bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SARS-CoV-2 Quasispecies Mediate Rapid Virus Evolution and Adaptation

M. Zeeshan Chaudhry[1,9,*], Kathrin Eschke[1,9], Martina Grashoff[2], Leila Abassi[1], Yeonsu Kim[1], Linda Brunotte[3], Stephan Ludwig[3], Željka Mačak Šafranko[4], Ivan-Christian Kurolt[4], Alemka Markotić[4], Andrea Kröger[2,5], Frank Klawonn[6,7], Luka Cicin-Sain[1,8,*]

[1]Department of Vaccinology and Applied Microbiology, Helmholtz Centre for Infection Research, Braunschweig, Germany.

[2]Research Group Innate Immunity and Infection, Helmholtz Centre for Infection Research, Braunschweig, Germany.

[3]Institut für Virologie (IVM), Westfälische Wilhelms-Universität Münster, Münster, Germany.

[4]University Hospital for Infectious Diseases “Dr. Fran Mihaljević”, Zagreb, Croatia.

[5]Institute of Medical Mircobiology and Hospital Hygiene, Otto von Guericke University, Magdeburg, Germany

[6]Biostatistics Group, Helmholtz Centre for Infection Research, Braunschweig, Germany.

[7]Department of Computer Science, Ostfalia University, Wolfenbüttel, Germany.

[8]Centre for Individualized Infection Medicine (CIIM), Hannover, Germany.

[9]These authors contributed equally.

*Correspondence: zeeshan.chaudhry@helmholtz-hzi.de (M.Z.C), luka.cicin-sain@helmholtz-hzi.de (L.C.-S.)

Keywords

SARS-CoV-2; Quasispecies; Evolution; Furin Cleavage Site; Heparan Sulfate Binding; Spike Mutation; Deep Sequencing; Mutation Rate

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

The pandemic spread of SARS-CoV-2 and the resulting global healthcare emergency warrants a better understanding of its biology. The potential of SARS-CoV-2 evolution to create novel dangerous variants remains underexplored. Thus, we passaged SARS-CoV-2 in defined conditions and determined its genomic adaptation dynamics. We demonstrate the presence of remarkably stable SARS-CoV-2 quasispecies. We further show that the quasispecies nature of the virus population ensured rapid adaptation of the spike PRRARS motif upon passaging in Vero cells. On the other hand, SARS-CoV-2 replication in TMPRSS2 expressing cells led to a reverse mutation at the same site. We observed the emergence of novel mutations in envelope protein upon virus culture in Calu-3 and Caco-2 cells. Finally, we show that the heparan sulfate-binding motif (PRRARS) of the SARS-CoV-2 S protein acted as a determinant of negative growth selection. Overall, our research has far-reaching implications for development of antiviral strategies, suggesting viral quasispecies may facilitate rapid emergence of escape mutants under selection pressure, such as the treatment with antivirals against SARS-CoV-2.

INTRODUCTION

A new infectious disease, now termed as COVID-19, emerged from Wuhan city in December 2019 and rapidly spread over the world. The pandemic spread of the disease that affected all human societies was found to be caused by a novel beta coronavirus, called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is closely related to Severe acute respiratory syndrome coronavirus (SARS-CoV) (Wang et al., 2020; Zhu et al., 2020). Since the start of the pandemic, there has been over 15,000,000 confirmed cases and more than 600,000 deaths attributed to SARS-CoV-2 (WHO, 2020). In the light of the ongoing global pandemic, it is of great importance to understand the biology of this newly emerged pathogen and identify cellular factors that influence the spread and replication of SARS-CoV-2, which would lead to development of better prophylactic and therapeutic interventions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SARS-CoV-2 is an enveloped, single-stranded, positive strand RNA virus that causes fever and severe respiratory illness in humans. Coronaviruses (CoVs) are composed of four major structural proteins: the nucleoprotein (N), the envelope (E), the transmembrane matrix glycoprotein (M) and the spike glycoprotein (S). The typical, crown-like appearance of the CoVs is attributed to homotrimers of spike (S) protein (Neuman et al., 2006), which mediate binding to the host cell receptors and subsequent membrane fusion. The proteolytic cleavage of S is essential for membrane fusion and subsequent virus entry (Shulla and Gallagher, 2009). The CoV S protein can be subdivided into two functional subunits: S1 and S2 (Tortorici and Veesler, 2019). In case of SARS-CoV-2, the S1 subunit directly binds to angiotensin-converting enzyme 2 (ACE2), which has been identified as the cellular receptor for SARS-CoV and SARS-CoV-2 (Hoffmann et al., 2020b; Li et al., 2005). Upon ACE2 binding, S needs to be primed by cellular proteases that leads to cleavage at the S1/S2 junction and the S2' site (Hoffmann et al., 2020b; Millet and Whittaker, 2015). This exposes the fusion peptide in the S2 subunit, allowing the fusion of viral and cellular membranes.

CoVs can enter the cell by fusing the virus envelope with either the endosomal or the cytoplasmic membrane, depending on the localization and availability of the cellular proteases in the host cell. Hence, cellular proteases have a direct impact on the cellular tropism and pathogenesis of CoVs. Different CoVs have evolved multiple ways to achieve proteolytic priming. A wide diversity of cellular proteases, including trypsin, endosomal cathepsins, transmembrane serine proteases (e.g., TMPRSS2) and furin are known to be involved in the priming of CoV spike (Millet and Whittaker, 2015). The SARS-CoV-2 S protein harbors a mutibasic cleavage site at S1/S2, as opposed to the monobasic site that is present in the SARS-CoV S (Hoffmann et al., 2020b). The multibasic site at S1/S2 also constitutes a putative furin cleavage site (RRAR). The consensus sequence for furin cleavage site is widely described as R-X-K/R-R↓ (Klimstra et al., 1999; Nakayama, 1997). However, the presence of consensus pattern is not always associated with cleavage by furin; for example, a mutated form of Sindbis Virus PE2 protein contains the pattern RXRR (RSKRLV), but is not efficiently cleaved by furin (Klimstra et al., 1999). On the other hand, the human C-type natriuretic peptide

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

precursor RLLR↓DL can be efficiently cleaved by furin (Wu et al., 2003). Thus, R-X-X-R is also regarded as minimal recognition motif, although R-X-R/K-R is highly favorable for efficient cleavage. Moreover, basic arginine at the P1 and P4 positions are strictly required due to furin's binding pocket containing complementary charged residues (Henrich et al., 2003). The S protein of numerous betacoronaviruses can be activated by furin at S1/S2 site; here, furin typically recognizes and cleaves at RRXRR motif (de Haan et al., 2008; Millet and Whittaker, 2015). SARS-CoV and murine hepatitis virus strain 2 (MHV-2) are two prominent exceptions of the betacoronavirus family that lack a furin cleavage site and need to be cleaved by other cellular proteases (Millet and Whittaker, 2015).

Interestingly, the furin recognition sequence present in SARS-CoV-2 S is identical (however, in opposite orientation) to the putative heparan sulfate (HS) interaction consensus sequences (XBBXBX, where B is a basic amino acid) (Cardin and Weintraub, 1989). It has been shown that cell culture adapted Sindbis virus can mediate attachment to cell surface HS via furin cleavage site (Klimstra et al., 1999). Similarly, upon *in vitro* culturing, CoVs often exhibit a tradeoff between the furin cleavability of S protein and HS binding (de Haan et al., 2008; de Haan et al., 2005). Thus, some CoVs can adapt to use HS as entry receptor in cultured cells (de Haan et al., 2005). Recently, SARS-CoV-2 mutants with deletions at the S1/S2 junction have been described to emerge upon culturing the virus in Vero E6 cells (Lau et al., 2020). Rapid adaptation of RNA viruses is a well-known phenomenon that can be partially explained by viral quasispecies traveling in a viral swarm. SARS-CoV quasispecies have been reported in previous reports (Xu et al., 2004). Clinical SARS-CoV-2 isolates, identified during the course of the current pandemic, have shown remarkably few mutations in the consensus sequence (van Dorp et al., 2020). However, less is known about the presence of RNA quasispecies in SARS-CoV-2 isolates and their dynamics of adaptation to a changing environment. A rapid adaptation of SARS-CoV-2 would complicate the virus control by antiviral compounds or biologicals, targeting a defined molecular site in the virus proteome, as it may facilitate the emergence of escape and resistant mutants.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In order to understand the dynamics of SARS-CoV-2 adaptation, we serially passaged virus strains in defined cell types and controlled environment. We show that SARS-CoV-2 rapidly adapts to culture conditions, by natural selection of pre-existing subdominant quasispecies. Furthermore, we show that the loss of the furin cleavage motif in Vero cells is due to a selection for virus variants that lose the overlapping HS binding motif. However, viruses with the intact motif are retained as subdominant population in the swarm. Therefore, re-passaging of such mutated swarms on cells expressing TMPRSS2 results in a prompt reversion to the original genotype observed in low-passage clinical isolates.

RESULTS

SARS-CoV-2 rapidly mutates upon culture on Vero E6 cells

Sequence analysis of SARS-CoV-2 Zagreb (Zg) strain passaged on Vero E6 cells revealed point mutation at virus genome position 23,607 (Figure 1A-B), which results in the loss of the furin cleavage site PRRARS at S1/S2 junction region (Table 1). These observations are in line with previous publications (Davidson et al., 2020; Lau et al., 2020). Moreover, a 15 base pair in-frame deletion was found at genomic position 23,582 in spike (S) protein (Figure 1B; Table S1). To understand how fast such mutations occur, we sequenced genomes of different SARS-CoV-2 isolates that were serially passaged on Vero E6 cells. Overall, the SARS-CoV-2 genomes were stable after multiple passages. However, we observed a very rapid onset of mutations in the furin cleavage site upon passaging in Vero E6 cells (Figure 1C-D). The Ischgl (NK) strain developed multiple point mutations at the furin cleavage site, resulting in loss of arginine in S protein position 682 and 685 (Figure 1C, 1E). We used the local Braunschweig (Br) isolate to compare the dynamics of mutation, as we had access to earlier passages of the Br strain. Here, a dramatic change in the composition of viral population was observed, where >85% of passage 3 (P3) genomes showed R682W mutation in S protein (Table 1; Figure S1A). We used the South Tyrol (FI) strain to compare the nature of mutation at the furin cleavage site and observed a single point mutation (R682W) in the RRAR site (Figure S1C), which was similar to Zg and Br strain. Rapid mutation in the furin site was accompanied with retention of

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a small fraction of genomes with wildtype furin site in all strains (Table 1). For instance, NK strain showed ~8.35% genomes with intact wildtype furin site after 12 passages in Vero E6 (Figure S2).

A mathematical model was used to fit a non-linear curve to observed mutation frequency data and calculate mutation rate at the furin cleavage site. The NK strain showed presence of multiple sub-species including three major variants with mutation at position 23,606, 23,607 and 23,616 (Figure S1D). Thus, we summed the mutation frequencies of these three positions to model the mutation rate at furin cleavage site for NK strain. The mutation rate for the NK strain was calculated as 0.3 (change in mutation frequency/day). The Br strain mutated at a much higher mutation rate of 0.81 per day and showed a reverse mutation rate of 0.03.

Overall, our data demonstrated that the furin cleavage site was rapidly mutated upon SARS-CoV-2 culture in Vero E6 cells, but it was not completely lost. Virus genomes with wildtype furin site were retained as subdominant variant in the population.

High passage SARS-CoV-2 strains show improved growth on Vero E6 but impaired growth on Calu-3 cells

In order to understand if the observed mutations in S protein influence viral growth, we compared the cell-to-cell spread on Vero E6 cells by assessing plaque sizes. At 3 days post infection (dpi), a significant increase in plaque sizes of high-passage virus in three independently passaged virus strains was observed (Figure 2A-B). To validate the results independently, we performed virus growth kinetics of two different virus isolates in Vero E6 cells and noticed that in both cases the high-passage virus showed significantly higher titers at 24h post infection (Figure 2C). We considered the possibility that the high-passage virus could be more efficiently released from infected cells, and compared virus titers in supernatants of infected cells (Figure 2D) and their lysates (Figure 2E). In both conditions, the high-passage viruses showed higher titers early upon infection. Therefore, we concluded that the growth advantage of high-passage virus was a result of an early event that occurred before virus release.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To ascertain the growth properties of Vero-passaged virus on cells expressing TMPRSS2, we tested the growth of virus strains on Calu-3 cells, and observed a mirror image of growth properties observed in Vero E6 cells: low-passage viruses grew substantially better at early time points post infection and the effects were observable equally in cell lysates and in the supernatants (Figure 2F-H). Our results agree with the previous report that a multibasic cleavage site in the S protein is necessary for priming by TMPRSS2 (Hoffmann et al., 2020a).

SARS-CoV-2 rapidly restores the furin cleavage site upon passage on TMPRSS2 competent cells

The low-passage virus growth data in Calu-3 cells suggest that the furin cleavage site confers growth advantage in these cells and may result in selection of SARS-CoV-2 genomes with wildtype furin site. We, therefore, took the NK strain that had been passaged 6 times in Vero E6 cells (NK6), and re-passaged it on Calu-3 cells. SARS-CoV-2 genomes with wildtype furin site were present as a subdominant fraction (~15%) in the NK6 viral population. Upon four passages in Calu-3 cells, we observed a substantial reversion to the wildtype sequence with the intact furin site (>97%) (Figure 3A). A similar reversion was observed upon passaging the virus on Caco-2 cells: the sequence changed to a functional furin cleavage site in the vast majority of viral genomes (>95%). While high-passage viruses showed growth disadvantage in Calu-3 cells early upon infection, the virus titers rapidly recovered to low-passage levels (Figure 2G-H). The data suggest that high-passage viruses rapidly mutated back to wildtype in Calu-3 cells. Thus, to understand the kinetics of the reversion, the virus was sequenced after each sequential passages on Calu-3 cells and the reverse mutation rate was calculated. Interestingly, the majority of the reversion occurred already by passage 1 (Figure 3B), but a fraction of genomes with the mutated cleavage site was retained as subdominant population in the virus swarm. The composition of NK6 virus quasispecies was reshaped with a reverse mutation rate of 0.63 (change in mutation frequency/day) and mutation rate of 0.01 (Figure 3C). The composition of virus genomes in the swarm of low-passage viruses was unaltered by their passaging on Calu-3 or Caco-2 cells (Figure 3D; Figure S3A). While furin-site mutant genomes were present at a low level, their frequency did not change. Remarkably, all viruses

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

passaged on Calu-3 cells showed a highly similar and stable composition of quasispecies, regardless of the underlying mutations induced by prior passaging on Vero cells. On the other hand, hitherto unreported mutations in the envelope (E) protein occurred upon virus passaging in Calu-3 (Figure 3E) and in Caco-2 cells (Figure S3B-C; Table S2). SARS-CoV-2 E protein showed L27S and T30I mutations in transmembrane domain (Table S2). Mutations in E protein were not detected in viruses passaged on Vero E6 cells. We tested FI2-P4 mutant to observe if E mutation modulate the growth of the virus in different cell lines. The E-T30I mutation conferred growth advantage in Calu-3 cells, but slow growth in Vero E6 cells (Figure 3E). We concluded that the furin site mutations in both directions were a result of targeted evolution, rather than drift.

Role of the SARS-CoV-2 furin cleavage site in cell entry

Cellular proteases were found to be critical for SARS-CoV-2 cell entry and their availability affects the virus tropism (Hoffmann et al., 2020b). Therefore, to test the role of different proteases, we infected Vero E6, Calu-3 and Caco-2 cells with low and high-passage viruses in presence of a TMPRSS2 inhibitor (camostat), a furin inhibitor (FI), or a broad-spectrum inhibitor of endosomal proteases (E-64d). The growth of all virus strains was inhibited by E-64d on Vero E6 cells, but not by camostat or FI, alone or in combination (Figure 4A). The combined use of E-64d and other protease inhibitors did not depress virus titers over the values observed in cells treated with E-64d alone (Figure S4A). This suggests that endosomal proteases, but not the cell-surface ones, were crucial for SARS-CoV-2 replication in Vero E6 cells. Importantly, no difference in titer between the high and the low-passage virus was observed in the presence of E-64d, arguing that the effect of the mutation in the high-passaged virus might occur during virus entry.

An inverse phenotype was observed in Calu-3 cells, where the TMPRSS2 inhibitor reduced the growth of all tested viruses, whereas E-64d and FI showed a less pronounced effect (Figure 4B). The differences in virus titers, seen in untreated controls, were erased in the presence of camostat. Finally, virus growth in presence of all three inhibitors was entirely abrogated in

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Calu-3 cells, arguing for a redundant function of these proteases in Calu-3 cells (Figure S4A). Virus replication in Caco-2 cells was not affected by inhibition of one type of protease alone, but combinations of protease inhibitors resulted in reduction of virus titers (Figure 4C). Although at 48 hpi the growth of low and high-passage viruses was similar in untreated controls, the pattern of virus growth reduction in the presence of protease inhibitors was similar to 24 hpi (Figure S3B-D).

In sum, our results indicated that the inhibition of virus entry into cells by protease inhibitors abrogated the growth advantages conferred by differences at the furin recognition site. This was consistent with the function of the spike protein during binding and fusion to cells.

Heparan sulfate binding influences SARS-CoV-2 growth in Vero E6

It has been previously reported that CoVs may trade furin cleavage sites in cell culture for a heparan sulfate (HS) binding site, thus improving virus binding to the cell surface (de Haan et al., 2008; de Haan et al., 2005). However, in the case of SARS-CoV-2, the furin cleavage site PRRARS has a composition that overlaps with the HS binding motif XBBXBX. The observed mutations in HS binding motif (R682W and R685H) decreased the positive charge of the site and increased its hydrophobicity. To understand if the growth advantage was due to the loss of HS binding site in the S protein, we compared the growth of the low-passage virus to the high-passage mutants in presence of the HS antagonist surfen in Vero E6 cells. At 24 hpi, we observed a 10-fold median increase in virus titers of the low-passage virus in the presence of surfen (Figure 5B). This effect was not observed in the viruses with the mutation in the heparan sulfate-binding site. To validate this by an independent experimental setting, we measured the size of virus plaques in Vero E6 cells in the presence or absence of surfen, comparing the low-passage virus to the high-passage one. We noticed that surfen increased the size of the plaques, but that the effects were more pronounced for low-passage viruses (Figure 5A). We observed a significant increase in the plaque size of two independent low-passage strains in the presence of surfen, but no significant surfen-related effects on the high-passage viruses. Likewise, absorbing the virus on soluble or immobilized heparin resulted in a reproducible

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

decrease in virus infectiousness of the low-passage virus, but no effects on the high-passage one.

Therefore, our results strongly suggest that the loss of the HS binding site on the S protein is a novel principle of virus selection that determines the growth of SARS-CoV-2 and its adaptation to the target host cell.

DISCUSSION

RNA virus populations are composed of a cloud of different genome variants known as quasispecies (Andino and Domingo, 2015; Lauring and Andino, 2010) that arise due to erroneous proof-reading of RNA-dependent RNA polymerase (RdRp), and is essential for adaptive evolution and fitness of RNA viruses (Borderia et al., 2016). Our deep sequencing data showed that SARS-CoV-2 exist as a quasispecies that exhibited remarkable genome stability overall. This can be attributed to a complex RNA synthesizing machinery and presence of exoribonuclease (ExoN) activity that confers genetic stability to the relatively large RNA genome (Gorbalenya et al., 2006). It has been reported that the SARS-CoV ExoN is important for maintaining genome stability, and virus with mutated ExoN shows more than 10-fold increase in mutation frequency (Eckerle et al., 2010). SARS-CoV and SARS-CoV-2 ExoN share 99% similarity at sequence level (Yoshimoto, 2020), thus explaining the observed stability of the SARS-CoV-2 genome. However, this stability does not hinder the ability of the virus to adapt quickly. We showed that SARS-CoV-2 could rapidly mutate the furin cleavage site upon passage in Vero cells. However, this rapid loss cannot be explained by random genetic drift as the presence of the furin cleavage site offers growth advantage on cells that express enzymes to cleave this site, but the loss of it should not be a selection criterion in the cells that lack these enzymes.

Although loss of furin site led to virus growth advantage in Vero cells, genomes with wildtype motif were not completely lost. The virus population accumulated the mutations at the furin cleavage site in an asymptotic manner, stabilizing after 4-5 passages and maintaining a subdominant fraction of population with the intact motif. Mutation dynamics in Calu-3 cells

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

followed a similar, albeit reverse, pattern and viruses with mutated PRRARS motif were retained in the virus swarm. Moreover, we observed unique mutations arising in E proteins upon virus culture in Calu-3 and Caco-2 cells, which is likely due to emergence of selection mutants in these cell lines. This further highlights that no single cell line is optimal for generating SARS-CoV-2 stocks and cell culture acquired mutations in high passage stocks might complicate our understanding of the virus biology. These observations suggest that SARS-CoV-2 maintains a considerable diversity in the quasispecies, facilitating natural selection and a rapid virus adaptation to changing environment and conditions. The imperfect fitness at individual level adds to the ability of the population to quickly adapt, thus contributing to the overall fitness of the swarm. While virus quasispecies are known to occur in numerous virus species, we are not aware of reports of such rapid adaptation due to swarm-like behavior of SARS-CoV-2.

The growth advantage on Vero cells was surprising, because the furin cleavage site should have provided an advantage to viruses on cells that express furin or TMPRSS2 enzymes (Calu-3 or Caco-2), but its absence should not confer growth advantage in cells that lack these enzymes. This can be explained by the fact that the SARS-CoV-2 furin cleavage site PRRARS has a composition that overlaps with the HS binding motif XBBXBX, and that HS binding in the absence of cleavage enzymes might exert selection pressure. The polybasic motifs binding heparan sulfate on virus surface proteins, as well as their acquisition upon *in vitro* virus culture, are well described across numerous virus families (Cagno et al., 2019; de Haan et al., 2008; de Haan et al., 2005; Klimstra et al., 1999; Mandl et al., 2001). However, the true importance of these receptors *in vivo* is debated as many viruses acquire HS binding upon cell culture adaptation (Cagno et al., 2019). The SARS-CoV-2 S protein has been shown to bind HS (Kim et al., 2020) and heparin was shown to inhibit SARS-CoV-2 infection *in vitro* (Kwon et al., 2020). Furthermore, Kim et al. has identified PRRARS site in the SARS-CoV-2 S protein as a putative HS binding site using an unbiased ligand-docking model (Kim et al., 2020). Interestingly, we report that the putative HS binding motif of the SARS-CoV-2 S protein act as a determinant of negative growth selection in cell culture, while all available evidence has

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

uniformly shown these sites as a positive selection marker, improving virus attachment on the surface of cells and thus virus cell entry and replication (Cagno et al., 2019). The most likely molecular basis of virulence attenuation is an adaptation of basic amino acid motifs in virus attachment to negatively charged heparin sulfate (Rostand and Esko, 1997). We observed that the mutation of arginine at position 682 in S protein leads to loss of the putative HS binding site in Vero-passaged Zg, Br and FI. However, in addition to ~35% genomes with mutated arginine at position 682, NK-P7 strain had ~48% genomes with mutation at arginine at position 685 that changed to histidine. Although the consensus sequence of heparin binding is widely described as XBBXBX, arginine and lysine are most common basic residues in heparin binding site (Munoz and Linhardt, 2004). Furthermore, arginine binds heparin ~2.5 time more tightly than lysine (Fromm et al., 1995) and forms more stable hydrogen bonds as well as stronger electrostatic interactions (Pearson, 1963). While our data do not define the reason for the adverse effects of HS binding on the PRRARS motif, we show that this is not due to effects occurring at the virus release from cells. Therefore, it is tempting to speculate that this is linked with virus entry and the unique circumstances of SARS-CoV-2 entry into Vero cells. One possibility is that the HS binding retains the virus on the cell surface and impairs its transition to the endosomal compartment. Another one would be a competitive effect of the HS binding with endosomal proteases recognizing the monobasic residue at the S1/S2 junction (and thus a steric inhibition of their activity). Since all of these effects are quantitative and subtle, their exact elucidation will require additional targeted research that is beyond the scope of the current study.

It is reasonable to assume that the PRRARS availability for proteolytic cleavage by cell-membrane associated proteases is a dominant function over HS binding for SARS-CoV-2, because the virus promptly regained the overlapping motifs upon passaging on cells expressing TMPRSS2 and/or furin. Therefore, it is not surprising that all known clinical isolates contain this motif. Furthermore, initial reports show that viruses lacking this motif are attenuated *in vivo* in the hamster model (Lau et al., 2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

One prediction from our research would be that antiviral substance administered as monotherapies may induce a rapid selection of subdominant species that may be resistant to the specific mode of action of the compound. Also, oligoclonal immune responses, focused on few epitopes, may result in the emergence of escape virus mutants. This aspect needs to be considered in future clinical studies, especially in the light of the oligoclonal immune responses in old age (Cicin-Sain et al., 2010), and the severe COVID19 presentation in this age group.

In conclusion, we demonstrate here a high potential of SARS-CoV-2 rapid adaptation, due to its swarm-like replication in tissue culture conditions. Future research in animal models shall explore the potential for the *in vivo* adaptation of SARS-CoV-2, the quasispecies diversity and the bottlenecks imposed on the virus spread. Our research indicates that the presence of subdominant genetic variants within SARS-CoV-2 isolates needs to be considered as a potential determinant of their virulence.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MATERIALS AND METHODS

Cell cultures and Viruses

Vero E6 (ATCC CRL-1586) cells were maintained in DMEM medium supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 100 IU/mL penicillin and 100 µg/mL streptomycin. Calu-3 (ATCC HTB-55) and Caco-2 (ATCC HTB-37) were cultured in Eagle's Minimum Essential Medium (EMEM) supplemented with 10% FCS, 2 mM L-glutamine, 100 IU/mL penicillin, 100 µg/mL streptomycin and 1x non-essential amino acid solution (Gibco MEM Non-Essential Amino Acids Solution 100X). All incubations of cells and virus were at 37 °C in a 5% CO_2 atmosphere.

The SARS-CoV-2 strains used in the study are Braunschweig isolate (hCoV-19/Germany/Br-ZK-1/2020, GISAID database ID: EPI_ISL_491115), South Tyrol isolate (hCoV-19/Germany/Muenster_FI1103201/2020, GISAID database ID: EPI_ISL_463008), Ischgl isolate (hCoV-19/Germany/NK1103201/2020) and Zagreb isolate (hCoV-19/Croatia/ZG-297-20/2020, GISAID database ID: EPI_ISL_451934). All work with infectious viruses was performed in biosafety level 3 facility by personnel wearing positive-pressure air-purifying respirators.

SARS-CoV-2 passage and virus stock generation

Braunschweig isolate (Br) was derived from an oropharyngeal swab using Vero E6 cells. Zagreb isolate (Zg) was isolated in Zagreb and received as passage 2 after propagation in Vero E6 cells. Ischgl (NK) and South Tyrol (FI) strains were isolated by Stephan Ludwig lab in Muenster. Passage 2 of NK, FI and Zg strains were further propagated by passaging twice in Vero E6 cells at low MOI (multiplicity of infection) to obtain working virus stocks. For serial passaging of SARS-CoV-2 strains, Vero E6, Calu-3 and Caco-2 cells were seeded in T25 or T75 flask one day before infection. Confluent monolayers of the cells were infected with virus and virus supernatant was collected three days post infection (dpi) to further passage the virus.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Virus stocks were generated in a two-step protocol, where seeding stock was generated by infecting one T75 flask of Vero E6 cells. The seeding stock virus supernatant was collected 3 dpi and used to infect 10 T75 flasks for virus stock generation. The virus stock supernatant was collected from all flasks at 3 dpi and spun at 3000 g for 10 minutes (min) to remove cell debris. Then, the virus supernatant was concentrated using Vivaspin 20 concentrators (Sartorius Stedim Biotech) by spinning at 6000 g for 30 min. The resulting virus stock was aliquoted and stored at -80°C until further use.

SARS-CoV-2 titration

SARS-CoV-2 was titrated with virus plaque assay on Vero E6 cells. Virus stocks were serially diluted in virus titration media (VTM, DMEM supplemented with 5% FCS, 2 mM L-glutamine, 100 IU/mL penicillin and 100 µg/mL streptomycin) and titrated on 24-well plates. The virus inoculum was added on the Vero E6 cells and incubated at 37°C. After 1 h, the inoculum was removed and the cell were overlaid with VTM supplemented with 1.5% carboxymethylcellulose (medium viscosity, C9481, Sigma-Aldrich) and incubated at 37°C. The overlay was removed from the cell at 3 dpi and the plates were fixed by submerging them in a tank of 6% formaldehyde (methanol stabilized) for at least 1 h. The cell monolayer was stained with crystal violet and the plaques were quantified by visual inspection with microscope. Virus supernatants were titrated in a similar fashion on 96-well plates.

Growth kinetics and plaque area measurement

Vero E6, Calu-3 or Caco-2 were seeded in a well of a 96-well plate. Confluent monolayers were infected with at an MOI of 0.001. After 1 h of infection at 37°C, the inoculum was removed and replaced by normal media. Supernatants and infected cell lysates were collected at different time points post infection and virus titer were determined as described above.

To analyze cell-to-cell spread of SARS-CoV-2, the virus plaque area was determined. Vero E6 seeded in a well of a 24-well plate were infected with 30 PFU of SARS CoV 2. After 1 h of infection, the inoculum was removed and cells were overlaid with VTM supplemented with 1.5% carboxymethylcellulose (medium viscosity, C9481, Sigma-Aldrich) and incubated at

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

37°C. For testing the effect of surfen, cells were pretreated with 10 µM surfen for one hour at 37°C before infection, and infection was performed in the presence of surfen. Finally, after 1 h, the virus inoculum was removed and monolayer was overlaid with VTM supplemented with 1.5% carboxymethylcellulose and 10 µM surfen. Cells were fixed with 6% formaldehyde 3 dpi and stained with crystal violet. The plaque area was quantified using a Zeiss LSM 980 Microscope. The plaque area was measured with help of ZEISS ZEN lite 3.0 (Blue edition), and plotted using GraphPad Prism.

SARS-CoV-2 infection and protease inhibition

SARS-CoV-2 cell entry inhibition was tested by seeding Vero E6, Calu-3 or Caco-2 cells in 96-well plates, one day before infection. On the day on infection, cell were treated with protease inhibitors (E-64d, 10µM; camostat, 50 10µM; furin inhibitor, 10µM) for 1 h at 37°C before virus infection. The infection was performed in the presence of protease inhibitors and after 1 h infection at 37°C, the virus inoculum was removed and normal media supplemented with protease inhibitors was given to the cells. The virus supernatant was collected at 1 and 2 dpi and titrated on Vero E6 cells as described in SARS-CoV-2 titration section.

For analyzing the heparin binding, SARS-CoV-2 suspension containing ~1000 PFU was incubated with heparin sodium salt (100 µg/mL) for 1 h at 37°C. After incubation, the suspension was titrated on Vero E6 cells in a 24-well plate format. Similarly, for testing the ability of immobilized heparin to bind SARS-CoV-2, Heparin−biotin sodium salt was immobilized on Pierce™ Streptavidin Coated Plates (Thermo Fisher Scientific). The volume used to immobilize Heparin-biotin (100 µg/mL) was 100 µL per well. Plates were incubated at 37°C for 15 minutes and then washed three times with 1x PBS. The virus suspension (1000 PFU) was added to the heparin-coated wells along with empty streptavidin coated wells and incubated for 1 h at 37°C. Afterwards, the suspension was titrated on Vero E6 cells in a 24-well plate format.

SARS-CoV-2 RNA isolation and sequencing

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For the inactivation of SARS-CoV-2, the virus suspension was mixed 1:1 with peqGOLD TriFast (VWR). Following, RNA was extracted using the innuPREP Virus RNA Kit (Analytik Jena) according to the manufacturer's instructions. The RNA quality was confirmed with Bioanalyzer (Agilent Technologies) and Qubit (Thermo Fisher Scientific). The sequencing libraries were prepared using NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs) including ERCC RNA Spike-in control. The sequencing was performed with NovaSeq 6000 SP Reagent Kit (100 cycles) on NovaSeq 6000 System (Illumina).

Raw fastq files were processed to trim adaptor and low quality reads (fastp 0.20.0). The resulting sequence reads were aligned with reference (BWA 0.7.8). The alignment file was sorted (SAMtools version 1.10) and realigned (GATK IndelRealigner version 3.7). Finally, pileup file (SAMtools) was created and mutation were detected with VarScan (version 2.3.9). Mutation frequency for whole genome was calculated by extracting total read depth and reference base reads for each nucleotide from pileup file using SAMtools, where mutation frequency in the population is given as ratio of reference reads to total reads at the position. Mutation frequencies were converted to percent values and plotted with ggplot2 (R/Bioconductor 4.0). Alignment files were inspected with IGV (version 2.8.3) and Tablet (version 1.19.09.03) to access if nearby mutations in the furin cleavage site are present on same reads.

Mathematical modelling

The mathematical model, to fit the curves for the mutation, assumes that the proportion of viruses with a mutation on the considered chromosome (SARS-CoV-2 genome positions 23606, 23607 and 23616) at time t depends on the initial proportion of mutated viruses $y(0)$, the mutation rate a and the reverse mutation rate b. This leads to the ordinary differential equation $y' = a(1-y) - by$ with the general solution $y(t) = ce^{-(a+b)t} + \frac{a}{a+b}$. Assuming that there are no mutations at time zero, i.e. $y(0) = 0$, then $c = -\frac{a}{a+b}$ follows. If we start with a mutated strain, i.e. $y(0) = 1$, this implies $c = \frac{b}{a+b}$. The mutation rates a and b of the

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

corresponding non-linear curve are estimated based on the data by the optimization procedure optim provided by the R statistics software, minimizing the sum of absolute errors. R script used is provided in supplemental data.

QUANTIFICATION AND STATISTICAL ANALYSIS

Kruskal-Wallis with Dunn's comparison and two-way analysis of variance (ANOVA) with Dunnett posttest were used to test for statistical significance. Where appropriate, two-way ANOVA with Sidak's multiple comparison was used. P-values < 0.05 were considered significant (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$), ****$p < 0.0001$, $p > 0.05$ not significant (n.s.). For all statistical analyses, the GraphPad Prism 7 software package was used (GraphPad Software).

DATA AND CODE AVAILABILITY

The SARS-CoV-2 deep sequencing data generated during this study are available at NCBI sequence read archive. (SRA accession: PRJNA650134; https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA650134). The R code used in this study are provided in supplemental data (Data S1).

MATERIALS AVAILABILITY

Further information and requests for resources and reagents can be directed to M. Zeeshan Chaudhry (Zeeshan.chaudhry@helmholtz-hzi.de) and Luka Cicin-Sain (luka.cicin-sain@helmholtz-hzi.de). The materials will be made available upon receipt of a material transfer agreement (MTA).

Acknowledgements

We thank HZI Genome Analytics team for support, especially Michael Jarek. We kindly acknowledge Susanne Talay and Markus Hoffmann for helpful discussion. We thank Ayse Barut and Inge Hollatz-Rangosch for technical assistance.

Author Contributions

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conceptualization, M.Z.C.; Methodology, M.Z.C., K.E. and F.K.; Software and Formal Analysis, M.Z.C. and F.K.; Investigation, M.Z.C., K.E., M.G., L.A. and Y.K.; Data Curation, M.Z.C.; Writing – Original Draft, M.Z.C., K.E. and L.C.-S.; Writing – Review & Editing, M.Z.C., K.E., A.K., F.K. and L.C.-S.; Supervision, A.K. and L.C.-S.; Resources, L.B., S.L., Z.M.S., I.-C.K., and A.M.; Funding Acquisition, L.C.-S.

Declaration of Interests

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Andino, R., and Domingo, E. (2015). Viral quasispecies. Virology *479-480*, 46-51.

Borderia, A.V., Rozen-Gagnon, K., and Vignuzzi, M. (2016). Fidelity Variants and RNA Quasispecies. Curr Top Microbiol Immunol *392*, 303-322.

Cagno, V., Tseligka, E.D., Jones, S.T., and Tapparel, C. (2019). Heparan Sulfate Proteoglycans and Viral Attachment: True Receptors or Adaptation Bias? Viruses *11*.

Cardin, A.D., and Weintraub, H.J. (1989). Molecular modeling of protein-glycosaminoglycan interactions. Arteriosclerosis *9*, 21-32.

Cicin-Sain, L., Smyk-Pearson, S., Currier, N., Byrd, L., Koudelka, C., Robinson, T., Swarbrick, G., Tackitt, S., Legasse, A., Fischer, M.*, et al.* (2010). Loss of naive T cells and repertoire constriction predict poor response to vaccination in old primates. J Immunol *184*, 6739-6745.

Davidson, A.D., Williamson, M.K., Lewis, S., Shoemark, D., Carroll, M.W., Heesom, K., Zambon, M., Ellis, J., Lewis, P.A., Hiscox, J.A.*, et al.* (2020). Characterisation of the transcriptome and proteome of SARS-CoV-2 using direct RNA sequencing and tandem mass spectrometry reveals evidence for a cell passage induced in-frame deletion in the spike glycoprotein that removes the furin-like cleavage site. 2020.2003.2022.002204.

de Haan, C.A., Haijema, B.J., Schellen, P., Wichgers Schreur, P., te Lintelo, E., Vennema, H., and Rottier, P.J. (2008). Cleavage of group 1 coronavirus spike proteins: how furin cleavage is traded off against heparan sulfate binding upon cell culture adaptation. J Virol *82*, 6078-6083.

de Haan, C.A., Li, Z., te Lintelo, E., Bosch, B.J., Haijema, B.J., and Rottier, P.J. (2005). Murine coronavirus with an extended host range uses heparan sulfate as an entry receptor. J Virol *79*, 14451-14456.

Eckerle, L.D., Becker, M.M., Halpin, R.A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R.L., Baric, R.S., Stockwell, T.B.*, et al.* (2010). Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog *6*, e1000896.

Fromm, J.R., Hileman, R.E., Caldwell, E.E., Weiler, J.M., and Linhardt, R.J. (1995). Differences in the interaction of heparin with arginine and lysine and the importance of these basic amino acids in the binding of heparin to acidic fibroblast growth factor. Arch Biochem Biophys *323*, 279-287.

Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., and Snijder, E.J. (2006). Nidovirales: evolving the largest RNA virus genome. Virus Res *117*, 17-37.

Henrich, S., Cameron, A., Bourenkov, G.P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W., and Than, M.E. (2003). The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. Nat Struct Biol *10*, 520-526.

Hoffmann, M., Kleine-Weber, H., and Pohlmann, S. (2020a). A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. Mol Cell *78*, 779-784 e775.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A.*, et al.* (2020b). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell *181*, 271-280 e278.

Kim, S.Y., Jin, W., Sood, A., Montgomery, D.W., Grant, O.C., Fuster, M.M., Fu, L., Dordick, J.S., Woods, R.J., Zhang, F.*, et al.* (2020). Characterization of heparin and severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) spike glycoprotein binding interactions. Antiviral Res, 104873.

Klimstra, W.B., Heidner, H.W., and Johnston, R.E. (1999). The furin protease cleavage recognition sequence of Sindbis virus PE2 can mediate virion attachment to cell surface heparan sulfate. J Virol *73*, 6299-6306.

Kwon, P.S., Oh, H., Kwon, S.J., Jin, W., Zhang, F., Fraser, K., Hong, J.J., Linhardt, R.J., and Dordick, J.S. (2020). Sulfated polysaccharides effectively inhibit SARS-CoV-2 in vitro. Cell Discov *6*, 50.

Lau, S.Y., Wang, P., Mok, B.W., Zhang, A.J., Chu, H., Lee, A.C., Deng, S., Chen, P., Chan, K.H., Song, W.*, et al.* (2020). Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. Emerg Microbes Infect *9*, 837-842.

Lauring, A.S., and Andino, R. (2010). Quasispecies theory and the behavior of RNA viruses. PLoS Pathog *6*, e1001005.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Li, F., Li, W., Farzan, M., and Harrison, S.C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science *309*, 1864-1868.
Mandl, C.W., Kroschewski, H., Allison, S.L., Kofler, R., Holzmann, H., Meixner, T., and Heinz, F.X. (2001). Adaptation of tick-borne encephalitis virus to BHK-21 cells results in the formation of multiple heparan sulfate binding sites in the envelope protein and attenuation in vivo. J Virol *75*, 5627-5637.
Millet, J.K., and Whittaker, G.R. (2015). Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. Virus Res *202*, 120-134.
Munoz, E.M., and Linhardt, R.J. (2004). Heparin-binding domains in vascular biology. Arterioscler Thromb Vasc Biol *24*, 1549-1557.
Nakayama, K. (1997). Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. Biochem J *327 (Pt 3)*, 625-635.
Neuman, B.W., Adair, B.D., Yoshioka, C., Quispe, J.D., Orca, G., Kuhn, P., Milligan, R.A., Yeager, M., and Buchmeier, M.J. (2006). Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. J Virol *80*, 7918-7928.
Pearson, R.G.J.J.o.t.A.C.s. (1963). Hard and soft acids and bases. *85*, 3533-3539.
Rostand, K.S., and Esko, J.D. (1997). Microbial adherence to and invasion through proteoglycans. Infect Immun *65*, 1-8.
Shulla, A., and Gallagher, T. (2009). Role of spike protein endodomains in regulating coronavirus entry. J Biol Chem *284*, 32725-32734.
Tortorici, M.A., and Veesler, D. (2019). Structural insights into coronavirus entry. Adv Virus Res *105*, 93-116.
van Dorp, L., Acman, M., Richard, D., Shaw, L.P., Ford, C.E., Ormond, L., Owen, C.J., Pang, J., Tan, C.C.S., Boshier, F.A.T.*, et al.* (2020). Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. Infect Genet Evol *83*, 104351.
Wang, C., Horby, P.W., Hayden, F.G., and Gao, G.F. (2020). A novel coronavirus outbreak of global health concern. Lancet *395*, 470-473.
WHO (2020). Coronavirus disease (COVID-19) Situation Report – 175.
Wu, C., Wu, F., Pan, J., Morser, J., and Wu, Q. (2003). Furin-mediated processing of Pro-C-type natriuretic peptide. J Biol Chem *278*, 25847-25852.
Xu, D., Zhang, Z., and Wang, F.S. (2004). SARS-associated coronavirus quasispecies in individual patients. N Engl J Med *350*, 1366-1367.
Yoshimoto, F.K. (2020). The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the Cause of COVID-19. Protein J *39*, 198-216.
Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R.*, et al.* (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med *382*, 727-733.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE LEGENDS

Figure 1. SARS-CoV-2 rapidly adapts upon passage in cultured cells.

(A) Mutations acquired by the Zagreb (Zg) strain passage 5 (P5) upon culturing in Vero E6 cells are shown by plotting mutation frequency across whole genome. Each symbol represents the mutation frequency (y-axis) of an individual nucleotide of SARS-CoV-2 at the indicated genomic position (x-axis). Increased mutation frequency is highlighted by symbol color shift from blue to red. Red arrow highlights the position of the furin cleavage site in the genome.

(B) Mutation frequency of the S protein from Zg strain P5, genomic range 23,575 to 23,610.

(C-D) Serial passages of Ischgl (NK) and Braunschweig (Br) strain were analyzed with deep sequencing to assess the composition of viral quasispecies upon passaging in Vero E6 cells. Each symbol represents the mutation frequency (y-axis) of individual nucleotide of SARS-CoV-2 at a given genomic position (x-axis). Red arrows highlight the position of the furin cleavage site in the genome.

(E) The mutation frequency of nucleotides at furin cleavage site in SARS-CoV-2 S are plotted on y-axis for NK and Br strains at different passage in Vero E6 cells. Three bars represent the nucleotide codon triplet for each corresponding amino acid at the furin cleavage site from genomic position 23,603 to 23,620 (x-axis).

(F) The sum of mutation frequency at the furin site position 23,606, 23,607 and 23,616 of NK strain serially passaged on Vero E6 cells is plotted against time. The mutation and reverse mutation rate corresponds to change in mutation frequency per day.

The genomic positions correspond to Wuhan-Hu-1 isolate (GenBank accession no: NC_045512).

Figure 2. Growth properties of passaged SARS-CoV-2 strains in different cell lines.

(A-B) Vero E6 cells were infected with 30 PFU of different passages of SARS-CoV-2 strains. The plaque size was measured 3 dpi. Panel A shows the area of virus plaque on Vero E6 cells,

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and panel B shows representative images of two different wells infected with NK-P4 and NK-P7.

(C-E) Virus growth kinetics on Vero E6 cells were established by infecting the cells at an MOI of 0.001. Supernatants (panel C-D) and cell lysates (panel E) were collected at indicated time points and titrated on Vero E6 cells.

(F-H) Calu-3 cells were infected at an MOI of 0.001. Supernatants (panel F-G) and cell lysates (panel H) were collected at indicated time post infection and titrated on Vero E6 cells.

Data are representative of two independent experiments. Mean and ±SEM are plotted. Each symbol in panel A represent one plaque and data are pooled from multiple infected wells of two independent experiments. Panel C-H symbols represent biological replicates. Statistical significance was calculated using Kruskal-Wallis test with Dunn's posttest.

Figure 3. SARS-CoV-2 quasispecies composition at furin cleavage site reverts to low passage virus levels upon passage in TMPRSS2 competent cells.

The mutation frequency of genomic position 23,603 to 23,620 (x-axis) in SARS-CoV-2 genome corresponding to amino acid codon triplets at furin cleavage site are plotted in Panel A, B and D.

(A) NK strain passaged 6 time in Vero E6 cells (NK6) was serially passaged independently on Calu-3 and Caco-2 cells. P0 vero represents the genome of input virus and P4 calu-3 and P4 caco-2 are the viruses cultured on respective cells lines for 4 passages.

(B) The change in SARS-CoV-2 furin site mutations is shown by plotting the nucleotide mutation frequency of NK6 virus upon serial passaging in Calu-3 cells.

(C) The sum of mutation frequency at furin site position 23,606, 23,607 and 23,616 of NK6 strain serially passaged on Calu-3 cells is plotted against time. The mutation and reverse mutation rate correspond to change in mutation frequency per day.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(D) NK strain passaged 4 time in Vero E6 cells (NK4) was serially passaged on Calu-3 and Caco-2 cells. P0 vero represents the genome of low passage input virus and P4 calu-3 and P4 caco-2 are the viruses cultured on Calu-3 and Caco-2 cell lines for 4 serial passages.

(E) Mutation frequency of NK4 strain serially passage on Calu-3 cells for 4 passages. Each symbol represents mutation frequency (y-axis) of individual nucleotide of SARS-CoV-2 at given genomic position (x-axis). Increased mutation frequency is highlighted by color shift from blue to red. Red arrow highlights the position of the envelope protein in the genome.

Figure 4. Inhibition of low and high-passage SARS-CoV-2 strain growth by different protease inhibitors.

Vero E6 (A), Calu-3 (B) and Caco-2 (C) cells were infected at an MOI of 0.01 in the presence or absence of different protease inhibitors. The supernatant from infected cells was collected at 24 hpi and titrated on Vero E6 cells. Data is representative of two independent experiments and error bars represent mean ± SEM of three biological replicates. Statistical significance was calculated using two-way ANOVA with Dunnett posttest, where untreated control cells served as reference.

Figure 5. Heparan sulfate antagonist enhances growth of low passage SARS-CoV-2 strains.

(A) Vero E6 cells were infected with low or high-passage NK or FI strains. The cells were overlaid with methylcellulose supplemented with or without 10 µM surfen. The virus plaque size was quantified 3 dpi. Each symbol represents one plaque and data is pooled from multiple infected wells of two independent experiments. Statistical significance was calculated using Kruskal-Wallis test with Dunn's posttest.

(B) Vero E6 cells were infected with different SARS-CoV-2 passaged strains at an MOI of 0.001 in the presence of absence of 10 µM Surfen. The supernatant was collected 24 hpi and titrated on Vero E6 cells.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(C) Heparin binding was tested by incubating NK-P4, NK-P7 and FI-P8 virus suspension with Heparin salt or Biotin-Heparin salt immobilized on Streptavidin coated plate. After one-hour incubation at 37°C, virus was titrated on Vero E6 cells.

Pane B and C data is representative of two independent experiments, and each symbol represents biological replicate. Mean and ±SEM are plotted.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLES

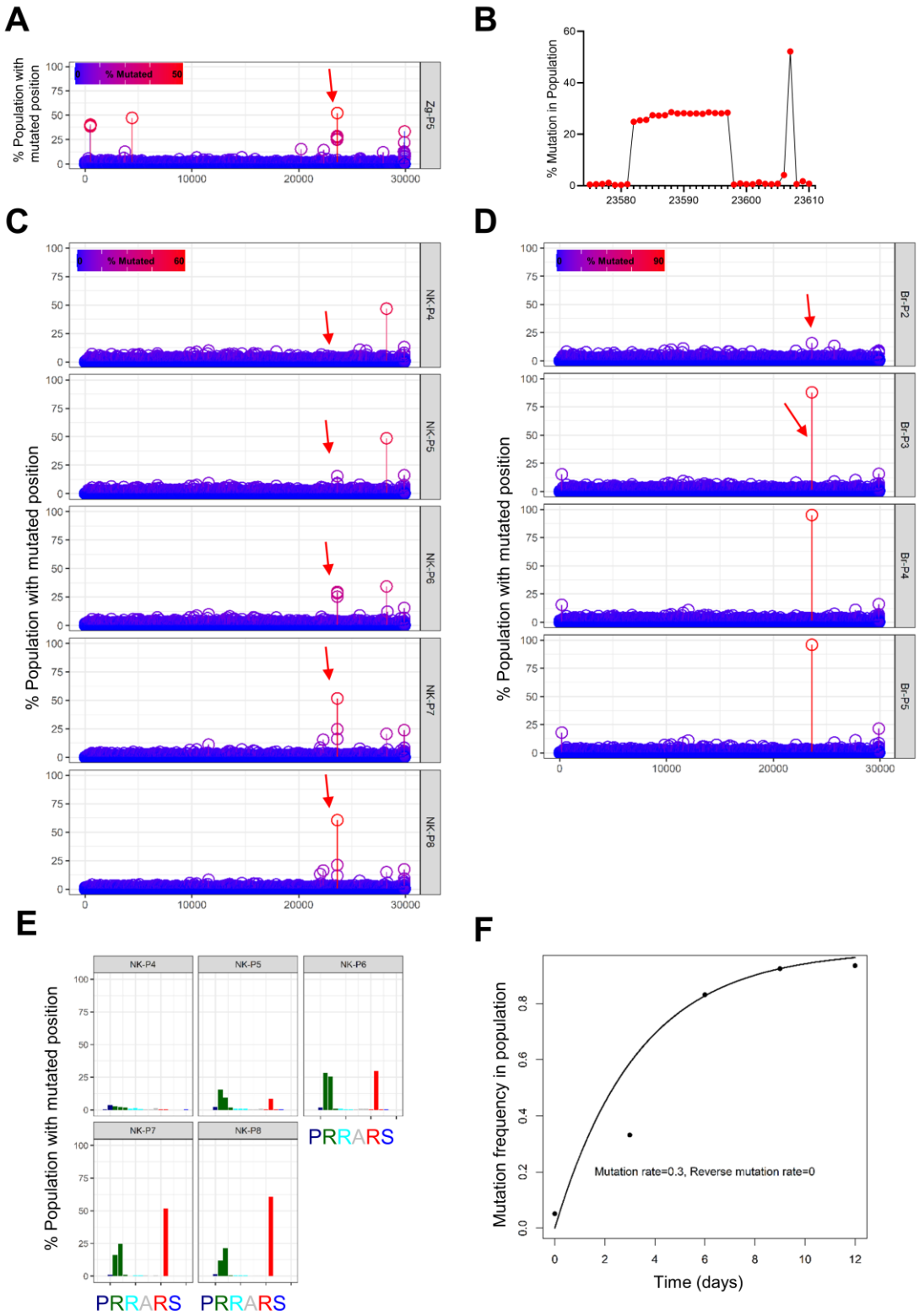
Table 1. Mutations (SNPs) in SARS-CoV-2 S protein upon passage in Vero E6 cells.

SARS-CoV-2 Strain	Passage #[a]	Mutation (position)[b]		Mutated Population
		Nucleotide[c]	Amino Acid[d]	
Zagreb-Zg	P4	C->A (22338)	T->K (259)	28.28%
		G->A (23607)	R->Q (682)	24.22%
Zagreb-Zg	P5	C->A (22338)	T->K (259)	13.79%
		G->A (23607)	R->Q (682)	52.10%
Braunschweig-Br	P2	C->T (23606)	R->W (682)	15.51%
Braunschweig-Br	P5	C->T (23606)	R->W (682)	95.62%
South Tyrol-FI	P4	C->T (23606)	R->W (682)	2.58%
South Tyrol -FI	P8	C->T (23606)	R->W (682)	89.00%
Ischgl-NK	P4	C->T (23606)	R->W (682)	2.06%
		G->A (23607)	R->Q (682)	1.80%
Ischgl-NK	P7	C->T (23606)	R->W (682)	15.62%
		G->A (23607)	R->Q (682)	19.79%
		G->A (23616)	R->H (685)	48.10%

[a] Number of passages in Vero E6 cells.
[b] Positions with Mutation frequency above 1% are shown.
[c] Numbers are according to Wuhan-Hu-1 sequence (GenBank accession no: NC_045512).
[d] Numbers start from the amino terminus of SARS-CoV-2 S protein.

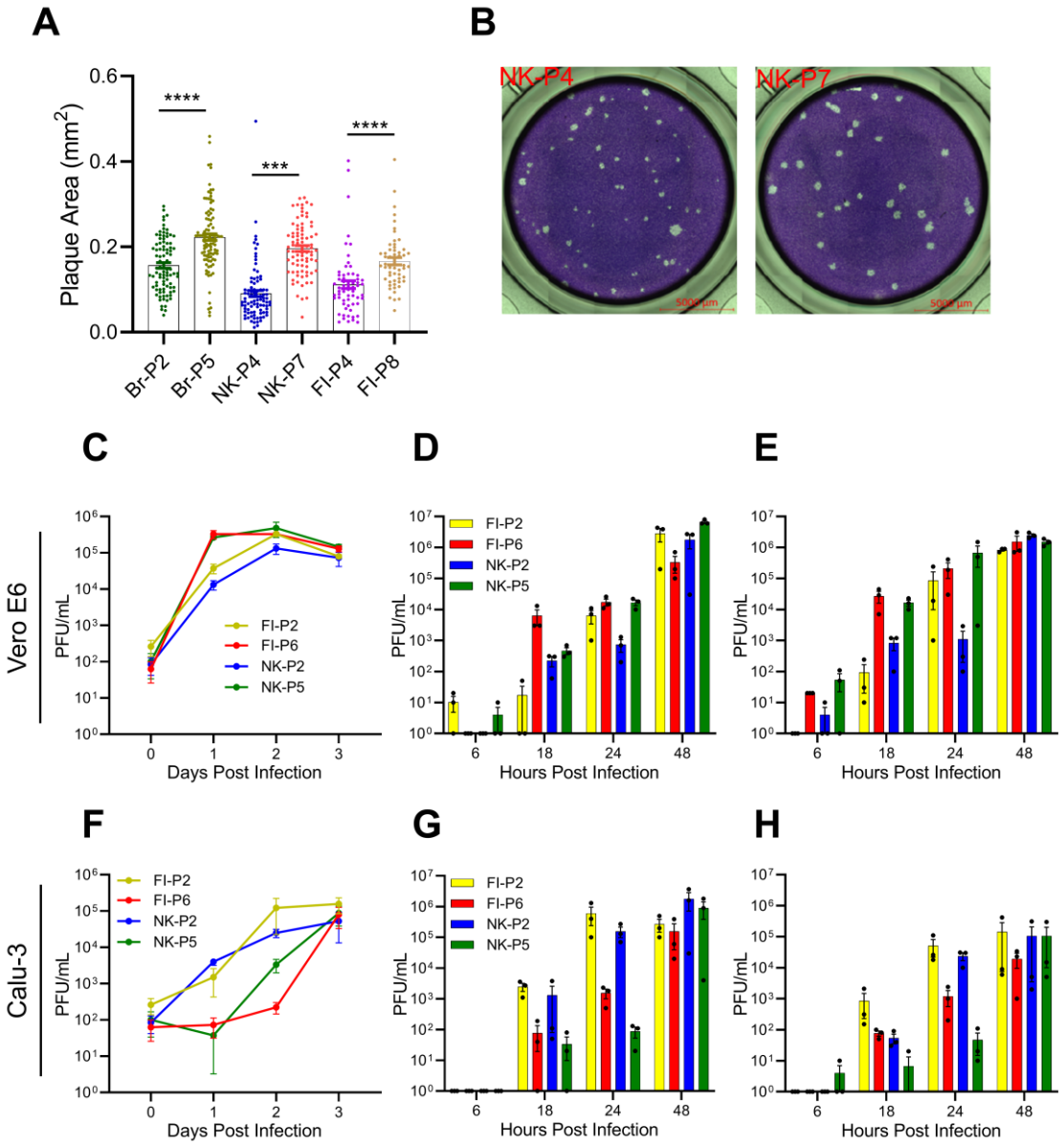
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1



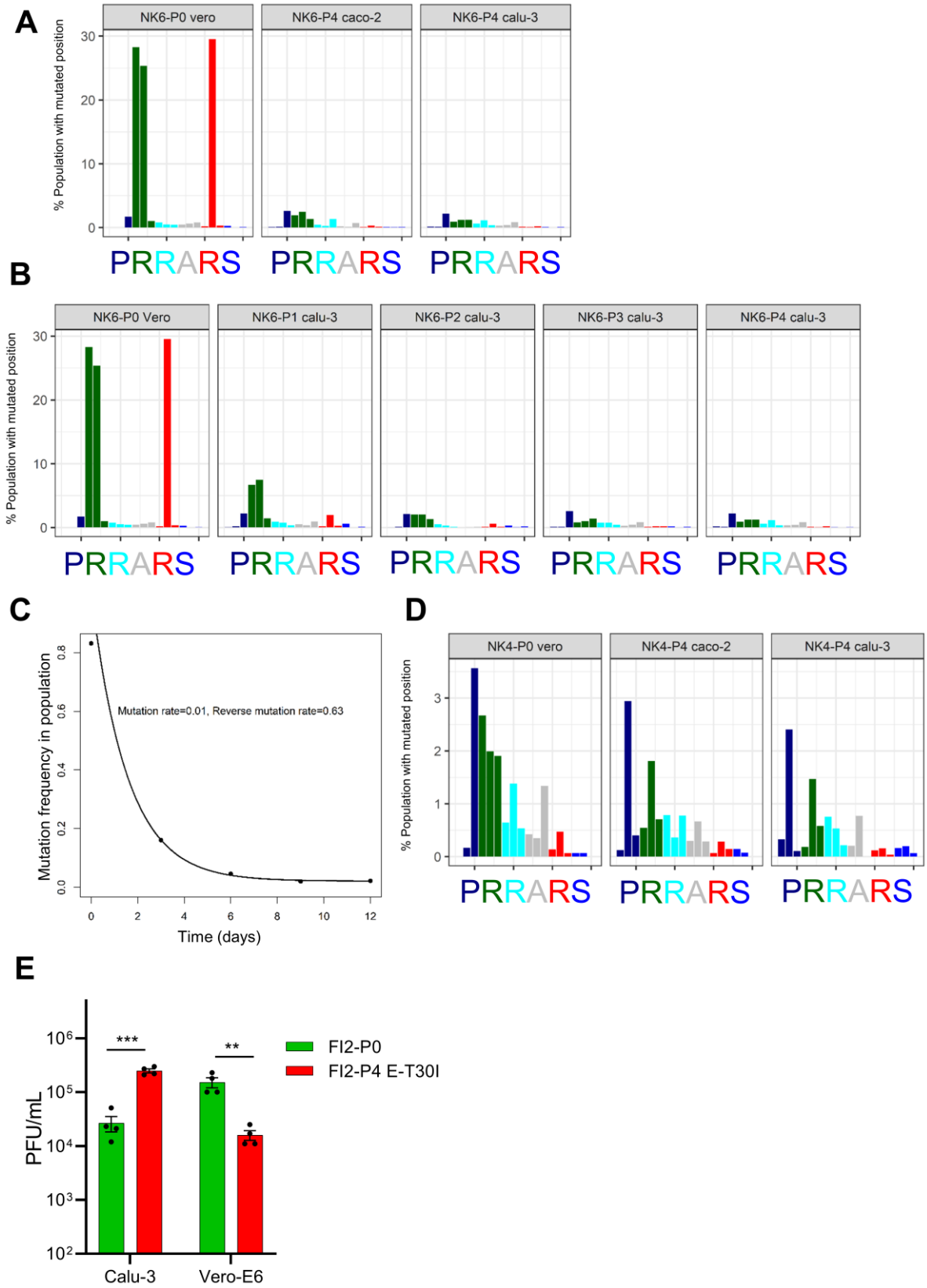
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2



bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3



bioRxiv preprint doi: https://doi.org/10.1101/2020.08.10.241414; this version posted August 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4

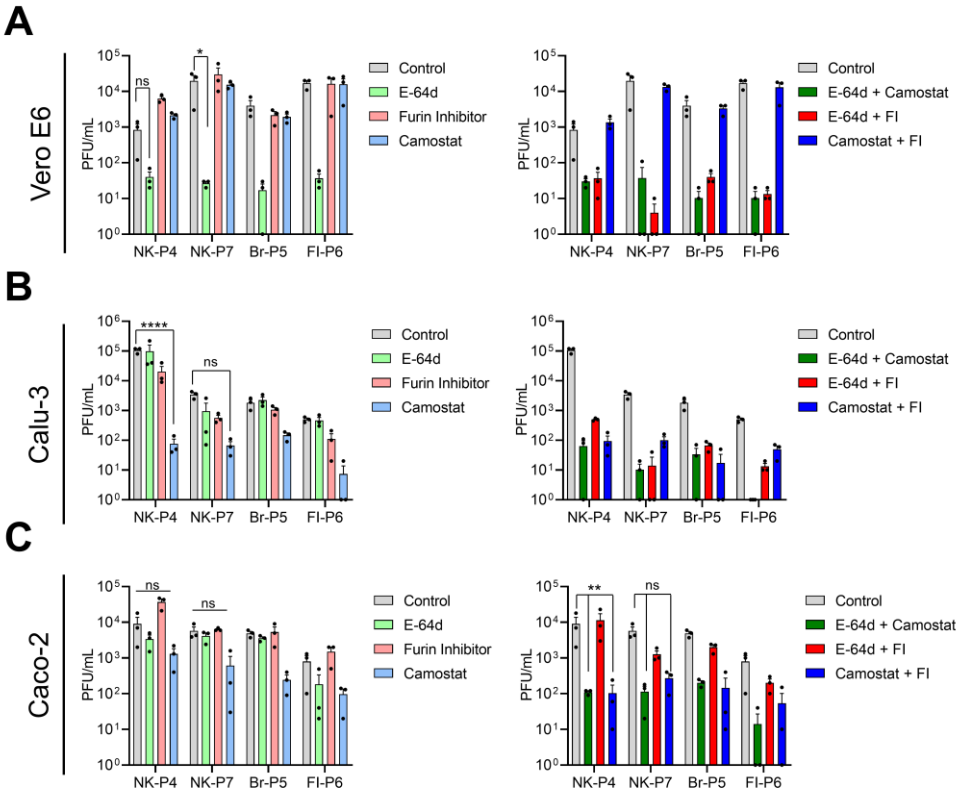
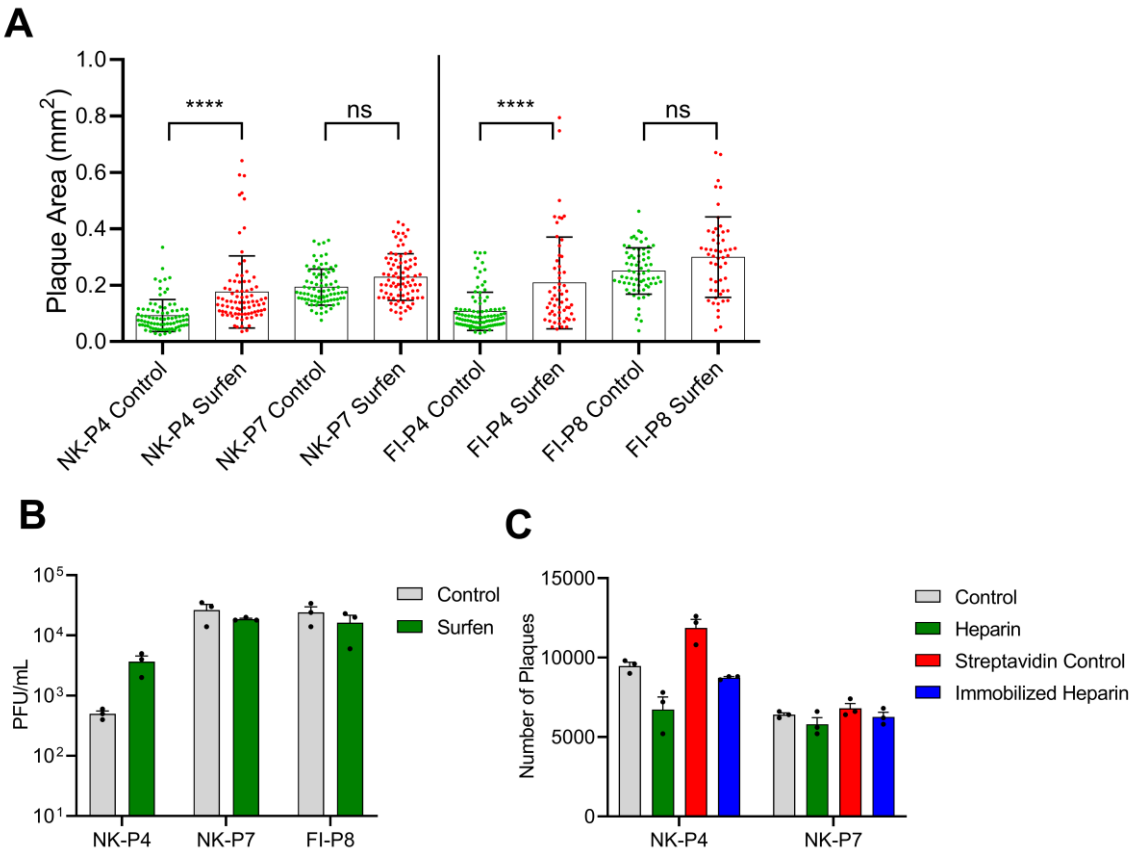



Figure 5