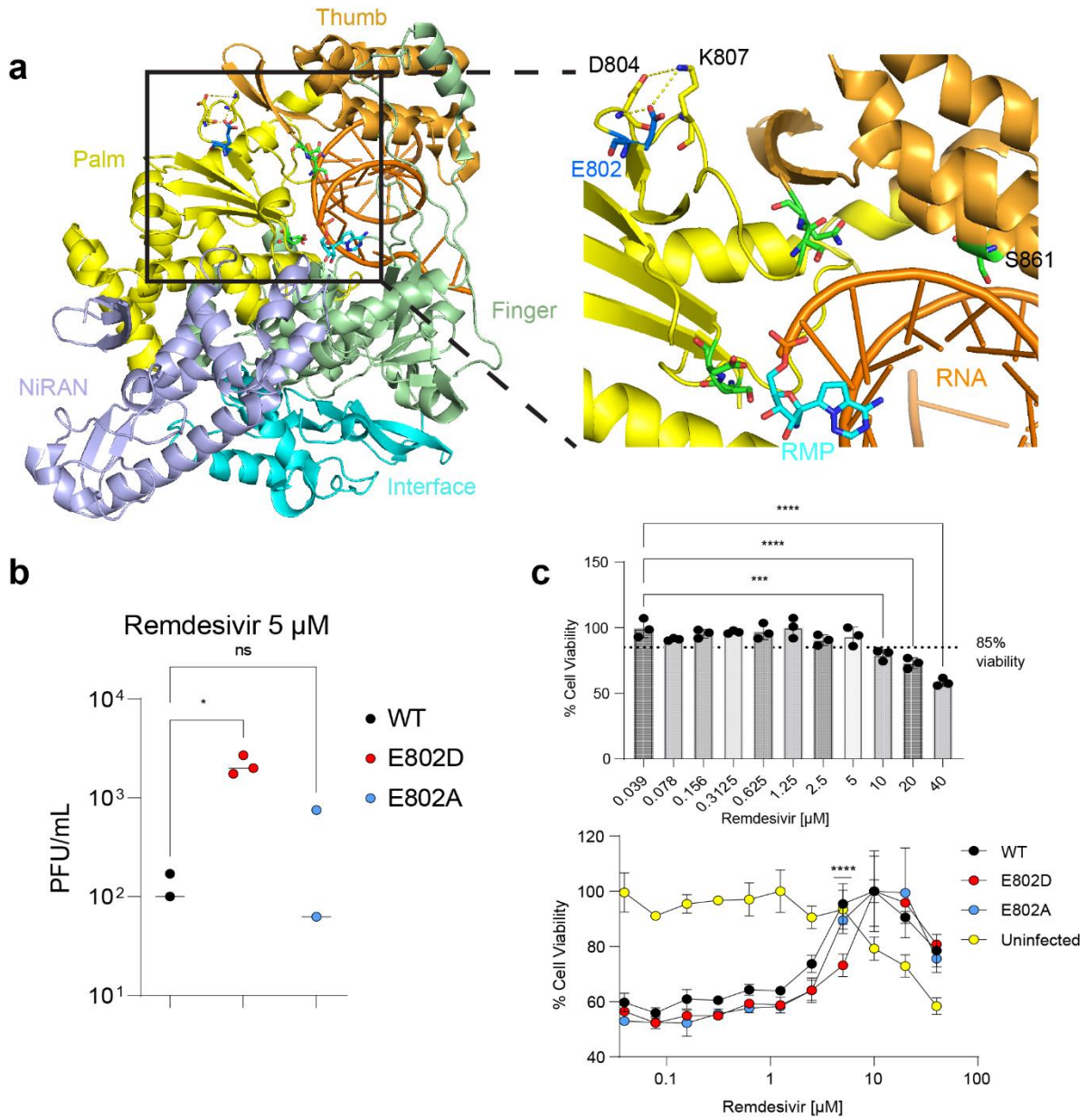


Fig. S2: Validation of a remdesivir resistance mutation

a) Structure depicting E802 in relationship to nsp12 (PDB:7BV2) b) Viral titers (PFU) assessed under 5 µM RDV treatment. Significance was assessed by an ANOVA (* $p < 0.05$) with 3 biological replicates. c) Cell viability experiment with RDV performed using CellTiter Glo. Significance of infected cells with SARS-CoV-2 WT (top panel) or E802D/A (bottom panel)

was assessed by two-way ANOVA corrected for multiple comparisons (**** $p < 0.0001$) with 3 biological replicates.

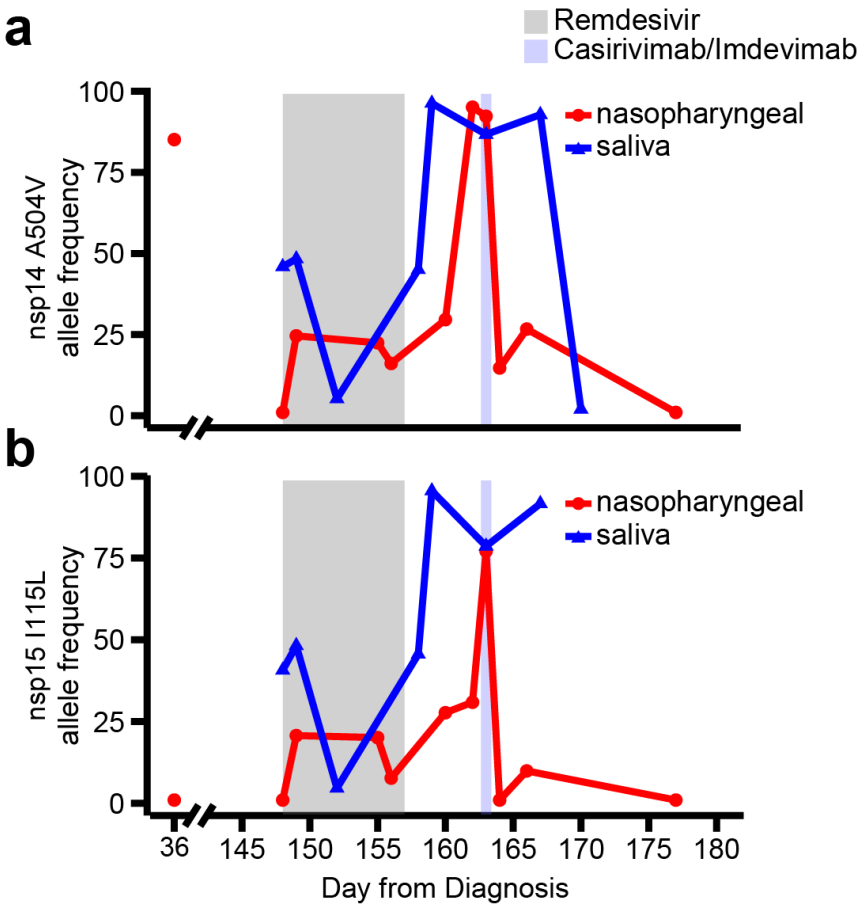


Fig S3: Identification of additional mutations temporally associated with remdesivir treatment

Longitudinal allele frequencies of A504V in nsp14 (a) and I115L (b) by whole genome sequencing of in relationship to RDV and casirivimab/imdevimab

SNP	Protein	AA	NP 36	NP 148	SL 148
G521T	leader protein	V86F			x
G569A	leader protein	E102K		x	
C1059T	nsp2	T85I	x	x	x
A4229G	nsp3	T504A		x	
C4230G	nsp3	T504R			x
C5178T	nsp3	T820I	x		x
G5180A	nsp3	D821N		x	
C9344T	nsp4	L264F			x
C9870T	nsp4	T439M		x	
G9929A	nsp4	D459N			x
A10323G	3C-like_proteinase	K90R		x	
G11083T	nsp6	L37F			x
A12519G	nsp8	D143G		x	
C12525T	nsp8	T145I			x
A12759G	nsp9	D25G		x	
A12759C	nsp9	D25A			x
C14408T	RdRp	P323L	x	x	x
C19550T	3'-to-5' exonuclease	A504V	x		
A23403G	surface glycoprotein	D614G	x	x	x
C23525T	surface glycoprotein	H655Y			x
C23997T	surface glycoprotein	P812L		x	
C25511T	ORF3a_protein	S40L			x
G25563T	ORF3a_protein	Q57H	x	x	x
C25714T	ORF3a_protein	L108F		x	
G25912T	ORF3a_protein	G174C		x	
A26162del	ORF3a_protein	N257fs	x	x	x
T26767C	membrane glycoprotein	I82T		x	
A28281G	nucleocapsid	D3G		x	
T28829C	nucleocapsid	S186P			x
C28472A	nucleocapsid	P67T	x	x	x
G29151A	nucleocapsid	R293K	x		
A29567G	ORF10_protein	I4V			x

Fig. S4: Variants identified in persistently infected host

Non-synonymous (NS) mutations identified at days 36 (NP) and 148 (NP and saliva) of infection. During early infection (day 36), whole genome sequencing identified 9 NS mutations from reference strain Wuhan-1 (NC_045512). Six of those persisted through day 148 (green background). Abbreviations: single nucleotide polymorphism (SNP); amino acid (AA); nasopharyngeal (NP); saliva (SL)

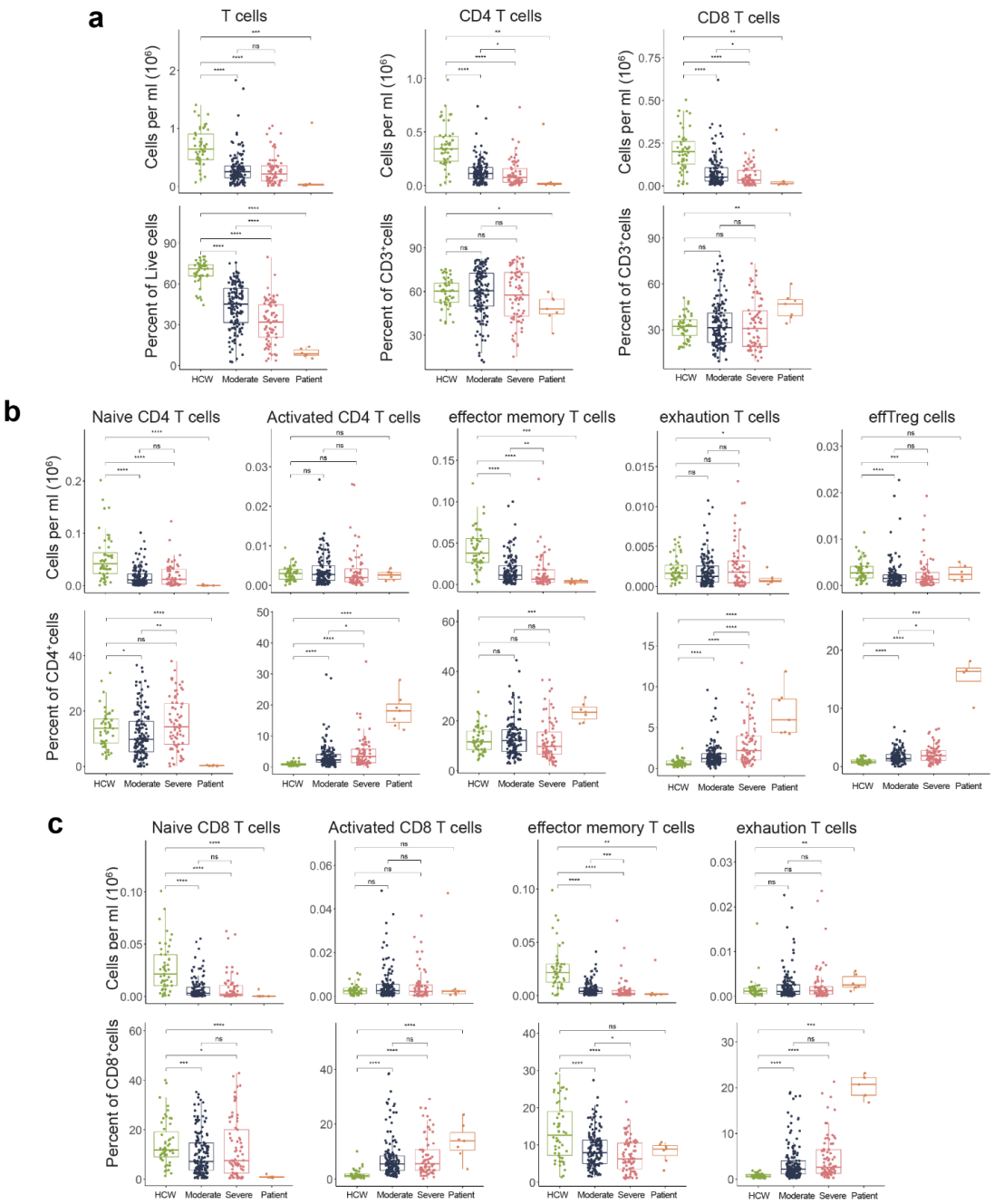
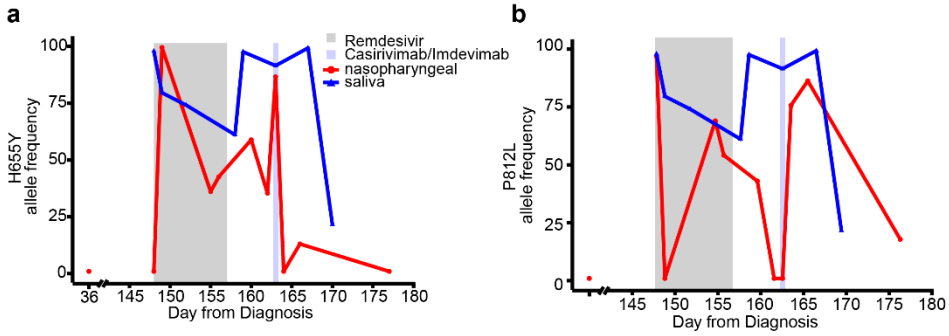


Fig. S5: Serial lymphocyte profiling reveals T-cell exhaustion

T-cell profiling of uninfected healthcare workers (HCW), patient's with moderate and severe SARS-CoV-2 infection. a) Global, $CD4^+$, and $CD8^+$ T-cells. $CD4^+$ (b) and $CD8^+$ (c) subsets demonstrates a phenotype consistent with T-cell exhaustion. Population definitions: naïve

CD4/CD8 subset: CD4/CD8+CD45RA+CD127+CCR7+PD-1-; activated CD4/CD8: CD4/CD8+CD38+HLA-DR+;effector memory CD4/CD8:CD4/CD8CD45RA-CD127+CCR7-; exhaustion CD4/CD8: CD4/CD8CD45RA-Tim-3+PD-1+; effTreg: CD4+CD45RA-CD127-CD25hiHLA-DR+. Gating strategy depicted in Fig S7.



c

Source	Day	D80Y	L141fs	A348S	D614G	V635G	H655Y	V656I	P812L	G1124V
NP	36	<5	<5	<5	100	<5	<5	<5	<5	<5
NP	148	<5	<5	<5	99.9	<5	<5	<5	96.9	<5
SL	148	<5	<5	<5	99.9	<5	97.9	<5	<5	<5
NP	149	<5	<5	<5	99.9	<5	99.5	<5	<5	<5
SL	149	<5	<5	<5	99.9	5.6	79.6	<5	20.7	<5
SL	152	<5	<5	8.8	99.8	<5	74.1	<5	<5	<5
NP	155	<5	<5	<5	99.8	<5	36.1	7.9	69	<5
NP	156	<5	<5	<5	99.8	<5	42.6	<5	54.2	<5
SL	158	<5	<5	<5	99.9	6.2	61.1	<5	20.9	<5
NP	160	<5	<5	<5	99.8	<5	58.9	<5	42.3	<5
NP	162	62.9	<5	<5	99.8	<5	35.4	24.3	<5	65.5
NP	163	13.9	<5	<5	99.8	<5	86.6	59.2	<5	15.8
SL	163	<5	<5	<5	100	5.3	91.5	11.3	10.4	<5
NP	164	17.9	12.6	<5	99.9	<5	<5	<5	75.7	17.3
NP	166	<5	<5	<5	100	<5	13.1	13.2	86.3	8.6
SL	167	<5	<5	<5	100	5.4	99.3	7.4	<5	<5
SL	170	<5	5.6	75	100	5.4	21.6	<5	<5	<5
NP	177	<5	<5	86.5	100	<5	<5	<5	17.9	<5

d

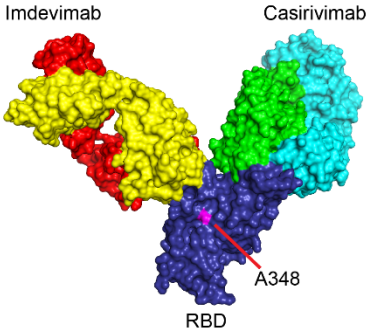


Fig. S6: Evolution of Spike protein in patient samples.

a) H655Y and b) P812L allele frequency over time. c) Allele frequencies of non-synonymous Spike protein mutations identified on more than one occasion by whole genome sequencing. Allele frequencies >50% are denoted by green background. d) Crystral structure of RBD (PDB:6XDG) depicting residue A348 in relationship to the imdevimab (heavy chain: yellow;

light rain: red) and casirivimab (heavy chain: green; light chain: blue) binding sites.

Abbreviations: nasopharyngeal (NP); saliva (SL); receptor binding domain (RBD).

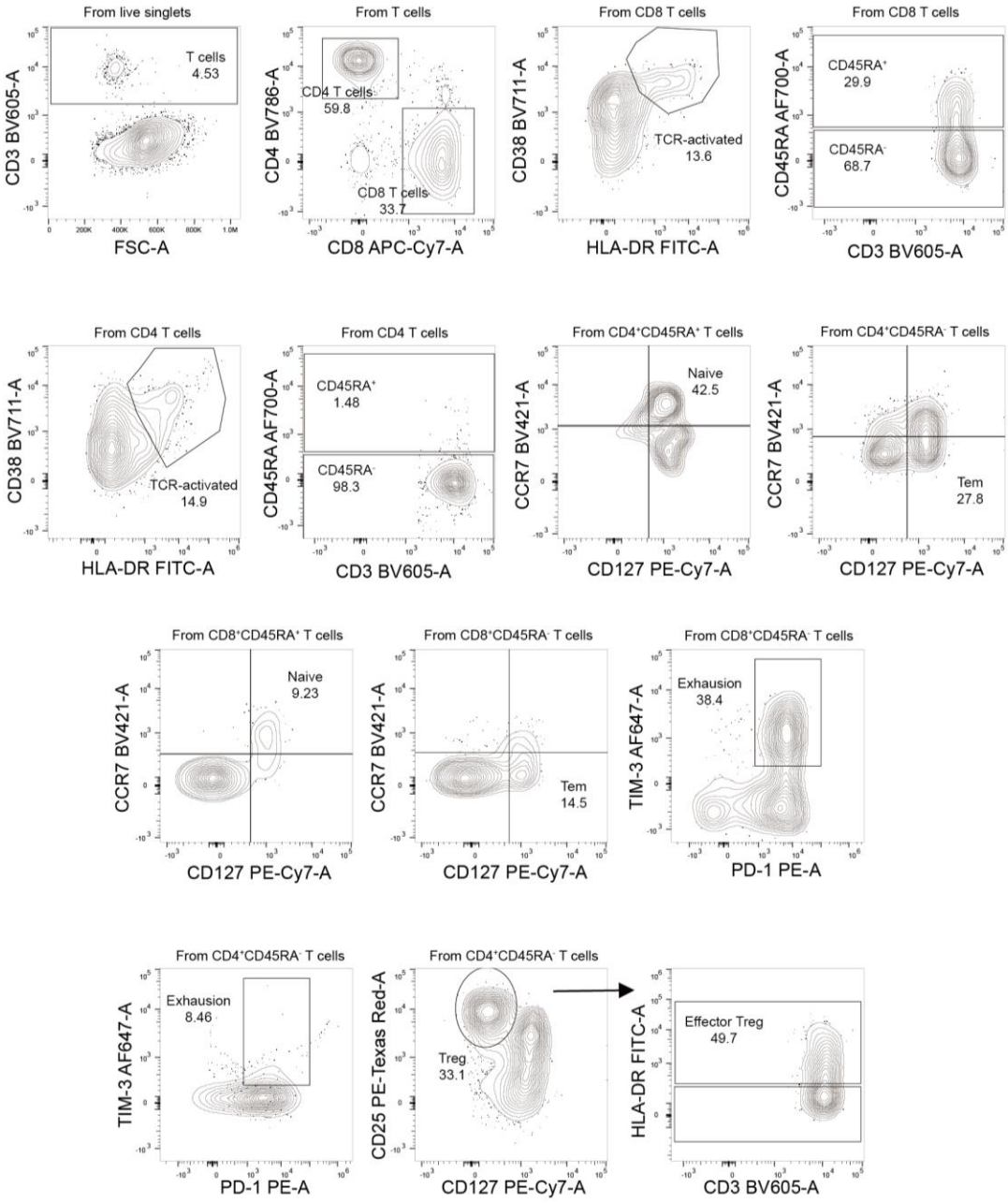


Figure S7: Gating strategy for flow cytometry

Gating strategies are shown for the key T cell populations described in Figure S5. The T cell surface staining gating strategy to identify CD8 & CD4 T cells, naïve T cells, TCR-activated T

cells, exhaustion T cells, effector memory T cells (Tem), and effector regulatory T cells (eff Treg) are depicted.

Sample Name	SRA Accession	GenBank Accession
INP-0411-NP-148	SRR16576689	OL307721
INP-0411-NP-149	SRR16576688	OL307722
INP-0411-NP-152*	SRR16576677	
INP-0411-NP-155	SRR16576668	OL307723
INP-0411-NP-156	SRR16576667	OL307724
INP-0411-NP-158	SRR16576666	OL307725
INP-0411-NP-160	SRR16576665	OL307726
INP-0411-NP-162	SRR16576664	OL307727
INP-0411-NP-163	SRR16576663	OL307728
INP-0411-NP-164	SRR16576662	OL307729
INP-0411-NP-166	SRR16576687	OL307720
INP-0411-NP-177	SRR16576686	OL307730
INP-0411-NP-36	SRR16576685	OL307731
INP-0411-SL-148	SRR16576684	OL307732
INP-0411-SL-149	SRR16576683	OL307733
INP-0411-SL-152	SRR16576682	OL307734
INP-0411-SL-156	SRR16576681	OL307735
INP-0411-SL-158	SRR16576680	OL307736
INP-0411-SL-159	SRR16576679	OL307737
INP-0411-SL-163	SRR16576678	OL307738
INP-0411-SL-167	SRR16576676	OL307739
INP-0411-SL-170	SRR16576675	OL307740
INP-0411-ST-149	SRR16576674	OL307741
INP-0411-ST-152	SRR16576673	OL307742
INP-0411-ST-156	SRR16576672	OL307743
INP-0411-ST-158	SRR16576671	OL307744
INP-0411-ST-159	SRR16576670	OL307745
INP-0411-ST-163	SRR16576669	OL307746

*depth was sufficent to calculate allele frequency at position E802 in nsp12 but coverage was insufficent to generate consenus genome

Supplemental Table 2: Accession numbers for submitted samples

We gratefully acknowledge the following Authors from the Originating laboratories responsible for obtaining the specimens, as well as the Submitting laboratories where the genome data were generated and shared via GISAID, on which this research is based.

All Submitters of data may be contacted directly via www.gisaid.org

Authors are sorted alphabetically.

Accession ID	Originating Laboratory	Submitting Laboratory	Authors
EPI_ISL_3354875	CT-Dr. Katherine A. Kelley State Public Health Lab	Centers for Disease Control and Prevention Division of Viral Diseases, Pathogen Discovery	Alex Burgin; Ben Rambo-Martin; Clinton Paden; Dakota Howard; Dave Wentworth; Dhwani Batra; Jasmine Padilla; Justin Lee; Krista Queen; Kristen Knipe; Kristine Lacek; Mark Burroughs; Matthew Schmerer; Meghan Bentz; Mili Sheth; Peter Cook; Sam Shepard; Sarah Nobles; Suxiang Tong; Vivien Dugan; Yvette Unoarumhi
EPI_ISL_454642, EPI_ISL_454645	CT-Dr. Katherine A. Kelley State Public Health Lab	Pathogen Discovery, Respiratory Viruses Branch, Division of Viral Diseases, Centers for Disease Control and Prevention	Anna Uehara; Bettina Bankamp; Clinton R. Paden; Haibin Wang; Jing Zhang; Krista Queen; Suxiang Tong; Yan Li; Ying Tao; Zachary Weiner
EPI_ISL_729735, EPI_ISL_729736, EPI_ISL_729741, EPI_ISL_729743, EPI_ISL_729744, EPI_ISL_729745, EPI_ISL_729746, EPI_ISL_729747, EPI_ISL_729748, EPI_ISL_729749, EPI_ISL_729750, EPI_ISL_729752, EPI_ISL_729753, EPI_ISL_729754, EPI_ISL_729755, EPI_ISL_729756, EPI_ISL_729758, EPI_ISL_729759, EPI_ISL_729760, EPI_ISL_729765, EPI_ISL_729772, EPI_ISL_729773, EPI_ISL_729778, EPI_ISL_729779, EPI_ISL_729780, EPI_ISL_729786, EPI_ISL_729789, EPI_ISL_729792			
see above	Connecticut Department of Health	Grubaugh Lab - Yale School of Public Health	Adam Moore; Akiko Iwasaki; Albert Ko; Alice Lu; Allison Nelson; Anderson Brito; Anne Wyllie; Annie Watkins; Arnau Casanovas; Catherine Muenker; Chaney Kalinich; Chantal Vogels; Charlese Dela Cruz; Isabel Ott; Joseph Fauver; Maria Tokuyama; Mary Petrone; Nathan Grubaugh; Patrick Wong; Peiwen Lu; Richard Martinello; Saad Omer; Shelli Farhadian; Tara Alpert
EPI_ISL_527738, EPI_ISL_527761, EPI_ISL_527762, EPI_ISL_527763, EPI_ISL_527764, EPI_ISL_527765, EPI_ISL_527766, EPI_ISL_527767, EPI_ISL_527768, EPI_ISL_527769, EPI_ISL_527770, EPI_ISL_527771, EPI_ISL_527772, EPI_ISL_527773, EPI_ISL_527774, EPI_ISL_527775, EPI_ISL_527776, EPI_ISL_527777, EPI_ISL_527778, EPI_ISL_527779, EPI_ISL_527780, EPI_ISL_527781, EPI_ISL_527782, EPI_ISL_527783, EPI_ISL_527784, EPI_ISL_527785			
see above	Connecticut State Department of Public Health	Grubaugh Lab - Yale School of Public Health	Albert Ko; Anderson Brito; Anthony Muyombwe; Chantal Vogels; Jafar Razeq; Joseph Fauver; Mary Petrone; Nathan Grubaugh; Tara Alpert
EPI_ISL_738284, EPI_ISL_738285, EPI_ISL_738289, EPI_ISL_738298, EPI_ISL_738299, EPI_ISL_738300, EPI_ISL_738301, EPI_ISL_738302, EPI_ISL_738307, EPI_ISL_738308, EPI_ISL_738309, EPI_ISL_738310, EPI_ISL_738311, EPI_ISL_738312, EPI_ISL_1017585, EPI_ISL_1017589, EPI_ISL_1017592, EPI_ISL_1017593, EPI_ISL_1017602, EPI_ISL_1017604, EPI_ISL_1017627, EPI_ISL_1017628, EPI_ISL_1017636, EPI_ISL_1017637, EPI_ISL_1017639, EPI_ISL_1017643, EPI_ISL_1017652, EPI_ISL_1017675, EPI_ISL_1067646, EPI_ISL_1067647, EPI_ISL_1067686, EPI_ISL_1067687, EPI_ISL_1080455, EPI_ISL_1080462, EPI_ISL_1091839, EPI_ISL_1091869, EPI_ISL_1239847			
see above	Connecticut Veterans' Affairs Hospital	Grubaugh Lab - Yale School of Public Health	Chantal Vogels; Danielle Plank; Ellen Foxman; Isabel Ott; Joseph Fauver; Mary Petrone; Nathan Grubaugh; Shaili Gupta; Tara Alpert
EPI_ISL_802136	MSHS Clinical Microbiology Laboratories	MSHS Pathogen Surveillance Program	Adolfo García-Sastre; Adriana van de Guchte; Ajay Obla; Alberto Paniz-Mondolfi; Ana S. Gonzalez-Reiche; Andrew Kasarskis; Angela Amoako; Ashley S. Salimbangon; Betsaida Salom Melo; Bremy Alburquerque; Brianne Ciferri; Charles Gleason; Deena R. Altman; Denise Jurczyszak; Elena Hirsch; Emilia Mia Sordillo; Emily Ferreri; Gintaras Deikus; Giulio Kleiner; Gopi Patel; Hala Alshammary; Harm van Bakel; Irina Oussenko; Jayeeta Dutta; Juan Soto; Katherine Beach; Kathryn Twyman; Kayla Russo; Komal Srivastava; Levy Sominsky; Mahmoud Awawda; Marta Luksza; Matthew M. Hernandez; Melissa Gitman; Michael D. Nowak; Mitchell J. Sullivan; Nancy Francoeur; Rachel Chernet; Robert Sebra; Sarah Schaefer; Shelcie Fabre; Shwetha Hara Sridhar; Viviana Simon; Ying-Chih Wang; Zenab Khan
EPI_ISL_430401	NYU Langone Health	Departments of Pathology and Medicine, New York University School of Medicine	Adriana Heguy; Amy Rapkiewicz; Andre M. Ribeiro-dos-Santos; Andrew Lytle; Antonio Serrano; Brendan Belovarac; Christian Marier; Dacia Dimartino; Emily Guzman; Emily Huang; Gael Westby; George Jour; Guomiao Shen; Iman Osman; Jared Pinnell; John Cadley; John Chen; Lawrence H. Lin; Ludovic Boytard; Margaret Black; Maria Aguero-Rosenfeld; Marie Samanovic-Golden; Mark J. Mulligan; Matija Snuderl; Matthew T. Maurano; Megan Hogan; Nick Vulpescu; Paolo Cotzia; Paul Zappile; Peter Meyn; Raquel Ordonez Ciriza; Raven Luther; Sitharam Ramaswami; Tatyana Gindin; Theodore Vougiouklakis; Vanessa Raabe; Xiaojun Feng; Yutong Zhang
EPI_ISL_937036, EPI_ISL_937037	Quest Diagnostics	Quest Diagnostics	Anderson, B.; D.F.; Gerasimova, A.; Hua, M.; K.E.; Kagan; Lacbawan, F.; Liu Y.; Livingston; Owen, R.; R.M.; Rosenthal; S.H.; Shalhout
EPI_ISL_429640, EPI_ISL_429641, EPI_ISL_429642, EPI_ISL_429645, EPI_ISL_429646, EPI_ISL_430907, EPI_ISL_430908, EPI_ISL_430917, EPI_ISL_430919, EPI_ISL_430921, EPI_ISL_430922, EPI_ISL_430923, EPI_ISL_430925, EPI_ISL_430926, EPI_ISL_430927, EPI_ISL_430960, EPI_ISL_430961, EPI_ISL_430962, EPI_ISL_430963, EPI_ISL_430965, EPI_ISL_430971, EPI_ISL_430972, EPI_ISL_430973, EPI_ISL_437865, EPI_ISL_437866, EPI_ISL_437867, EPI_ISL_437868, EPI_ISL_437869, EPI_ISL_437870, EPI_ISL_437871, EPI_ISL_437872, EPI_ISL_461408, EPI_ISL_461409, EPI_ISL_461411, EPI_ISL_461412, EPI_ISL_461414, EPI_ISL_461416, EPI_ISL_461417, EPI_ISL_461420, EPI_ISL_461435, EPI_ISL_461441, EPI_ISL_570070, EPI_ISL_570076, EPI_ISL_570092, EPI_ISL_570093, EPI_ISL_570094, EPI_ISL_570095, EPI_ISL_570104, EPI_ISL_570107, EPI_ISL_570109, EPI_ISL_570130, EPI_ISL_570131, EPI_ISL_570132, EPI_ISL_570133, EPI_ISL_570135, EPI_ISL_570137, EPI_ISL_570141, EPI_ISL_570145, EPI_ISL_570146, EPI_ISL_570148, EPI_ISL_570166, EPI_ISL_570178, EPI_ISL_570179, EPI_ISL_570180, EPI_ISL_570181, EPI_ISL_570182, EPI_ISL_570184, EPI_ISL_570185			
see above	UW Virology Lab	UW Virology Lab	Alexander Greninger; Amin Addetia; Hong Xie; Keith Jerome; Keith R Jerome; Lasata Shrestha; Meei-Li Huang; Pavitra Roychoudhury; Truong Nguyen; Victoria M Rachleff
EPI_ISL_484885	University of Wisconsin-Madison AIDS Vaccine Research Laboratories	University of Wisconsin-Madison AIDS Vaccine Research Laboratories	Gage Moreno; Katarina Braun; et al. AIDS Vaccine Research Laboratories
EPI_ISL_1168354, EPI_ISL_1168356, EPI_ISL_1168357, EPI_ISL_1168363, EPI_ISL_1168368, EPI_ISL_1168377, EPI_ISL_1168384, EPI_ISL_1168391, EPI_ISL_1168396, EPI_ISL_1168406, EPI_ISL_1168426, EPI_ISL_1168427, EPI_ISL_1168428, EPI_ISL_1168431, EPI_ISL_1168435, EPI_ISL_1168437			
see above	VA Connecticut Healthcare System	Grubaugh Lab - Yale School of Public Health	Chantal Vogels; Danielle Plank; Isabel Ott; Joseph Fauver; Mary Petrone; Nathan Grubaugh; Shaili Gupta; Tara Alpert
EPI_ISL_428384, EPI_ISL_428385, EPI_ISL_428386, EPI_ISL_428389, EPI_ISL_428391, EPI_ISL_428392, EPI_ISL_428393, EPI_ISL_428394, EPI_ISL_428395, EPI_ISL_428396, EPI_ISL_428397, EPI_ISL_428398, EPI_ISL_428399, EPI_ISL_428401, EPI_ISL_428402, EPI_ISL_428403, EPI_ISL_428404, EPI_ISL_428405, EPI_ISL_428746, EPI_ISL_428747, EPI_ISL_428748, EPI_ISL_428749, EPI_ISL_428750, EPI_ISL_428751, EPI_ISL_428752, EPI_ISL_431081, EPI_ISL_431082, EPI_ISL_431083, EPI_ISL_431084, EPI_ISL_431085, EPI_ISL_431086, EPI_ISL_431087, EPI_ISL_431088, EPI_ISL_431089, EPI_ISL_431090, EPI_ISL_431091, EPI_ISL_431092, EPI_ISL_431093, EPI_ISL_431094, EPI_ISL_431095, EPI_ISL_431096, EPI_ISL_435702, EPI_ISL_435703, EPI_ISL_435704, EPI_ISL_435705, EPI_ISL_435706, EPI_ISL_435707, EPI_ISL_435708, EPI_ISL_452143, EPI_ISL_452144, EPI_ISL_452145, EPI_ISL_452146, EPI_ISL_452147, EPI_ISL_594133, EPI_ISL_730052			
see above	Yale COVID-19 Biorepository	Grubaugh Lab - Yale School of Public Health	Adam Moore; Akiko Iwasaki; Albert Ko; Alice Lu; Allison Nelson; Anderson Brito; Anne Wyllie; Annie Watkins; Arnau Casanovas; Catherine Muenker; Chaney Kalinich; Chantal Vogels; Charlese Dela Cruz; Cole Jensen; Isabel Ott; Joseph Fauver; Maria Tokuyama; Mary Petrone; Nathan Grubaugh; Patrick Wong; Peiwen Lu; Richard Martinello; Saad Omer; Shelli Farhadian; Tara Alpert
EPI_ISL_463167, EPI_ISL_463168, EPI_ISL_463169, EPI_ISL_463170, EPI_ISL_463171, EPI_ISL_463172, EPI_ISL_463173, EPI_ISL_463174, EPI_ISL_469171, EPI_ISL_469172, EPI_ISL_469173, EPI_ISL_469174, EPI_ISL_469175, EPI_ISL_469176, EPI_ISL_469177, EPI_ISL_469178, EPI_ISL_469179, EPI_ISL_469180, EPI_ISL_469181, EPI_ISL_469182, EPI_ISL_469183, EPI_ISL_469184, EPI_ISL_469185, EPI_ISL_469186, EPI_ISL_469187, EPI_ISL_469188, EPI_ISL_469189, EPI_ISL_469190, EPI_ISL_469191, EPI_ISL_469192, EPI_ISL_469193, EPI_ISL_469194, EPI_ISL_469195, EPI_ISL_469196, EPI_ISL_469197, EPI_ISL_469198, EPI_ISL_469199, EPI_ISL_469200, EPI_ISL_469201, EPI_ISL_469202, EPI_ISL_469203, EPI_ISL_469204, EPI_ISL_469205, EPI_ISL_469206, EPI_ISL_469207, EPI_ISL_469208, EPI_ISL_476547, EPI_ISL_476548, EPI_ISL_476549, EPI_ISL_476550, EPI_ISL_476551, EPI_ISL_476552, EPI_ISL_476553, EPI_ISL_476554, EPI_ISL_476555, EPI_ISL_476556, EPI_ISL_476557, EPI_ISL_730062, EPI_ISL_730063, EPI_ISL_730064, EPI_ISL_730065, EPI_ISL_730066, EPI_ISL_730067, EPI_ISL_730068, EPI_ISL_730069, EPI_ISL_730070, EPI_ISL_730071, EPI_ISL_730072, EPI_ISL_730073, EPI_ISL_730074, EPI_ISL_730075, EPI_ISL_730076, EPI_ISL_730077, EPI_ISL_730078, EPI_ISL_730079, EPI_ISL_730080, EPI_ISL_730081, EPI_ISL_730082, EPI_ISL_730083, EPI_ISL_730084			
see above	Yale Clinical Virology Laboratory	Grubaugh Lab - Yale School of Public Health	Adam Moore; Akiko Iwasaki; Albert Ko; Alice Lu; Allison Nelson; Anderson Brito; Anne Wyllie; Annie Watkins; Arnau Casanovas; Catherine Muenker; Chaney Kalinich; Chantal Vogels; Charlese Dela Cruz; Cole Jensen; Isabel Ott; Joseph Fauver; Maria Tokuyama; Mary Petrone; Nathan Grubaugh; Patrick Wong; Peiwen Lu; Richard Martinello; Saad Omer; Shelli Farhadian; Tara Alpert
EPI_ISL_738290, EPI_ISL_738291, EPI_ISL_738292, EPI_ISL_738293, EPI_ISL_738294, EPI_ISL_738295, EPI_ISL_738296, EPI_ISL_738297			
see above	Yale Pathology Lab	Grubaugh Lab - Yale School of Public Health	Chantal Vogels; Danielle Plank; Ellen Foxman; Isabel Ott; Joseph Fauver; Mary Petrone; Nathan Grubaugh; Shaili Gupta; Tara Alpert

Supplemental Table 3: GISAID acknowledgements for phylogenetic analysis