

# Genetic consequences of effective and suboptimal dosing with mutagenic drugs in a hamster model of SARS-CoV-2 infection

Christopher J. R. Illingworth<sup>1,§</sup>, Jose A. Guerra-Assuncao<sup>2,3,\*</sup>, Samuel Gregg<sup>3\*</sup>, Oscar Charles<sup>2,3\*</sup>, Juanita Pang<sup>3</sup>, Sunando Roy<sup>3</sup>, Rana Abdelnabi<sup>4,5</sup>, Johan Neyts<sup>4,5</sup>, Judith Breuer<sup>2,3</sup>

<sup>1</sup>MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom

<sup>2</sup>Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

<sup>3</sup>Infection, Immunity and Inflammation Research and Teaching Department, University College London, London, UK

<sup>4</sup>KU Leuven Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy, B-3000 Leuven, Belgium.

<sup>5</sup>The VirusBank Platform, Gaston Geenslaan, B-3000 Leuven, Belgium.

\*Authors contributed equally

§Corresponding author: Email: christopher.illingworth@glasgow.ac.uk

## Abstract

Mutagenic antiviral drugs have shown promise against multiple viruses, but concerns have been raised about whether their use might promote the emergence of new and harmful viral variants. Recently, genetic signatures associated with molnupiravir use have been identified in the global SARS-COV-2 population. Here, we examine the consequences of using favipiravir and molnupiravir to treat SARS-CoV-2 infection in a hamster model, comparing viral genome sequence data collected from (i) untreated hamsters, and (ii) from hamsters receiving effective

and suboptimal doses of treatment. We identify a broadly linear relationship between drug dose and the extent of variation in treated viral populations, with a high proportion of this variation being composed of variants at frequencies of less than one per cent, below typical thresholds for variant calling. Treatment with an effective dose of antiviral drug was associated with a gain of between 7 and 10 variants per viral genome relative to drug-free controls: Even after a short period of treatment a population founded by a transmitted virus could contain multiple sequence differences to that of the original host. Treatment with a suboptimal dose of drug showed intermediate gains of variants. No dose-dependent signal was identified in the numbers of single nucleotide variants reaching frequencies in excess of 5%. We did not find evidence to support the emergence of drug resistance or of novel immune phenotypes. Our study suggests that where onward transmission occurs, a short period of treatment with mutagenic drugs may be sufficient to generate a significant increase in the number of viral variants transmitted.

## Introduction

Molnupiravir (EIDD-2801/MK-4482) is an oral prodrug of the nucleoside analogue  $\beta$ -D-N4-hydroxycytidine (EIDD-1931, NHC) with broad inhibitory activity against RNA dependent RNA polymerases of positive and negative strand RNA viruses (RdRps) including SARS-CoV-2 and other coronaviruses<sup>1-6</sup>. Following oral administration, molnupiravir is metabolized and phosphorylated into the active drug NHC 5'-triphosphate (NHC-TP). During viral replication NHC-TP is substituted by SARS-CoV-2 RdRp for cytidine or uridine triphosphate into newly synthesized viral RNA, resulting in incorporation of either guanosine or adenosine when the RNA is copied<sup>3,7</sup>. The accumulation of random C to U or G to A transition errors in the viral RNA genome is associated with lethal mutagenesis resulting in loss of viral fitness as evidenced by reduced viral loads, reduced viral infectivity and reduced lung pathology in animal models<sup>8,9</sup>. Favipiravir, a nucleoside analogue, has also been associated with C to U and G to A transition

errors and lethal mutagenesis<sup>8,10</sup>. A meta-analysis of small clinical trials suggested that favipiravir led to a statistically significant improvement over seven days in hospitalised SARS-CoV-2 patients as measured on the WHO Ordinal Scale<sup>11</sup>, although the data collected to the date of that study were insufficient to show a significant reduction in viral load or in rates of mortality or transfer to intensive care units<sup>12</sup>.

There is a substantial history of theoretical work examining the robustness of viruses to mutation and the potential for mutagenic drugs to treat viral infection<sup>13-20</sup>. On the basis of evidence gathered from other viruses, and theoretical insights, the potential use of mutagenic antiviral drugs was suggested early in the SARS-CoV-2 pandemic<sup>21</sup>. When administered in a randomised phase 3 clinical trial to SARS-CoV-2 positive individuals who were symptomatic for five days or fewer, molnupiravir taken for five days was reported to reduce rates of hospitalisation and death<sup>22</sup>. Against a background of concerns about efficacy, it has been licensed by the Food and Drug Administration (FDA) for use in mild to moderate coronavirus disease (COVID-19) where the person was at increased risk for progression to severe disease, and under similar conditions for use in the UK<sup>23-25</sup>. A recent clinical trial has shown that molnupiravir increased the transition:transversion mutation ratio in SARS-CoV-2<sup>26</sup>. The potential use of molnupiravir in combination with favipiravir has also been suggested<sup>27</sup>. However, concerns have been raised that the mutagenic action of molnupiravir could drive the emergence of drug-resistance mutations against itself or other antiviral therapies<sup>28,29</sup>, while a recent manuscript has suggested that molnupiravir usage may have left a mutational signature upon the global SARS-CoV-2 viral population<sup>30</sup>. Particular concerns have been raised about the use of sublethal doses of mutagenic drugs, whereby treatment induces viral mutagenesis without significantly reducing viral load<sup>31</sup>. In this context there is a need for a better understanding of how the use of mutagenic drugs can affect viral populations.

A combination of evolutionary experimentation and genome sequencing has been applied in various contexts to study antiviral mutagenesis. We previously used a Syrian hamster model to demonstrate the efficacy of short-term (4 days) molnupiravir treatment on reduction of SARS-CoV-2 viral load, infectivity and lung pathology<sup>32</sup>. In this model, hamsters were infected with an intranasal inoculum of 50µl, containing  $2 \times 10^6$  TCID<sub>50</sub> SARS-CoV-2 (day 0). Treatment was initiated within an hour of the infection being founded, with animals being treated with either favipiravir, molnupiravir, a combination of the two drugs, or no drug treatment. A dose of 200 mg/kg BID (twice per day) of molnupiravir was shown to be effective in reducing infectious virus in the lung by 1.8 to 2.0 log<sub>10</sub> TCID<sub>50</sub>/mg (median Tissue Culture Infectious Dose) lung tissue, depending on the SARS-CoV-2 variant used for challenge, with evidence of significantly reduced lung pathology in treated animals. In similar experiments, treatment with 300 mg/kg BID of favipiravir for four consecutive days achieved a reduction of nearly 2 log<sub>10</sub> TCID<sub>50</sub>/mg in infectious virus titres in the lungs. On day 4 animals were euthanised for sampling of viral material from the lungs; during the intervening time the viral population may have undergone around 8-10 cycles of replication<sup>33</sup>. Data describing viral loads are reported in the original experimental papers<sup>32,34,35</sup>. The short period of replication under drug treatment mirrors the use of molnupiravir in humans, with the advantage of the increased replicability afforded by an animal study.

Here, using viral sequence data from these hamster studies, we examine the impact of effective and suboptimal dosing with molnupiravir (200mg/kg BID and <200 mg/kg BID respectively)<sup>35</sup> and favipiravir (300 mg/kg BID and <300 mg/kg BID respectively)<sup>34</sup> upon viral genetic diversity. We assess different metrics for evaluating sequence diversity in treated and untreated populations. By comparing data from animals treated with different mutagenic drugs, we consider the potential for sequence-based statistics to be used as a biomarker for clinical

efficacy. We discuss the potential influence of mutagenic drugs upon a transmitted viral population.

## Results

Sequence data described the viral population in each hamster four days after the initiation of infection. Examining these data, a statistically significant relationship was identified between the dose of either molnupiravir or favipiravir and a measure of the diversity of the viral population at the time of sample collection. For each experimental population we calculated the sum of all variant allele frequencies with respect to the consensus of the viral inoculum, irrespective of their frequency. We denote this statistic by  $q$ . It provides an estimate of the mean number of variants per virus. We found significant positive correlations between the dose of drug and the value of  $q$ ; these correlations existed for both favipiravir and molnupiravir ( $p$ -values  $< 2 \times 10^{-4}$ ) (Figure 1A, B). Our results suggest a link between the reduced fitness of the treated viral populations, seen in the reduced infectious viral titres associated with treatment, and the accumulation of mutations in the viral genomes. Our analysis suggests that molnupiravir may be more potent than favipiravir as a mutagenic agent, with a greater number of variants induced per mg of drug, although this result did not achieve statistical significance ( $p=0.063$ , Student's  $t$ -test). Samples collected following treatment with a combination of 150mg/kg BID molnupiravir and 300mg/kg BID favipiravir had a greater value of  $q$  than did populations treated with either drug individually ( $p$ -value 0.021 for molnupiravir, 0.011 for favipiravir). The increased sum of variants in this case suggests that that the two drugs may operate in a cooperative manner, with the addition of a second drug leading to a significant additional increase in the rate of viral mutation.

When evaluated in terms of the mean number of mutations per virus, the thresholds at which drug dosing became effective for SARS-CoV-2 treatment in the hamsters (i.e. 300 mg/kg BID

for favipiravir and 200 mg/kg BID for molnupiravir)<sup>32,34</sup> were similar, with estimated values of  $q$  equal to 20.8 mutations per genome (relative to the initial consensus) for favipiravir, and 21.9 mutations per genome for molnupiravir in our regression model (Figure 1A,B); these values represent a gain of between 7 and 10 variants per viral genome versus the untreated control, associated with effective drug treatment.

Our linear regression analysis excluded one data point, collected from a population treated with 500mg/day BID of molnupiravir; our regression model would predict this to have a higher number of mutations per genome than were observed. A possible reason for this observation is that under a high dose of treatment viruses lost their potential for replication before a higher burden of mutations could accumulate. Alternatively, the higher dose of drug administered may not have led to a correspondingly higher dose in blood plasma<sup>36</sup>. The collection of only a single data point at this dose prevented further investigation of this result.

Examining the frequencies of minor variants, we observed that the bulk of variation in viral populations was comprised of low-frequency variants, with, in most populations, more than half of the total variation consisting of variants at frequencies less than 1% (Figure 1C, D). A regression model showed a significant positive relationship between the drug dose and the sum of variants at frequencies of less than 1% (Supplementary Figure 1,  $p$ -values < 0.004). The preponderance of low-frequency mutations may reflect the combination of mutagenic treatment with a short period of infection, with limited time for de novo mutations to reach higher frequencies, even under positive selection<sup>37</sup>. Sequencing error likely contributes to our overall sum of variant frequencies, and complex patterns of error have been identified in sequence data describing SARS-CoV-2 infection<sup>38</sup>. However, an analysis of sequence data found no correlation between read depth and the value of  $q$ , suggesting that our results were not purely artefactual (Supplementary Figure 2).

In addition to changes in the amount of variation in viral populations, the use of favipiravir and molnupiravir were associated with significant change in the composition of genetic variants (Figure 2). Examining the mutational spectra of variants with frequencies less than 1%, we identified significantly higher proportions of C to T and G to A mutations in samples from treated, compared to untreated hamsters ( $p < 2 \times 10^{-6}$ , Student's t-test). A previous analysis of the effect of favipiravir during long-term influenza B infection examined the same statistic for variants with frequencies of less than 5%<sup>10</sup>; repeating the analysis with this higher frequency cut-off produced very similar results (Supplementary Figure 3).

Considering variants at higher frequencies, we identified a total of 306 *de novo* variants at frequencies of 5% or greater within the treated populations (Figure 3). Of these variants, 207 were non-synonymous with 98 synonymous and one nonsense mutation. Accounting for the proportion of sites that could engender synonymous or non-synonymous mutations, these values indicate a bias towards synonymous variation, with a  $\pi_N/\pi_S$  ratio equal to 0.63. An equivalent calculation for variants in untreated populations gave significantly a lower  $\pi_N/\pi_S$  ratio of 0.25 ( $p < 0.02$ , likelihood ratio test, Figure 4), indicating a greater bias against the emergence of nonsynonymous variants in untreated populations.

Among the variants identified at frequencies of 5% or greater we did not identify a strong signal of repeated evolution, as might be expected if the virus was channelled into specific pathways of adaptation. Among the variants observed in the 30 treated populations, only six were observed in more than one of the populations. As a potential exception to this, the A21804G variant, which codes for the N81S mutation in the Spike protein, was observed as a consensus-level substitution (i.e. with a frequency in excess of 50%) in two populations that were treated with favipiravir. Variants at frequencies 5% or greater were roughly uniformly distributed across the

genome, following a pattern seen in the general evolution of SARS-CoV-2 in the human population<sup>39</sup>. In contrast to our identification of a link between the sum of variant allele frequencies and drug dose, there was no significant relationship between drug dose and the number of variants observed at frequencies of 5% or greater (Supplementary Figure 4), suggesting that the expansion of new variants to these frequencies was dominated by stochastic events.

Among variants observed at frequencies of 5% or greater, we did not find evidence of adaptation that could confer resistance against the antiviral drugs used in these experiments. The variants we observed did not replicate known remdesivir-resistance variants observed in SARS-CoV-2<sup>40-42</sup>. Two consensus-level variants (i.e. > 50% with respect to the original consensus) were observed in RdRp from hamsters given treatment with favipiravir, but nothing was observed at these frequencies following treatment with molnupiravir (Figure 5). We did not identify mutations in the RdRp binding site, or in sites that have been associated with drug resistance for other RdRp inhibitors (Supplementary Table 1). Within the Spike protein consensus-level variants were observed in populations treated with favipiravir, two of which coded for the N81S substitution, while a single consensus-level variant was observed for populations treated with molnupiravir (Supplementary Figure 5, Supplementary Table 2). One variant, I692V, observed at a frequency of 8% in a population treated with molnupiravir, was among the set of variants which defined the B.1.1.198 variant of concern, identified early in the pandemic<sup>43</sup>. Although it is possible that treatment with mutagenic drugs could promote the emergence of new variants of concern, our data were not sufficient to draw firm conclusions on this hypothesis.

Taken together our results in combination suggest that the primary effect of short-term favipiravir and molnupiravir treatment upon SARS-CoV-2 genomes was an accumulation of low-



frequency mutations, biased towards the accumulation of C to U and G to A variants, and without significant gain of variants at higher frequencies. Lower doses of antiviral drug, which did not have a significant effect upon viral load, had a clear, but smaller, effect upon viral populations. The bias in our result towards the accumulation of low-frequency variants stands in contrast with cases that have described the extended use of mutagenic drugs. For example, a study of the prolonged treatment of norovirus infection with favipiravir showed a progressive accumulation of variants at higher frequencies<sup>44</sup>. It is likely that the shorter timespan of both infection and treatment in our experimental model underpin this difference.

## Discussion

Both molnupiravir and favipiravir are mutagenic antiviral drugs. When used against SARS-CoV-2 they interfere with viral replication, leading to an increase in the mutation rate of the virus. Increasing the mutation rate raises the possibility that consequent increased genetic variation will result in the evolution of drug resistance mutations, or of mutations that could make a virus better adapted to a new host. Previously, a SARS-CoV-2 infection model in Syrian hamsters showed a dose-dependent reduction in viral loads, infectivity and lung pathology in molnupiravir or favipiravir treated animals<sup>32,34,35</sup>. Taking this further, we here used viral genome sequence data to explore how favipiravir and molnupiravir shaped the viral populations in these animals. Studying samples from hamsters provides an increased degree of replicability versus cases of human infection, and the short period of treatment (following initiation of infection with a large number of viruses) provides a reasonable model of the use of antivirals against a recently-established case of human SARS-CoV-2 infection.

We identified statistically significant correlations between the dose of favipiravir or molnupiravir given to the animals and a statistic, calculated as the sum of the variant allele frequencies in the viral genome. The increase in our statistic was linked to an increase in low-frequency variation,

comprising variants at <1% frequency, which in most populations made up the majority of the total variation. When the mutational spectra of favipiravir and molnupiravir, were matched, significant increases were seen in the proportion of low-frequency variation that was comprised of C to U and G to A mutations.

Given the mutagenic action of favipiravir and molnupiravir, our result is not unexpected, but it provides a novel perspective upon how the use of mutagenic drugs could contribute to the global evolution of the virus. Our result is not unexpected in so far as an increase in the viral mutation rate would be expected to increase the extent of genetic variation in viral genomes. Given a short period of infection, there is limited time for novel mutations to reach high frequencies, so that most additional variation occurs at low frequency. Our result provides a novel perspective on the consequences of mutagenic drug use by means of our inclusion of low-frequency variation. A common procedure in the evaluation of viral sequence data is the identification of single nucleotide polymorphisms (SNPs), which on the basis of statistical analysis can be unambiguously identified in a population<sup>26,45</sup>. This identification usually requires the collection of a minimum number of reads, and for the variant to have been observed at a cut-off frequency, often of around 2%<sup>46</sup>. While this approach identifies unambiguous individual variants, it potentially neglects changes in the viral population at lower frequencies. Although our statistic  $q$  will be affected by sequencing error, the consistent experimental procedures used in our experiment imply that differences in the statistic between samples should be reliable.

Our estimate that effective treatment with favipiravir or molnupiravir led to between 7 and 10 additional variants per viral genome sequence provides some insight into the potential for mutagenic drugs to affect the global viral population. Transmission bottlenecks for SARS-CoV-2 are usually tight, with few viruses initiating new infections<sup>38,47</sup>. Where a transmission bottleneck involves a single viral particle, any variants carried by the founder virus will be fixed

in the recipient population. In this way, while selection could play a role in reducing the number of variants transmitted, and while treatment with mutagenic drugs will potentially reduce the mean viral load and hence the potential for transmission, the potential exists, if transmission does take place, for multiple variants to accumulate in the viral lineage in a short period of time. Long branches in the tree of SARS-CoV-2 sequences could potentially be created without the need for a long period of infection, or a need for variants to reach high frequencies in the donor individual.

Our study identified that an effective dose of treatment for both favipiravir or molnupiravir led to a viral population with close to 20 mutations per virus. Our statistic describing the number of variants per viral genome may have some utility as a clinical biomarker for evaluating the effectiveness of mutagenic drugs. However, several barriers remain to be overcome ahead of its use. Firstly, a greater understanding would be required of how sequencing error affects the absolute value of this statistic: Different sequencing approaches may have distinct error profiles. Secondly, in our dataset identical doses of drug led to variable levels of viral load; further investigation of the extent of this variation would be required. Thirdly, our samples were collected after four days. In the absence of multiple samples per individual we could not assess whether this statistic had reached an equilibrium value. Finally, hamsters likely represent a more homogeneous population than human patients. Even if a dose of drug were to produce a consistent fitness reduction, the intrinsic viral fecundity in the absence of treatment may vary according to the specific immune response of the individual and the genotype of the virus.

By contrast to our results for low-frequency variation, we did not identify a meaningful signal in the numbers of higher-frequency variants in populations treated with a higher drug dose. This result may also arise from the short time span of the experiment: Variants at higher frequencies take longer to emerge. The general lack of repeated variants in treated populations is consistent

with a process whereby variants rose to high frequency by chance, via a random process of mutation and genetic drift. Against this background, we cannot exclude the possibility of having observed some genuinely beneficial mutations, whether conferring adaptation to the hamster host or some form of drug resistance. Adaptive viral evolution in multiple directions has previously been observed in experimental cases of animal infection<sup>48</sup>. Thus, a process in which distinct beneficial mutations emerged in different hosts cannot entirely be ruled out. Our identification of a clear and statistically significant signal in one metric of genetic variation, but the lack of a significant signal in another, highlights the need for care in using sequence data to evaluate the effects of mutagenic drugs on viral populations.

Our result concerning animals treated with suboptimal doses of drug validates suggestions that such treatment could increase viral genetic diversity without a significant reduction in viral load. A complex relationship likely exists between the potential for variant transmission and drug dose, with higher doses creating more variants in the viral population, but reducing the probability of transmission due to the reduction in viral load.

Despite their limitations, our results suggest that treatment with mutagenic drugs increased the numbers of mutations accumulating in SARS-CoV-2 in a roughly continuous manner, with the increase in load being proportional to the level of antiviral dosing. To the extent that mutagenic drugs result in a loss of viral infectivity, it can be concluded that the burden of mutations they impose upon the virus is deleterious, with successive gains of mutations being to the detriment of the viral population. On this basis we would expect that, while suboptimal dosing could lead to the evolution of more mutated viruses, the emerged viruses would mostly be of reduced fitness compared to those transmitted from an untreated population. Our study did not find evidence for either the emergence of antiviral resistance, or the systematic emergence of beneficial variants that would militate against the clinical use of mutagenic drugs. In the

absence of specific evidence of negative effects, clinical studies describing the efficacy (or otherwise) of treatment should be given priority in determining the use of mutagenic antiviral drugs.

## Methods

### *Virus samples*

The data for this study was generated from a Syrian hamster model designed to demonstrate the efficacy of short term (4 days) molnupiravir and favipiravir in reducing SARS-CoV-2 viral load, infectivity and lung pathology. Details of these experiments have been described in previous publications<sup>32,34,35</sup>. Large numbers of viral genomes were recovered from these animals (Supplementary Figure S6).

### *Sequence Analysis*

Fastq files were processed with Trimmomatic<sup>49</sup> using default settings, retaining reads for which both reads in a pair passed filtering. The remaining reads were aligned to the MN908947.3 reference sequence using bwa<sup>50</sup>, before using samtools<sup>51</sup> to remove reads that did not align to the primary alignment. Further processing was conducted using the SAMFIRE package<sup>52</sup>. Filtering was conducted to remove low quality base pairs (PHRED score <30), following which variants were identified genome-wide. Samples were kept for analysis if the mean coverage of the viral genome was at least 150x. Applying this criteria gave us samples from a total of seven untreated hamsters, 17 hamsters treated with favipiravir (3 at 150mg, 8 at 300mg, 1 at 450mg, and 5 at 500mg), and 10 hamsters treated with molnupiravir (5 at 75mg, 2 at 150mg, 2 at

200mg 1 at 500mg), alongside two hamsters treated with both favipiravir and molnupiravir. A full list of samples is provided in Supplementary Table 3.

### *Number of variants per genome*

An estimate of the mean number of variants per viral genome was calculated using the aligned sequence data. We assessed this statistic by calculating the sum of allele frequencies across the genome. We first identified consensus alleles as the most common nucleotides present in the inoculum population. At position  $i$  in the genome we then calculated the frequency  $q_{ia}$  of the allele  $a$  for each non-consensus allele  $a$ , considering sites with a coverage of at least 150 reads but with no filter on variant allele frequencies. We then calculated the sum

$$q = \sum_i \sum_a q_{ia}$$

The sum in this case encompasses all variants, irrespective of their frequency. In this calculation, requiring 150 reads at a site removes the potential for sites with low read depth and a small number of erroneous reads to unduly bias our measurement of  $q$ , but comes at the cost of potentially excluding genuine variation at poorly covered sites. Reducing the threshold number of reads required to 50 or 10, or removing this requirement altogether did not produce dramatic changes in the statistics obtained (Supplementary Figure S7), with extremely strong linear relationships identified between values obtained under different thresholds.

Our use of linear regression to analyse these data was motivated by a crude model of variant evolution at low frequencies. Making the assumption that a site in a genome may contain only two possible alleles, and ignoring linkage between variants, the frequency of a deleterious allele

is expected to tend, under mutation-selection balance, to an equilibrium frequency of  $x=\mu/s$ , where  $\mu$  is the mutation rate and  $s$  is the magnitude of selection against the allele<sup>53</sup>, according to the equation

$$\frac{dx}{dt} = \mu - sx$$

Furthermore, assuming mutation rates from one allele to the other are equal, the frequency of a neutral allele is expected to tend to a frequency of one half, according to the equation

$$\frac{dx}{dt} = \mu(1 - 2x)$$

In both of these formulas, the change in the frequency  $x$  is proportional to  $\mu$  when  $x$  is small. Although positive selection and linkage disequilibrium will affect the evolution of the viral population, we nevertheless sought to fit a linear model to our data.

#### *Analysis of variant composition*

The proportion of low-frequency variation of distinct mutational classes was measured using variants at frequencies of 1% or below. Briefly, the calculation of  $q$  was repeated, considering exclusively low-frequency variants, following which the proportion of this sum that is comprised of each of the 12 mutational classes was calculated.

#### *Non-synonymous and synonymous variation*

In order to explore the potential adaptive evolution of the viral populations, we identified variants which had reached a frequency of 5% or more. To analyse the composition of these variants we calculated  $\pi_N/\pi_S$ , defined as

$$\frac{\pi_N}{\pi_S} = \frac{cN/oN}{cS/oS}$$

Where  $cN$  and  $cS$  were the genome-wide counts of non-synonymous and synonymous variants reaching a frequency of 5% or more, and  $oN$  and  $oS$  are the number of potential non-synonymous and synonymous variants that could arise given the various reading frames in the SARS-CoV-2 genome, using the consensus sequence of the initial viral population as a reference. A value of this statistic less than one suggests that the emergence of variants to a frequency of 5% or greater was biased by selection against nonsynonymous variants.

To calculate a p-value for the difference between  $\pi_N/\pi_S$  values, we calculated likelihoods for the underlying proportions  $p_U$  and  $p_T$  of variants in the untreated and treated populations which were non-synonymous, relative to the opportunity, and based upon the observations made. Where we observed  $cN_U$  nonsynonymous variants and  $cS_U$  synonymous variants in the untreated population, and  $cN_T$  nonsynonymous variants and  $cS_T$  synonymous variants in the treated population, a likelihood for the values  $p_U$  and  $p_T$  is given by

$$L(p_U, p_T) = \binom{cN_U + cS_U}{cN_U} (p_U)^{cN_U} (1 - p_U)^{cS_U} \binom{cN_T + cS_T}{cN_T} (p_T)^{cN_T} (1 - p_T)^{cS_T}$$



This distribution is shown in Supplementary Figure S8. Defining the hypothesis  $H_1: p_U < p_T$ , and  $H_0: p_U \geq p_T$ , and integrating over the corresponding regions of this distribution, we obtained the p-value given in the main text.

#### *Identification of spike mutations of interest*

As the hamsters were infected with Wuhan-like virus, we assumed mutations of interest would be those identified in the human population in variants of concern / variants of interest. Mutations arising were looked up in tables of mutations with antibody escape information from a previous publication and referenced material<sup>43</sup>.

#### *Visualization of protein structures*

Images of protein structures were made with the Visual Molecular Dynamics software package<sup>54</sup>.

#### **Author contributions:**

Conceptualisation: CI, JG-A, SG, JB. Data Curation: CI, JG-A, SG. Formal Analysis: CI, JG-A, SG, OC. Funding Acquisition: JB. Investigation: CI, JG-A, SG, OC, JP, SR. Methodology: CI, JG-A, SG, RA. Project Administration: CI, RA, JN, JB. Software: CI, JG-A, SG. Supervision: JG-A, SG, RA, JN, JB. Validation: CI, JG-A, RA, JN. Visualisation: CI, JG-A, SG, OC. Writing - Original Draft Preparation: CI, JG-A, SG, JB. Writing - Review and Editing: CI, JG-A, OC, RA, JN, JB.

## Acknowledgements:

This work was supported by the UK Medical Research Council (MC\_UU\_00034/6).

## Data availability:

Sequence data is available from the Sequence Read Archive with accession number PRJNA935666.

## References

1. Reynard, O. *et al.* Identification of a New Ribonucleoside Inhibitor of Ebola Virus Replication. *Viruses* **7**, 6233–6240 (2015).
2. Painter, G. R. *et al.* The prophylactic and therapeutic activity of a broadly active ribonucleoside analog in a murine model of intranasal venezuelan equine encephalitis virus infection. *Antiviral Research* **171**, 104597 (2019).
3. Kabinger, F. *et al.* Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nat Struct Mol Biol* **28**, 740–746 (2021).
4. Wahl, A. *et al.* SARS-CoV-2 infection is effectively treated and prevented by EIDD-2801. *Nature* **591**, 451–457 (2021).
5. Agostini, M. L. *et al.* Small-Molecule Antiviral  $\beta$ -D-N<sup>4</sup>-Hydroxycytidine Inhibits a Proofreading-Intact Coronavirus with a High Genetic Barrier to Resistance. *J Virol* **93**, e01348-19 (2019).

6. Yoon, J.-J. *et al.* Orally Efficacious Broad-Spectrum Ribonucleoside Analog Inhibitor of Influenza and Respiratory Syncytial Viruses. *Antimicrob Agents Chemother* **62**, e00766-18 (2018).
7. Malone, B. & Campbell, E. A. Molnupiravir: coding for catastrophe. *Nat Struct Mol Biol* **28**, 706–708 (2021).
8. Sheahan, T. P. *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.* **12**, eabb5883 (2020).
9. Fischer, W. A. *et al.* A phase 2a clinical trial of molnupiravir in patients with COVID-19 shows accelerated SARS-CoV-2 RNA clearance and elimination of infectious virus. *Sci. Transl. Med.* **14**, eabl7430 (2022).
10. Lumby, C. K. *et al.* Favipiravir and Zanamivir Cleared Infection with Influenza B in a Severely Immunocompromised Child. *Clinical Infectious Diseases* **71**, e191–e194 (2020).
11. Marshall, J. C. *et al.* A minimal common outcome measure set for COVID-19 clinical research. *The Lancet Infectious Diseases* **20**, e192–e197 (2020).
12. Hassanipour, S. *et al.* The efficacy and safety of Favipiravir in treatment of COVID-19: a systematic review and meta-analysis of clinical trials. *Sci Rep* **11**, 11022 (2021).
13. Lynch, M., Bürger, R., Butcher, D. & Gabriel, W. The Mutational Meltdown in Asexual Populations. *Journal of Heredity* **84**, 339–344 (1993).
14. Lynch, M., Conery, J. & Burger, R. Mutation Accumulation and the Extinction of Small Populations. *The American Naturalist* **146**, 489–518 (1995).
15. Bull, J. J., Sanjuán, R. & Wilke, C. O. Theory of Lethal Mutagenesis for Viruses. *JVI* **81**, 2930–2939 (2007).
16. Luring, A. S., Frydman, J. & Andino, R. The role of mutational robustness in RNA virus evolution. *Nat Rev Microbiol* **11**, 327–336 (2013).

17. Stern, A. *et al.* Costs and Benefits of Mutational Robustness in RNA Viruses. *Cell Reports* **8**, 1026–1036 (2014).
18. Matuszewski, S., Ormond, L., Bank, C. & Jensen, J. D. Two sides of the same coin: A population genetics perspective on lethal mutagenesis and mutational meltdown. *Virus Evolution* **3**, (2017).
19. Pénisson, S., Singh, T., Sniegowski, P. & Gerrish, P. Dynamics and Fate of Beneficial Mutations Under Lineage Contamination by Linked Deleterious Mutations. *Genetics* **205**, 1305–1318 (2017).
20. Jensen, J. D., Stikeleather, R. A., Kowalik, T. F. & Lynch, M. Imposed mutational meltdown as an antiviral strategy. *Evolution* **74**, 2549–2559 (2020).
21. Jensen, J. D. & Lynch, M. Considering mutational meltdown as a potential SARS-CoV-2 treatment strategy. *Heredity* **124**, 619–620 (2020).
22. Jayk Bernal, A. *et al.* Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. *N Engl J Med* **386**, 509–520 (2022).
23. Thorlund, K., Sheldrick, K., Meyerowitz-Katz, G., Singh, S. & Hill, A. Making Statistical Sense of the Molnupiravir MOVE-OUT Clinical Trial. *The American Journal of Tropical Medicine and Hygiene* **106**, 1301–1304 (2022).
24. Mahase, E. Covid-19: UK becomes first country to authorise antiviral molnupiravir. *BMJ* **375**, (2021).
25. Dyer, O. Covid-19: FDA expert panel recommends authorising molnupiravir but also voices concerns. *BMJ* **375**, (2021).
26. Donovan-Banfield, I. *et al.* Characterisation of SARS-CoV-2 genomic variation in response to molnupiravir treatment in the AGILE Phase IIa clinical trial. *Nat Commun* **13**, 7284 (2022).
27. Eloy, P., Le Grand, R., Malvy, D. & Guedj, J. Combined treatment of molnupiravir and favipiravir against SARS-CoV-2 infection: One + zero equals two? *eBioMedicine* **74**, 103663 (2021).

28. Singh, A. K., Singh, A., Singh, R. & Misra, A. Molnupiravir in COVID-19: A systematic review of literature. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **15**, 102329 (2021).
29. Swanstrom, R. & Schinazi, R. F. Lethal mutagenesis as an antiviral strategy. *Science* **375**, 497–498 (2022).
30. Sanderson, T. *et al.* A molnupiravir-associated mutational signature in global SARS-CoV-2 genomes. *Nature* (2023) doi:10.1038/s41586-023-06649-6.
31. Nelson, Chase W & Otto, Sarah P. Mutagenic antivirals: the evolutionary risk of low doses. *Virological* <https://virological.org/t/mutagenic-antivirals-the-evolutionary-risk-of-low-doses/768> (2021).
32. Abdelnabi, R. *et al.* Molnupiravir Inhibits Replication of the Emerging SARS-CoV-2 Variants of Concern in a Hamster Infection Model. *The Journal of Infectious Diseases* **224**, 749–753 (2021).
33. Bar-On, Y. M., Flamholz, A., Phillips, R. & Milo, R. SARS-CoV-2 (COVID-19) by the numbers. *eLife* **9**, e57309 (2020).
34. Kaptein, S. J. F. *et al.* Favipiravir at high doses has potent antiviral activity in SARS-CoV-2-infected hamsters, whereas hydroxychloroquine lacks activity. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 26955–26965 (2020).
35. Abdelnabi, R. *et al.* The combined treatment of Molnupiravir and Favipiravir results in a potentiation of antiviral efficacy in a SARS-CoV-2 hamster infection model. *eBioMedicine* **72**, 103595 (2021).
36. Nguyen, T. H. T. *et al.* Favipiravir pharmacokinetics in Ebola-Infected patients of the JIKI trial reveals concentrations lower than targeted. *PLoS Negl Trop Dis* **11**, e0005389 (2017).
37. Morris, D. H. *et al.* Asynchrony between virus diversity and antibody selection limits influenza virus evolution. *eLife* **9**, e62105 (2020).

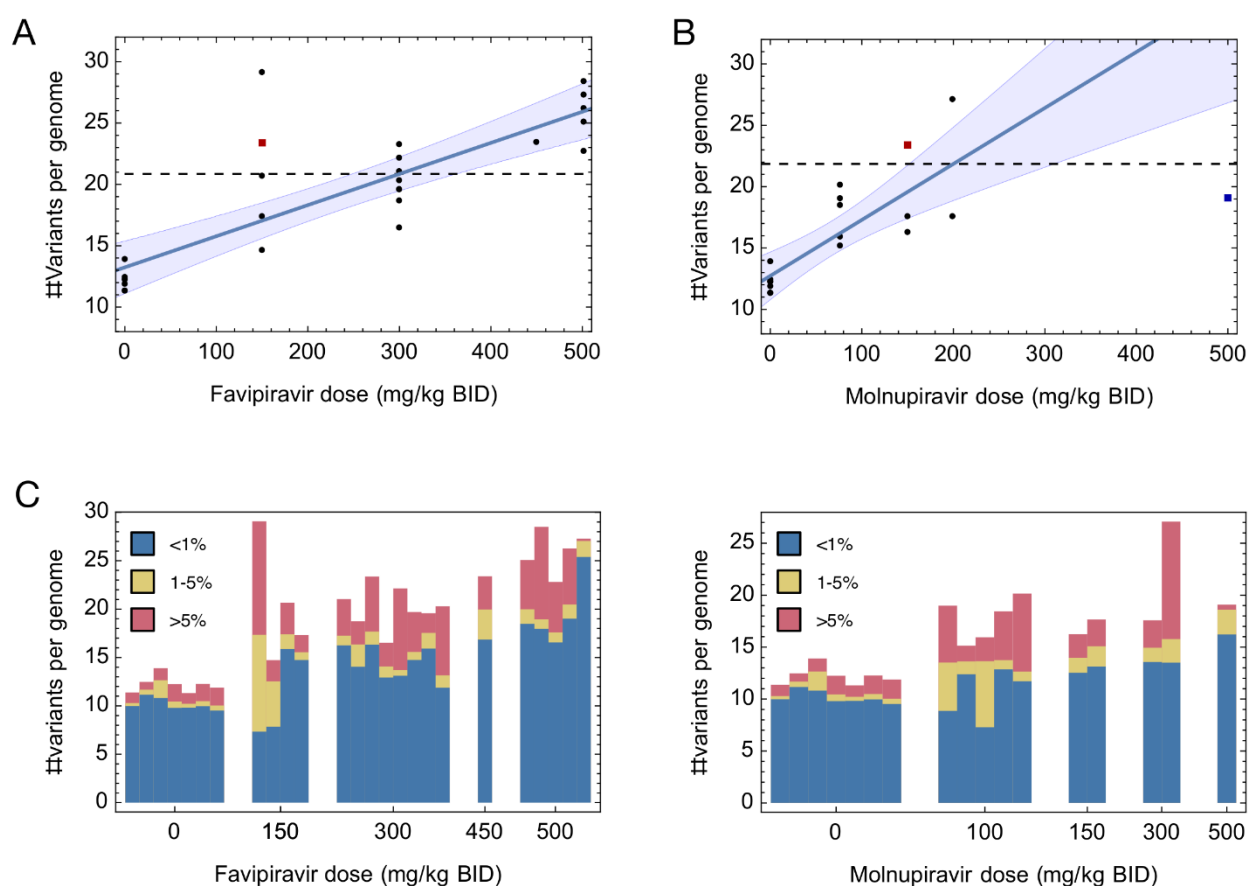
38. Lythgoe, K. A. *et al.* SARS-CoV-2 within-host diversity and transmission. *Science* eabg0821 (2021) doi:10.1126/science.abg0821.
39. Harari, S. *et al.* Drivers of adaptive evolution during chronic SARS-CoV-2 infections. *Nat Med* **28**, 1501–1508 (2022).
40. Gandhi, S. *et al.* De novo emergence of a remdesivir resistance mutation during treatment of persistent SARS-CoV-2 infection in an immunocompromised patient: a case report. *Nat Commun* **13**, 1547 (2022).
41. Stevens, L. J. *et al.* Mutations in the SARS-CoV-2 RNA-dependent RNA polymerase confer resistance to remdesivir by distinct mechanisms. *Sci. Transl. Med.* **14**, eabo0718 (2022).
42. Szemiel, A. M. *et al.* In vitro selection of Remdesivir resistance suggests evolutionary predictability of SARS-CoV-2. *PLoS Pathog* **17**, e1009929 (2021).
43. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* **19**, 409–424 (2021).
44. Ruis, C. *et al.* Mutagenesis in Norovirus in Response to Favipiravir Treatment. *n engl j med* **4** (2018).
45. Alteri, C. *et al.* A proof-of-concept study on the genomic evolution of Sars-Cov-2 in molnupiravir-treated, paxlovid-treated and drug-naïve patients. *Commun Biol* **5**, 1376 (2022).
46. McCrone, J. T. *et al.* Stochastic processes constrain the within and between host evolution of influenza virus. *eLife* **7**, e35962 (2018).
47. Bendall, E. E. *et al.* Rapid transmission and tight bottlenecks constrain the evolution of highly transmissible SARS-CoV-2 variants. *Nat Commun* **14**, 272 (2023).
48. Illingworth, C. J. R. Fitness Inference from Short-Read Data: Within-Host Evolution of a Reassortant H5N1 Influenza Virus. *Mol Biol Evol* **32**, 3012–3026 (2015).
49. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

50. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
51. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
52. Illingworth, C. J. R. SAMFIRE: multi-locus variant calling for time-resolved sequence data. *Bioinformatics* **32**, 2208–2209 (2016).
53. Haldane, J B S. The Effect of Variation on Fitness. *The American Naturalist* **71**, (1937).
54. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).
55. Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **368**, 779–782 (2020).
56. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e6 (2020).

## Figure Legends

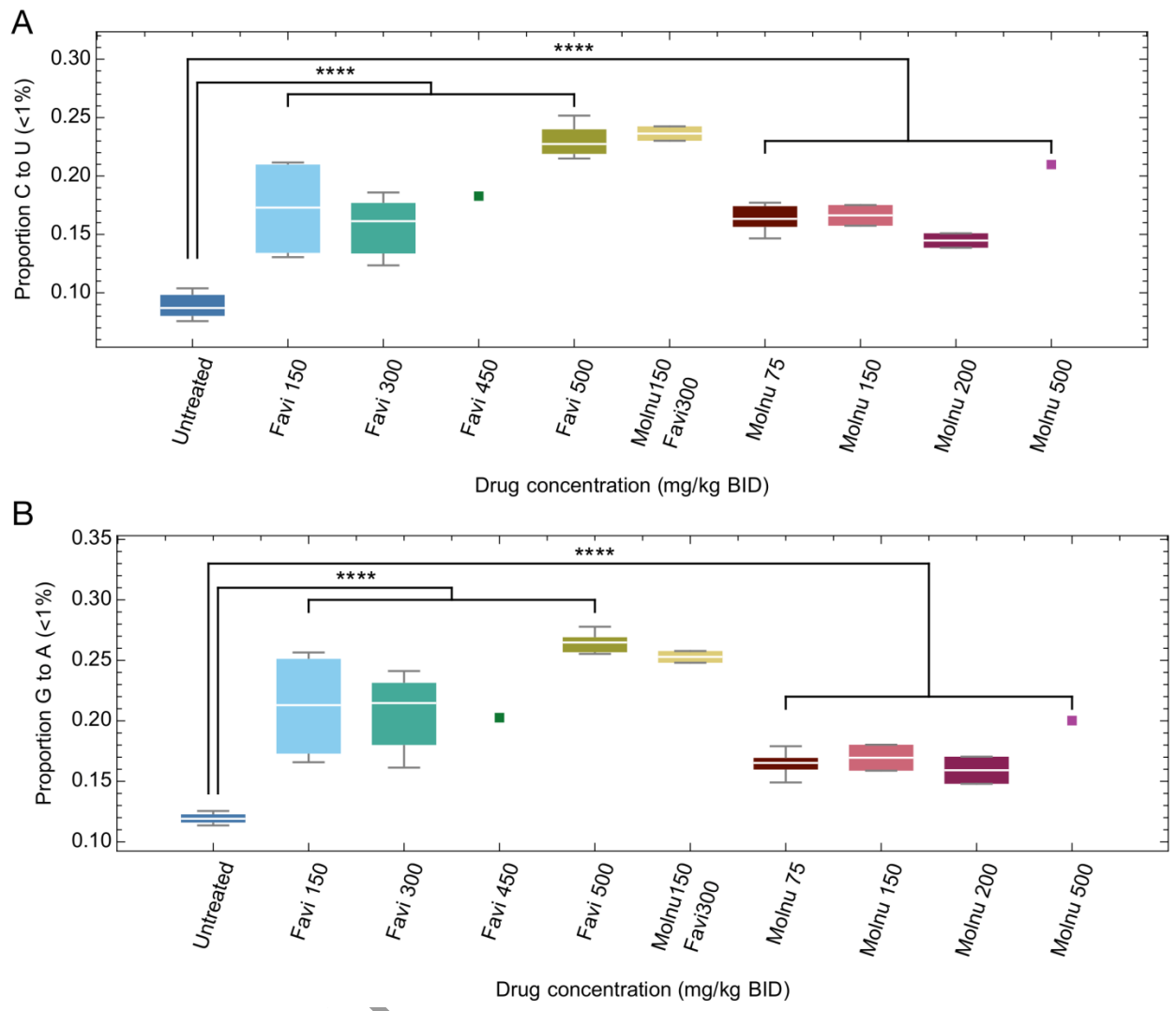
**Figure 1 : Statistics of virus diversity in treated and untreated populations. A.** Increases in the mean number of variants per viral genome were observed in populations treated with increasing doses of favipiravir. Black dots show genome-wide statistics calculated from populations: Each dot represents the viral population in a single hamster. The red square shows the mean value for populations treated with 300mg favipiravir and 150mg molnupiravir. The blue square at 500mg molnupiravir shows a datapoint that was omitted from the regression calculation; this point is discussed in the main text. Shading indicates a 95% confidence interval for the regression line. A horizontal dashed line indicates the model threshold at the point of effective treatment. **B.** Increases in the mean number of variants per viral genome per virus were observed in populations treated with increasing doses of molnupiravir. **C.** Composition of

variants denoted by variant frequency, ordered by favipiravir dose. Bars show the total number of variants per genome (plotted as points in Figure 1A), decomposed into the contributions to this total from variants of different frequencies. Each bar describes statistics for the viral population in a single hamster. **D.** Composition of variants denoted by variant frequency, ordered by molnupiravir dose. Bars show the total number of variants per genome (plotted as points in Figure 1A), decomposed into the contributions to this total from variants of different frequencies.



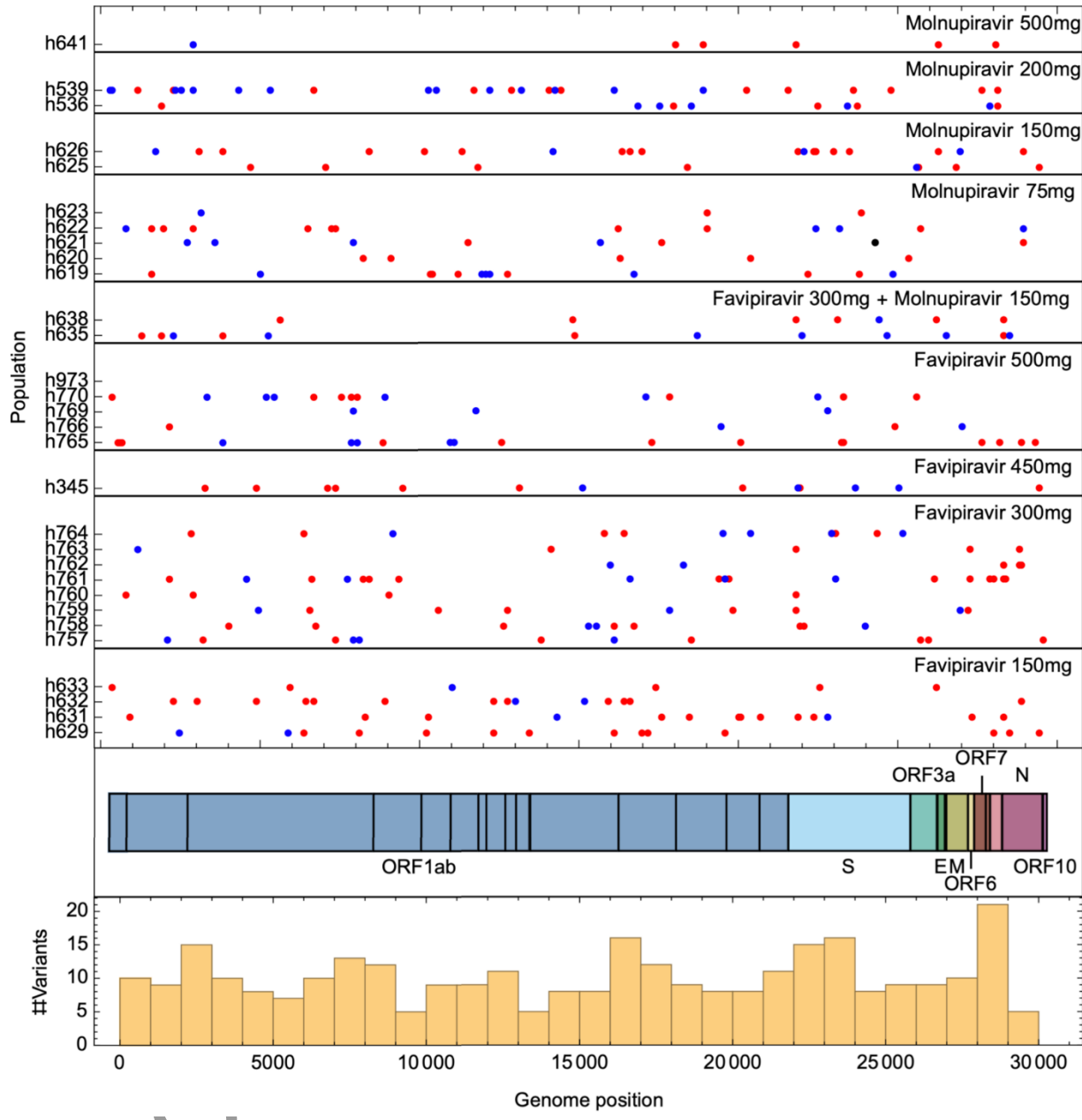
**Figure 2. Mutational spectra of low-frequency variants.** **A.** The proportion of low frequency variation (variant frequency <1%) that was comprised of C to U mutations was higher for treated populations than for untreated populations. Squares show treatments for which a single data point was collected. **B.** The proportion of low frequency variation that was comprised of G to A mutations was higher for treated populations than for untreated populations.



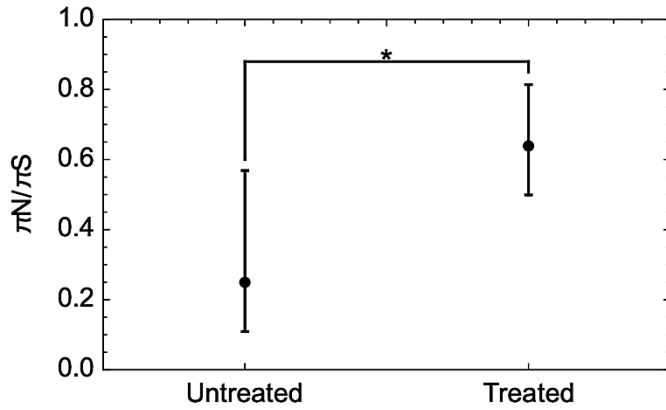


**Figure 3: Genomic locations of variants in treated viral populations.** The locations of variants which reached a frequency of at least 5% in the viral population are shown in red (non-synonymous variant), blue (synonymous variant) or black (nonsense variant). With the exception of variants transmitted through standing variation, very little replication of variants between treated populations was observed.

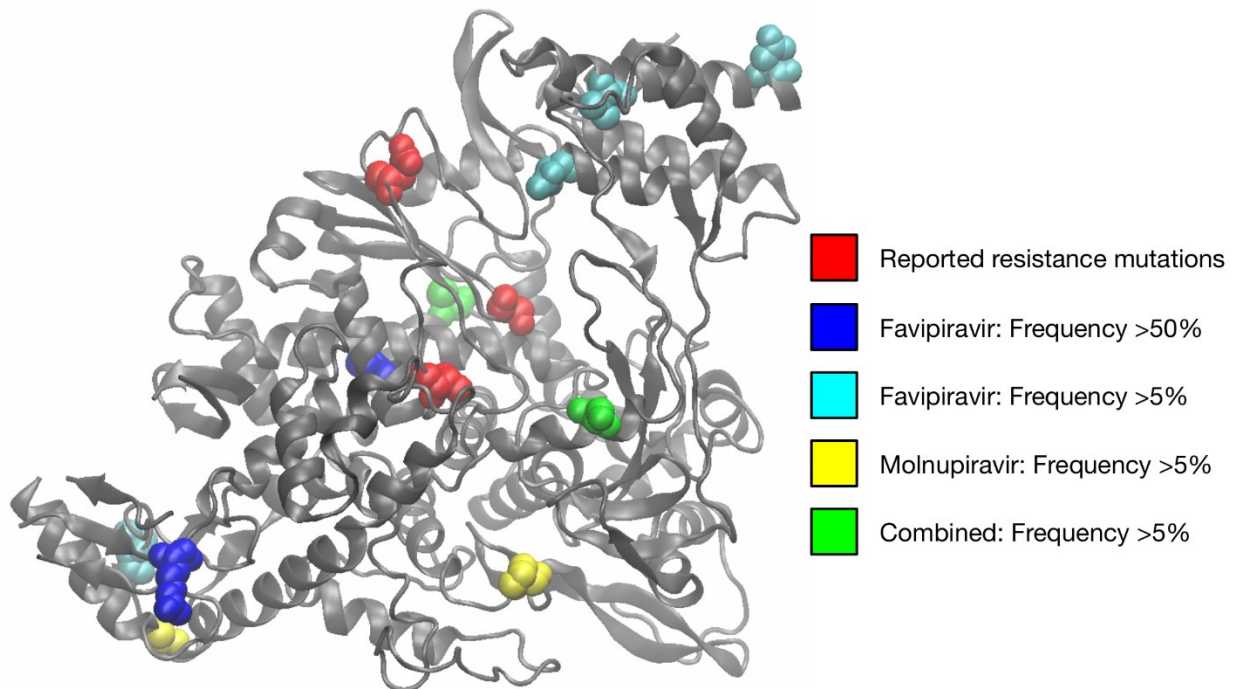
ACCEPTED



**Figure 4:**  $\pi_N/\pi_S$  was lower in untreated than in treated populations. Error bars were calculated using a likelihood model.



**Figure 5: Variants observed in RdRp at frequencies of 5% or greater.** Sites reported to convey resistance to remdesivir are shown in red vdW representation<sup>40–42</sup>. Other colours show variants observed in hamster populations at frequencies of 5% or more. Image created in VMD<sup>54</sup> based on the PDB structure 6m71<sup>55</sup>.



## Supplementary Figure Legends

**Supplementary Figure S1:** Increases in the sum of genetic variants with frequencies less than 1% were associated with increases in the dose of drug administered. Significant correlations ( $p$ -value  $< 0.004$ ) were observed for each dataset. Black dots show genome-wide statistics calculated for individual viral populations: A single dot describes the viral population in a single hamster. The solid blue line shows a linear regression model fitted to the data, with surrounding 95% confidence interval.

**Supplementary Figure S2:** No correlation was observed between the mean read depth of a sample and the number of mutations per genome ( $p$ -value 0.60,  $T$ -test)

**Supplementary Figure S3: Mutational spectra of low-frequency variants (<5%).** **A.** The proportion of low frequency variation (variant frequency  $< 5\%$ ) that was comprised of C to U mutations was higher for treated populations than for untreated populations. Squares show treatments for which a single data point was collected. **B.** The proportion of low frequency variation that was comprised of G to A mutations was higher for treated populations than for untreated populations.

**Supplementary Figure S4:** **A.** Numbers of de novo variants reaching  $> 5\%$  frequency in populations under each treatment. **B.** Numbers of de novo consensus changes ( $> 50\%$  frequency) in populations under each treatment.

**Supplementary Figure S5: Variants observed in Spike at frequencies of 5% or greater.** Sites reported to convey resistance to remdesivir are shown in red vdW representation<sup>40–423</sup>. Other colours show variants observed in hamster populations at frequencies of 5% or more.

Image created in VMD<sup>54</sup> based on the PDB structure 6vxx . One nonsynonymous mutation, N81S, was observed in two independent populations treated with favipiravir and is shown in black.

**Supplementary Figure S6:** Samples which produced sequence data with a mean genome-wide read-depth in excess of 150x coverage arose from populations with high numbers of genomes per ml of collected material.

**Supplementary Figure S7:** Effect of different approaches to filtering data upon the calculation of the total variants per genome. Our default approach calculates the sum of all minor allele frequencies across all sites with a read depth of at least 150 reads. Reducing the threshold to a lower number of reads, or removing the filtering by read depth produces broadly similar results across the dataset, with a strong correlation between values of the statistic  $q$ .

**Supplementary Figure S8:** The fraction of non-synonymous variants reaching a frequency of at least 5% was lower in untreated than in treated populations. Joint likelihood function  $L(p_U, p_T)$  for the underlying proportion of non-synonymous variants in untreated ( $p_U$ ) and treated populations ( $p_T$ ). The relative likelihood, describing the likelihood as a fraction of the maximum likelihood value, is shown. The orange region denotes the section of space for which  $p_U < p_T$ , while the blue region denotes that for which  $p_T \geq p_U$ .