Nucleocapsid mutations in SARS-CoV-2 augment replication and pathogenesis.

Authors: Bryan A. Johnson[1], Yiyang Zhou[2], Kumari G. Lokugamage[1], Michelle N. Vu[1], Nathen Bopp[3], Patricia A. Crocquet-Valdes[3], Craig Schindewolf[1], Yang Liu[2], Dionna Scharton[1,5], Jessica A. Plante[1,5], Xuping Xie[2], Patricia Aguilar[3, 6], Scott C. Weaver[1,4,5, 6], Pei-Yong Shi[2,4,5, 6,7], David H. Walker[3,6], Andrew L. Routh[2,4], Kenneth S. Plante[1,5], Vineet D. Menachery[1,4, 6*]

Affiliations:
[1]Department of Microbiology and Immunology, University of Texas Medical Branch; Galveston, TX, USA
[2]Department of Biochemistry and Molecular Biology, University of Texas Medical Branch; Galveston, TX, USA
[3]Department of Pathology, University of Texas Medical Branch; Galveston, TX, USA
[4]Institute for Human Infection and Immunity, University of Texas Medical Branch; Galveston, TX, USA
[5]World Reference Center of Emerging Viruses and Arboviruses, University of Texas Medical Branch; Galveston, TX, USA
[6]Center for Biodefense and Emerging Infectious Diseases, University of Texas Medical Branch; Galveston, TX, USA
[7]Institute for Drug Discovery, University of Texas Medical Branch; Galveston, TX, USA

Corresponding Author: Vineet D. Menachery
Address: University of Texas Medical Branch, 301 University Blvd, Route #0610 Galveston, TX 77555
Email: Vimenach@utmb.edu

Keywords: Coronavirus, 2019-nCoV, SARS-CoV-2, COVID-19, Variant of concern, Nucleocapsid, SR Region

Abstract: While SARS-CoV-2 continues to adapt for human infection and transmission, genetic variation outside of the spike gene remains largely unexplored. This study investigates a highly variable region at residues 203-205 in SARS-CoV-2 nucleocapsid protein. Recreating the alpha variant mutation in an early pandemic (WA-1) background, we found that the R203K/G204R mutation is sufficient to enhance replication, fitness, and pathogenesis of SARS-CoV-2. Importantly, the R203K/G204R mutation increases nucleocapsid phosphorylation, providing a molecular basis for these phenotypes. Notably, an analogous alanine substitution mutant also increases SARS-CoV-2 fitness and phosphorylation, suggesting that infection is enhanced through ablation of the ancestral 'RG' motif. Overall, these results demonstrate that variant mutations outside spike are also key components in SARS-CoV-2's continued adaptation to human infection.

One-Sentence Summary: A mutation in the nucleocapsid gene of the SARS-CoV-2 alpha variant is found to enhance replication, fitness, and pathogenesis.

Main Text:

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS)-CoV-2 is the most significant infectious disease event of the 21st century (*1, 2*). Since its initial expansion, SARS-CoV-2 has continued to adapt for human infection and transmission, resulting in several variants of concern (*3*). While most mutations occur within a single lineage, a small number are shared across multiple variants (*4*). Spike mutations have dominated SARS-CoV-2 variant research, owing to concerns that they enhance replication, augment transmission, or allow escape from immunity (*4*). However, less attention has been focused on mutations outside spike, despite the existence of other "mutational hotspots" in the genome (*4*). The SARS-CoV-2 nucleocapsid (N) gen is one hotspot for coding mutations, particularly at amino acid residues 203-205 within its serine rich (SR) domain (*5*). Three prominent mutations occur in this region including R203K/G204R, a double substitution (KR mt) present in the alpha and gamma variants; T205I present in the beta variant; and R203M that occurs in the kappa and delta variants (*6, 7*). Together, this genetic variation and convergent evolution in the 203-205 amino acid region suggests positive selection in this motif of N.

Utilizing our reverse genetic system (*8, 9*), we generated the KR nucleocapsid mutation in the ancestral WA-1 strain of SARS-CoV-2. This change alone was sufficient to increase viral replication in respiratory tract infection models and exhibited enhanced fitness in direct competition studies with wild-type (WT) SARS-CoV-2. In the hamster model, the KR mutant (mt) enhanced pathogenesis and outcompeted the WT in direct competition. We subsequently found that the KR mt increased nucleocapsid phosphorylation relative to WT SARS-CoV-2; similar increases in N phosphorylation were also observed in the alpha and kappa variants. Finally, an analogous alanine double substitution mutant (AA mt) also increased fitness and augmented

phosphorylation relative to WT SARS-CoV-2. Together, these results suggest that disruption of the ancestral "RGT" motif in nucleocapsid augments infection, fitness, and pathogenesis of SARS-CoV-2.

Results:

Genetic Analysis of a highly variable motif of SARS-CoV-2 Nucleocapsid

Using SARS-CoV-2 genomic data from the GISIAD database (*6*), we binned each sequence by month of collection and performed an *in silico* search for variation at residues 203-205 within nucleocapsid (**Fig. 1a**). Three prominent mutations emerged from this analysis. The first is the R203K/G204R double substitution (KR mt), present in the alpha, gamma, and lambda variants (**Fig. 1d**). Historically, the KR mt has been the most abundant mutation in this region, emerging early in the pandemic and peaking at 73% of reported sequences in April of 2021 (**Fig. 1a, table S1**). The second prominent mutation, T205I, is present in the beta, eta, and mu lineages (**Fig. 1d**). While also emerging early in the pandemic, T205I is a minority variant which peaked at 9% in February 2021 (**Fig. 1a, table S1**). The third prominent variant mutation is R203M, currently present in the delta and kappa variants (**Fig. 1d**). Interestingly, while the R203M mutation was first detected in March of 2020, it persisted as a rare (<1%) variant until April of 2021 when it began expanding rapidly, reaching 91% of all reported sequences in July 2021 (**Fig. 1a, table S1**). Together, these data reveal a complex pattern of genetic variation and convergent evolution for residues 203-205 of SARS-CoV-2 N.

To determine if mutations in this variable motif have the potential to enhance infection, we evaluated the replication kinetics of SARS-CoV-2 variants. Two cell models were selected for this analysis: Vero E6 (commonly used for propagation and titration of SARS-CoV-2) and Calu-3 2b4 (a respiratory cell line used to study coronavirus and influenza infection) (*10, 11*). Briefly, Vero E6 or Calu-3 2b4 cells were inoculated at a low multiplicity of infection (MOI) of 0.01 plaque forming units/cell with the early pandemic Washington-1 (WA-1) strain or a SARS-CoV-2 variant and replication kinetics monitored for 48 hours post infection (hpi). In Vero E6 cells, while the

alpha and beta variants replicated to equivalent endpoint titers compared to WA-1, both variants had slightly lower titer at 24 hpi (**Fig. 1b, fig. S1a**). In contrast, the kappa variant replicated to ~15-fold lower titer than WA-1 throughout infection (**fig. S1a**), potentially due to processing mutations in SARS-CoV-2 spike shared with the delta variant (*12*). Interestingly, in Calu-3 2b4 cells, the alpha variant replicated to a 5.6 fold higher endpoint titer compared to WA-1 (**Fig. 1c**). In contrast, the beta variant replicated to a lower (2.7 fold) mean endpoint titer compared to WA-1 while the kappa variant showed no significant differences with WA-1 (**fig. S1b**). Together, these data suggest that SARS-CoV-2 variants harboring N mutations may correspond to altered replication kinetics.

The KR mt alone is sufficient to increase viral replication

Because the alpha variant exhibited enhanced replication in Calu-3 2b4 cells, we selected the KR mt for further examination. To study its effects of the KR mt in isolation, we utilized a SARS-CoV-2 reverse genetic system to recreate the KR mt in a WA-1 background (**Fig. 1e**) (*8, 9*). In addition, due to gain-of-function concerns, the accessory protein ORF7 was replaced with mNeonGreen (mNG), which reduces but does not eliminate disease in golden Syrian hamsters (**fig. S2**). After recovery of recombinant virus, we evaluated the KR mt's effects on SARS-CoV-2 replication. Both Vero E6 and Calu-3 2b4 cells were infected at a low MOI (0.01) with either SARS-CoV-2 WA-1 harboring the mNG reporter (herein referred to as WT) or the KR mt and viral titer monitored for 48 hpi. Like the alpha variant, the KR mt grew to a lower titer at 24 hpi, but had an equivalent endpoint titer in Vero E6 cells (**Fig. 1f**). Notably, in Calu-3 2b4 cells, the KR mt had increased viral titer at both 24 and 48 hpi compared to WT SARS-CoV-2 (**Fig. 1g**). These data suggest the KR mt in nucleocapsid alone is sufficient to enhance viral replication.

We next examined replication of the KR mutant on air-liquid interface (ALI) cultures. Grown on raised membranes with their apical surfaces exposed to air, ALI cultures trigger differentiation of cells into a pseudostratified epithelium, inducing cell polarization, mucus production, and tight junction formation, which are key features of the airway epithelium absent in traditional monolayer cultures (*13*). To confirm the effects of the KR mt in physiologically relevant systems, ALI cultures of both Calu-3 2b4 and primary Human Airway Epithelia (HAE) cells were infected with WT SARS-CoV-2 or the KR mt and replication kinetics evaluated. For Calu-3 2b4 ALIs, the KR mt grew to a higher mean titer at 24 hpi, but had a lower endpoint titer than WT SARS-CoV-2 (**fig. S3a**). Coupled with the Calu3 2B4 monoculture data, these results suggest that the KR mt enhances SARS-CoV-2 replication in human respiratory cells.

The KR mt enhances SARS-CoV-2 fitness during direct competition

We next determined if the KR mt increases SARS-CoV-2 fitness using competition assays, which offers increased sensitivity compared to individual culture experiments (*14*). WT SARS-CoV-2 and the KR mt were directly competed by infecting Vero E6 and Calu-3 2b4 cells at a 1:1 plaque forming unit ratio. Twenty-four hpi, total cellular RNA was harvested and the ratio of WT to KR mt genomes determined by next generation sequencing (NGS) (*15*). Consistent with the kinetic data, WT outcompeted the KR mt at a ratio of ~4:1 in Vero E6 cells (**Fig. 1h**). In contrast, the KR mt outcompeted WT at a ratio of ~10:1 in Calu-3 2b4 cells (**Fig. 1h**). These data indicate that the KR mt has a fitness advantage over WT SARS-CoV-2 in Calu-3 2b4, but not Vero E6 cells.

The KR mt increases pathogenesis and fitness in vivo

We next determined the effects of the KR mt *in vivo* using a golden Syrian hamster model of SARS-CoV-2 infection (*16*). Three- to four-week-old male hamsters underwent intranasal inoculation with either PBS (mock) or 10^4 plaque forming units (PFU) of WT SARS-CoV-2 or the

KR mt and weight loss was monitored for 10 days post infection (dpi, **Fig. 2a**). On days 2 and 4 dpi, a cohort of five animals from each group underwent nasal washing, were euthanized, and trachea and lung tissue harvested for measurement of viral loads and histopathologic analysis. On day 10, surviving animals were euthanized and lung tissue harvested for histopathological analysis. Strikingly, animals infected with the KR mt had increased weight loss compared to WT throughout the experiment (**Fig. 2b**). Curiously, weight loss changes did not correlate with increased viral loads, as no significant viral titer difference in the lung or trachea was observed between WT and the KR mt (**Fig. 2c-d**). Furthermore, the KR mt caused a small, but significant, decrease in titer in nasal washes at day 2, but not day 4 (**Fig. 2e**). Contrasting the titer data, histopathologic analysis of lungs revealed that the KR mt had more severe lesions compared to WT SARS-CoV-2 (**fig. S5**). Compared to mock (**fig. S5a**), both WT and the KR mt had bronchiolitis and interstitial pneumonia; however, larger more diffuse pulmonary lesions were observed in the KR mt at day 4 (**fig. S5b-c**). In addition, the KR mt had cytopathic alveolar pneumocytes as well as alveoli containing both mononuclear cells and red blood cells (**fig. S5c**). By day 10, both WT and the KR mutants showed signs of recovery, but maintained interstitial pneumonia adjacent to the bronchi absent in mock infected animals (**fig. S5d-f**). Notably, the KR mutant had evidence of cytopathic effect in bronchioles, perivascular edema, and immune infiltration of the endothelium. Together, these data demonstrate that the KR mt increases disease following SARS-CoV-2 infection *in vivo*.

We next evaluated the KR mt's effects on SARS-CoV-2 fitness and transmission *in vivo*. Singly-housed three- to four-week-old male donor hamsters were intranasally inoculated with 10^4 PFU of WT SARS-CoV-2 and the KR mt at a ratio of 1:1 (**Fig. 2f**). On day 1 of infection, each donor hamster was co-housed with a recipient for 8 hours to allow transmission. Hamsters were then separated and SARS-CoV-2 present in the nasal cavities of the donors sampled by nasal wash.

On day 2, each recipient hamster underwent nasal washing to sample transmitted SARS-CoV-2. Donor and recipient hamsters then underwent nasal washing and harvesting of lung tissue on days 4 and 5, respectively, and the ratio of WT to KR mt in all samples was determined by NGS (*15*). Neither WT nor the KR mt consistently dominated in the donor washes on day 1; similarly, the day 2 nasal washes from the recipients showed no distinct advantage between KR mt and WT for transmission (**Fig. 2g-h**). However, at day 4 and 5 in the donor and recipient, respectively, the KR mt was slightly more predominant in the nasal wash (**Fig. 2g-h**). Notably, the KR mt dominated the SARS-CoV-2 population found in lungs of both donor and recipient animals on days 4 and 5. Together, the results suggest that the KR mt outcompetes WT *in vivo* independent of transmission.

The KR mt results in increased phosphorylation of SARS-CoV-2 N

The KR mt is located at the C-terminus of the SR domain of SARS-CoV-2 N, a hyperphosphorylated motif targeted by the SRPK, GSK3, and Cdk1 kinases (**Fig. 1d**) (*5, 17-22*) Given the proximity to a key priming residue required for GSK3 phosphorylation (*22*), we hypothesized that the KR mt alters nucleocapsid phosphorylation. To overcome the lack of phospho-specific antibodies for nucleocapsid, we used phosphate-affinity SDS-PAGE (PA SDS-PAGE). PA SDS-PAGE utilizes a divalent Zn^{2+} compound (Phos-Tag™) within acrylamide gels that selectively binds to phosphorylated serine, threonine, and tyrosine residues; the bound Zn^{2+} decreases electrophoretic mobility of a protein proportionally with the number of phosphorylated amino acids (*23*). Importantly, if a protein exhibits multiple phosphorylation states, this will cause a laddering effect, with each phospho-species appearing as a distinct band (**Fig. 3a**).

To assess the KR mt's effects on N-phosphorylation, we infected Calu-3 2b4 cells at a MOI of 0.01 and harvested whole cell lysates 48 hpi. Lysates then underwent PA SDS-PAGE followed by western blotting with an N-specific antibody. When analyzed by PA SDS-PAGE, WT SARS-

CoV-2 displayed a two-band pattern consisting of a faint upper and prominent lower band, corresponding to a highly phosphorylated and a less phosphorylated species, respectively (**Fig. 3b**, lane 1). In contrast, the KR mt displayed four dark bands of progressively slower mobility, indicating a substantially different phosphorylation pattern (**Fig. 3b**, lane 2). Importantly, all four bands migrated more slowly than the prominent WT band, indicating an overall increase in phosphorylation in the KR mt. Together, these data indicate that the KR mt modulates and increases the relative level of SARS-CoV-2 nucleocapsid phosphorylation.

Given the KR mt's effects, we next determined if SARS-CoV-2 variants had altered N-phosphorylation. Calu-3 2b4 cells were infected at a MOI of 0.01 with WA-1 or the alpha, beta, or kappa variants and whole cell lysates harvested at 48 hpi. When analyzed by PA SDS-PAGE, WA-1 had a two-band pattern similar to WT SARS-CoV-2, while the alpha variant displayed a four-band pattern with slower mobility similar to that of the KR mt (**Fig. 3b**, lanes 3-4). Interestingly, the mobilities of the alpha variant bands were decreased compared to the KR mt indicating an even higher level of phosphorylation, potentially due to the additional nucleocapsid mutations at D3L and S235F that alpha carries (*6, 7*). While both the beta (T205I) and kappa (R203M) variants also displayed slower electrophoretic mobility compared to WA-1, the beta variant displayed a two-band pattern reminiscent of WA-1 while kappa displayed a laddered pattern similar to the KR mt (**Fig. 3b**, lanes 5-6). Together, these data suggest variant mutations in N at residues 203-205 result in increased phosphorylation of N.

An alanine double substitution mimics the KR mt.

Given that variant mutations at residues 203-205 are diverse in sequence, we assessed the importance of the specific R→K and G→R mutations to the KR mt's enhancement of infection. To do so, we made an R203A/G204A double alanine substitution mutant (AA mt) in the WA-1

mNG background (**fig. S6**). After recovery of recombinant SARS-CoV-2, we assessed replication in Vero E6 and Calu-3 2b4 monolayer cultures. In contrast with the KR mt, the AA mt had no significant effect on titers in Vero E6 cells (**Fig. 3c**). Nevertheless, the AA mt increased viral titers over WT in Calu-3 2b4 cells throughout infection, mimicking the augmented replication of the KR mt (**Fig. 3d**). We next confirmed these results in Calu-3 2b4 cells grown as ALI cultures. Again phenocopying the KR mt, the AA mt had a higher viral titer at 24 hpi, but decreased endpoint titer when compared to WT (**fig. S3b**). We next tested the fitness of the AA mt by direct competition with WT SARS-CoV-2. Vero E6 and Calu-3 2b4 monolayer cultures were infected with WT and the AA mt at a 1:1 ratio, whole cell RNA harvested 24 hpi, and the ratio of WT to AA mt determined by NGS. In Vero E6 cells, the AA mt had a small but consistent advantage over WT. In contrast, the AA mt outcompeted WT with approximately 5-fold frequency increase after infection of Calu-3 2b4 cells (**fig. S4**). Overall, the similarities in replication and fitness between the KR and AA mts in Calu-3 2b4 cells suggest that the KR mt enhances infection primarily by ablating the ancestral "RG" motif.

Given that the AA mt mimicked the KR mt's enhancement of *in vitro* infection, we hypothesized that the AA mt also increased N phosphorylation. Calu-3 2b4 cells were infected at a MOI of 0.01 with WT SARS-CoV-2 or the AA mt and whole cell lysates were harvested 48 hpi. Lysates then underwent PA SDS-PAGE followed by Western blotting for nucleocapsid. Interestingly, the AA mt exhibited several dark bands with laddered mobility absent in WT, suggesting an increase in phosphorylation similar to the KR mt. However, contrasting with the KR mt, in the AA mt the electrophoretic mobility of the lowest band was similar to WT (**Fig. 3e**). These data suggest that while the KR and AA mts both alter nucleocapsid phosphorylation, their

effects are not identical. Nevertheless, the changes in phosphorylation induced by the KR and AA mt are both sufficient to enhance SARS-CoV-2 replication.

Discussion:

In this manuscript, we investigate a highly variable motif in SARS-CoV-2 nucleocapsid located at residues 203-205 and characterized the effects of the alpha variant's R203K/G204R double substitution mutation (KR mt) on SARS-CoV-2 infection. Inserting the KR mt in the WA-1 background demonstrated that the KR mt alone is sufficient to increase titer and fitness in respiratory models of infection. Similarly, in a hamster model, the KR mt is sufficient to increase disease in infected animals and enhances fitness during direct competition studies in the lung. Importantly, we demonstrate that the KR mt, the alpha variant, and variants with analogous mutations have increased nucleocapsid phosphorylation compared to WT SARS-CoV-2, providing a molecular basis for the observed phenotypic changes. Notably, an analogous double alanine substitution (AA mt) also showed increased replication, fitness, and altered nucleocapsid phosphorylation. Together, these data suggest that the KR mt and similar mutations in the nucleocapsid enhance SARS-CoV-2 infection by disrupting the ancestral "RGT" motif.

The KR mt occurs within the SR domain of nucleocapsid and has a complex role during SARS-CoV-2 infection. This region of N is hyper-phosphorylated intracellularly, but unphosphorylated within the mature virion (*24, 25*). Several studies suggest that phosphorylation of the SR region acts as a biophysical switch, regulating nucleocapsid function through phase separation (*21, 26, 27*). In the proposed model, unphosphorylated nucleocapsid forms gel-like condensates with viral RNA and the SARS-CoV-2 membrane (M) protein to facilitate genome packaging and virus assembly. In contrast, phosphorylated nucleocapsid forms distinct liquid-like condensates to promote N binding to SARS-CoV-2 nsp3 (*21*), G3P1 in stress granules (*26*), and (presumably) other nucleocapsid functions (*28*). While not tested in the context of infection, this model is consistent with studies demonstrating interactions between N and M (*29, 30*), G3BP1 and

G3BP2 within stress granules (*31-33*), and phosphorylated N and nsp3 to promote the synthesis of long subgenomic RNAs and full genomes (*24, 34-36*). Within this context, the KR mt may optimize this biomolecular switch for human infection, increasing the amount of phosphorylated nucleocapsid and shifting the overall function of N during infection. Alternatively, the KR mt may impact the interaction between nucleocapsid and host 14-3-3 proteins which bind nucleocapsid at several sites in the SR region of N in a phosphorylation dependent manner (*37*). This interaction is required for cytoplasmic localization of nucleocapsid in SARS-CoV (*38*). Notably, one 14-3-3 binding site encompasses the 203-205 motif examined in this study (*37*) and the KR mt may enhance infection by altering this interaction.

Overall, in this study we establish that the KR mt enhances SARS-CoV-2 infection relative to WT, increasing viral fitness in vitro and in vivo, which along with the N501Y mutation (*39*), likely selected for the emergence of the alpha variant. We also find that the KR mt increases nucleocapsid phosphorylation, suggesting a molecular basis for the mutant's effects. Importantly, we show that other variant mutations in this motif increase nucleocapsid phosphorylation, which may have aided in the emergence of their respective lineages. When taken with the prevalence of mutations in the SR region, these data suggest that increasing nucleocapsid phosphorylation through mutations in residues 203-205 is a common feature in circulating variants and represents a positive selection event for SARS-CoV-2 during human infection. Importantly, our work highlights that mutations outside of SARS-CoV-2 spike have significant effects on infection and must not be overlooked while characterizing mechanism of variant emergence.

References and Notes

1. International Monetary Fund Research Dept., "World Economic Outlook, April 2020 : The Great Lockdown," *World Economic Outlook No. No. 2020/001* (International Monetary Fund, 2020).
2. Word Health Organization, Coronavirus disease (COVID-19) pandemic. **2021**, (2021).
3. Word Health Organization, Tracking SARS-CoV-2 variants. **2021**, (2021).
4. J. A. Plante *et al.*, The variant gambit: COVID-19's next move. *Cell Host Microbe*, (2021).
5. Q. Ye, A. M. V. West, S. Silletti, K. D. Corbett, Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Sci* **29**, 1890-1901 (2020).
6. S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall* **1**, 33-46 (2017).
7. J. L. Mullen *et al.*, outbreak.info. **2021**, (2020).
8. X. Xie *et al.*, Engineering SARS-CoV-2 using a reverse genetic system. *Nature Protocols* **16**, 1761-1784 (2021).
9. X. Xie *et al.*, An Infectious cDNA Clone of SARS-CoV-2. *Cell Host & Microbe* **27**, 841-848.e843 (2020).
10. J. Harcourt *et al.*, Severe Acute Respiratory Syndrome Coronavirus 2 from Patient with Coronavirus Disease, United States. *Emerg Infect Dis* **26**, 1266-1273 (2020).
11. V. D. Menachery *et al.*, Pathogenic influenza viruses and coronaviruses utilize similar and contrasting approaches to control interferon-stimulated gene responses. *mBio* **5**, e01174-01114 (2014).
12. Y. Liu *et al.*, Delta spike P681R mutation enhances SARS-CoV-2 fitness over Alpha variant. *bioRxiv*, 2021.2008.2012.456173 (2021).
13. X. Cao *et al.*, Invited review: human air-liquid-interface organotypic airway tissue models derived from primary tracheobronchial epithelial cells-overview and perspectives. *In Vitro Cell Dev Biol Anim* **57**, 104-132 (2021).
14. J. A. Plante *et al.*, Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**, 116-121 (2021).
15. E. Jaworski *et al.*, Tiled-ClickSeq for targeted sequencing of complete coronavirus genomes with simultaneous capture of RNA recombination and minority variants. *bioRxiv*, (2021).
16. M. Imai *et al.*, Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proceedings of the National Academy of Sciences* **117**, 16587 (2020).
17. M. Bouhaddou *et al.*, The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell* **182**, 685-712 e619 (2020).
18. A. D. Davidson *et al.*, Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Medicine* **12**, 68 (2020).
19. K. Klann *et al.*, Growth Factor Receptor Signaling Inhibition Prevents SARS-CoV-2 Replication. *Mol Cell* **80**, 164-174.e164 (2020).
20. T. M. Yaron *et al.*, The FDA-approved drug Alectinib compromises SARS-CoV-2 nucleocapsid phosphorylation and inhibits viral infection in vitro. *bioRxiv*, 2020.2008.2014.251207 (2020).
21. C. R. Carlson *et al.*, Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions. *Mol Cell* **80**, 1092-1103.e1094 (2020).
22. X. Liu *et al.*, Targeting the coronavirus nucleocapsid protein through GSK-3 inhibition. *Proc Natl Acad Sci U S A* **118**, (2021).
23. E. Kinoshita, E. Kinoshita-Kikuta, T. Koike, The Cutting Edge of Affinity Electrophoresis Technology. *Proteomes* **3**, 42-55 (2015).
24. C. H. Wu, P. J. Chen, S. H. Yeh, Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription. *Cell Host Microbe* **16**, 462-472 (2014).
25. C. H. Wu *et al.*, Glycogen synthase kinase-3 regulates the phosphorylation of severe acute respiratory syndrome coronavirus nucleocapsid protein and viral replication. *J Biol Chem* **284**, 5229-5239 (2009).
26. S. Lu *et al.*, The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat Commun* **12**, 502 (2021).
27. J. Cubuk *et al.*, The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat Commun* **12**, 1936 (2021).
28. R. McBride, M. van Zyl, B. C. Fielding, The coronavirus nucleocapsid is a multifunctional protein. *Viruses* **6**, 2991-3018 (2014).

29. R. He *et al.*, Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus. *Virus Res* **105**, 121-125 (2004).
30. K. R. Hurst *et al.*, A major determinant for membrane protein interaction localizes to the carboxy-terminal domain of the mouse coronavirus nucleocapsid protein. *J Virol* **79**, 13285-13297 (2005).
31. D. E. Gordon *et al.*, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459-468 (2020).
32. J. Li *et al.*, Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis. *Med (N Y)*, (2020).
33. Z. Q. Zheng, S. Y. Wang, Z. S. Xu, Y. Z. Fu, Y. Y. Wang, SARS-CoV-2 nucleocapsid protein impairs stress granule formation to promote viral replication. *Cell Discov* **7**, 38 (2021).
34. Y. Cong *et al.*, Nucleocapsid Protein Recruitment to Replication-Transcription Complexes Plays a Crucial Role in Coronaviral Life Cycle. *J Virol* **94**, (2020).
35. M. H. Verheije *et al.*, The coronavirus nucleocapsid protein is dynamically associated with the replication-transcription complexes. *J Virol* **84**, 11575-11579 (2010).
36. S. Zuniga *et al.*, Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription. *J Virol* **84**, 2169-2175 (2010).
37. K. V. Tugaeva *et al.*, The Mechanism of SARS-CoV-2 Nucleocapsid Protein Recognition by the Human 14-3-3 Proteins. *Journal of Molecular Biology* **433**, 166875 (2021).
38. M. Surjit *et al.*, The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. *J Virol* **79**, 11476-11486 (2005).
39. Y. Liu *et al.*, The N501Y spike substitution enhances SARS-CoV-2 transmission. *bioRxiv*, (2021).
38. L. Josset et al., Cell host response to infection with novel human coronavirus EMC predicts potential antivirals and important differences with SARS coronavirus. mBio 4, e00165-00113 (2013).
39. A. Vanderheiden et al., Development of a Rapid Focus Reduction Neutralization Test Assay for Measuring SARS-CoV-2 Neutralizing Antibodies. Curr Protoc Immunol 131, e116 (2020).
40. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359 (2012).
41. T. Smith, A. Heger, I. Sudbery, UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res 27, 491-499 (2017).

Acknowledgments:

We would like to thank Shinji Makino for gifting a nucleocapsid antibody.

Funding:

Research was supported by grants from NIAID of the NIH to (AI153602 and 1R21AI145400 to VDM; R24AI120942 (WRCEVA) to SCW. ALR and SCW supported by Centers for Disease Control and Prevention (Contract 200-2021-11195). ZY and ALR was supported by an Institute of Human Infection and Immunity at UTMB COVID-19 Research Fund. Research was also supported by STARs Award provided by the University of Texas System to VDM, and trainee funding provided by the McLaughlin Fellowship Fund (BAJ) and the UTMB Center for Tropical Disease (BAJ) at UTMB. NIH training grants supported MNV(T32-AI060549) and CS (5T32AI007526-20).

Author contributions:

Conceptualization: BAJ, VDM

Formal analysis: BAJ, YZ, JAP, ALR

Funding acquisition: BAJ, YZ, SCW, P-YS, ALR, VDM

Investigation: BAJ, YZ, KGL, MNV, NB, CS, PACV, YL, DS, JAP, XX, DW, KSP

Methodology: BAJ, YZ, PACV, JAP, ALR, SCW, VDM

Project Administration: BAJ, VDM

Supervision: PA, SCW, P-YS, KSP, ALR, VDM

Visualization: BAJ, JAP, DW

Writing – original draft: BAJ, VDM

Writing – review and editing: BAJ, P-YS, PACV, JAP, YL, XX, SCW, DW, ALR, CS, KSP, VDM,

Competing interests: XX, P-YS, and VDM have filed a patent on the reverse genetic system and reporter SARS-CoV-2. Other authors declare no competing interests

Data and materials availability: All raw sequencing data will be available in the NCBI Small Read Archive with BioProject. Raw data available upon request.

Supplementary Materials:

Materials and Methods

Figures S1 to S6

Table S1

References (38 - 41)

Figure Legends

Figure 1: The KR mt enhances SARS-CoV-2 replication. (**A**) Amino acid frequencies for nucleocapsid residues 203-205 in SARS-CoV-2 sequences reported to the GISAID, binned by month of collection and graphed as percent of total sequences reported during that period. (**B-C**) Viral titer from Vero E6 (B) or Calu-3 2b4 cells (C) infected with WA-1 (black) or the alpha variant (red) at an MOI of 0.01 (n≥6). (**D**). Schematic of the SR region of SARs-CoV-2 nucleocapsid. Variable residues are displayed as red text within the sequence of their corresponding lineages. Phosphorylated residues are indicated by a †. (**E**) Schematic of the SARS-CoV-2 genome, showing the creation of the KR mutation and the replacement of ORF7 with the mNG reporter protein. (**F-G**) Viral titer of Vero E6 (F) or Calu-3 2b4 (G) cells infected with WT (black) or the KR mt (red) at a MOI of 0.01 (n=9). (**H**) HAE cells infected with WT (black) or the KR mt (red) at an MOI of 0.01 (n ≥2). Data are the mean ± s.d. Statistical significance was determined by two-tailed students T-test with p≤0.05 (*), p≤0.01 (**), and p≤ 0.001 (***). Dotted lines are equal to LOD.

Figure 2: The KR mt enhances SARS-CoV-2 pathogenesis and fitness. (**A**) Schematic of the infection of hamsters with SARS-CoV-2. (**B-E**) Three to four-week-old male hamsters were mock infected (gray) or inoculated with 10^4 PFU of WT SARS-CoV-2 (black) or the KR mt (red). Animals were then monitored for weight loss (B). On days 2 and 4 post infection, viral titers in the lung (C), trachea (D), and from nasal washes (E) were determined. (**F**) Schematic of competition/transmission experiment. (**G-H**) Three to four-week-old male donor hamsters were inoculated with 10^4 PFU of WT SARS-CoV-2 and the KR mt at a 1:1 ratio. On day 1 of the experiment, donor and recipient hamsters were co-housed for 8 hours, then separated and the donor hamsters underwent nasal washing. On day 2, recipient hamsters were nasal washed. Hamsters

were then monitored and nasal washes and lung tissue harvested on days 4 (donors) and 5 (recipients). The ratio of WT (gray) to KR mt (red) was then determined by NGS of all Donor (G) and recipient (H) samples. For weight loss data, mean percent weight loss was graphed ± s.e.m. For titer data, individual weights were graphed with means ± s.d. indicated by lines. For competition studies, individual replicates are graphed while bars represent the mean. Significance was determined by students T-Test with $p \leq 0.05$ (*), $p \leq 0.01$ (**), and $p \leq 0.001$ (***). Dotted lines are equal to LOD.

Figure 3: The KR mt increases N phosphorylation to enhance infection through the disruption of the ancestral RG sequence. (**A**) Schematic of phosphate-affinity (PA) SDS-PAGE. (**B**) Whole cell lysates from Calu-3 2b4 cells infected with SARS-CoV-2 WA-1-mNG (WT), the KR mt, WA-1, alpha, beta, and kappa variants were analyzed by PA SDS-Page (top) or standard SDS-PAGE (bottom) followed by blotting with an N-specific antibody. (**C-D**) Viral titer from Vero E6 (C) or Calu-3 2b4 (D) cells were infected with WT or the AA mt at an MOI of 0.01 (n=9). (**E**) Whole cell lysates from Calu-3 2b4 cells infected with WT SARS-CoV-2 or the AA mt were analyzed by PA SDS-PAGE (top) or standard SDS-PAGE (bottom) followed by blotting for nucleocapsid. Blots are representative of three independent experiments. Graphs represent mean titer ± s.d. Significance was determined by two-tailed student's t-test with ≤ 0.05 (*), $p \leq 0.01$ (**), and $p \leq 0.001$ (***). Dotted lines are equal to LOD.

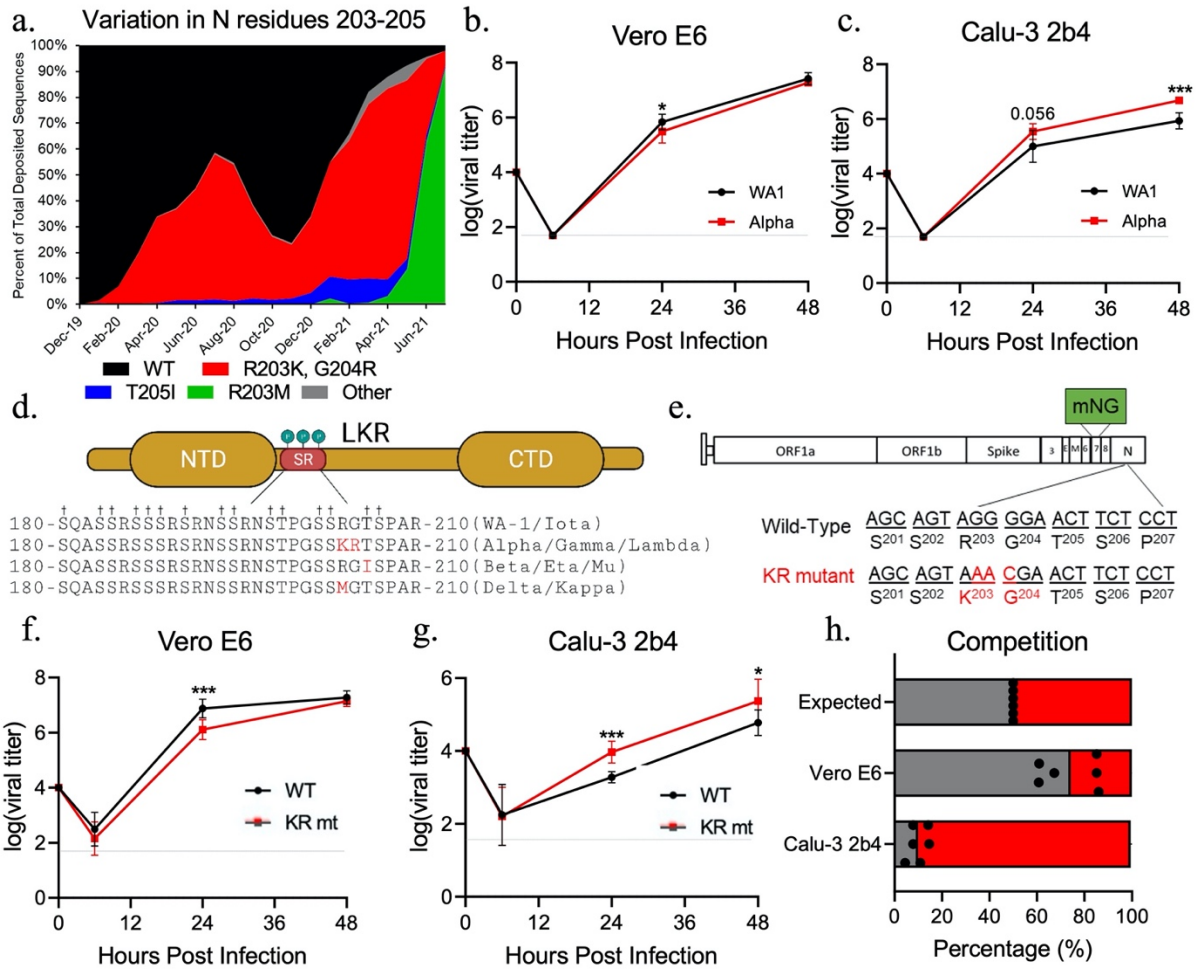

Figure 1: The KR mt enhances SARS-CoV-2 replication. (**A**) Amino acid frequency for nucleocapsid residues 203-205 in SARS-CoV-2 sequences reported to the GISAID, binned by month of collection and graphed as percent of total sequences reported during that period. (**B-C**) Viral titer from Vero E6 (B) or Calu-3 2b4 cells (C) infected with WA-1 (black) or the alpha variant (red) at an MOI of 0.01 (n≥6). (D). Schematic of the SR region of SARs-CoV-2 nucleocapsid. Variable residues are displayed as red text within the sequence of their corresponding lineages. Phosphorylated residues are indicated by a †. (**E**) Schematic of the SARS-CoV-2 genome, showing the creation of the KR mutation and the replacement of ORF7 with the mNG reporter protein. (**F-G**) Viral titer of Vero E6 (F) or Calu-3 2b4 (G) cells infected with WT (black) or the KR mt (red) at a MOI of 0.01 (n=9). (**H**) Competition assay between WT (gray) and KR mt (red) in Vero E6 or Calu-3 2b4 cells infected at a 1:1 input ratio with an MOI of 0.01 (n=6). Titer data are the mean ± s.d. For competition, the percentage WT is graphed as points for individual replicates while the mean percentage of each virus is displayed as a bar graph. Statistical significance was determined by two-tailed students T-test with p≤0.05 (*), p≤0.01 (**), and p≤ 0.001 (***). Dotted lines are equal to LOD.

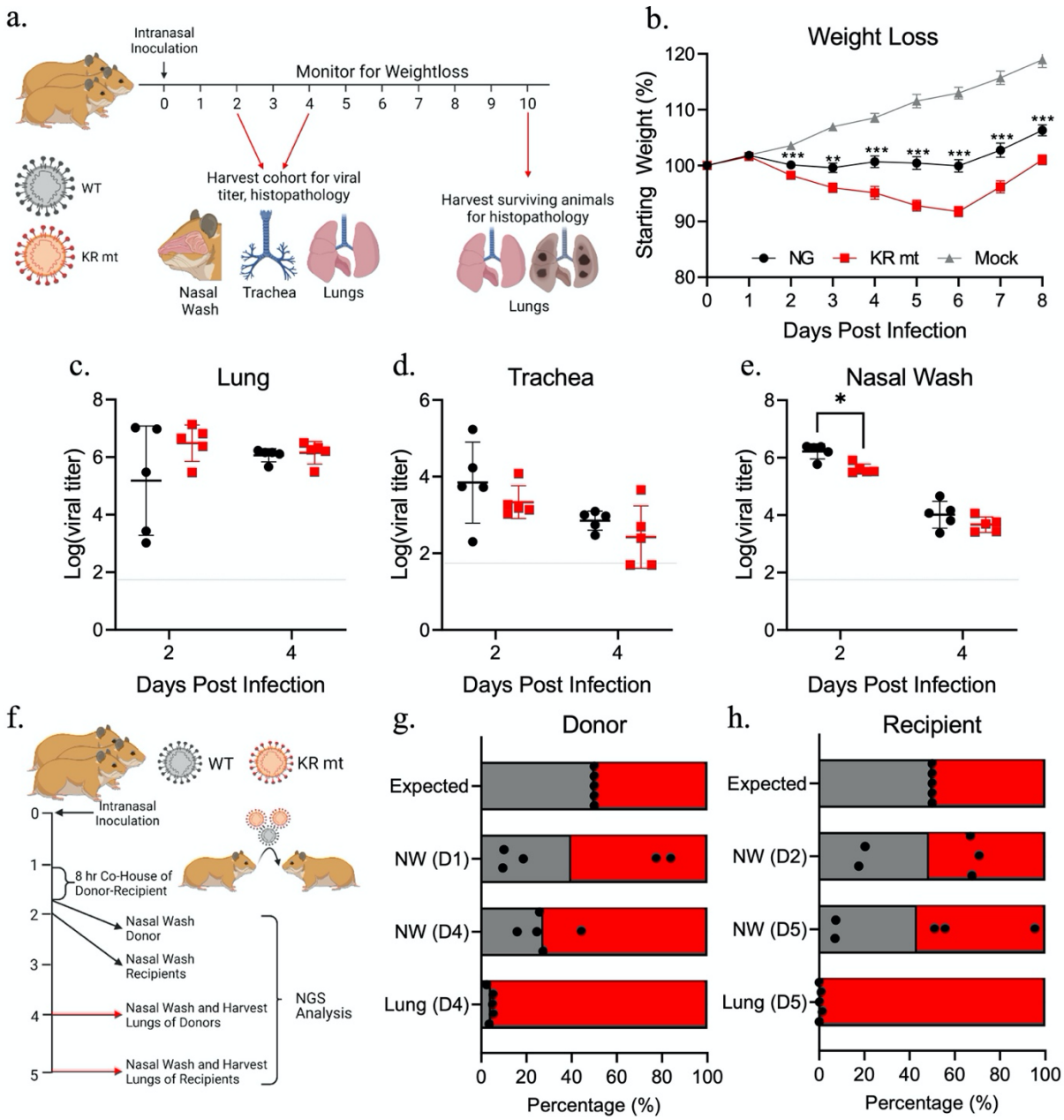

Figure 2: The KR mt enhances SARS-CoV-2 pathogenesis and fitness. (**A**) Schematic of the infection of hamsters with SARS-CoV-2. (**B-E**) Three to four week old male hamsters were mock infected (gray) or inoculated with 10^4 PFU of WT SARS-CoV-2 (black) or the KR mt (red). Animals were then monitored for weight loss (B). On days 2 and 4 post infection, viral titer in the lung (C), trachea (D), and from nasal washes (E) was determined. (**F**) Schematic of competition/transmission experiment. (**G-H**) Three to four week old male donor hamsters were inoculated with 10^4 PFU of WT SARS-CoV-2 and the KR mt at a 1:1 ratio. On day 1 of the experiment, donor and recipient hamsters were co-housed for 8 hours, then separated and the donor hamsters underwent nasal washing. On day 2, recipient hamsters were nasal washed. Hamsters were then monitored and nasal washes and lung tissue harvested on days 4 (donors) and 5 (recipients). The ratio of WT (gray) to KR mt (red) was then determined by NGS of all Donor (G) and recipient (H) samples. For weight loss data, mean percent weight loss was graphed ± s.e.m. For titer data, individual weights were graphed with means ± s.d. indicated by lines. For competition studies, individual replicates are graphed while bars represent the mean. Significance was determined by students T-Test with $p \leq 0.05$ (*), $p \leq 0.01$ (**), and $p \leq 0.001$ (***). Dotted lines are equal to LOD.

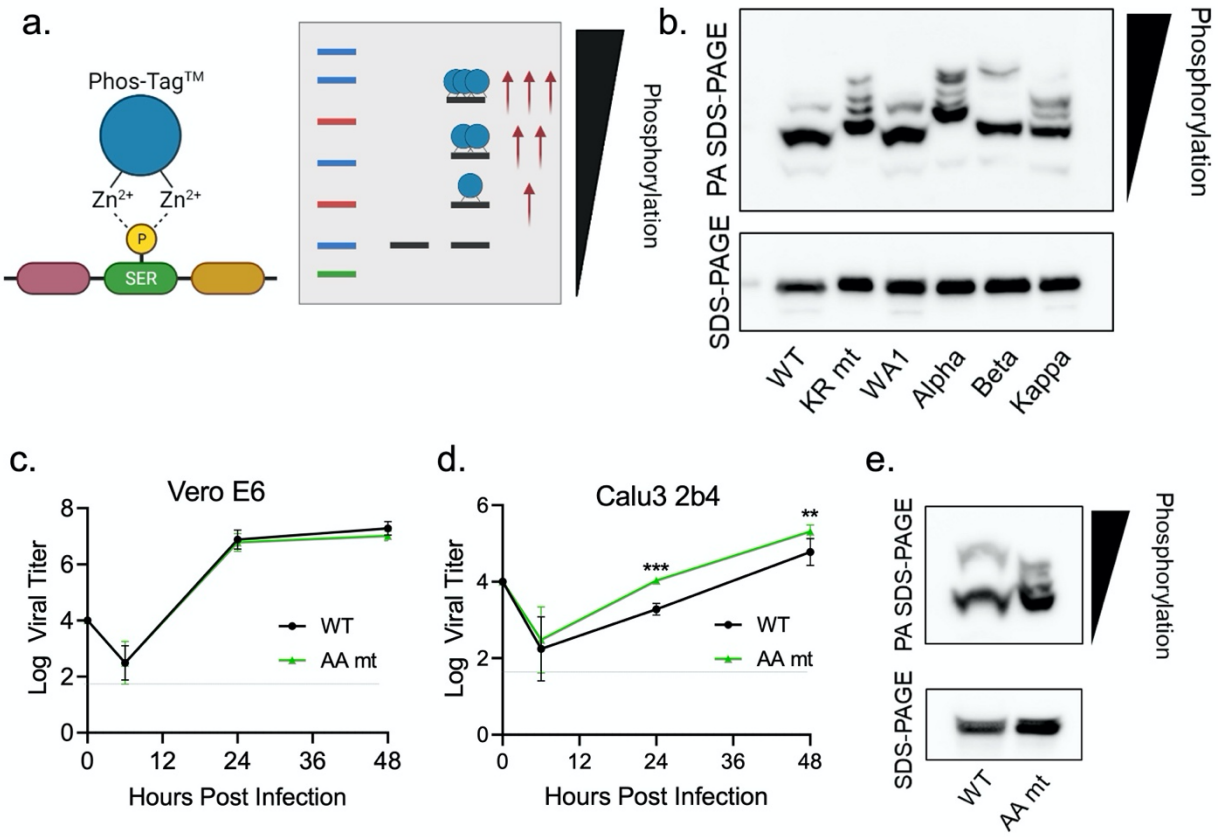

Figure 3: The KR mt increases N phosphorylation to enhance infection through the disruption of the ancestral RG sequence. (**A**) Schematic of phosphate-affinity (PA) SDS-PAGE. (**B**) Whole cell lysates from Calu-3 2b4 cells infected with SARS-CoV-2 WA-1-mNG (WT), the KR mt, WA-1, alpha, beta, and kappa variants were analyzed by PA SDS-Page (top) or standard SDS-PAGE (bottom) followed by blotting with an N specific antibody. (**C-D**) Viral titer from Vero E6 (C) or Calu-3 2b4 (D) cells were infected with WT or the AA mt at an MOI of 0.01 (n=9). (**E**) Whole cell lysates from Calu-3 2b4 cells infected with WT SARS-CoV-2 or the AA mt were analyzed by PA SDS-PAGE (top) or standard SDS-PAGE (bottom) followed by blotting for nucleocapsid. Blots are representative of three independent experiments. Graphs represent mean titer ± s.d. Significance was determined by two-tailed student's t-test with ≤0.05 (*), p≤0.01 (**), and p≤ 0.001 (***). Dotted lines are equal to LOD.

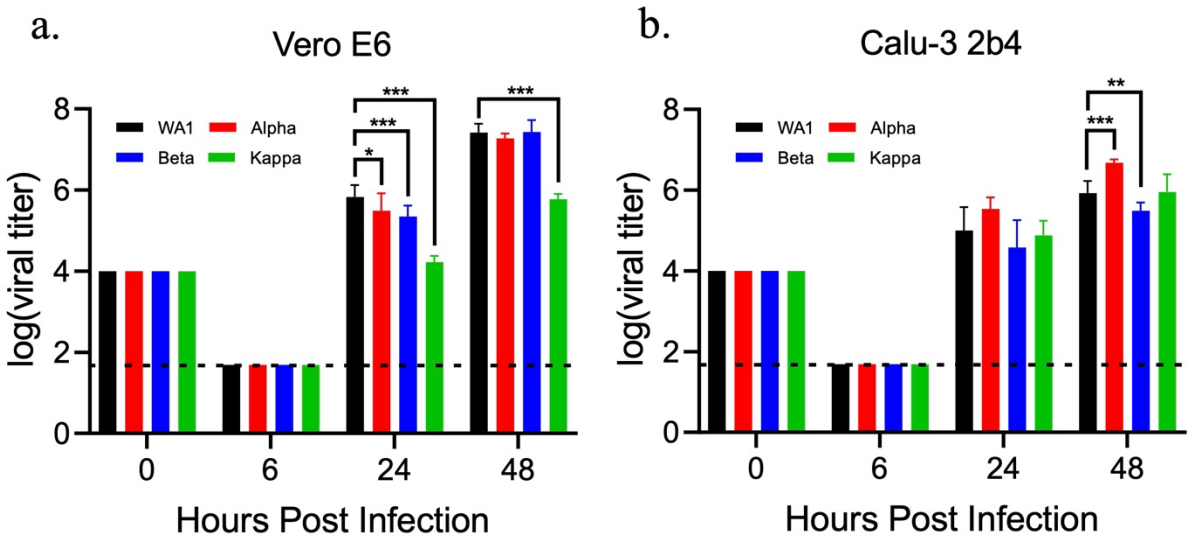

Figure S1: Replication of SARS-CoV-2 variants. Viral titer from Vero E6 (**A**) or Calu-3 2b4 cells (**B**) inoculated with SARS-CoV-2 WA-1 (black) or the alpha (red), beta (blue) or kappa (green) variant at a MOI of 0.01. Graphed data represent the mean ± s.d. Statistical significance was determined by two-tailed students T-test with $p \leq 0.05$ (*), $p \leq 0.01$ (**), and $p \leq 0.001$ (***). Dotted lines are equal to LOD.

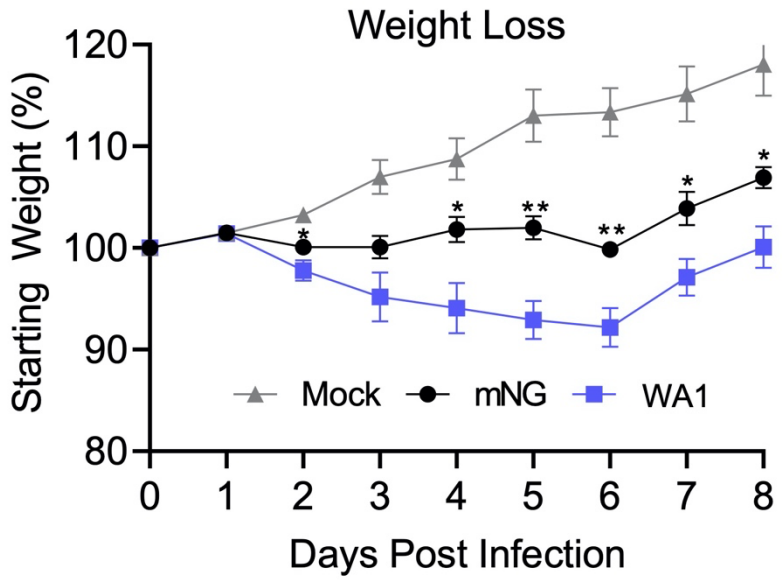

Figure S2: *In vivo a*ttenuation of the SARS-CoV-2 mNeonGreen reporter virus. Golden Syrian hamsters were intranasally inoculated with PBS alone (gray) or 10^4 PFU of WT SARS-CoV-2 (blue) or SARS-CoV-2 mNG (black). Graphed data represent the mean weight loss ± s.e.m. Statistical significance determined by two-tailed students T-test with $p \leq 0.05$ (*) and $p \leq 0.01$ (**).

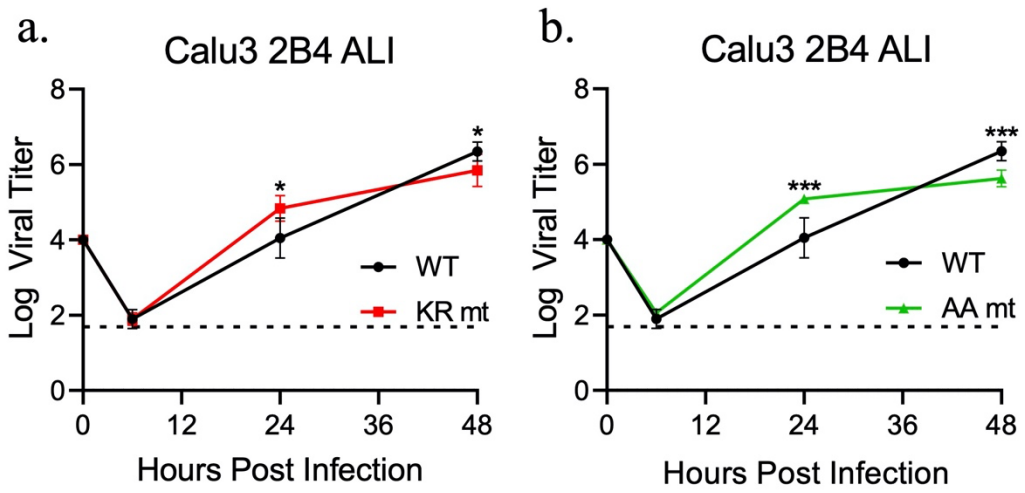

Figure S3: Both the KR mt and AA mt augment viral replication at 24 hpi in Calu-3 2b4 grown as ALI. (**A**) Viral titer from Calu-3 2b4 ALI cultures were infected with WT (black) or the KR mt (red) at an MOI of 0.01. (**B**) Viral titerCalu-3 2b4 ALI cultures were infected with WT (black) or the AA mt at an MOI of 0.01. Graphed data represent the mean ± s.d. Statistical significance was determined by two-tailed students T-test with $p \leq 0.05$ (*), $p \leq 0.01$ (**), and $p \leq 0.001$ (***). Dotted lines are equal to LOD.

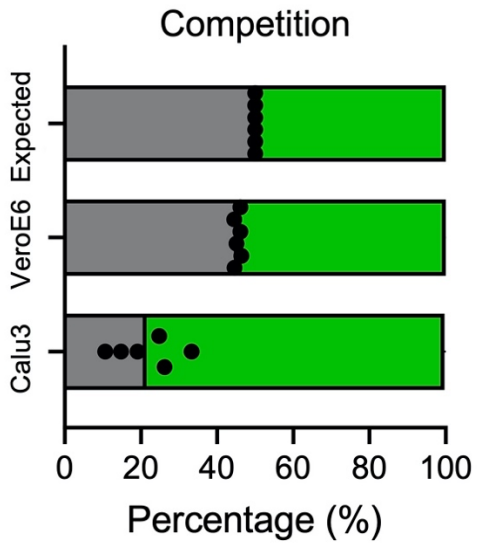

Figure S4: The KR and AA mutants augment viral fitness. Competition assay between WT (gray) and AA mt (green) in Vero E6 or Calu-3 2b4 cells infected at a 1:1 input ratio with an MOI of 0.01 (n=6). The percentage WT is graphed as points for individual replicates while the mean percentage of each virus is displayed as a bar graph.

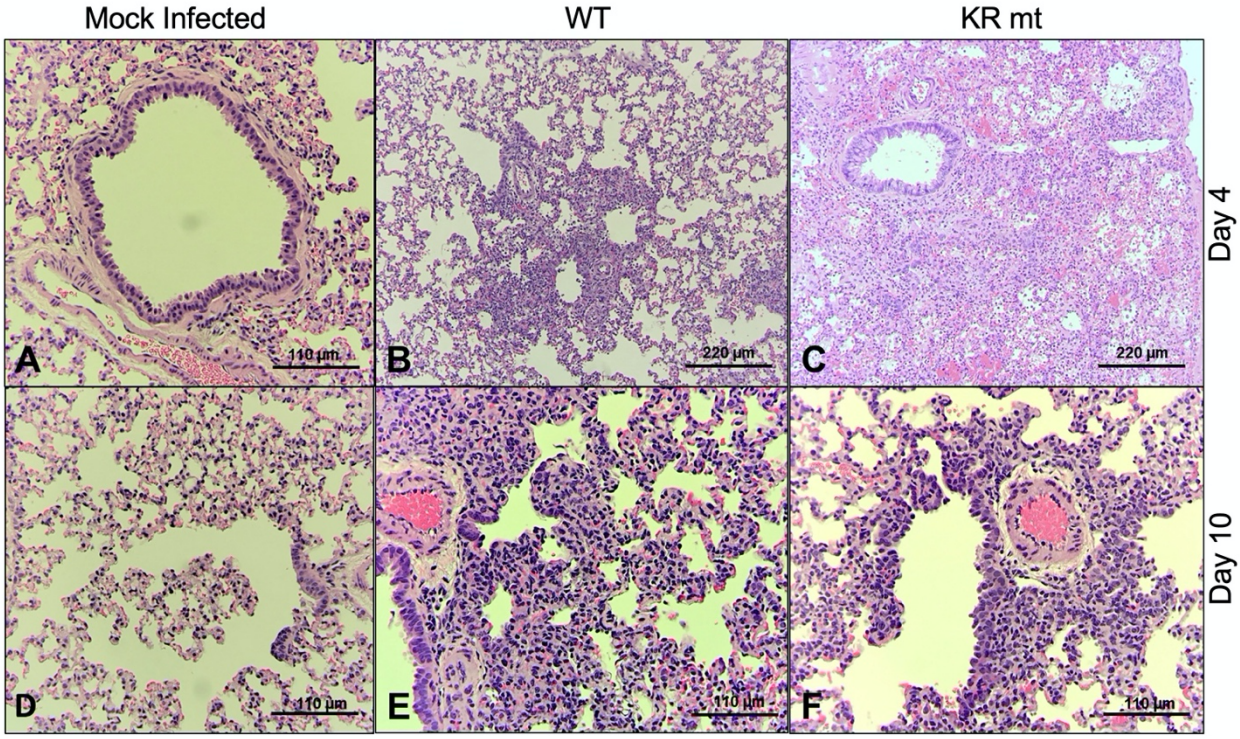

Figure S5: Lung histopathology in hamsters infected with WT and KR mt SARS-CoV-2. Lung tissue was harvested, fixed, and 5 µm sections cut from mock, WT SARS-CoV-2, or the KR mt-infected hamsters and stained with hematoxylin and eosin. (**A**) Normal bronchus, pulmonary artery, and alveoli in mock infection on day 4 (20X). (**B**) Bronchiolitis, peribronchiolitis, interstitial pneumonia, and edema surrounding branch of the pulmonary artery at day 4 in hamsters infected with WT virus (10X). (**C**) Severe bronchiolar cytopathic effect, interstitial pneumonia, cytopathic alveolar pneumocytes, alveoli containing mononuclear cells and red blood cells at day 4 in hamsters infected with KR mt. This lesion extended over numerous fields (10X). (**D**) Normal respiratory bronchiole, alveolar ducts, and alveolar sacs in mock infection on day 10 (20X). (**E**) Interstitial pneumonia adjacent to a bronchus at day 10 in hamsters infected with WT (20X). (**F**) Bronchiolar epithelial cytopathic effect, peribronchiolitis, focal interstitial pneumonia, branch of pulmonary artery with surrounding edema and mononuclear cell infiltration of endothelium at day 10 in a hamster infected with KR mt (20X). Shown are representative images typical of data gathered from 5 animals from each group.

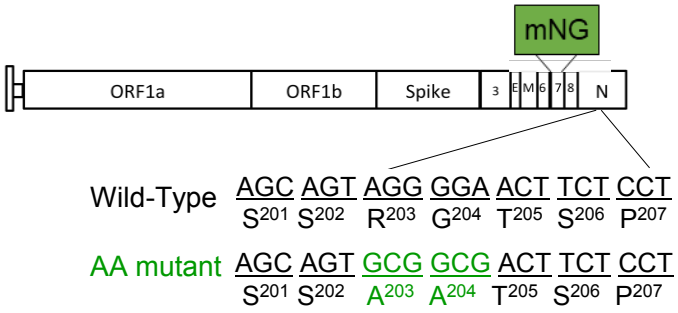

Figure S6: Schematic representation of the AA mt. Schematic shows the creation of the AA mutation within the SARS-CoV-2 genome and the replacement of ORF7 with the mNeonGreen reporter.

a.

Of Sequences

Collection Period	Total # of Sequences	Wild Type	R203M	R203K/G204R	T205I	Other
Dec-19	24	24	0	0	0	0
Jan-20	636	627	0	5	4	0
Feb-20	1,776	1,654	0	118	4	0
Mar-20	51,789	41,743	34	9,849	110	53
Apr-20	52,515	34,698	17	17,492	225	83
May-20	26,722	16,797	12	9,478	365	70
Jun-20	30,279	16,755	13	12,992	438	81
Jul-20	33,132	13,702	51	18,630	574	175
Aug-20	37,196	16,805	28	19,578	484	301
Sep-20	39,886	24,574	20	14,203	871	218
Oct-20	65,901	48,270	18	16,184	1,117	312
Nov-20	95,070	72,554	53	19,843	2,162	458
Dec-20	130,225	86,181	67	37,771	5,709	497
Jan-21	231,857	104,432	5,480	102,299	19,646	0
Feb-21	253,458	86,343	650	135,877	23,757	6,831
Mar-21	400,236	71,259	3,225	269,759	36,535	19,458
Apr-21	395,799	47,657	12,634	291,466	25,294	18,748
May-21	290,394	22,032	39,441	200,989	11,291	16,641
Jun-21	217,989	9,358	137,296	64,233	5,220	1,882
Jul-21	195,146	3,828	178,242	11,290	1,786	0

b.

% Of Sequences

Collection Period	Total # of Sequences	Wild Type	R203M	R203K/G204R	T205I	Other
Dec-19	N/A	100.00	0.00	0.00	0.00	0.00
Jan-20	N/A	98.58	0.00	0.79	0.63	0.00
Feb-20	N/A	93.13	0.00	6.64	0.23	0.00
Mar-20	N/A	80.60	0.07	19.02	0.21	0.10
Apr-20	N/A	66.07	0.03	33.31	0.43	0.16
May-20	N/A	62.86	0.04	35.47	1.37	0.26
Jun-20	N/A	55.34	0.04	42.91	1.45	0.27
Jul-20	N/A	41.36	0.15	56.23	1.73	0.53
Aug-20	N/A	45.18	0.08	52.63	1.30	0.81
Sep-20	N/A	61.61	0.05	35.61	2.18	0.55
Oct-20	N/A	73.25	0.03	24.56	1.69	0.47
Nov-20	N/A	76.32	0.06	20.87	2.27	0.48
Dec-20	N/A	66.18	0.05	29.00	4.38	0.38
Jan-21	N/A	45.04	2.36	44.12	8.47	0.00
Feb-21	N/A	34.07	0.26	53.61	9.37	2.70
Mar-21	N/A	17.80	0.81	67.40	9.13	4.86
Apr-21	N/A	12.04	3.19	73.64	6.39	4.74
May-21	N/A	7.59	13.58	69.21	3.89	5.73
Jun-21	N/A	4.29	62.98	29.47	2.39	0.86
Jul-21	N/A	1.96	91.34	5.79	0.92	0.00

Table S1: Frequency of mutations in residues 203-205 in SARS-CoV-2 nucleocapsid. (**A-B**) Frequency of WT, R203M, R203K/G204R, T205I, or all other genotypes binned by month of collection, represented as the raw totals (A) or as a percentage of total sequences in a given month (B).