



RESEARCH ARTICLE

In vitro selection of Remdesivir resistance suggests evolutionary predictability of SARS-CoV-2

Agnieszka M. Szemiel¹[✉], Andres Merits², Richard J. Orton¹[✉], Oscar A. MacLean¹[✉], Rute Maria Pinto¹[✉], Arthur Wickenhagen¹[✉], Gauthier Lieber¹[✉], Matthew L. Turnbull¹[✉], Sainan Wang², Wilhelm Furnon¹, Nicolas M. Suarez¹[✉], Daniel Mair¹[✉], Ana da Silva Filipe¹, Brian J. Willett¹[✉], Sam J. Wilson¹[✉], Arvind H. Patel¹[✉], Emma C. Thomson¹[✉], Massimo Palmarini¹[✉], Alain Kohl¹, Meredith E. Stewart¹[✉]*

1 MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom, **2** Institute of Technology, University of Tartu, Tartu, Estonia

 These authors contributed equally to this work.

 Current address: Institut für Medizinische Virologie, Zurich, Switzerland

* Meredith.Stewart@glasgow.ac.uk



OPEN ACCESS

Citation: Szemiel AM, Merits A, Orton RJ, MacLean OA, Pinto RM, Wickenhagen A, et al. (2021) *In vitro* selection of Remdesivir resistance suggests evolutionary predictability of SARS-CoV-2. PLoS Pathog 17(9): e1009929. <https://doi.org/10.1371/journal.ppat.1009929>

Editor: Shin-Ru Shih, Chang Gung University, TAIWAN

Received: April 2, 2021

Accepted: August 30, 2021

Published: September 17, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.ppat.1009929>

Copyright: © 2021 Szemiel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: SARS-CoV-2Eng12 was supplied under a MTA between The University of Glasgow and Public Health England. Consensus sequences and raw FASTQ files have been

Abstract

Remdesivir (RDV), a broadly acting nucleoside analogue, is the only FDA approved small molecule antiviral for the treatment of COVID-19 patients. To date, there are no reports identifying SARS-CoV-2 RDV resistance in patients, animal models or *in vitro*. Here, we selected drug-resistant viral populations by serially passaging SARS-CoV-2 *in vitro* in the presence of RDV. Using high throughput sequencing, we identified a single mutation in RNA-dependent RNA polymerase (NSP12) at a residue conserved among all coronaviruses in two independently evolved populations displaying decreased RDV sensitivity. Introduction of the NSP12 E802D mutation into our SARS-CoV-2 reverse genetics backbone confirmed its role in decreasing RDV sensitivity *in vitro*. Substitution of E802 did not affect viral replication or activity of an alternate nucleoside analogue (EIDD2801) but did affect virus fitness in a competition assay. Analysis of the globally circulating SARS-CoV-2 variants (>800,000 sequences) showed no evidence of widespread transmission of RDV-resistant mutants. Surprisingly, we observed an excess of substitutions in spike at corresponding sites identified in the emerging SARS-CoV-2 variants of concern (i.e., H69, E484, N501, H655) indicating that they can arise *in vitro* in the absence of immune selection. The identification and characterisation of a drug resistant signature within the SARS-CoV-2 genome has implications for clinical management and virus surveillance.

Author summary

The emergence of SARS-CoV-2 has led to a worldwide pandemic with significant morbidity and mortality. Remdesivir is the only antiviral with FDA approval for treatment. Antivirals use comes at a risk, as viruses may acquire mutations to overcome the inhibition. We identified a mutation in the virus polymerase responsible for decreased sensitivity to

uploaded to GenBank under BioProject number PRJNA692078. We used the publicly available CoV-Glue database (<http://cov-gluе.cvr.gla.ac.uk/#/home>) to examine for replacements at specific sites observed in the GISAID hCoV-19 sequences. All relevant data within the manuscript and its supporting information are available on Enlighten: Research Data (doi.org/10.5525/gla.researchdata.1184).

Funding: This work was supported by the UK Medical Research Council (MC_UU_12014/8, MC_UU12014/2 and MC_UU_12014/12). RJO was funded by MC_UU_12014/12, AM & SW were funded by European Regional Development Fund, Centre of Excellence in Molecular Cell Engineering, Estonia (2014-2020.4.01.15-013), AMS was funded by UKRI/DHSC (BB/R019843/1) and MES by European Commission Horizon2020 (H2020-EU.3.2.1.1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Remdesivir. A change at this conserved site was not predicted, and the mutation did not cause a replication advantage or change in sensitivity to another antiviral drug. Importantly, this change occurred at very low frequency globally. Unexpectedly, passage of SARS-CoV-2 led to an accumulation of mutations in spike. A number occurred at the same sites but to different residues as those in emerging variants of concern indicating they arise in the absence of immune pressure. Our data indicates low-level Remdesivir resistance in SARS-CoV-2 is different to other RNA viruses and monitoring changes *in vitro* provides insight into general virus adaptation of newly emerging viruses.

Introduction

The COVID-19 pandemic has caused more than 4 million deaths by July, 2021 [1] and placed the global economy under considerable strain [2]. The global effort to repurpose antiviral inhibitors and anti-inflammatory compounds to stem virus replication and clinical pathology identified Remdesivir (RDV or Veklury), a broadly acting nucleoside analogue, as a frontline treatment for patients hospitalized with severe acute respiratory syndrome virus-2 (SARS-CoV-2), the infectious agent underlying COVID-19. It is the only small molecular inhibitor that has FDA approval for the treatment of hospitalized SARS-CoV-2 patients. RDV exhibits a potent ability to restrict highly pathogenic human coronaviruses including severe acute respiratory syndrome virus (SARS-CoV), Middle eastern respiratory virus (MERS-CoV) as well as SARS-CoV-2 and Ebola virus (EBOV) replication *in vitro* [3,4] and in pre-clinical animal models [4–6]. Three randomized trials [7–9] demonstrated that RDV treatment reduced recovery time by 31% and demonstrated a non-significant trend towards lower mortality, thus reducing long-term healthcare costs. This trend of reduced hospitalization time and decreased morbidity was further supported by smaller non-randomized studies [10]. Conversely, a larger trial conducted by WHO (Solidarity Therapeutics Trial) reported no effect on patient survival [11]. The timing of administration of RDV appeared to be critical for its efficacy [4–6]. Despite these findings, countries including the USA and UK routinely use RDV for the treatment of hospitalized SARS-CoV-2 patients requiring oxygen who are still within the virological phase of infection (<10 days of illness). RDV is often prescribed in combination with dexamethasone, a steroid treatment, which reduces mortality in ventilated patients [12,13]. However, RDV and dexamethasone have yet to be trialed in combination. Recently RDV use in combination with Baricitinib has shown great promise [14].

Viruses often adapt and mutate to become resistant to antiviral therapy and this can affect patient and disease management. This is exemplified by viruses including human immunodeficiency virus type 1, hepatitis C virus, and influenza A which have all shown the ability to develop resistance during single drug use therapies [15–18]. Currently, there are no reports of circulating RDV-resistant strains of SARS-CoV-2. We are reliant on models based on studies in murine hepatitis virus (MHV), severe acute respiratory syndrome virus (SARS-CoV) and Ebola virus (EBOV) [19–21] in order to predict the changes in amino acid residues that could confer drug resistance. Given the global threat presented by SARS-CoV-2, it is important to determine whether SARS-CoV-2 can become resistant to RDV, identify which mutations confer resistance, monitor the emergence of such variants in the population and adapt treatments in COVID-19 patients.

Results

Change in SARS-CoV-2 RDV susceptibility after serial passage

After determining optimal culture conditions (S1 Fig), SARS-CoV-2_{Engl2} was passaged serially in either 1 μ M or 2.5 μ M RDV-supplemented media for 13 passages (SARS-CoV-2_{Engl2} was isolated in February 2020; Fig 1A). Viruses serving as controls were passaged in parallel in either DMSO or media to monitor for mutations due to cell culture adaptation unrelated to RDV resistance. In total, we passaged SARS-CoV-2_{Engl2} in parallel in 24 distinct cultures with different selective pressures (4 different conditions and 2 different virus inputs; Fig 1A). We monitored for cytopathic effect (CPE) during passaging of the cultures. CPE was observed in 7 of the 12 lineages passaged in RDV, with the loss of 5 lineages between p1 and p4 (Fig 1A and 1B). We observed general adaptation of the viruses to VeroE6 cells, with an increase in overall viral titers by 0.5 to 1 log₁₀ (Fig 1B) as well as a change in plaque phenotype (S2A Fig) after 13 passages.

Next, the replication kinetics and change in the half maximal effective concentration (EC₅₀) of RDV in a subset of passaged virus populations (Rem2.5p13.5, DMSOp13.5 and Mediap13.4) were assessed. Rem2.5p13.5 alone actively replicated in the presence of 7.5 μ M RDV (Fig 1C). However, titers in the presence of RDV were lower than those grown in the absence of RDV. Titers of control viruses, DMSOp13.5 and Mediap13.4 were consistently 5 log₁₀ lower when grown in the presence of RDV (Fig 1C). The Rem2.5p13.5, DMSOp13.5 and Mediap13.4 lineages displayed similar replication kinetics when cultured in the absence of RDV (Fig 1C). Furthermore, Rem2.5p13.5 lineage required a greater concentration of RDV to protect infected cells (S2B Fig) corresponding to a 2- to 2.5-fold increase in RDV EC₅₀ over a range of virus inputs in comparison with DMSOp13.5, and Mediap13.4 (S2C Fig). The partial resistance to a nucleoside analogue was specific for RDV, as we observed a minimal change in EC₅₀ of a second nucleoside analogue (EIDD2801), when comparing Rem2.5p13.5 (EC₅₀ ~9.1 μ M) to SARS-CoV-2_{Engl2} (EC₅₀ ~8.9 μ M) (Figs 1D and S3C and S3D and S1 Text).

Subsequent analyses identified a second lineage, Rem1p13.5 with reduced sensitivity to RDV (Fig 1E). The EC₅₀ of Rem1p13.5 (~0.83 μ M) was comparable to Rem2.5p13.5 (~0.88 μ M), this was a 3.5- to 3.7-fold increase in EC₅₀ from the parental virus (EC₅₀~0.23 μ M). The RDV EC₅₀ for virus passaged in either media alone (EC₅₀~0.29–0.32 μ M) or DMSO (EC₅₀~0.12–0.22 μ M) corresponded with EC₅₀ for the parental stock virus (Fig 1F). The changes in RDV sensitivity paralleled those previously reported for MHV, SARS-CoV and EBOV resistant viruses [4,20].

Common mutations in two virus populations with partial resistance to RDV

Direct comparison of the consensus sequences from a subset of viruses passaged in different conditions with original SARS-CoV-2_{Engl2} revealed two fixed non-synonymous mutations in lineages with decreased RDV susceptibility of at least 2-fold change in EC₅₀ in two independently generated populations (Rem1p13.5 & Rem2.5p13.5). These mutations were not present in viruses either passaged in absence of RDV, or the input virus (SARS-CoV-2_{Engl2}) or SARS-CoV-2_{Wu1} (MN908947) (S1 Data). The first mutation was identified as glutamine to aspartate at amino acid 802 (E802D) in the RNA-dependent RNA polymerase (RdRp) NSP12 (Fig 2A). A glutamate at this position is highly conserved between all betacoronaviruses including SARS-CoV, MERS-CoV and unclassified sarbecoviruses (Fig 2B and S2A Data). The E802 mutation occurs within the palm sub-domains (T680 to Q815; Fig 2A), and in proximity to amino acids predicted to interact with newly synthesized RNA (C813, S814 and Q815 [22];

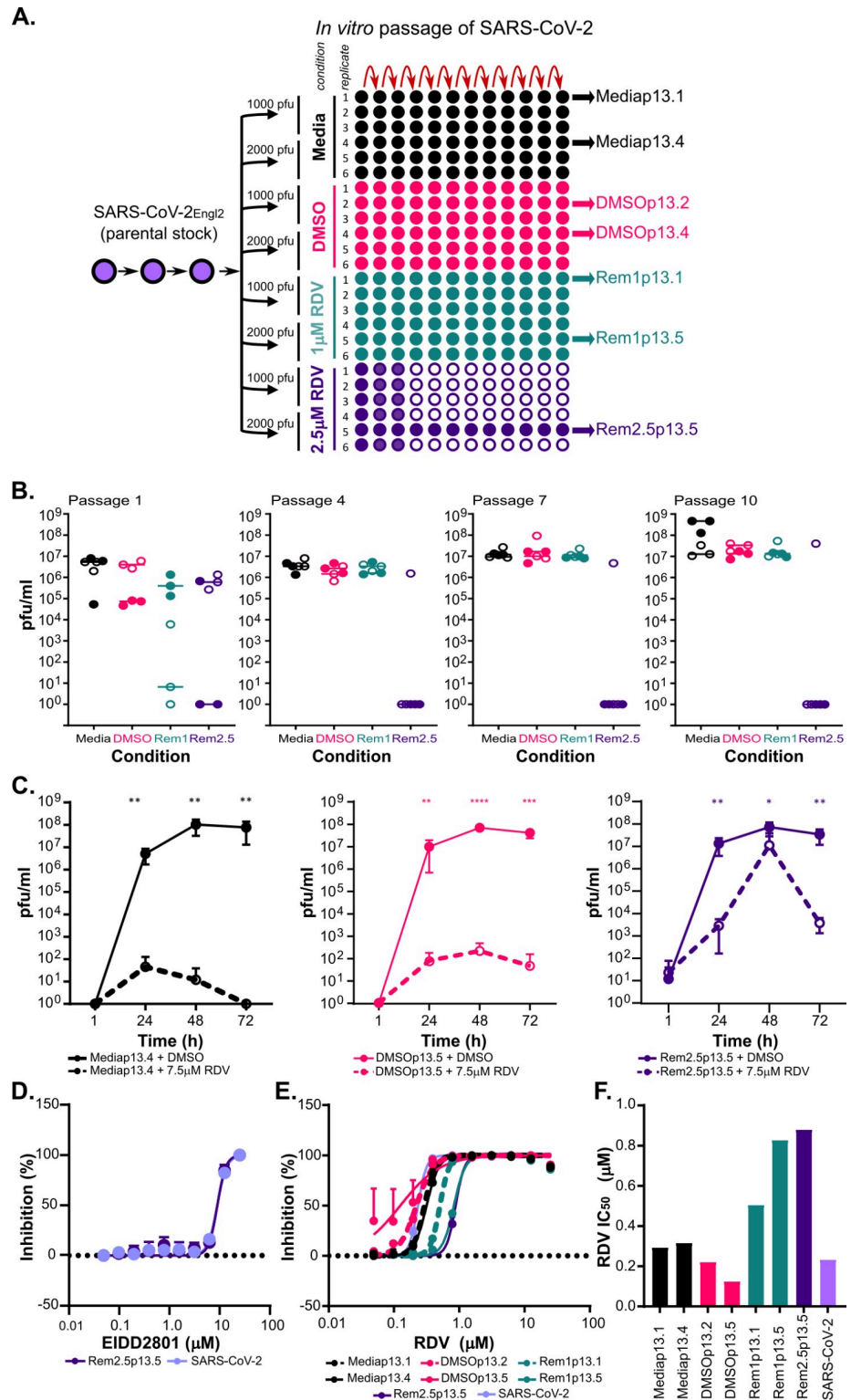


Fig 1. Continuous passage of SARS-CoV-2_{Engl2} in RDV selects for partial resistant populations. (A) Schematic of the experimental layout to select for RDV resistant viruses. The passage (p) number of the input virus SARS-CoV-2_{Engl2} is given. We consider p1 as the stock supplied by Public Health England, it was passed twice in VeroE6 prior to selection in RDV. SARS-CoV-2_{Engl2} p3 is the input stock that was sequenced and used for analysis. Each condition is given a different colour and the amount of virus used at the start of the experiment is indicated. Populations that failed to amplify (no CPE or/and virus titre) are indicated by the empty circles. All lineages sequenced are indicated. (B)

Virus titers (pfu/ml) at p1, p4, p7 and p10. 6 lineages per condition and two different virus inputs; 1000 pfu (solid circle) and 2000 pfu (open circle). Median for each is shown. (C) Virus growth kinetics in VeroE6 in the presence (dashed line) or absence (solid line) of 7.5µM RDV for 3 different virus populations. Data is from 2 independent experiments with 3 replicates. Error bars represent SEM. Unpaired t-tests (Holm-Šidák method; *, P < 0.05; **, P < 0.01; ***, P < 0.001. ****, P < 0.0001). (D) EIDD2801 dose dependency curve. EIDD2801 treated VeroE6-ACE2-TMPRSS2 infected with 8400 pfu/ml of each virus. (E) RDV dose dependency curves determined in A549NPro-ACE2 infected with 8400 pfu/ml of each virus. (F) Bar graph of RDV EC₅₀ for different viruses in A549NPro-ACE2 with 8400 pfu/well. For all panels, error bars represent SEM.

<https://doi.org/10.1371/journal.ppat.1009929.g001>

Fig 2A). We propose that the E802D mutation changes the steric interactions between amino acid side chains within this region resulting in minor structural changes, (Fig 2A), thereby influencing binding of nt+3 during synthesis of template RNA and allowing elongation when the active form of RDV is incorporated into the RNA. The mutation identified in NSP12

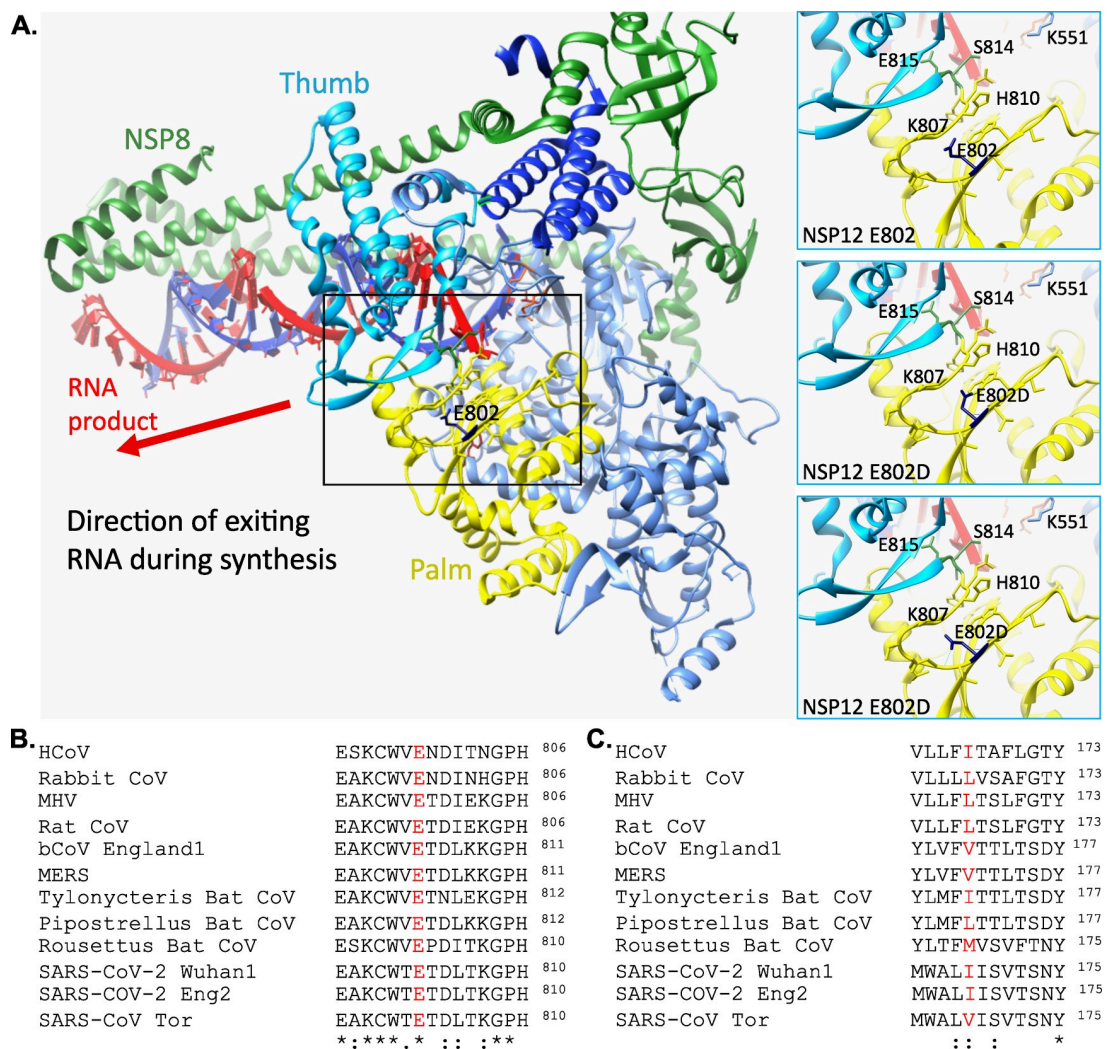


Fig 2. Common mutations in partial RDV resistance populations. (A) Location of E802 within structure of SARS-CoV-2 NSP12 in association with NSP7 and NSP8 (PDB ID 6YYT). Three focused panels are WT (upper) and two potential confirmations of E802D. H-bonds are indicated by light blue line. (B) Conservation of E802 amino acid across coronaviruses. Accession numbers for the coronavirus sequences are in the materials and methods. (C) NSP6 I168 amino acid is not conserved across coronaviruses.

<https://doi.org/10.1371/journal.ppat.1009929.g002>

differs from amino acid residues involved with decreased RDV sensitivity in EBOV, other betacoronaviruses, (MHV, & SARS-CoV), and sites predicted in SARS-CoV-2 [19–21].

The second mutation was an isoleucine to threonine substitution (I168T) in NSP6, a highly conserved protein involved in restricting autophagosome expansion [23]. This site is not conserved across coronaviruses with either an isoleucine (SARS-CoV-2), or valine (SARS-CoV & MERS-CoV) or leucine (MHV) in this position (Fig 2C and S2 Data). We predict this mutation may alter the structure of the transmembrane and extracellular domains (S4 Fig).

A single mutation in NSP12 confer partial resistance to RDV in rSARS-CoV-2_{Wu1}

To ascertain whether a mutation of NSP12 E802 was sufficient to mediate partial RDV resistance, we introduced either an E802D or E802A mutation at this site into SARS-CoV-2_{Wu1} by reverse genetics and recovered infectious virus. Whilst we thought it was unlikely to play a role, we also recovered virus with I168T mutation in NSP6 either alone or in combination with the NSP12 mutations (E802D or E802A). All rescued viruses with specific point mutations replicated similarly to the rescued wild-type virus (rSARS-CoV-2) in human lung cells, Calu-3, with comparable replication kinetics and achieving similar peak virus titers (Fig 3A; rSARS-CoV-2 vs rNSP12-E802D, $p_{\text{adj}} = 0.44$; rSARS-CoV-2 vs rNSP12-E802A, $p_{\text{adj}} = 0.41$; rSARS-CoV-2 vs rNSP6-I168T, $p_{\text{adj}} = 0.51$; rSARS-CoV-2 vs rNSP6-I168T+NSP12-E802A, $p_{\text{adj}} = 0.50$, rSARS-CoV-2 vs rNSP6-I168T+NSP12-E802D, $p_{\text{adj}} = 0.50$; RM One-Way Anova with Dunnett's Test). This indicates that these mutations did not provide a replicative advantage in a representative cell-line without drug exposure. In our drug screen assays, both the E802D and E802A mutations in NSP12 recapitulated partial resistance observed in the virus populations continually passaged in RDV (Fig 3B). We observed a 2.14- to 2.54-fold change in RDV EC_{50} ; from 2.61 μM for rSARS-CoV-2 to 5.58 μM and 6.62 μM for the E802A and E802D mutants, respectively (Fig 3B and S1 Table). This change in RDV sensitivity was evident over a range of virus inputs for both NSP12 mutants (S5A Fig). NSP6 I168T substitution did not confer decreased sensitivity to RDV (Fig 3B), with comparable EC_{50} values to rSARS-CoV-2 (S1 Table).

Indeed, viruses bearing both the NSP12 and NSP6 mutations were more sensitive to RDV in comparison to NSP12 single mutant viruses (rNSP6-I168T+NSP12-E802D, EC_{50} 3.21 μM ; rNSP6-I168T+NSP12-E802A, EC_{50} 3.89 μM). Importantly, introduction of NSP6 and/or NSP12 mutations did not significantly affect sensitivity to an alternate nucleoside analogue EIDD2801 (S1 Table and S5B Fig). These data confirm results obtained with other viruses demonstrating EIDD2801 sensitivity was not influenced by mutations conferring decreased RDV sensitivity [20,24]. We further assessed the anti-viral activity of RDV in Calu-3. While a RDV dose-dependent reduction in titer for all viruses was observed, rNSP12-E802D and rNSP12-E802A titers were consistently higher than wild-type and rNSP6-I168T at 24h and 48h pi (Fig 3C) over a range of concentrations. Interestingly, at 24h pi, a slight shift in an increase rNSP6-I168T infectious titer was observed in comparison with wild type, though this effect disappeared by 48h.

To examine whether the viruses containing the NSP12 E802D mutation would persist and outcompete the WT in the absence of RDV, an *in vitro* co-infection competition assay over multiple passages was undertaken. VeroE6-ACE2-TMPRSS2 were infected with two different MOI ratios of rSARS-CoV-2 to rNSP12-E802D (1:9 and 9:1), samples were harvested every 24 h and used to infect fresh monolayers of cells. This was repeated for 3 passages. While both stocks had high infectious titres (rSARS-CoV-2, $\sim 2.5 \times 10^7$ pfu/ml and rNSP12-E802D, $\sim 8.13 \times 10^7$ pfu/ml), there was a difference in the equivalent genome copies (GE)/ml, with

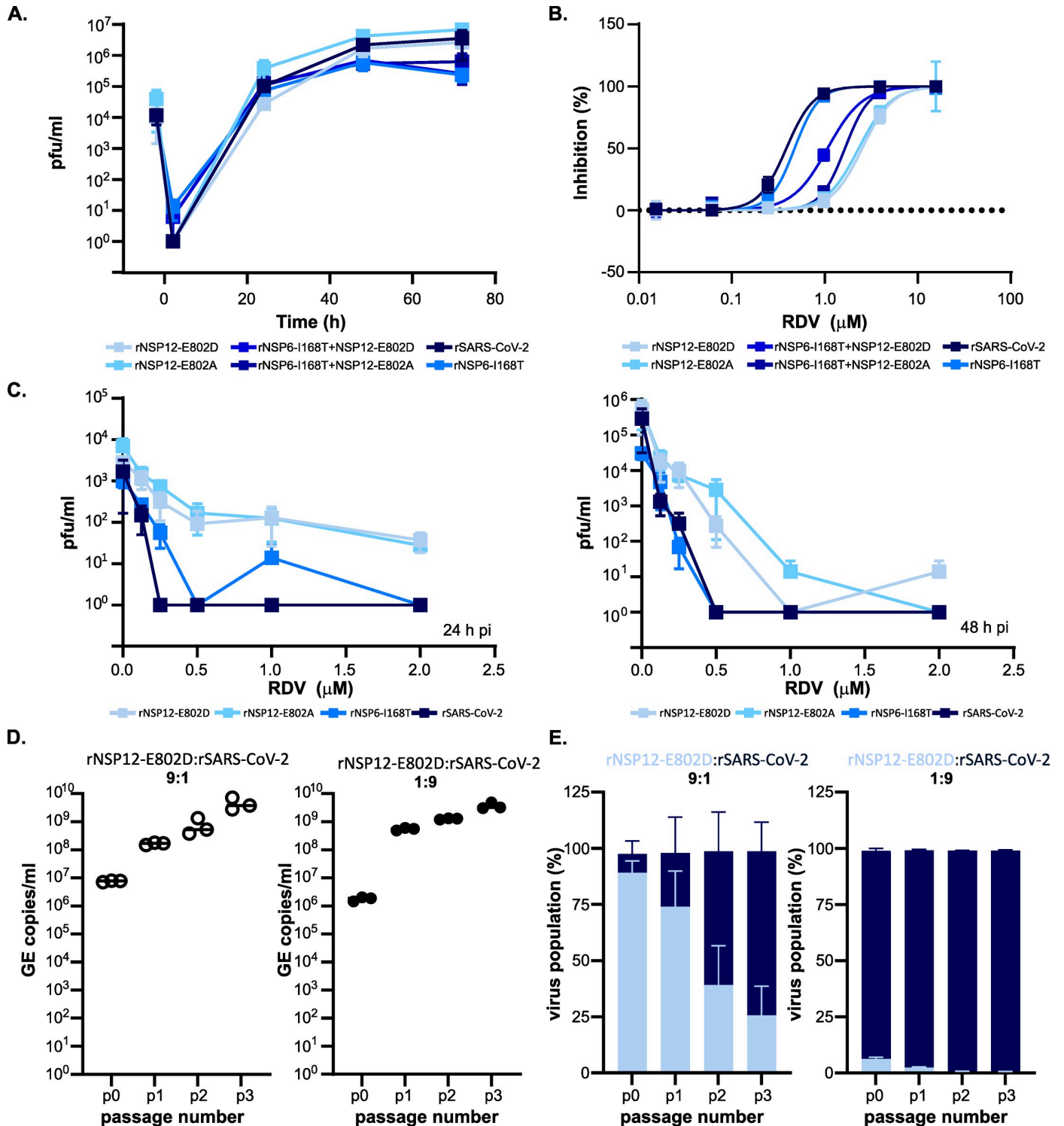


Fig 3. NSP12 E802 mutation recapitulates change in RDV susceptibility. All viruses were derived by reverse genetics and have a SARS-CoV-2_{Wu1} backbone with specific point mutations as indicated. (A) Virus replication kinetics of rescued viruses with single mutation in either NSP12, NSP6 or both in Calu-3. Data is from 3 independent experiments with 3 replicates, there was no significant difference between growth of the mutants versus the wild-type rSARS-CoV-2. (B) RDV dose-dependent inhibition for each mutant virus. Post-treatment with RDV, VeroE6-ACE2-TMPRSS2 were infected with 500 pfu/ml of each virus. Mutations in NSP12 decrease the sensitivity to RDV. (C) RDV dose effect on virus titers at 24 h (left) and 48 h (right). Calu-3 were treated with decreasing doses of RDV and infected with an MOI~0.01. Data from 2 independent virus stocks with 2 replicates except for rSARS-CoV-2 and rNSP12-E802A. All error bars are SEM. (D) Co-infection competition assay virus titres for the input (p0) and each subsequent passage (p1, p2 and p3) expressed as genome copies/ml. VeroE6-ACE2-TMPRSS2 were

infected with different ratios of rNSP12-E802D to rSARS-CoV-2. The assay does not discriminate between rSARS-CoV-2 and rNSP12-E802D. Data from 3 independent infections. (E) Co-infection competition assay with different input ratios (1:9 and 9:1) of rSARS-CoV-2 WT and rNSP12E802D in VeroE6-ACE2-TMPRSS2. The percentage of each virus population over successive passages is shown (Data from 3 independent infections).

<https://doi.org/10.1371/journal.ppat.1009929.g003>

rNSP12-E802D a $\sim 0.5 \log_{10}$ higher (rSARS-CoV-2, 7.16×10^8 GE/ml and rNSP12-E802D, 2.90×10^9 GE/ml). Therefore, we used different ratios of MOI rather than GE/ml equivalents for the initial infection. We monitored for virus replication by qRT-PCR and found an increase in titres at each passage (Fig 3D). We next examined the change in the ratio of wild type to E802D in the viral populations at each passage using the GridION (Oxford Nanopore Technologies). Regardless of input ratio, rSARS-CoV-2 out competed viruses containing the rNSP12-E802D mutation (Fig 3E). This would indicate the E802D mutation incurs a fitness cost in the absence of RDV and is unlikely to be sustained in the circulating population.

Prevalence of RDV mutations in circulating populations

We next examined the available SARS-CoV-2 genome sequences in the GISAID database (n = 844376 as of March 2021) to determine the frequency of replacements at NSP12 E802 and NSP6 I168. We only identified 36 (0.004%) viral sequences with a mutation at E802; E802D was the predominant replacement (n = 24), with the remaining sequence either E802A (n = 4), E802G (n = 3), E802Q (n = 1) or E802X (n = 1). The sequences containing these mutations were mostly geographically and temporally dispersed suggesting they did not share homology, though there was a small cluster of sequences (E802D, n = 9) from Germany within the same time frame. The 4 sequences with E802A substitution from May 2020 from the same geographic region (United Arab Emirates) but there is no evidence of further transmission. As one of these sequences, hCoV-19/Scotland/CVR2716/2020 was isolated from a patient who was not treated with RDV, this suggests mutation of E802 can arise in the community in the absence of drug selection.

The same pattern of very low observed global frequency of the E802 substitutions was also found at sites known to confer partial RDV resistance in other coronaviruses [20]. There were a handful of sequences with changes at either F480 (n = 13) or V557 (n = 35), and as with the E802 mutations, the sequences were mainly geographically and temporally dispersed. There was a cluster of sequences containing the V557I substitution observed in the Netherlands within a week but there was no information on whether it was the same patient or whether they were treated with RDV. When examining these mutations in respect to the global amino acid diversity along in NSP12 sequence, the region around E802 shows greater diversity than the area around F480 or V557 (S6A Fig). This suggests that strong purifying selection is acting in the regions of these sites, and they maybe more constrained by other dimensions of selection *in vivo*.

Replacement of NSP6 I168 occurred in 347 sequences with isoleucine replaced with threonine, valine, leucine, glycine or methionine. These data indicate that in absence of widespread selective pressure, substitutions of either NSP12 E802 or NSP6 I168 are currently rare global events. However, the identification of these sequences in the genome databases demonstrate that these viruses are viable and could potentially acquire a resistant phenotype when a stronger selective pressure is applied.

In vitro mutations provide insight into spike adaption

We next focused on those mutations arising in the *in vitro* passaged virus populations that were likely not directly linked to RDV resistance. The consensus sequences of a subset of the

passed stocks displayed a total of 41 distinct non-synonymous mutations and 10 synonymous mutations across the genome compared to the parental SARS-CoV-2_{Engl2} sequence (Fig 4A). Importantly, we did not observe any previously identified mutations in the proof-reading ExoN (NSP14) that would change the sensitivity of the virus to RDV (Fig 4A and S1 Data). Deletions of ExoN have been demonstrated to increase RDV sensitivity for other coronaviruses [20]. While there was clear positive selection pressure across the entire genome (S2 Table), there were no major differences in the number of mutations that accumulated in any specific population, and in the ratio or type of transition vs transversion change (Figs 4B and S6B). Although, Rem2.5p13.5 displayed a slight elevation in non-synonymous changes (Fig 4C), we are unable to draw conclusions on the effect of RDV concentration on virus mutation rate due to small number recovered populations selected in RDV (S1 Text). Further comparative analysis over with the SARS-CoV-2_{Wu1} sequence identified a further 2 synonymous and 8 non-synonymous mutations present in the original SARS-CoV-2_{Engl2} population (Fig 4A). SARS-CoV-2_{Engl2} was a 50:50 mix of two virus populations with 5 of the mutations present at a frequency ~50%, all but one of these became fixed in all passaged populations by p13 (Fig 4A).

Most of the collective 22 mutations observed in all sequences occurred within the spike (S) open reading frame. All of the *in vitro* non-synonymous substitutions appeared within multiple lineages the circulating viruses (GISAID data accessed March 2021, S3 Data) with the exception of I68R, N709H and D985G. Unlike observations in other studies [25–27], the furin-like cleavage site was preserved in all but one of the passaged populations. DMSOp13.2 displayed a 24nt deletion of the entire furin-like cleavage site at a high but unfixed frequency of 77% after 13 passages (S1 Data). Importantly, in our *in vitro* passaged viruses we observed both synonymous and non-synonymous substitutions occurring at the same sites within spike (H69R, E484D, N501T, H655Y, P681P) as those identified in the emerging SARS-CoV-2 variants of concern (Alpha (B.1.1.7): Δ69/70, N501Y, P681H; Gamma (P.1): E484K, N501Y, H655Y; Beta (B.1.351): E484K, N501Y) (Fig 4). Except for synonymous P681P, these substitutions were not present in SARS-CoV-2_{Engl2} (S1 Data and S1 Text).

It is important to stress these emerging variants of concern, collectively, share a combination of three amino acid mutations in spike receptor binding domain (RBD): N501Y common to all and K417N and E484K in the Brazil and South African variants. We observe substitutions at 2 out of 3 of these sites in our *in vitro* evolution lines. The probability of such a large overlap (5 codons) between the substitutions observed *in vitro*, and variants of concern defining mutations without a common selective pressure driving convergence, was exceptionally small ($P = 3.1 \times 10^{-5}$; S6C Fig). This demonstrates commonality in the fitness landscape that these *in vitro* populations and the circulating lineages are evolving under (driving evolutionary convergence). We further examined the global distribution of all circulating amino acid replacements within spike to determine whether our *in vitro* substitutions occurred within hot spots for change. There were 1384 replacements observed in a minimum of 5 sequences ($n = 242865$ sequence up dated 14th December 2020), many of these were clustered into certain regions within spike, creating visible hot spots of diversity (Fig 4E). For example, the window surrounding amino acid E484 appears to be a relative hot spot for replacement.

Of note are the four spike substitutions which occurred within the receptor binding domain (RBD) at either consensus (E484D, & N501T) or sub-consensus (G413R & Q498H) frequency. The E484 substitution appeared in the consensus sequence of Rem2.5p13.5 and Remp1p13.1 (Fig 4A), but it was present at a frequency of 20–40% in all the other viruses apart from DMSOp13.2 (S1 Data). The N501 substitution was present in one virus at consensus (Mediap13.1) and the sub-consensus of a second (Rem2.5p13.5) (Fig 4D). We noted that some *in vitro* substitutions in spike were changes to amino acids associated with SARS-CoV-1 spike

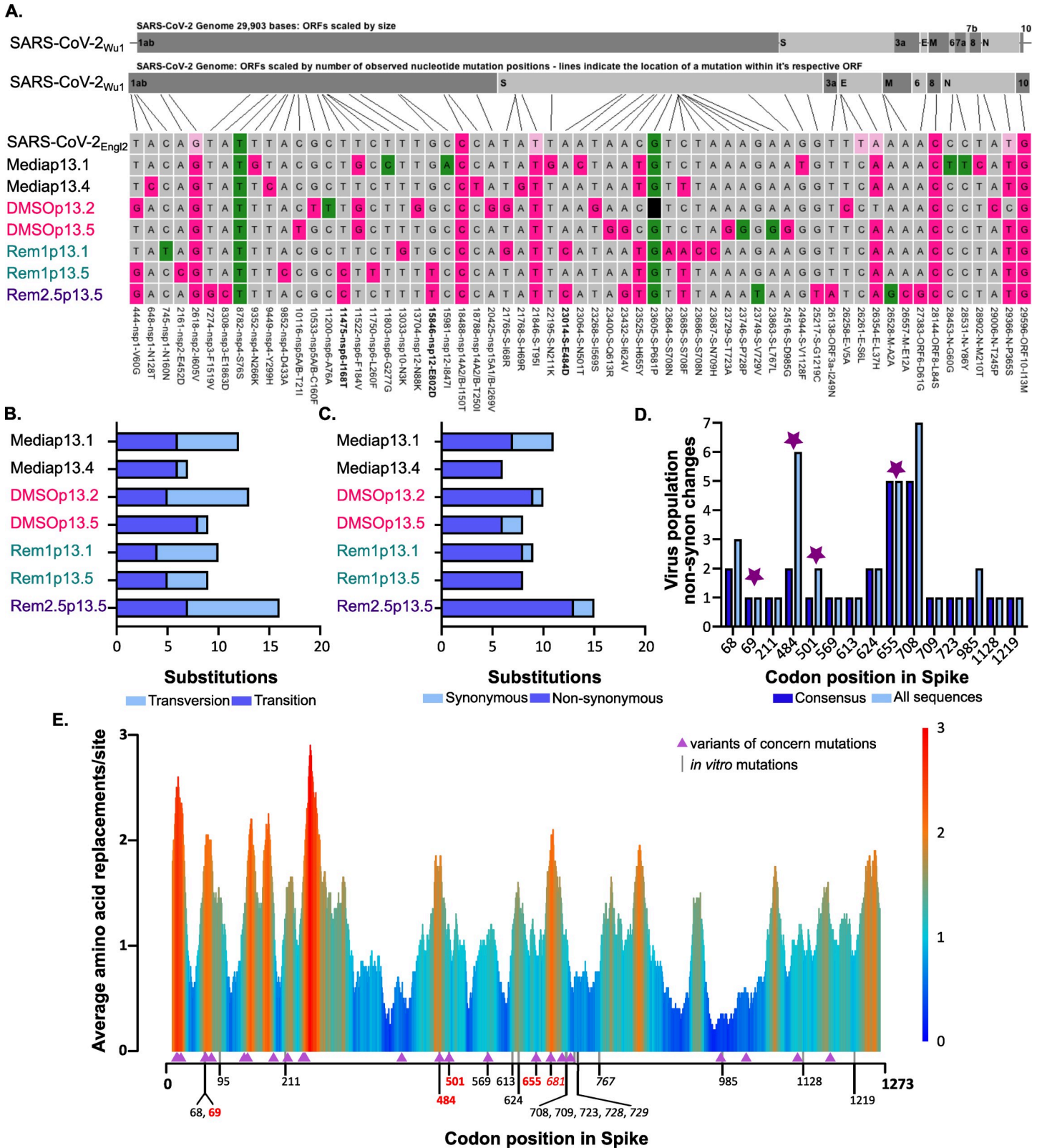


Fig 4. Sequence analysis of continuous *in vitro* passaged SARS-CoV-2_{Engl2}. (A) Alignment of serially passaged viruses and SARS-CoV-2_{Engl2} to SARS-CoV-2_{Wu1}. Non-synonymous (pink) and synonymous (green) changes from Wuhan-1 are highlighted. Light pink are sites fixed at 50% in SARS-CoV-2_{Engl2} and black box is a deletion mutation in serially passaged virus. Positions of mutations are indicated, and mutations only found in RDV selected populations are in bold. (B) Synonymous vs non-

synonymous changes observed in continually passaged virus populations compare to input SARS-CoV-2_{Engl2}. (C) Transversion vs transitional changes observed in continually passaged virus populations compare to input SARS-CoV-2_{Engl2}. (D) Number of *in vitro* passaged viruses with non-synonymous changes in Spike in comparison to SARS-CoV-2_{Engl2}. Mutation fixed in the consensus genomes (dark blue) are compared to the total number of viruses with evidence of the mutation at sub-consensus levels (light blue). Amino acid residues in common with the emerging variants of concern (Alpha (UK, B.1.1.7), Gamma (Brazil, P.1); and Beta (South Africa, B.1.351) are highlighted by a star. (E) Worldwide diversity of Spike protein sites of circulating SARS-CoV-2 variants. The average number of different substitutions at each codon is calculated along a 20 amino acid residues wide sliding windows. Data was calculated using only substitutions observed in a minimum of 5 sequences from the publicly available SARS-CoV-2 genomes (n = 1384). The position of the amino acid substitutions in the *in vitro* passaged viruses are indicated at the bottom, residues are red are shared with variants of concerns, black are specific to *in vitro* virus, residues in italics were synonymous. Mutations in amino acid residues that are also mutated in the variants of concern Alpha (UK, B.1.1.7), Gamma (Brazil, P.1), & Beta (South Africa, B.1.351) are shown in purple triangles.

<https://doi.org/10.1371/journal.ppat.1009929.g004>

(Q498H & N501T) adaptation. Our data shows these mutations arise *in vitro* in the absence of any immune selection suggesting intrinsic fitness increases unrelated to immune evasion.

Discussion

To our knowledge there are no reports identifying signatures within the genome of SARS-CoV-2 which lead to full, or even partial resistance to RDV. The change in sensitivity to RDV we observed was similar to single NSP12 mutation in either MHV or SARS-CoV [20] but lower than observed for the EBOV RdRp mutation [21]. It is of note that our partially-resistant RDV populations arose rapidly, with at least one lineage appearing within 4 passages (Rem2.5 lineage) and a second by passage 13 rather than 30 passages as observed for MHV, respectively [20,21]. However, RDV resistance mutations appeared in the EBOV genome by passage 9 but did not fix until passage 23 [21].

The potential for resistance to rapidly occur in RDV non-responding patients may be an issue that needs to be examined in order to discern whether it is due to a genomic mutation or drug tolerance by synchronization [28]. In a clinical setting, there has not been any evidence of RDV resistance arising in patients. Two studies showed clearance of viraemia after the second treatment but not after the first [29,30]. In the study reported by Kemp et al., [31] SARS-CoV-2 virus was never cleared or suppressed by any of the multiple treatments used and they speculated that a V157 mutation may be responsible for the resistance. This mutation has not been observed in coronaviruses including SARS-CoV-2 passaged in RDV *in vitro* or *in vivo* [20,21,29,32]. These studies followed the 10-day treatment RDV regime for the clearance of SARS-CoV-2 and the development of resistance may require longer treatment periods in patients before this phenotype becomes noticeable.

The *in vitro* mutations associated with partial RDV resistance for SARS-CoV-2 as well as those identified in MHV and SARS-CoV-1, circulate at extremely low frequency. This could in part be attributed to the decrease in fitness of coronaviruses with these RdRp mutations, as observed in competition assays results [20,21]. Indeed, we found that SARS-CoV-2 NSP12 E802D were outcompeted by wild type virus in our *in vitro* assays. This is unlike RdRp mutations in EBOV that conferred partial resistance and where the frequency was maintained in competition assays [21].

The action of nucleoside analogues to restrict coronaviruses is often thwarted by the unique proofreading activity encoded by the virus [20,33,34]. It has been demonstrated RDV functions by stalling the incorporation into the extending RNA product due to a steric clash with amino acid residue S861 *in vitro* [3,35,36]. However, it is possible for template-arrest caused by the incorporation of the nucleoside analogue and extension of the genome containing multiple RDV residues to occur [3,37]. The interplay between balance of naturally occurring dinucleotide concentration [3,37] and the cell's ability to metabolize RDV to its active form [4] are likely to influence whether break-through replication will occur. The SARS-CoV-2 NSP12 E802D mutation, similar to the NSP12 F480 mutation in MHV and SARS-CoV-1 does not fall

within the active sites, and we propose that this amino acid substitution results in minor structural modification of NSP12 that potentially influences the position of C813 and S814. These amino acids residues have a direct interaction with the newly formed RNA product [22,38,39] and recently have been shown to be critical for maintaining structural integrity of the NTP entry tunnel and reverse motion of the enzyme complex [40].

The SARS-CoV_{Engl2} RDV-resistant (Rem1p13.5 and Rem2.5p13.5) and the reverse-genetic derived SARS-CoV-2_{Wu1} NSP12 mutants (rNSP12_E802D & rNSP12_E802A) increased the EC₅₀ by at least 2-fold regardless of the cell type used for the experiments. Thus, we are confident that the change in EC₅₀ was not due to cellular drug metabolism or differences in virus entry and replication between wild-type and RDV-resistant viruses. We also noted that the cell-culture adaptation in viruses passaged in the absence of RDV resulted in a shift in EC₅₀ in VeroE6 based assays in comparison with input SARS-CoV-2_{Engl2} (S3A and S3B Fig) but this shift was not as predominant as the RDV-selected viruses. Difference in EC₅₀ due to adaptation, availability of receptors and ability to metabolize RDV is widely acknowledged [4,41]. We hypothesize that this was due to more efficient virus entry and spread as many of the mutations observed occurred within the spike protein.

Coronaviruses correct most errors that occur during genome replication. Therefore, if mutations arise and are fixed there must be a general benefit that could include inhibitor escape, fitness advantage [42–44], antigenic change to escape antibody restriction [45,46] and adaptation to the host [47]. There was no evidence within our data that RDV acts as a mutagen driving spontaneous mutations nor is it associated with increased rate of genome mutation with MHV and Ebola [20,21]. Other nucleoside analogues including ribavirin and EIDD2801 have been shown to act as mutagens [48,49], whereas the template stalling action of the RDV limits the spontaneous mutations and therefore the changes we observed *in vitro* are an evolutionary fine tuning of the viral proteins to a new host.

In our study, we also observe Spike mutations which evolved *in vitro* occurred disproportionately in the same codon sites as the variants of concern defining mutations. Additionally, many were found in localized hotspots of global diversity. We accessed the deep mutation scanning published data by Starr et al., [43] to examine whether these changes would influence RBD and affinity to ACE-2. In this study, only N501T and Q498Y changes observed in our data were associated with an increased affinity for ACE-2, while the E484D substitutions has a negative effect on ACE-2 binding in this system. Whereas, in an analysis using the entire spike protein there was an increase the free energy of the RBD-ACE-2 complex with the E484D substitution, resulting in a more infectious virus [50].

There is strong evidence based on circulating strains and *in vitro* sequences that the RDB of spike is under strong positive selection pressure [51,52]. Our *in vitro* observations suggest the predictability of evolution of the SARS-CoV-2 genome and could provide greater insight into the evolutionary pathway the virus is undergoing to adapt to its new host. The high ratio of non-synonymous to synonymous mutations we observed in spike (S2 Table) was driven by positive selection and not neutral evolution. Recently published studies have also shown the emergence of mutations in SARS-CoV-2 spike during *in vitro* passage at the same residues we identified including E484 and H655 [53,54]. Our data suggests that the parallel emergence of some mutations in the Spike in the variants of concern were not necessarily due to immune-based selection pressure. This is further highlighted by two mutations that arise *in vitro*; N501T and H655Y, the first within the RDB and the latter outside of the RDB and toward the S1-S2 cleavage site. Mutation of N501 increased affinity to human ACE-2 [43,45] and is involved in host tropism (i.e., SARS-CoV-1 N501 equivalent T487 mutation increase spike binding from civet to human ACE-2 [55,56]). H655Y has been observed by others *in vitro* [57], associated with adaptation to cats and hamsters [54,58] antibody escape [59] and, identified

in the sequence of super spreaders [60] as well as arising by convergent evolution in circulating lineages (eg., P1 and A.27 lineages). While the phenotypic importances of H655 mutation is unknown, Braun et al., [54] postulated it is unlikely H655Y improves spike fusion efficiency and entry. This may indeed be the case, as the H655 substitution was not present in the only passaged virus with the multi-basic cleavage site deletion (DMSOp13.2) present in our data set. While not in the scope of this paper, functional studies related to change observed during *in vitro* evolution experiments can help to identify candidate amino acid residues likely to drive future viral adaptation are recommended.

Unlike other studies using VeroE6 [25,26], we did not observe the rapid loss of the multi-basic (furin cleavage) site in SARS-CoV-2 spike *in vitro*. The cleavage of spike into the S1/S2 subunit is critical for virus entry and can be processed by multiple proteases including furin, serine protease (TMPRSS2/4) and the cysteine proteases Cathepsin B/L [61–63]. The pathways used for entry by coronaviruses is dependent on either TMPRSS2 or Cathepsin L expression [63–65]. VeroE6 are generally TMPRSS2 deficient but over express and process Procathepsin L to the active cathepsin L in abundance [64]. We postulate that this over expression of Cathepsin L allows for the maintenance of the multi-basic cleavage site in our continually passaged SARS-CoV-2 lineages. Other studies have also shown that the proteolytic cleavage site is maintained *in vitro* in continual passage for SARS-CoV-2 [66–68].

In summary, we have identified in our *in vitro* evolution studies a genome signature in SARS-CoV-2 which leads to replicative advantage in the presence of RDV. In the US, RDV treatment is currently prescribed to at least half of all hospitalized SARS-CoV-2 patients (as of January 2021 [69]) and was reportedly widely used during the second wave in India. Our data demonstrates that selection of RDV resistance in SARS-CoV-2 can occur but there is no evidence of global spread of RDV-resistant strains as the RDV resistance mutation identified in our study provides a cost to viral fitness. In addition, we have also shown that key amino acid residues that have been identified in emerging variants of concerns in three different continents can occur *in vitro* in the absence of immune pressure. Overall, our study offers new perspectives for the surveillance of new SARS-CoV-2 variants and the clinical management of patients treated with RDV.

Materials & methods

Cell culture

African Green monkey kidney cells VeroE6 expressing ACE-2 (VeroE6-ACE2) alone and with TMPRSS2 (VeroE6-ACE2-TMPRSS2), BHK-21, and A549NPro expressing ACE-2 (A549NPro-ACE2) were grown in DMEM-Glutamax supplemented with 10% fetal calf serum (FCS; Gibco) and non-essential amino acids (NEAA; Gibco). VeroE6 were a kind gift from Professor M. Bouloy (Institut Pasteur, France), BHK were sourced from the ATCC (ATCC CCL-10) and, VeroE6-ACE2, and VeroE6-ACE2-TMPRSS2 have been described elsewhere [66] A549NPro cells (a kind gift from Prof. R. Randall, University of St. Andrews, UK) were transduced to stably express ACE-2 (A549NPro-ACE2) as described [66]. Calu-3, a human lung epithelial cell line was grown in DMEM Glutamax, supplemented with 20%(v/v) FCS and NEAA was sourced from the ATCC (ATCC HTB-55). All cells were maintained at 37°C in 5% (v/v) CO₂, humidified incubator.

Viruses

All initial work with SARS-CoV-2 was undertaken with the clinical isolate SARS-CoV-2_{Engl2} strain (hCoV-19/England/02/2020 (England-02), GISAID accession: EPI_ISL_407073) kindly supplied by Public Health England. Validation of the mutations was undertaken in rescued

SARS-CoV-2 based on Wuhan-Hu-1 (MN908947) sequence. All virus stocks were grown in VeroE6 cells as described, supplemented with 4%(v/v) FCS, and for the drug screens regardless of cell type the final concentration of FCS was 2%(v/v). For virus growth curves, FCS concentration was decreased to 2%(v/v) in VeroE6 while in Calu-3 FCS concentration was 10%(v/v).

Plaque assays

SARS-CoV-2 titers were determined by plaque assay in either VeroE6 or VeroE6-A-CE2-TMPRSS2. Cells were seeded at 2.5×10^5 cells/well in 12-well plates and plaque assays performed as described [66]. The plates were fixed in 8% [w/v] formaldehyde in PBS, stained with Coomassie Blue staining solution (0.1%(w/v) Coomassie Brilliant Blue R-250; 45%(v/v) methanol; 10%(v/v) glacial acetic acid) and imaged with a photo scanner (Epson Expression 1680 Pro).

Chemical inhibitors

RDV (GS-5734; HY-104077) and EIDD-2801 (HY-135853) were purchased from MedChem-Express, diluted to final concentration of 10 mM stock solutions in 100% (v/v) dimethyl sulfoxide (DMSO, Sigma) and stored at -80C.

Plasmids, virus rescue and validation of RDV resistant viruses

Reverse genetics plasmids used to rescue SARS-CoV-2_{Wu1} have been described previously [66]. Briefly, specific point mutations in SARS-CoV-2 cDNAs in fragment 3 (nt11475 T to C for NSP6 I168T) and fragment 5 (nt 15845 A to C and nt 15846 G to C for NSP12 either E802A, or nt15846 G to C for NSP12 E802D) were synthesized by Biobasic and assembled into the full-length molecular clone to give pCCI-4K-SARS-CoV-2-NSP12_E802D, pCCI-4K-SARS-CoV-2-NSP12_E802A, pCCI-4K-SARS-CoV-2-NSP6_I168T, pCCI-4K-SARS-CoV-2-NSP6_I168T-NSP12_E802D-NSP6, pCCI-4K-SARS-CoV-2-NSP6_I168T-NSP12_E802D-NSP6. The pCCI-4K based plasmids containing the full-length SARS-CoV-2_{Wu1} sequence with the mutations were propagated in TransforMax™ EPI300™ Electrocompetent *E. coli* and purified by using a Macherey Nagel Endotoxin free plasmid purification kit. For virus rescue, BHK-21 cells were transfected with 3 ug of plasmids using Lipofectamine LTX reagent as per manufacturer's recommendation. Cell culture supernatant was harvested 48h and 72h post transfection (p0) and used to infect VeroE6 to generate p1 stock of the recovered viruses.

SARS-CoV-2 compound sensitivity studies and EC₅₀ calculations

To test for a change in sensitivity to RDV of our continually passaged SARS-CoV-2_{Engl2} viruses and reverse genetic derived viruses (rNSP12-E802D, rNSP12-E802A, rNSP6-I168T, rNSP6-I168T-NSP12-E802D, rNSP6-I168T-NSP12-E802A and rSARS-CoV), two different compound screen layouts were tested (S7 Fig). Cells were seeded at 1.2×10^4 cells/well for VeroE6, VeroE6-ACE2, and VeroE6-ACE2-TMPRSS2 and 1.3×10^4 cells/well for A549NPro-ACE2 cell lines in a 96-well format. Two-fold serial dilutions of RDV starting at either 25 or 50 μ M were added to the cells and incubated for a minimum of 2h prior to addition of virus. Depending on plate layout, either a set amount of virus or 2-fold serial dilutions of the virus were added, and plates incubated in a 37°C humidified incubator in 5% CO₂ for 72h. For the mixed array plate layout, each drug dilution and virus combination were in triplicate on each plate and repeated twice. Plates were fixed in 8%(w/v) formaldehyde for 1h and stained in Coomassie Brilliant Blue solution. Washed and airdried plates were scanned using the Celigo

(Nexcelcom) and the total intensity of each well was determined. The total intensity values from the Celigo scan were used to calculate the EC_{50} values using Prism GraphPad v8 and v9. Analysis programs used included transformation of concentration (\log_{10}), normalization of transformed data and non-linear regression ($\log(\text{agonist})$ vs. normalized response—variable slope).

Selection of RDV resistant viruses

Continual passages of SARS-CoV-2 were performed in the presence of RDV in VeroE6 cells. Assays were undertaken in a 12 well plate with each condition in triplicate; two RDV concentrations (1 μ M and 2.5 μ M) and two virus inputs (1000 & 2000 pfu) were used, and controls for cell culture adaption of the virus including media alone or 0.2%(v/v) DMSO were undertaken in parallel (Fig 1). These concentrations were kept constant throughout the experiment. Time between each amplification was dependent of visible CPE from passage 1 to passage 5 (p1 to p5), there was up to 7 days between each amplification. For these passages, 100 μ l of each amplification was used to generate the subsequent passage. It became evident after passage 4 (p4) that some of the conditions used resulted in loss of infectious virus. After passage 5, 10 μ l of culture media was used for passaging the virus in wells that showed evident CPE. Where CPE was not evident, 100 μ l of cell culture was blindly passaged. After each passage the cell media was harvested and stored at -80°C. Samples from p4, p7, p10 and p12 were also harvested for titration and cell monolayer was harvested in TrizolLS for future analysis. Prior to sequencing, each condition was tested to determine its ability to grow in the presence of RDV.

Virus replication kinetics

To compare the replication kinetics of the parental stock, cell culture adapted and RDV resistant mutant SARS-CoV-2, monolayers of VeroE6 cells were treated with 7.5 μ M RDV and infected at a MOI~0.1. The inoculum was removed after 1h, cells washed and fresh media containing either DMSO or RDV added. Growth curves in VeroE6 were undertaken twice in triplicate. rSARS-CoV-2_{Wu1} and mutant viruses virus growth was also assessed in Calu3 cells. For this, Calu-3 were seeded 7 days prior to infection, media was changed every 2–3 days prior to infection. The cells were infected with a MOI~0.01 (based on VeroE6-ACE2-TMPRSS2 titres) for 1h, inoculum then aspirated, cells washed, and fresh media added. All infections for virus growth curves were asynchronous. Cell culture media from each well were collected at 24, 48 and 72h post infection (p.i.) and cell-free virus titers were determined by plaque assay in VeroE6 cells. Each experiment was performed at least thrice with a triplicate of the virus each time (mutant viruses n = 3, wt n = 6). Virus titers at each time were determined by plaque assay as described above.

To assess the influence of RDV on the growth of rSARS-CoV-2 and mutant viruses, Calu-3 were seeded 7 days prior to infection in a 24 well format (2 x 10⁵ cells/well), media was changed every 2–3 days. Two hours prior to infection 2-fold serial dilutions of RDV from 2 μ M to 0.125 μ M were added to cells. Subsequently, cells were infected at a MOI of 0.01 asynchronously for 1h, inoculum aspirated, and media containing the appropriate amount of RDV added. Samples were taken 24h and 48h pi. Titers were determined by plaque assay. Data from 2 independent virus stocks with 2 replicates except we were missing one of the replicates for wild-type and a 48h pi replicated for rNSP12-E802A at 48h pi.

RNA purification from virus stock

Clarified supernatant from p13 was added to TrizolLS and total RNA was extracted as described [70] for sequence analysis.

Sequence of virus stock using Illumina MiSeq

Sequencing libraries composed of overlapping amplicons across the SARS-CoV-2 genome were prepared utilizing the ARTIC Network protocol (<https://artic.network/ncov-2019>). Amplicons were paired-end sequenced (2x250 nt) on an Illumina MiSeq as described previously [71]. Reads were trimmed with trim_galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and then mapped with Burrows–Wheeler Aligner [72] to the SARS-CoV-2_{Wu1} (MN908947) genome, followed by amplicon primer trimming and consensus calling with iVar [73] using a minimum read coverage of 10. Consensus sequences and raw FASTQ files have been uploaded to GenBank under BioProject number PRJNA692078.

Co-infection competition assay

VeroE6-ACE2-TMPRSS2 were co-infected with different ratios (1:9, and 9:1) of rNSP12E802D and rSARS-CoV-2 WT to assess virus fitness as described by Agostini et al., (2018) [20]. Briefly, a 1:9 ratio was equivalent of MOI ~ 0.01 to a MOI of ~0.09, cells were infected for 1 h at 37°C, media aspirated, cell washed, and fresh media added. Twenty-four h post-infection, virus supernatant was harvested, and 100 µl (10% of the volume) was used to infect new VeroE6-ACE2-TMPRSS2 and the infection procedure repeated for 3 passages. RNA was extracted from the clarified supernatant after heat inactivation at 56°C for 20 min using Beckman Coulter RNAdvance Blood kit and KingFisher automated extraction system as per manufacturer's protocols.

Quantification of SARS-CoV-2 genome copies

To examine the virus growth kinetic during the competition assay, the viral load was determined by quantitative RT-PCR based on copies of N RNA transcript. All samples were quantified using a NEB Luna Universal Probe One-Step RT-qPCR Kit (New England Biolabs, E3006) and a previously described CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time reverse transcription PCR Panel [74,75]. 2019-nCoV N1 forward primer 5'-GAC CCC AAA ATC AGC GAA AT-3', 2019-nCoV N1 reverse primer 5'-TCT GGT TAC TGC CAG TTG AAT CTG-3' and 2019-nCoV N1 probe 5'-FAM-ACC CCG CAT TAC GTT TGG TGG ACC-BHQ1-3' were part of the SARS-CoV-2 (2019-nCoV) CDC RUO Primers and Probes kit (IDT Technologies, 10006713). SARS-CoV-2 N RNA was quantified using a ten-fold serial dilutions of an in-house SARS-CoV-2 RNA standard to enable genome equivalent (GE) copies/ml as there is a high correlation between copy number of N and ORF1ab as targets [76] from SARS-CoV-2 RNA extracted from supernatant. All reactions were performed on ABI 7500 Fast thermal cycler.

Sequence of virus stock using GridION (Oxford Nanopore Technology)

The GridION system was used to sequence the virus populations used for the competition assay as previously described [71]. GridION reads were processed using the ARTIC bioinformatics pipeline (<https://github.com/artic-network/artic-ncov2019>). Briefly, reads were size filtered then aligned to the Wuhan-Hu-1 reference (MN908947) with Minimap2 [77] followed by post-alignment amplicon primer trimming and variant calling with Nanopolish (<https://github.com/jts/nanopolish>).

Modelling of mutations in SARS-CoV-2 NSP12

UCSF Chimera [78] to generate the ribbon structure and rotamer function [79] used to model the mutations into SARS-CoV-2_{Wu1} NSP12 (PDB 6YYT [39]). Protter [80] was used to visualize the effect of the I168T mutation on organization of the transmembrane domains in NSP6.

Statistics and bioinformatic

GraphPad Prism v8 & v9 software (La Jolla, CA, USA) was used to undertake statistical analysis (transformation, normalization, non-linear regression, RM one-way ANOVA (Dunn Dunnett's multiple comparisons test), two-way ANOVA, unpaired t-test (Holm-Šídák method)) and generate figures. EMBL-EBI Clustal omega [81] was used to align the sequences of NSP6 and NSP12 from SARS-CoV-2 Wuhan1 (MN908947), betacoronavirus England1 (bCoV England1; YP009944302 & YP009944297.1), SARS-CoV-2_{Engl2}, SARS-CoV Tor (NP_828849.7 & YP009944371.1), human coronavirus (hCoV, HKU1, YP459941.1 & YP009944274.1), middle eastern respiratory virus (MERS; YP009047223.1 & 009047234.1), murine hepatitis virus (MHV; YP009924352.1 & 009915678.1), rat coronavirus (Parker, YP009924378.1 & YP_009924378.1), Roussettus bat coronavirus (HKU9; YP009924393.1 & YP009924388.1), rabbit CoV (HKU14; YP009924419.1 & YP009924414.1), Tylonycteris bat CoV (HKU4 YP009944320.1 & YP009944315.1), and Pipostrellus bat CoV (HKU5; YP009944349.2 & YP009944344.1). The online database CoV-GLUE (<http://cov-gluе.cvr.gla.ac.uk/> [82]) was used to monitor for replacement in the SARS-CoV-2 genome. The database was accessed on 10th of January 2021 (n = 242865) for generating the diversity plot and further on 15th of March 2021 n = 426623). The dN/dS ratios were calculated for each sample using all ORFs concatenated together, as well as the Spike and entire ORF1ab sequences individually (with respect to the SARS-CoV-2_{Engl2} sample to monitor within passage evolution), using the program SNAP (<https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html>) [83].

The code used to calculate the null distribution and the probability as well as the global distribution of the mutations within in Spike can be assessed at <https://github.com/omaclean/CV19IV/>.

The receptor binding domain of Spike is from amino acids 319–541: B1.351 and P1 had 3 mutations in the RBD (K417N/T, E484K and N501Y) while Alpha (B.1.1.7) only had one (N501Y). We used the following amino acids replacements and deletions in Spike for each lineage-based COG-UK defined changes for each lineage; Beta (B.1.351; L18F, D80A, D215G, R246L, K417N, E484K, N501Y, A701V), Gamma (P.1; L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y H655Y, T1027I), Alpha (B.1.1.7; Δ69–70, D144, N501Y, A570D, P681H, T716L, S982A, D1118H). We further included the substitution observed in P1 (V1176F) and deletion in B1.351 (Δ242–245). There is phylogenetic uncertainty if V1176F was present in the ancestor preceding the formation of the Gamma (P.1) lineage, the deletion (Δ242–245; like other defining mutations) in Beta (B.1.351) is polymorphic. We erred on the side of inclusivity in the face of this uncertainty, making our analysis more conservative. All sequences available sequences for SARS-CoV-2 were assessed from CoV-Glu (n = 242865, last update 14th December 2020).

Supporting information

S1 Fig. RDV mixed array inhibitor screen with SARS-CoV-2_{Engl2} in VeroE6 cells. Two-fold serial dilutions of RDV from 50μM and either a two-fold (left) or three-fold (right) dilution of SARS-CoV-2. (A) Heatmap of total intensity of the stained plated generated by the Celigo. Lightest colors indicate clearance of the monolayer. (B) RDV dose dependency over a range of virus inputs. For each RDV concentration the survival (%) of the monolayer with different virus inputs per well is plotted. Values are from 1 replicate per condition. (TIF)

S2 Fig. Adaptation of SARS-CoV-2_{Engl2} to VeroE6 and selection of decreased RDV sensitivity. (A) Coomassie blue stained SARS-CoV-2 plaque assays in VeroE6. Change in plaque

morphology from passage 1 and passage 10. A subset of populations from the 4 different conditions were compared. **(B)** Celigo scans of Coomassie blue stained 96-well format mixed array drug for two control populations (Mediap13.4 and DMSOp13.5) and a RDV adapted population (Rem2.5p13.5). RDV concentration and virus input is indicated. **(C)** Bar graph of RDV EC₅₀ for an RDV adapted population and two control populations over a range of virus inputs.

(TIF)

S3 Fig. Antiviral activity of RDV and EIDD2801 in Vero-ACE-2 against three continually passed populations and the input SARS-CoV-2_{Engl2}. The EC₅₀ of RDV adapted population REM2.5p13.5 was compared to the input virus SARS-CoV-2_{Engl2} as well as DMSOp13.4 and Mediap13.4. **(A)** RDV dose dependency curve. **(B)** Bar graph of RDV EC₅₀ required to protect the monolayer. **(C)** EIDD2801 dose dependency curve. **(D)** Bar graph of EIDD2801 EC₅₀ required to protect the monolayer.

(TIF)

S4 Fig. Protter prediction transmembrane organisation of NSP6 WT (left) and NSP6 I168T (right) mutation. Position of I168 and I168T are indicated by an arrow. Extra-, intra-cellular and membrane are indicated, and putative glycosylation site is shown in green.

(TIF)

S5 Fig. Antiviral activity of RDV and EIDD2801 against NSP6 and NSP12 mutant viruses. All viruses have a SARS-CoV-2_{Wu1} backbone with specific point mutations as indicated. Assay undertaken in VeroE6-ACE2-TMPRSS2. **(A)** Bar graph of RDV EC₅₀ values for rNSP12E902A, rNSP12E802D and rSARS-CoV-2 over a range of virus inputs. **(B)** Bar graph of EIDD2801 EC₅₀ for each rescued virus. Error bars are SEM. Experiment from 2 independent virus stocks with 3 technical replicates.

(TIF)

S6 Fig. Sequence analysis of continually RDV selected SARS-CoV-2_{Engl2}. **(A)** Diversity occurring in SARS-CoV-2 NSP-12 in 20 amino acid sliding window. Calculated from 988 mutations observed in minimum of 5 sequences. Position of mutation associated with decrease RDV sensitivity in SARS-CoV-2 is highlighted in red and sites identified in murine hepatitis virus (MHV) and Ebola virus (EBOV) are indicated in black. **(B)** Heatmap of the distribution of transversion and transition depending on nucleotide. **(C)** Distribution of convergence. The null distribution of overlapping *in vitro* mutations hitting the same Spike codon positions as in the SARS-CoV-2 variants of concern. 21 *in vitro* mutations and 20 variant of concern mutations spread across 1271 codons. 100 million simulations were run.

(TIF)

S7 Fig. Schematic layout of 96-well plate clearance assay for drug screen. All layouts have controls for inhibitor toxicity, virus clearance and cell monolayer integrity (DMSO only). **(A)** Mixed array layout with range of inhibitor dilutions vs a range of virus dilutions. **(B)** Mixed plate layout inhibitor is diluted and a set amount of virus. Amount of virus added causes complete well clearance by 72 h pi. The mixed plate layout can also be modified to dilute inhibitor down the rows rather than across the columns.

(TIF)

S1 Table. RDV and EIDD2801 EC₅₀ and fold-change in VeroE6-ACE2-TMPRSS2. All viruses have a SARS-CoV-2_{Wu1} backbone with specific point mutations as indicated. Average values and SEM from independently rescued virus stocks, except for rNSP12 E802A. Experiment undertaken at least twice in triplicate by independent operators apart for rNSP6I168T in

the presence of EIDD2801 ($n = 1$) and $n = 3$ for rSARS-CoV-2, rNSP12E802D, rNSP12E802A and rNSP6I168T.

(XLSX)

S2 Table. dN/dS ratio for all SARS-CoV-2 open reading frames, ORF1ab and spike for the *in vitro* continually passaged SARS-CoV-2.

(XLSX)

S1 Data. Excel sheet summarizing mutations across different continually passaged virus populations in comparison to SARS-CoV-2_{Wuhan-1} (MN908947) and SARS-CoV-2_{Engl2}. (A) Key for the virus names (B) All the mutations occurring in the consensus sequence in the continually passaged viruses and in the input virus SARS-CoV-2_{Engl2} in comparison to SARS-CoV-2_{Wu1}. (C) The frequency of all mutations (consensus and subconsensus) that arose in continually passaged SARS-CoV-2_{Engl2} and (D) a list of the deletions and insertions.

(XLSX)

S2 Data. Conservation of NSP12 E802 and NSP6 I168 in coronaviruses. Excel file with the frequency of NSP12 E802 (A) and frequency of NSP6 I168 (B).

(XLSX)

S3 Data. SARS-CoV-2 spike mutation arising *in vitro* are prevalence in the circulating sequences. The mutation counts (total of 3588809) and lineage frequency are generated from >800 000 sequences available on GISAID. Frequency >0.5 is the number of lineages the mutation is >50% of the total sequences within each lineage, whereas frequency <0.5 is the number of lineages the mutation appears where the frequency of that mutation is less than 50%. Lineages that the mutations appear in are indicated.

(XLSX)

S4 Data. GISAID Acknowledgements. Authors and laboratories contributing to the genome data to GISAID database used to identify the NSP12 E802 mutation in SARS-CoV-2. All Submitters of data may be contacted directly via www.gisaid.org. Contributing authors and laboratories are sorted alphabetically. The accession number of each sequence is provided along with the originating laboratory.

(XLSX)

S1 Text. Supporting information contains additional information on the SARS-CoV-2 adaptation in VeroE6, expands on lack of mutations in the presence on RDV, and the pressure to evolve with greater focus on the Spike.

(DOCX)

Acknowledgments

We would like to thank all the global researchers who shared their data on GISAID (<https://www.gisaid.org>; S4 Data), COVID-19 Genomics UK Consortium (COG-UK) and Kyriaki Nomikou, Jenna Nichols, Yasmin Parr, Lily Tong and Natasha Johnson from the CVR Genomics Team.

Author Contributions

Conceptualization: Agnieszka M. Szemiel, Emma C. Thomson, Massimo Palmarini, Alain Kohl, Meredith E. Stewart.

Data curation: Agnieszka M. Szemiel, Richard J. Orton, Oscar A. MacLean, Meredith E. Stewart.

Formal analysis: Agnieszka M. Szemiel, Meredith E. Stewart.

Funding acquisition: Andres Merits, Ana da Silva Filipe, Brian J. Willett, Sam J. Wilson, Arvind H. Patel, Massimo Palmarini, Alain Kohl.

Investigation: Agnieszka M. Szemiel, Andres Merits, Richard J. Orton, Oscar A. MacLean, Gauthier Lieber, Matthew L. Turnbull, Sainan Wang, Daniel Mair, Meredith E. Stewart.

Methodology: Agnieszka M. Szemiel, Andres Merits, Oscar A. MacLean, Rute Maria Pinto, Wilhelm Furnon, Nicolas M. Suarez, Ana da Silva Filipe, Sam J. Wilson, Arvind H. Patel, Meredith E. Stewart.

Project administration: Agnieszka M. Szemiel, Alain Kohl, Meredith E. Stewart.

Resources: Arthur Wickenhagen, Ana da Silva Filipe, Sam J. Wilson, Arvind H. Patel, Emma C. Thomson, Massimo Palmarini, Alain Kohl.

Validation: Agnieszka M. Szemiel, Oscar A. MacLean, Meredith E. Stewart.

Visualization: Agnieszka M. Szemiel, Richard J. Orton, Oscar A. MacLean, Meredith E. Stewart.

Writing – original draft: Agnieszka M. Szemiel, Brian J. Willett, Massimo Palmarini, Alain Kohl, Meredith E. Stewart.

Writing – review & editing: Agnieszka M. Szemiel, Andres Merits, Richard J. Orton, Oscar A. MacLean, Rute Maria Pinto, Arthur Wickenhagen, Gauthier Lieber, Matthew L. Turnbull, Brian J. Willett, Emma C. Thomson, Massimo Palmarini, Alain Kohl, Meredith E. Stewart.

References

1. WHO Coronavirus (COVID-19) Dashboard. [cited 21 Jul 2021]. Available: <https://covid19.who.int>
2. Kaye AD, Okeagu CN, Pham AD, Silva RA, Hurley JJ, Arron BL, et al. Economic impact of COVID-19 pandemic on healthcare facilities and systems: International perspectives. *Best Pract Res Clin Anaesthesiol.* 2020 [cited 29 Dec 2020]. <https://doi.org/10.1016/j.bpa.2020.11.009> PMID: 34511220
3. Gordon CJ, Tcheshnokov EP, Woolner E, Perry JK, Feng JY, Porter DP, et al. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J Biol Chem.* 2020; 295: 6785–6797. <https://doi.org/10.1074/jbc.RA120.013679> PMID: 32284326
4. Pruijssers AJ, George AS, Schäfer A, Leist SR, Gralinski LE, Dinnon KH, et al. Remdesivir Inhibits SARS-CoV-2 in Human Lung Cells and Chimeric SARS-CoV Expressing the SARS-CoV-2 RNA Polymerase in Mice. *Cell Rep.* 2020; 32: 107940. <https://doi.org/10.1016/j.celrep.2020.107940> PMID: 32668216
5. Wit E de, Feldmann F, Cronin J, Jordan R, Okumura A, Thomas T, et al. Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection. *Proc Natl Acad Sci.* 2020; 117: 6771–6776. <https://doi.org/10.1073/pnas.1922083117> PMID: 32054787
6. Williamson BN, Feldmann F, Schwarz B, Meade-White K, Porter DP, Schulz J, et al. Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. *Nature.* 2020; 585: 273–276. <https://doi.org/10.1038/s41586-020-2423-5> PMID: 32516797
7. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. Remdesivir for the Treatment of Covid-19—Final Report. *N Engl J Med.* 2020 [cited 2 Jan 2021]. <https://doi.org/10.1056/NEJMoa2007764> PMID: 32445440
8. Spinner CD, Gottlieb RL, Criner GJ, Arribas López JR, Cattelan AM, Soriano Viladomiu A, et al. Effect of Remdesivir vs Standard Care on Clinical Status at 11 Days in Patients With Moderate COVID-19: A Randomized Clinical Trial. *JAMA.* 2020; 324: 1048. <https://doi.org/10.1001/jama.2020.16349> PMID: 32821939

9. Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *The Lancet*. 2020; 395: 1569–1578. [https://doi.org/10.1016/S0140-6736\(20\)31022-9](https://doi.org/10.1016/S0140-6736(20)31022-9) PMID: 32423584
10. Grein J, Ohmagari N, Shin D, Diaz G, Asperges E, Castagna A, et al. Compassionate Use of Remdesivir for Patients with Severe Covid-19. *N Engl J Med*. 2020; 382: 2327–2336. <https://doi.org/10.1056/NEJMoa2007016> PMID: 32275812
11. Repurposed Antiviral Drugs for Covid-19—Interim WHO Solidarity Trial Results | NEJM. [cited 2 Jan 2021]. Available: <https://www.nejm.org/doi/10.1056/NEJMoa2023184>
12. Matthay MA, Thompson BT. Dexamethasone in hospitalised patients with COVID-19: addressing uncertainties. *Lancet Respir Med*. 2020; 8: 1170–1172. [https://doi.org/10.1016/S2213-2600\(20\)30503-8](https://doi.org/10.1016/S2213-2600(20)30503-8) PMID: 33129421
13. RECOVERY Collaborative Group, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in Hospitalized Patients with Covid-19—Preliminary Report. *N Engl J Med*. 2020. <https://doi.org/10.1056/NEJMoa2021436> PMID: 32678530
14. Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19 | NEJM. [cited 4 Mar 2021]. Available: <https://www.nejm.org/doi/full/10.1056/NEJMoa2031994>
15. Duwe S. Influenza viruses—antiviral therapy and resistance. *GMS Infect Dis*. 2017; 5: Doc04. <https://doi.org/10.3205/id000030> PMID: 30671326
16. Ghany MG, Doo EC. Antiviral resistance and hepatitis B therapy. *HepatoL Baltim Md*. 2009; 49: S174–184. <https://doi.org/10.1002/hep.22900> PMID: 19399794
17. Ray SC, Thomas DL. 156—Hepatitis C. In: Bennett JE, Dolin R, Blaser MJ, editors. *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition)*. Philadelphia: W.B. Saunders; 2015. pp. 1904–1927.e9. <https://doi.org/10.1016/B978-1-4557-4801-3.00156-9>
18. Siddiqui S. Chapter 14—Resistance in Pathogenic Microorganisms. In: Gupta VK, Pandey A, editors. *New and Future Developments in Microbial Biotechnology and Bioengineering*. Amsterdam: Elsevier; 2019. pp. 183–191. <https://doi.org/10.1016/B978-0-444-63504-4.00014-1>
19. Padhi AK, Shukla R, Saudagar P, Tripathi T. High-Throughput Rational Design of the Remdesivir Binding Site in the RdRp of SARS-CoV-2: Implications for Potential Resistance. *iScience*. 2020;0. <https://doi.org/10.1016/j.isci.2020.101992> PMID: 33490902
20. Agostini ML, Andres EL, Sims AC, Graham RL, Sheahan TP, Lu X, et al. Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio*. 2018; 9. <https://doi.org/10.1128/mBio.00221-18> PMID: 29511076
21. Lo MK, Albariño CG, Perry JK, Chang S, Tchesnokov EP, Guerrero L, et al. Remdesivir targets a structurally analogous region of the Ebola virus and SARS-CoV-2 polymerases. *Proc Natl Acad Sci*. 2020; 117: 26946–26954. <https://doi.org/10.1073/pnas.2012294117> PMID: 33028676
22. Yin W, Mao C, Luan X, Shen D-D, Shen Q, Su H, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*. 2020; 368: 1499–1504. <https://doi.org/10.1126/science.abc1560> PMID: 32358203
23. Cottam EM, Whelband MC, Wileman T. Coronavirus NSP6 restricts autophagosome expansion. *Autophagy*. 2014; 10: 1426–1441. <https://doi.org/10.4161/auto.29309> PMID: 24991833
24. Sheahan TP, Sims AC, Zhou S, Graham RL, Puijssers AJ, Agostini ML, et al. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci Transl Med*. 2020; 12. <https://doi.org/10.1126/scitranslmed.abb5883> PMID: 32253226
25. Klimstra WB, Tilston-Lunel NL, Nambulli S, Boslett J, McMillen CM, Gilliland T, et al. SARS-CoV-2 growth, furin-cleavage-site adaptation and neutralization using serum from acutely infected hospitalized COVID-19 patients. *J Gen Virol*. 2020; 101: 1156–1169. <https://doi.org/10.1099/jgv.0.001481> PMID: 32821033
26. Davidson AD, Williamson MK, Lewis SD, Shoemark DK, Carroll MW, Heesom KJ, et al. Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med*. 2020; 12: 68. <https://doi.org/10.1186/s13073-020-00763-0> PMID: 32723359
27. Liu Z, Zheng H, Lin H, Li M, Yuan R, Peng J, et al. Identification of Common Deletions in the Spike Protein of Severe Acute Respiratory Syndrome Coronavirus 2. *J Virol*. 2020; 94. <https://doi.org/10.1128/JVI.00790-20> PMID: 32571797
28. Neagu IA, Olejarz J, Freeman M, Rosenbloom DIS, Nowak MA, Hill AL. Life cycle synchronization is a viral drug resistance mechanism. *PLOS Comput Biol*. 2018; 14: e1005947. <https://doi.org/10.1371/journal.pcbi.1005947> PMID: 29447150

29. Helleberg M, Niemann CU, Moestrup KS, Kirk O, Lebech A-M, Lane C, et al. Persistent COVID-19 in an Immunocompromised Patient Temporarily Responsive to Two Courses of Remdesivir Therapy. *J Infect Dis.* 2020; 222: 1103–1107. <https://doi.org/10.1093/infdis/jiaa446> PMID: 32702095
30. Buckland MS, Galloway JB, Fhogartaigh CN, Meredith L, Provine NM, Bloor S, et al. Treatment of COVID-19 with remdesivir in the absence of humoral immunity: a case report. *Nat Commun.* 2020; 11: 6385. <https://doi.org/10.1038/s41467-020-19761-2> PMID: 33318491
31. Kemp SA, Collier DA, Datir RP, Ferreira IATM, Gayed S, Jahun A, et al. SARS-CoV-2 evolution during treatment of chronic infection. *Nature.* 2021; 592: 277–282. <https://doi.org/10.1038/s41586-021-03291-y> PMID: 33545711
32. Boshier FAT, Pang J, Penner J, Hughes J, Parker M, Shepherd J, et al. Remdesivir induced viral RNA and subgenomic RNA suppression, and evolution of viral variants in SARS-CoV-2 infected patients. *medRxiv.* 2020; 2020.11.18.20230599. <https://doi.org/10.1101/2020.11.18.20230599>
33. Minskaia E, Hertzog T, Gorbalenya AE, Campanacci V, Cambillau C, Canard B, et al. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc Natl Acad Sci.* 2006; 103: 5108–5113. <https://doi.org/10.1073/pnas.0508200103> PMID: 16549795
34. Smith EC, Blanc H, Vignuzzi M, Denison MR. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. *PLOS Pathog.* 2013; 9: e1003565. <https://doi.org/10.1371/journal.ppat.1003565> PMID: 23966862
35. Bravo JPK, Dangerfield TL, Taylor DW, Johnson KA. Remdesivir is a delayed translocation inhibitor of SARS-CoV-2 replication. *Mol Cell.* 2021; S1097276521000551. <https://doi.org/10.1016/j.molcel.2021.01.035> PMID: 33631104
36. Kokic G, Hillen HS, Teginov D, Dienemann C, Seitz F, Schmitzova J, et al. Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat Commun.* 2021; 12: 279. <https://doi.org/10.1038/s41467-020-20542-0> PMID: 33436624
37. Tchesnokov EP, Gordon CJ, Woolner E, Kocinkova D, Perry JK, Feng JY, et al. Template-dependent inhibition of coronavirus RNA-dependent RNA polymerase by remdesivir reveals a second mechanism of action. *J Biol Chem.* 2020; 295: 16156–16165. <https://doi.org/10.1074/jbc.AC120.015720> PMID: 32967965
38. Wang Q, Wu J, Wang H, Gao Y, Liu Q, Mu A, et al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell.* 2020; 182: 417–428.e13. <https://doi.org/10.1016/j.cell.2020.05.034> PMID: 32526208
39. Hillen HS, Kokic G, Farnung L, Dienemann C, Teginov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. *Nature.* 2020; 584: 154–156. <https://doi.org/10.1038/s41586-020-2368-8> PMID: 32438371
40. Malone B, Chen J, Wang Q, Llewellyn E, Choi YJ, Olinares PDB, et al. Structural basis for backtracking by the SARS-CoV-2 replication-transcription complex. *Biochemistry.* 2021 Mar. <https://doi.org/10.1101/2021.03.13.435256> PMID: 33758867
41. Mumtaz N, Jimmerson LC, Bushman LR, Kiser JJ, Aron G, Reusken CBEM, et al. Cell-line dependent antiviral activity of sofosbuvir against Zika virus. *Antiviral Res.* 2017; 146: 161–163. <https://doi.org/10.1016/j.antiviral.2017.09.004> PMID: 28912011
42. Dicken SJ, Murray MJ, Thorne LG, Reuschl A-K, Forrest C, Ganeshalingham M, et al. Characterisation of B.1.1.7 and Pangolin coronavirus spike provides insights on the evolutionary trajectory of SARS-CoV-2. *bioRxiv.* 2021; 2021.03.22.436468. <https://doi.org/10.1101/2021.03.22.436468> PMID: 33791702
43. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell.* 2020; 182: 1295–1310.e20. <https://doi.org/10.1016/j.cell.2020.08.012> PMID: 32841599
44. Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinnon KH, et al. SARS-CoV-2 D614G variant exhibits efficient replication *ex vivo* and transmission *in vivo*. *Science.* 2020; eabe8499. <https://doi.org/10.1126/science.abe8499> PMID: 33184236
45. Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY, et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe.* 2021; 29: 463–476.e6. <https://doi.org/10.1016/j.chom.2021.02.003> PMID: 33592168
46. Liu Z, VanBlargan LA, Bloyet L-M, Rothlauf PW, Chen RE, Stumpf S, et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe.* 2021; 29: 477–488.e4. <https://doi.org/10.1016/j.chom.2021.01.014> PMID: 33535027
47. Qu X-X, Hao P, Song X-J, Jiang S-M, Liu Y-X, Wang P-G, et al. Identification of Two Critical Amino Acid Residues of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein for Its Variation in

- Zoonotic Tropism Transition via a Double Substitution Strategy*. *J Biol Chem.* 2005; 280: 29588–29595. <https://doi.org/10.1074/jbc.M500662200> PMID: 15980414
48. Gordon CJ, Tchesnokov EP, Schinazi RF, Götte M. Molnupiravir promotes SARS-CoV-2 mutagenesis via the RNA template. *J Biol Chem.* 2021; 100770. <https://doi.org/10.1016/j.jbc.2021.100770> PMID: 33989635
 49. Crotty S, Maag D, Arnold JJ, Zhong W, Lau JYN, Hong Z, et al. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat Med.* 2000; 6: 1375–1379. <https://doi.org/10.1038/82191> PMID: 11100123
 50. Chen J, Wang R, Wang M, Wei G-W. Mutations Strengthened SARS-CoV-2 Infectivity. *J Mol Biol.* 2020; 432: 5212–5226. <https://doi.org/10.1016/j.jmb.2020.07.009> PMID: 32710986
 51. Zahradník J, Marciano S, Shemesh M, Zoler E, Chiaravalli J, Meyer B, et al. SARS-CoV-2 RBD *in vitro* evolution follows contagious mutation spread, yet generates an able infection inhibitor. *bioRxiv.* 2021; 2021.01.06.425392. <https://doi.org/10.1101/2021.01.06.425392>
 52. Ou J, Zhou Z, Dai R, Zhao S, Wu X, Zhang J, et al. Emergence of SARS-CoV-2 spike RBD mutants that enhance viral infectivity through increased human ACE2 receptor binding affinity. *bioRxiv.* 2020; 2020.03.15.991844. <https://doi.org/10.1101/2020.03.15.991844>
 53. Ramirez S, Fernandez-Antunez C, Galli A, Underwood A, Pham LV, Ryberg LA, et al. Overcoming Culture Restriction for SARS-CoV-2 in Human Cells Facilitates the Screening of Compounds Inhibiting Viral Replication. *Antimicrob Agents Chemother.* 65: e00097–21. <https://doi.org/10.1128/AAC.00097-21> PMID: 33903110
 54. Braun KM, Moreno GK, Halfmann PJ, Hodcroft EB, Baker DA, Boehm EC, et al. Transmission of SARS-CoV-2 in domestic cats imposes a narrow bottleneck. *PLOS Pathog.* 2021; 17: e1009373. <https://doi.org/10.1371/journal.ppat.1009373> PMID: 33635912
 55. Wu K, Peng G, Wilken M, Geraghty RJ, Li F. Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. *J Biol Chem.* 2012; 287: 8904–8911. <https://doi.org/10.1074/jbc.M111.325803> PMID: 22291007
 56. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.* 2020; 181: 281–292.e6. <https://doi.org/10.1016/j.cell.2020.02.058> PMID: 32155444
 57. Wec AZ, Wrapp D, Herbert AS, Maurer DP, Haslwanter D, Sakharkar M, et al. Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science.* 2020; 369: 731–736. <https://doi.org/10.1126/science.abc7424> PMID: 32540900
 58. Chan JF-W, Zhang AJ, Yuan S, Poon VK-M, Chan CC-S, Lee AC-Y, et al. Simulation of the Clinical and Pathological Manifestations of Coronavirus Disease 2019 (COVID-19) in a Golden Syrian Hamster Model: Implications for Disease Pathogenesis and Transmissibility. *Clin Infect Dis.* 2020; 71: 2428–2446. <https://doi.org/10.1093/cid/ciaa325> PMID: 32215622
 59. Dieterle ME, Haslwanter D, Bortz RH, Wirchnianski AS, Lasso G, Vergnolle O, et al. A Replication-Competent Vesicular Stomatitis Virus for Studies of SARS-CoV-2 Spike-Mediated Cell Entry and Its Inhibition. *Cell Host Microbe.* 2020; 28: 486–496.e6. <https://doi.org/10.1016/j.chom.2020.06.020> PMID: 32738193
 60. Yang X, Dong N, Chan EW-C, Chen S. Identification of super-transmitters of SARS-CoV-2. *medRxiv.* 2020; 2020.04.19.20071399. <https://doi.org/10.1101/2020.04.19.20071399>
 61. Bestle D, Heindl MR, Limburg H, Van Lam van T, Pilgram O, Moulton H, et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci Alliance.* 2020; 3: e202000786. <https://doi.org/10.26508/lsa.202000786> PMID: 32703818
 62. Peacock TP, Goldhill DH, Zhou J, Baillon L, Frise R, Swann OC, et al. The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. *Nat Microbiol.* 2021; 6: 899–909. <https://doi.org/10.1038/s41564-021-00908-w> PMID: 33907312
 63. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell.* 2020; 181: 271–280.e8. <https://doi.org/10.1016/j.cell.2020.02.052> PMID: 32142651
 64. Koch J, Uckeley ZM, Doldan P, Stanifer M, Boulant S, Lozach P-Y. TMPRSS2 expression dictates the entry route used by SARS-CoV-2 to infect host cells. *EMBO J.* 2021;n/a: e107821. <https://doi.org/10.15252/emj.2021107821> PMID: 34159616
 65. Kawase M, Shirato K, Matsuyama S, Taguchi F. Protease-mediated entry via the endosome of human coronavirus 229E. *J Virol.* 2009; 83: 712–721. <https://doi.org/10.1128/JVI.01933-08> PMID: 18971274
 66. Rihn SJ, Merits A, Bakshi S, Turnbull ML, Wickenhagen A, Alexander AJT, et al. A plasmid DNA-launched SARS-CoV-2 reverse genetics system and coronavirus toolkit for COVID-19 research. *PLOS Biol.* 2021; 19: e3001091. <https://doi.org/10.1371/journal.pbio.3001091> PMID: 33630831

67. Chiem K, Morales Vasquez D, Park J-G, Platt RN, Anderson T, Walter MR, et al. Generation and Characterization of Recombinant SARS-CoV-2 Expressing Reporter Genes. Parrish CR, editor. *J Virol*. 2021; 95. <https://doi.org/10.1128/JVI.02209-20> PMID: 33431557
68. Moreno GK, Braun KM, Halfmann PJ, Prall TM, Riemersma KK, Haj AK, et al. Limited SARS-CoV-2 diversity within hosts and following passage in cell culture. *bioRxiv*. 2020; 2020.04.20.051011. <https://doi.org/10.1101/2020.04.20.051011>
69. Pharmabox.in. JPM: Gilead's Veklury, now used in half of hospitalized U.S. COVID-19 patients, delivers multibillion-dollar revenue bump. In: PHARMABOX.IN [Internet]. 12 Jan 2021 [cited 24 Jan 2021]. Available: <https://pharmabox.in/jpm-gileads-veklury-now-used-in-half-of-hospitalized-u-s-covid-19-patients-delivers-multibillion-dollar-revenue-bump/>
70. Shaw AE, Hughes J, Gu Q, Behdenna A, Singer JB, Dennis T, et al. Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. *PLOS Biol*. 2017; 15: e2004086. <https://doi.org/10.1371/journal.pbio.2004086> PMID: 29253856
71. da Silva Filipe A, Shepherd JG, Williams T, Hughes J, Aranday-Cortes E, Asamaphan P, et al. Genomic epidemiology reveals multiple introductions of SARS-CoV-2 from mainland Europe into Scotland. *Nat Microbiol*. 2021; 6: 112–122. <https://doi.org/10.1038/s41564-020-00838-z> PMID: 33349681
72. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009; 25: 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324> PMID: 19451168
73. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol*. 2019; 20: 8. <https://doi.org/10.1186/s13059-018-1618-7> PMID: 30621750
74. Fajnzylber J, Regan J, Coxen K, Corry H, Wong C, Rosenthal A, et al. SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat Commun*. 2020; 11: 5493. <https://doi.org/10.1038/s41467-020-19057-5> PMID: 33127906
75. Case JB, Bailey AL, Kim AS, Chen RE, Diamond MS. Growth, detection, quantification, and inactivation of SARS-CoV-2. *Virology*. 2020; 548: 39–48. <https://doi.org/10.1016/j.virol.2020.05.015> PMID: 32838945
76. Vogels CBF, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. *Nat Microbiol*. 2020; 5: 1299–1305. <https://doi.org/10.1038/s41564-020-0761-6> PMID: 32651556
77. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018; 34: 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191> PMID: 29750242
78. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25: 1605–1612. <https://doi.org/10.1002/jcc.20084> PMID: 15264254
79. Shapovalov MV, Dunbrack RL. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Struct Lond Engl* 1993. 2011; 19: 844–858. <https://doi.org/10.1016/j.str.2011.03.019> PMID: 21645855
80. Omasits U, Ahrens CH, Müller S, Wollscheid B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinforma Oxf Engl*. 2014; 30: 884–886. <https://doi.org/10.1093/bioinformatics/btt607> PMID: 24162465
81. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011; 7: 539. <https://doi.org/10.1038/msb.2011.75> PMID: 21988835
82. Singer J, Gifford R, Cotten M, Robertson D. CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation. 2020 [cited 14 Jan 2021]. <https://doi.org/10.20944/preprints202006.0225.v1>
83. Korber B. HIV Signature and Sequence Variation Analysis. *HIV Signat Seq Var Anal Comput Anal HIV Mol Seq*. 2000; 55–72.