# Natural and Recombinant SARS-CoV-2 Isolates Rapidly Evolve *In Vitro* to Higher Infectivity through More Efficient Binding to Heparan Sulfate and Reduced S1/S2 Cleavage

Nikita Shiliaev,[a] Tetyana Lukash,[a] Oksana Palchevska,[a] David K. Crossman,[b] Todd J. Green,[a] Michael R. Crowley,[b] Elena I. Frolova,[a] Ilya Frolov[a]

[a]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, USA
[b]UAB Genomics Core Facility, Department of Genetics, University of Alabama at Birmingham, Birmingham, Alabama, USA

**ABSTRACT** One of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virulence factors is the ability to interact with high affinity to the ACE2 receptor, which mediates viral entry into cells. The results of our study demonstrate that within a few passages in cell culture, both the natural isolate of SARS-CoV-2 and the recombinant cDNA-derived variant acquire an additional ability to bind to heparan sulfate (HS). This promotes a primary attachment of viral particles to cells before their further interactions with the ACE2. Interaction with HS is acquired through multiple mechanisms. These include (i) accumulation of point mutations in the N-terminal domain (NTD) of the S protein, which increases the positive charge of the surface of this domain, (ii) insertions into the NTD of heterologous peptides containing positively charged amino acids, and (iii) mutation of the first amino acid downstream of the furin cleavage site. This last mutation affects S protein processing, transforms the unprocessed furin cleavage site into the heparin-binding peptide, and makes viruses less capable of syncytium formation. These viral adaptations result in higher affinity of viral particles to heparin, dramatic increase in plaque sizes, more efficient viral spread, higher infectious titers, and 2 orders of magnitude higher infectivity. The detected adaptations also suggest an active role of NTD in virus attachment and entry. As in the case of other RNA-positive (RNA⁺) viruses, evolution to HS binding may result in virus attenuation *in vivo*.

**IMPORTANCE** The spike protein of SARS-CoV-2 is a major determinant of viral pathogenesis. It mediates binding to the ACE2 receptor and, later, fusion of viral envelope and cellular membranes. The results of our study demonstrate that SARS-CoV-2 rapidly evolves during propagation in cultured cells. Its spike protein acquires mutations in the NTD and in the P1′ position of the furin cleavage site (FCS). The amino acid substitutions or insertions of short peptides in NTD are closely located on the protein surface and increase its positive charge. They strongly increase affinity of the virus to heparan sulfate, make it dramatically more infectious for the cultured cells, and decrease the genome equivalent to PFU (GE/PFU) ratio by orders of magnitude. The S686G mutation also transforms the FCS into the heparin-binding peptide. Thus, the evolved SARS-CoV-2 variants efficiently use glycosaminoglycans on the cell surface for primary attachment before the high-affinity interaction of the spikes with the ACE2 receptor.

**KEYWORDS** ACE2, SARS-CoV-2, coronavirus, furin cleavage, heparin binding, recombinant virus, spike NTD, spike protein, viral evolution

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has recently emerged in Wuhan, China, and then demonstrated an unprecedented spread all over the world. To date, it is circulating on all continents, and the associated COVID-19 disease has already caused millions of deaths worldwide (1, 2).

SARS-CoV-2 is a member of the *Betacoronavirus* genus ($\beta$-CoV) in the *Coronaviridae* family (3). It is a spherical lipid envelope-containing virus. The SARS-CoV-2 genome (G RNA) is represented by a single-stranded RNA of positive polarity, ~30 kb in length. It mimics cellular mRNA by having a cap and a poly(A) tail at the 5′ and 3′ termini, respectively. Viral nonstructural proteins (nsps) are translated directly from G RNA as two overlapping polyproteins encoded by open reading frames 1a (ORF1a) and ORF1ab, which are further processed into individual nsp1 to nsp16 by the viral proteases. These nsps function in replication of the viral genome and synthesis of 8 subgenomic RNAs (SG RNAs), which encode virus-specific structural and accessory proteins (4). Functions of the structural N protein include packaging of viral G RNA during formation of infectious viral particles. It is also required for synthesis of G RNA during viral replication. M and E proteins are embedded into the viral lipid envelope and determine virion assembly and release. The structural S protein forms trimeric spikes on the surface of virions and is a major determinant of viral pathogenesis. It mediates binding to the ACE2 receptor and, later, fusion of the viral envelope and cellular membranes, resulting in release of G RNA-containing nucleocapsid into the cytoplasm. SARS-CoV-2 utilizes two entry mechanisms. In the first (early) pathway, the transmembrane serine protease(s), such as TMPRSS2, mediates cleavage of the S2′ site, which releases the fusion peptide, and this ultimately leads to fusion of viral envelope and the plasma membrane. In the cells lacking TMPRSS proteases, the S2′ processing is achieved by cathepsins, and fusion takes place in the endosomes (5, 6). Compared to SARS-CoV-1, the SARS-CoV-2-specific S protein contains an additional furin cleavage site, which mediates processing of this protein into S1/S2 subunits during particle release from infected cells (7).

Human ACE2 is a high-affinity receptor of SARS-CoV-2 and determines specificity of the infection. However, similar to numerous other viral infections, interaction with the receptor may not be the only determinant of virus attachment and entry (8). Other attachment factors at the plasma membrane may also play an important role(s). Despite the low affinity of this binding to additional factors at the membrane, the high abundance of such molecules can promote the primary attachment of viral particles before their further interactions with receptors. In many cases, these additional attachment factors have been identified as glycosaminoglycans (GAGs), most frequently heparin and heparan sulfate (HS). They are ubiquitously expressed at the plasma membrane of most vertebrate cell lines and in the extracellular matrix (9). It appears that SARS-CoV-2 is not an exception. The receptor-binding domain of its S protein was found to interact with HS (10). This binding changes the S protein receptor-binding domain (RBD) from a closed to an open conformation, thereby promoting more efficient binding to ACE2.

The ability of the viruses to interact with HS does not exclude their further evolution in developing a more efficient binding of their structural proteins to HS and other GAGs. For example, serial passaging of alphaviruses and flaviviruses in cultured cells during development of live attenuated vaccines results in accumulation of a few point mutations, which increase a number of positively charged amino acids (aa) on the glycoprotein surface (11–15). These aa substitutions increase affinity of the virus to HS and viral infectivity *in vitro* and, compared to natural isolates, strongly decrease the genome equivalent to PFU (GE/PFU) ratio. On the other hand, these mutations usually result in variants that are dramatically less pathogenic, but the mechanism of attenuation remains incompletely understood. However, in some cases, HS-binding alphaviruses exhibit high neurovirulence in mice (15, 16).

The results of our new study show that within very few passages *in vitro*, SARS-CoV-2 rapidly evolves, and variants demonstrating more efficient spread in cultured cells are selected. These variants acquire higher affinities to HS, which largely result from the insertion of a short positively charged peptide into the N-terminal domain (NTD) of the S protein. Another means of adaptation to cell culture is a partial inactivation of SARS-CoV-2-specific furin cleavage site (FCS) by substitution of a single aa. This mutation appears rapidly during virus rescue from infectious cDNA in Vero E6 cells and becomes
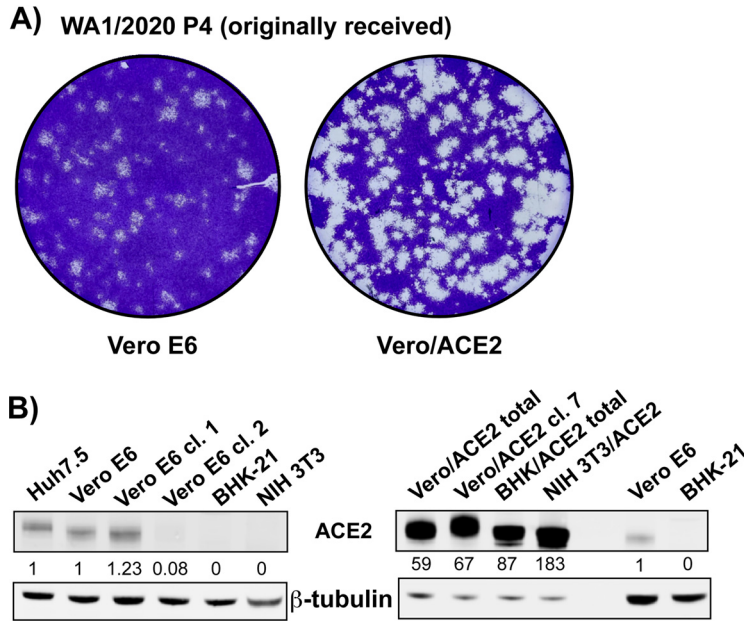
**FIG 1** The original sample of WA1/2020 is heterogeneous. (A) The original WA1/2020 (P4) isolate received was titrated in parallel on Vero E6 and Vero/ACE2 cells (see Materials and Methods for details). Cells were fixed and stained at 68 hpi. (B) Expression of ACE2 in different cell lines and cell clones. "Total" indicates that cells were not cloned after blasticidin selection.

prevalent. The positively charged amino acids in the uncleaved furin site also increase the affinity of the virus to HS. However, the decrease in cleavage efficiency does not prevent further evolution that results in accumulation of additional positively charged amino acids in the NTD. Both the insertion of the peptide and point mutation in the FCS strongly increase infectious titers and plaque size of SARS-CoV-2 on Vero cells and accelerate viral spread. They also decrease the GE/PFU ratio by several orders of magnitude.

## RESULTS

The original WA1/2020 isolate that we received from the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA) at the University of Texas Medical Branch at Galveston (UTMB) had already been passaged 4 times (P4). Similar to another previously described P4 stocks (17, 18), it produced heterogeneous plaques on Vero E6 cells (Fig. 1A), suggesting presence of multiple variants. In addition, plaques formed on Vero E6 cells were not clear because some cells in the plaque areas remained uninfected and continued to grow. Cloning of the cells showed that Vero E6 cells were heterogeneous in terms of ACE2 expression. Figure 1B presents an example where clone 1 of Vero E6 demonstrated a 15-fold higher level of ACE2 expression than clone 2 derived from the same cell line. To avoid a time-consuming selection of clones with the highest level of ACE2, we generated a cell line of Vero E6 cells which stably expressed human ACE2 (hACE2) in addition to the endogenously expressed homolog (see Materials and Methods for details). The increase in ACE2 expression in Vero/ACE2 cells (Fig. 1A) resulted in formation of noticeably larger and clearer plaques than those developed on the parental Vero E6 cells. One of the clones of Vero/ACE2 cells (cl. 7) (Fig. 1B) was used in most plaque assays. Similar ACE2-expressing stable cell lines were also generated on human Huh7.5, hamster BHK-21, and mouse NIH 3T3 cells. All these cell lines became highly susceptible to SARS-CoV-2 and, upon infection, rapidly developed a cytopathic effect (CPE). However, compared to Vero/ACE2, infection of these cells with WA1/2020 induced formation of larger syncytia with hundreds of cells involved. This resulted in relatively low virus production and, thus, limited application of the latter cell lines.

Attempts to isolate variants from the pinpoint and small plaques (Fig. 1A) developed by the WA1/2020 P4 virus on Vero/ACE2 cells were unsuccessful. As it had been previously

shown (18), plaque-purified variants immediately evolved during amplification in both Vero E6 and Vero/ACE2 cells, and harvested samples produced heterogeneous plaques. This was an indication of rapid evolution of SARS-CoV-2 during its passaging on Vero cells.

In order to better understand the mechanism of SARS-CoV-2 evolution in cultured cells, virus from the original stock was passaged 5 times on Vero E6 with gradual decrease of the inoculum volume used for infection. Infectious titers of the harvested stocks increased from $1.5 \times 10^6$ PFU/ml at starting passage 5 to $\sim 2 \times 10^8$ PFU/ml at passage 9. After this selection, the harvested viral pool became homogeneous in terms of plaque size (Fig. 2A), and, compared to the original stock, the plaques formed became dramatically larger (Fig. 1A and Fig. 2A). Comparison of particle concentrations released into the media (genome equivalents [GE] per milliliter) demonstrated that passaging did not result in a profound increase in virus release because concentrations of GE in the media at different passages were similar (Fig. 2B). This passaging drastically decreased the GE/PFU ratio, suggesting that the $\sim 100$-fold higher infectious titers resulted from a higher infectivity of the released virions. However, it remained unclear whether the more infectious variant(s) was present in the originally received heterogeneous pool of SARS-CoV-2 or there was viral evolution and selection during the above-described passaging.

**Evolution of the S protein the in WA1/2020 variant.** Next, the S protein-coding sequence of the P9 viral pool was amplified by reverse transcription-PCR (RT-PCR) and sequenced. The furin cleavage site (FCS), which was frequently deleted in prior studies (18), was found to be intact, indicating that deletion of this sequence is not the only means of increasing infectious titers *in vitro*. A major mutation identified in the S protein of passage 9 virus was an in-frame insertion of 21 nucleotides (nt) into the NTD-coding sequence (Fig. 2C). This insertion added a peptide GLTSKRN between aa 214 and 215. A second dominant mutation, H655Y, was in the CTD2 domain. It has been found in many natural isolates and became dominant in Brazilian lineage P.1 (https://www.gisaid.org/hcov19-variants).

Further analysis of viral pools, which were harvested at different passages by next-generation sequencing (NGS), demonstrated that the insertion was likely present in the originally received WA1/2020 P4 sample. It rapidly became dominant in the viral pool by passage 9 during further passaging (Fig. 2D). Thus, the insertion-containing variant was selected from the original heterogeneous pool, as it had higher infectivity for Vero E6 cells and thus was able to spread more efficiently than others. The distinguishing characteristic of the inserted peptide was the presence of 2 positively charged aa, lysine and arginine. Their presence could potentially increase affinity of the virus to GAG(s), which are abundantly present at the plasma membrane and mediate more efficient primary viral attachment to the cells before interaction with the ACE2 receptor. The H655Y mutation was already abundant in P5 virus and thus could not directly increase viral infectivity, at least while being alone. In addition, we have also detected an S247R substitution in all the samples of passaged WA1/2020 (Fig. 2D). It was less abundant than the above insertion and H655Y mutation and present in a smaller fraction of viral genomes, and its frequency was not increasing. However, the possible benefit of S247R for viral infectivity and spread cannot be completely ruled out.

**cDNA-derived SARS-CoV-2 rapidly evolves.** In another line of experiments, we analyzed evolution of SARS-CoV-2 derived from recombinant cDNA. This cDNA was synthesized according to the GenBank sequence accession no. NC_045512.2, but the ORF7a was replaced with a green fluorescent protein (GFP)-coding sequence (Fig. 3A). GFP expression was used to simplify monitoring of infection spread. SARS-CoV-2/GFP cDNA (CoV-2/GFP) (Fig. 3A) was propagated in two low-copy-number plasmids (see Materials and Methods for details). Rescuing of the virus was performed by electroporation of the *in vitro*-synthesized RNA into BHK-21 cells stably expressing the N protein of SARS-CoV-2 from the persistently replicating alphavirus replicon (see Materials and Methods for details). The transfected cells were seeded on a subconfluent monolayer of Vero/ACE2 cells, and the recombinant virus was harvested after development of cytopathic effect (CPE). The initially harvested CoV-2/GFP expressed high levels of GFP

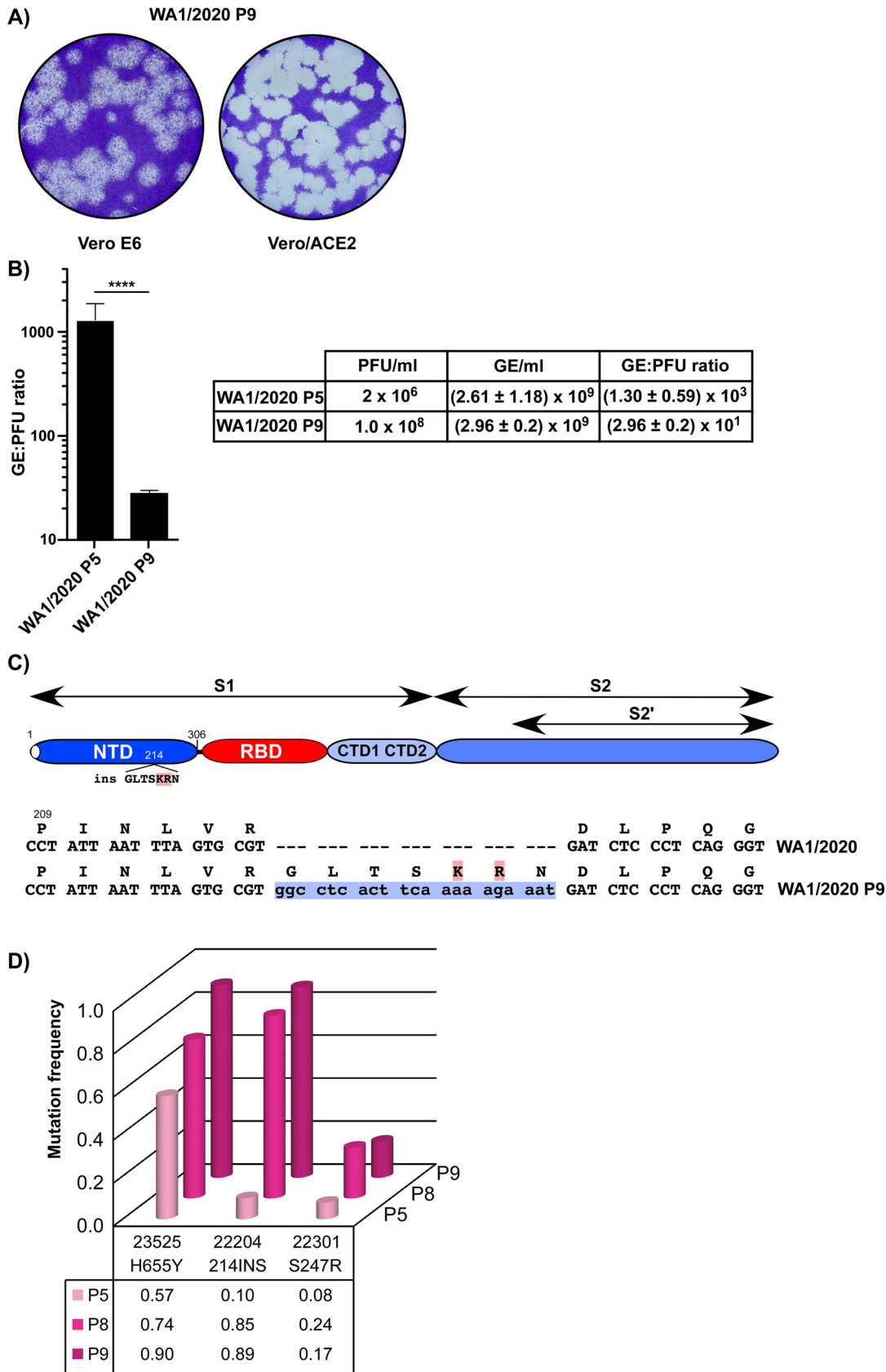**FIG 2** WA1/2020 isolate evolves to the large plaque phenotype and higher infectivity. (A) Plaques formed by WA1/2020 virus on Vero E6 and Vero/ACE2 cells. Plaques were stained by crystal violet at 68 hpi. (B) Concentrations of GE and PFU and GE/

(Fig. 3B) upon infection of Vero cells and, as in the case of WA1/2020 P4, formed heterogeneous small- and medium-size plaques (Fig. 3C). Its further passaging led to rapid evolution to the large plaque-forming phenotype (Fig. 3C). By passage 5, viral infectious titers increased by 2 orders of magnitude and approached $2 \times 10^8$ PFU/ml. As with the above-described experiments using WA1/2020, the detected evolution did not result in more efficient accumulation of released particles but, rather, increased their infectivity, which led to a strong decrease in the GE/PFU ratio (Fig. 3D).

Preliminary Sanger sequencing of the S gene in the high-passage-number viral pool identified a single nucleotide mutation resulting in S686G substitution, which was the first aa downstream of the FCS (Fig. 4A). Further NGS-based analysis demonstrated rapid accumulation of the variants, with the latter mutation in viral samples harvested at different passages (Fig. 4B). More than 60% of the variants in the P1 pool already contained this mutation. The P7 pool had the latter mutation in nearly 100% of the genomes. This suggested that the S686G substitution had a stimulatory effect on viral spread in Vero cells. There was also a gradual increase in the presence of additional mutations, mostly in the NTD of the S protein. Two of them, S68R and N74K, increased the positive charge of the surface of this domain. Another mutation in the FCS (R685H) was detectable at early passages and likely inactivated the furin-mediated cleavage. However, it was eliminated by passage 7, suggesting that the latter mutation was less advantageous than S686G for viral spread. In the course of this study, electroporation of the in vitro-synthesized SARS-CoV-2/GFP RNA was repeated multiple times, and there was always evolution of the rescued virus to a large plaque-forming phenotype within a few passages. In addition, we have also sequenced genomes in the virions pooled from three independent electroporations of the in vitro-synthesized CoV-2/GFP RNAs and found that rescued viruses also contained mutations in FCS site (S686G at 59% and R683P at 30%). Thus, as in the case of the above-described natural isolate WA1/2020, serial passaging of the cDNA-derived virus in vitro was also leading to rapid selection of variants demonstrating higher infectivity and formation of large plaques. However, this adaptation was achieved by accumulation of different mutations in the spike-coding gene. In the case of recombinant virus, the mutations in FCS became dominant by passage 2 and were likely selected from those randomly generated by the SP6 polymerase during RNA synthesis.

Next, we experimentally evaluated the effect of the S686G mutation on processing of the S protein in released viral particles. The viruses analyzed included (i) WA1/2020 P10, (ii) rescued CoV-2/GFP P2, and (iii) P6 of the latter virus. All the samples were produced in serum-free media (see Materials and Methods for details), and the released particles were harvested at 24 hours postinfection (hpi) before CPE development. Samples of the media were used directly for Western blotting. The results presented in Fig. 4C demonstrated that the insertion in the NTD and H655Y mutation identified in high-passage-number WA1/2020 did not abrogate furin-mediated processing of S protein. However, the S686G mutation, which is located in the P1′ position of the FCS, strongly affected the efficiency of cleavage, and most of the S protein in the released virions remained unprocessed.

**Mutations identified in the high-passage-number viruses strongly increase infectivity of the cDNA-derived recombinant CoV-2/GFP.** The above data suggested that serial passaging of SARS-CoV-2 in vitro leads to evolution of its spike protein that results in higher infectious titers. To experimentally support these data, we introduced the most prevalent mutations identified into cDNA of CoV-2/GFP. One of the variants encoded the above-described insertion in its NTD (CoV-2/GFP/ins), the second one had an S686G mutation downstream in P1′ position of FCS (CoV-2/GFP/G), and the third mutant (CoV-2/GFP/ins/G) contained both modifications in the S protein. Positive effects of

**FIG 2** Legend (Continued)
PFU ratios in the samples of WA1/2020 at P5 (passage 1 following receipt) and P9 (passage 5 following receipt). Significance of differences was determined by the unpaired t test with Welch's correction (****, $P < 0.0001$; $n = 3$). (C) Schematic presentation of the SARS-CoV-2 spike protein, position of the identified insertion, and its nt and aa sequences. (D) Evolution of WA1/2020 pool during virus replication in Vero E6 cells for 5 passages.
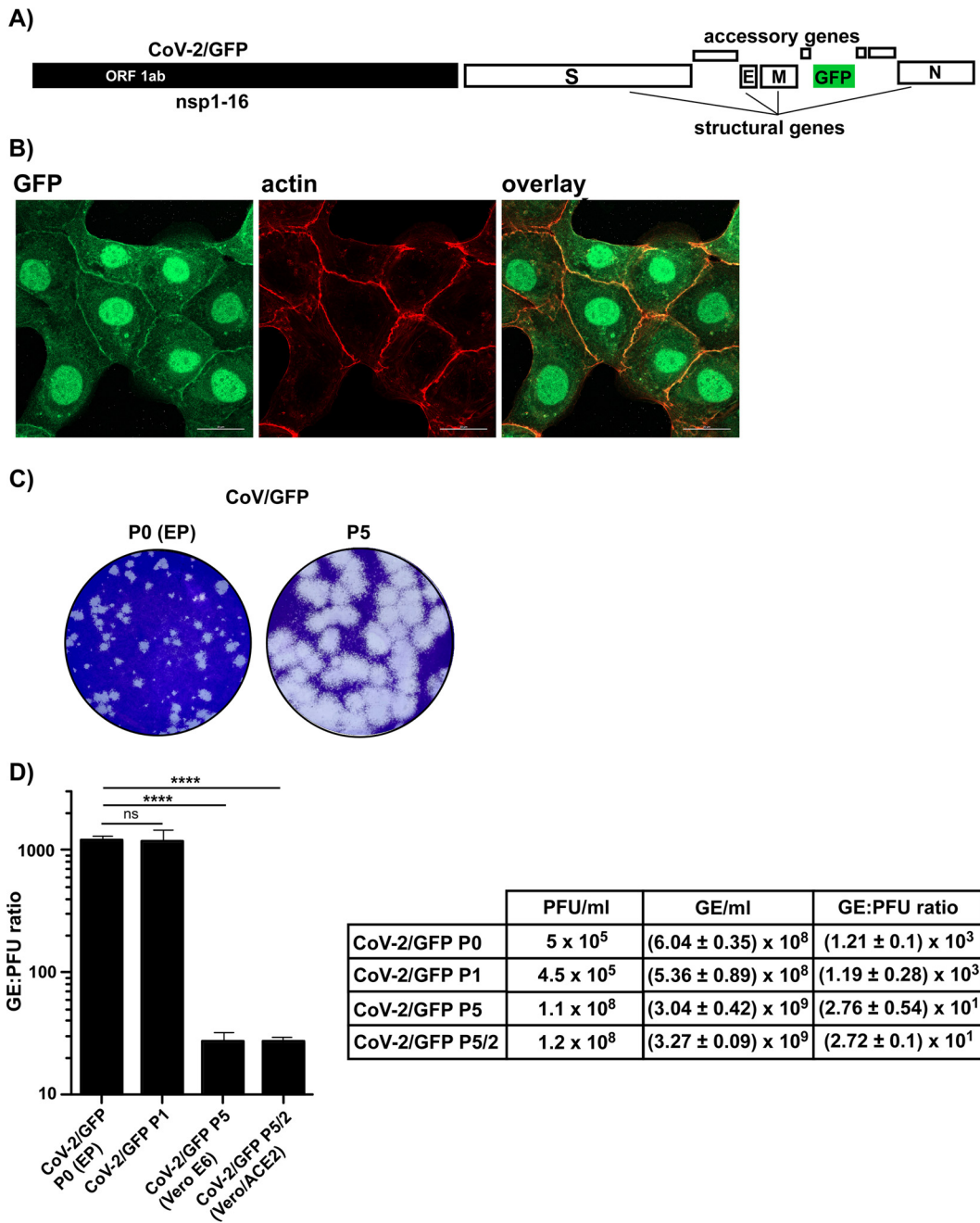
**A)**



**B)**

GFP     actin     overlay



**C)**

CoV/GFP

P0 (EP)     P5



**D)**



| | PFU/ml | GE/ml | GE:PFU ratio |
|---|---|---|---|
| CoV-2/GFP P0 | $5 \times 10^5$ | $(6.04 \pm 0.35) \times 10^8$ | $(1.21 \pm 0.1) \times 10^3$ |
| CoV-2/GFP P1 | $4.5 \times 10^5$ | $(5.36 \pm 0.89) \times 10^8$ | $(1.19 \pm 0.28) \times 10^3$ |
| CoV-2/GFP P5 | $1.1 \times 10^8$ | $(3.04 \pm 0.42) \times 10^9$ | $(2.76 \pm 0.54) \times 10^1$ |
| CoV-2/GFP P5/2 | $1.2 \times 10^8$ | $(3.27 \pm 0.09) \times 10^9$ | $(2.72 \pm 0.1) \times 10^1$ |

**FIG 3** The cDNA-derived CoV-2/GFP rapidly adapts during passaging in Vero cells to higher infectivity and large plaque phenotype. (A) Schematic presentation of the coding strategy of the recombinant CoV-2/GFP genome. (B) Vero/ACE2 cells were infected with CoV-2/GFP at an MOI of 10 PFU/cell and fixed at 8 hpi with paraformaldehyde (PFA) and stained with phalloidin, labeled with Alexa Fluor 546. Images were acquired on a Zeiss LSM 800 confocal microscope in Airyscan mode with a 63×/1.4 numerical-aperture (NA) PlanApochromat oil objective. (C) Plaques formed on Vero/ACE2 by CoV-2/GFP P0 (harvested after electroporation) and by the same virus after 5 passages on Vero/ACE2 cells. Staining was performed at the same time postinfection. (D) Concentrations of GE and PFU and GE/PFU ratios in the samples of CoV-2/GFP harvested at P0, P1, and P5 on Vero E6 cells. P5/2 indicates a sample harvested at passage 5 on Vero/ACE2 cells. Significance of differences was determined by the unpaired *t* test with Welch's correction (****, $P < 0.0001$; $n = 3$).

the mutations on viral spread were already detectable during virus rescue from the *in vitro*-synthesized RNAs. After electroporation, cells that received mutant RNAs developed profound CPE within 48 h, while the parental construct SARS-CoV-2/GFP produced CPE after 72 h. Sequencing of the S gene in the rescued viruses confirmed the presence of introduced mutations in the genomes. The recombinant viruses formed large plaques on
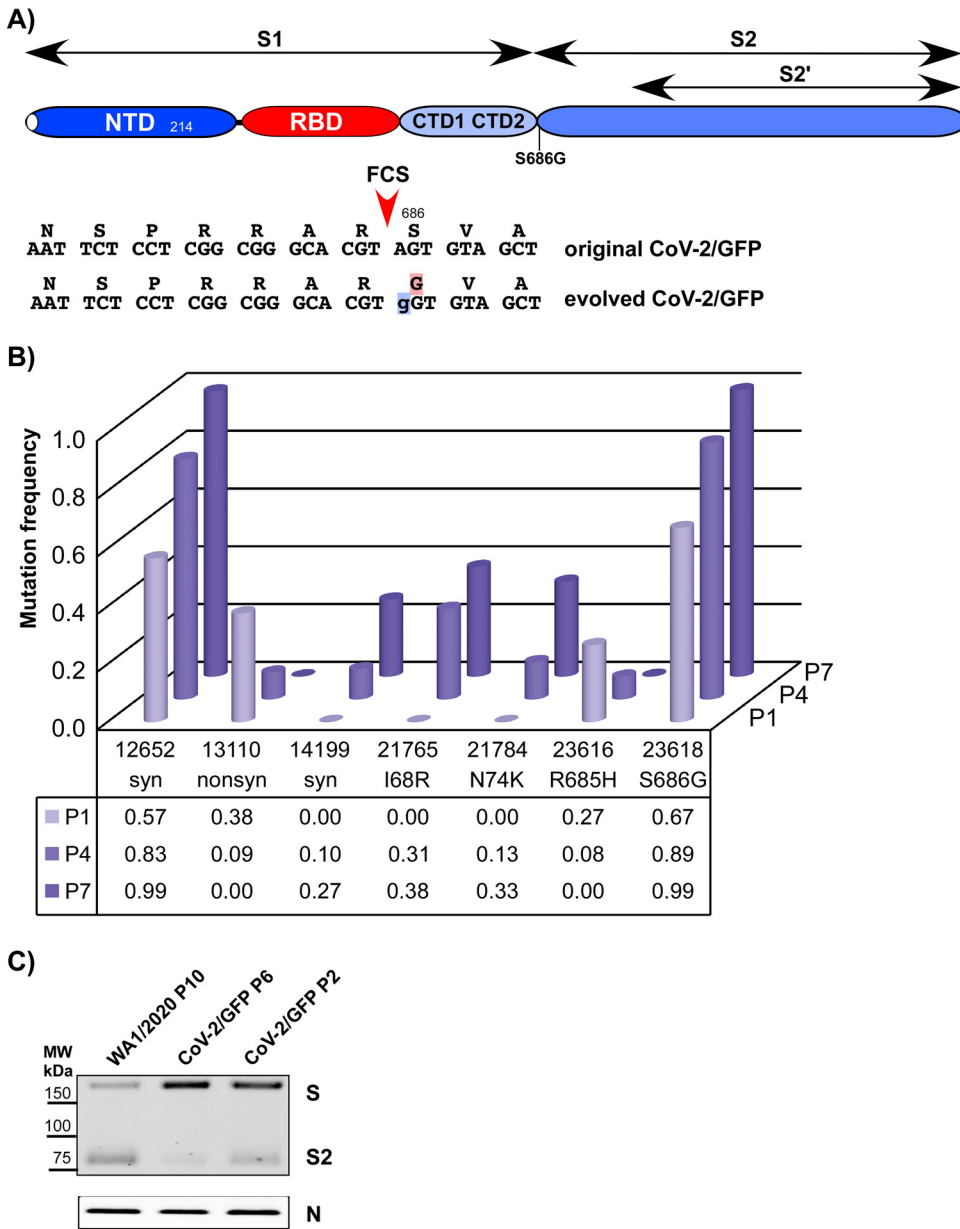
**A)**



**B)**



| | 12652 syn | 13110 nonsyn | 14199 syn | 21765 I68R | 21784 N74K | 23616 R685H | 23618 S686G |
|---|---|---|---|---|---|---|---|
| P1 | 0.57 | 0.38 | 0.00 | 0.00 | 0.00 | 0.27 | 0.67 |
| P4 | 0.83 | 0.09 | 0.10 | 0.31 | 0.13 | 0.08 | 0.89 |
| P7 | 0.99 | 0.00 | 0.27 | 0.38 | 0.33 | 0.00 | 0.99 |

**C)**



FIG 4 CoV-2/GFP acquires an S686G mutation in P1′ of FCS. (A) Schematic presentation of the SARS-CoV-2 S protein, position of the mutation, and nt and aa sequences of the gene fragment in the parental and Vero cells-adapted variant. (B) Evolution of CoV-2/GFP pool during virus replication in Vero E6 cells for 5 passages. (C) Western blot analysis of S protein in the particles released from cells infected with WA1/2020 and CoV-2/GFP harvested at different passages on Vero E6 cells (see Materials and Methods for details).

Vero/ACE2 cells (Fig. 5A), which were similar to those of the parental serially passaged viruses (Fig. 2A and Fig. 3C). Their larger size correlated with the strong decrease in GE/PFU ratios, suggesting their higher infectivity for Vero cells (Fig. 5B). Importantly, the combination of both mutations further reduced the GE/PFU ratio. Comparative analysis of viral replication rates in Vero/ACE2 cells (Fig. 5C) showed that at any times postinfection, the mutations led to higher infectious titers identical to those of the originally selected late-passage viruses. The recombinant viruses, but not CoV-2/GFP P0 and WA1/2020 P4, also caused complete CPE by 40 hpi. The double-mutant CoV-2/GFP/ins/G replicated more efficiently than the single mutants (Fig. 5C). Variants containing the S686G mutation (the late-passage CoV-2/GFP, CoV-2/GFP/G, and CoV-2/GFP/ins/G) demonstrated less efficient processing of the S protein (Fig. 5D). Taken together, the data
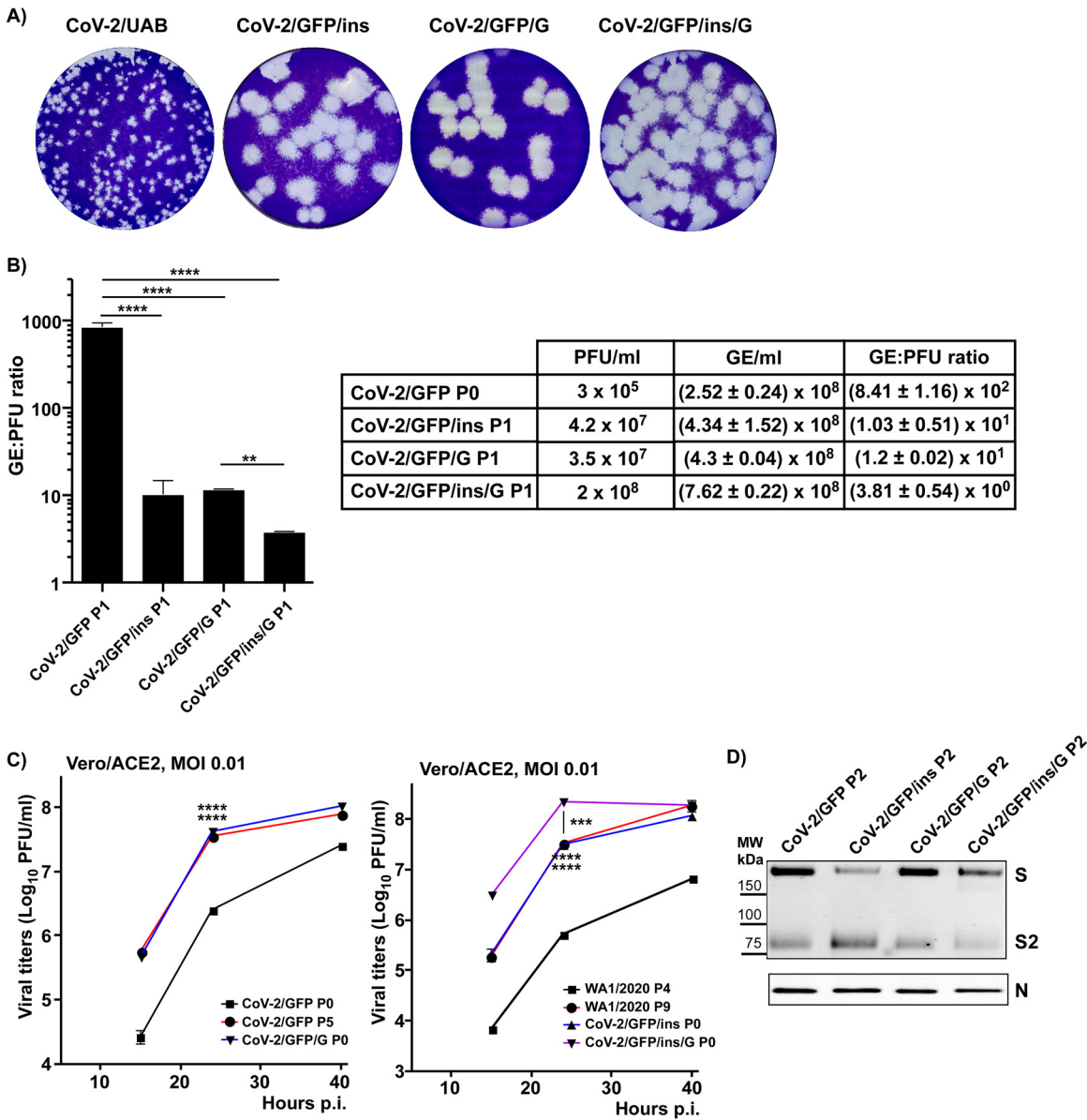
FIG 5 Mutations in the S protein-coding sequence confirm the large plaque phenotype and higher infectivities of the recombinant viruses to Vero cells. (A) Plaques formed by the early-passage (P1) isolate of CoV-2/UAB and the recombinant CoV-2/GFP-based variants, containing either the insertion in the NTD or S686G mutation or both modifications. In all cases, P1 viruses were assessed in the standard plaque assay, and plaques were stained at day 3 postinfection. (B) Concentrations of GE and PFU and GE/PFU ratios in the samples of CoV-2/GFP encoding the S protein with the indicated modifications. (C) Vero/ACE2 cells were infected with the indicated viruses at an MOI of 0.01 PFU/cell. After 1 h of infection, cells were washed 3 times with PBS and further incubated in the corresponding complete media. At the indicated time points, media were replaced, and viral titers were determined by plaque assay on Vero/ACE2 cells. For panels B and C, significance of differences was determined by the unpaired $t$ test with Welch's correction (**, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; $n = 3$). (D) Western blot analysis of S protein in the particles released from Vero cells infected with CoV-2/GFP, CoV-2/GFP/ins, CoV-2/GFP/G, and CoV-2/GFP/ins/G (see Materials and Methods for details).

indicated independent functions of the mutations introduced into the S protein in viral replication to higher infectious titers, at least in Vero cells. However, it remained unclear whether these changes in the S protein function via the same or different mechanisms.

**Mutations in the S protein of selected mutants enhance SARS-CoV-2 binding to heparin.** Next, we evaluated the interaction of different viral variants with heparin. Samples of all of the viruses indicated in Fig. 6 were prepared in serum-free media. This set included the UAB clinical isolate CoV-2/UAB P2, WA1/2020 P5, early-passage CoV-2/GFP P2, late-passage WA1/2020 P10, and CoV-2/GFP P7. Viral samples were diluted to 0.1 M concentration of NaCl, loaded to the heparin Sepharose column, and eluted using a step gradient
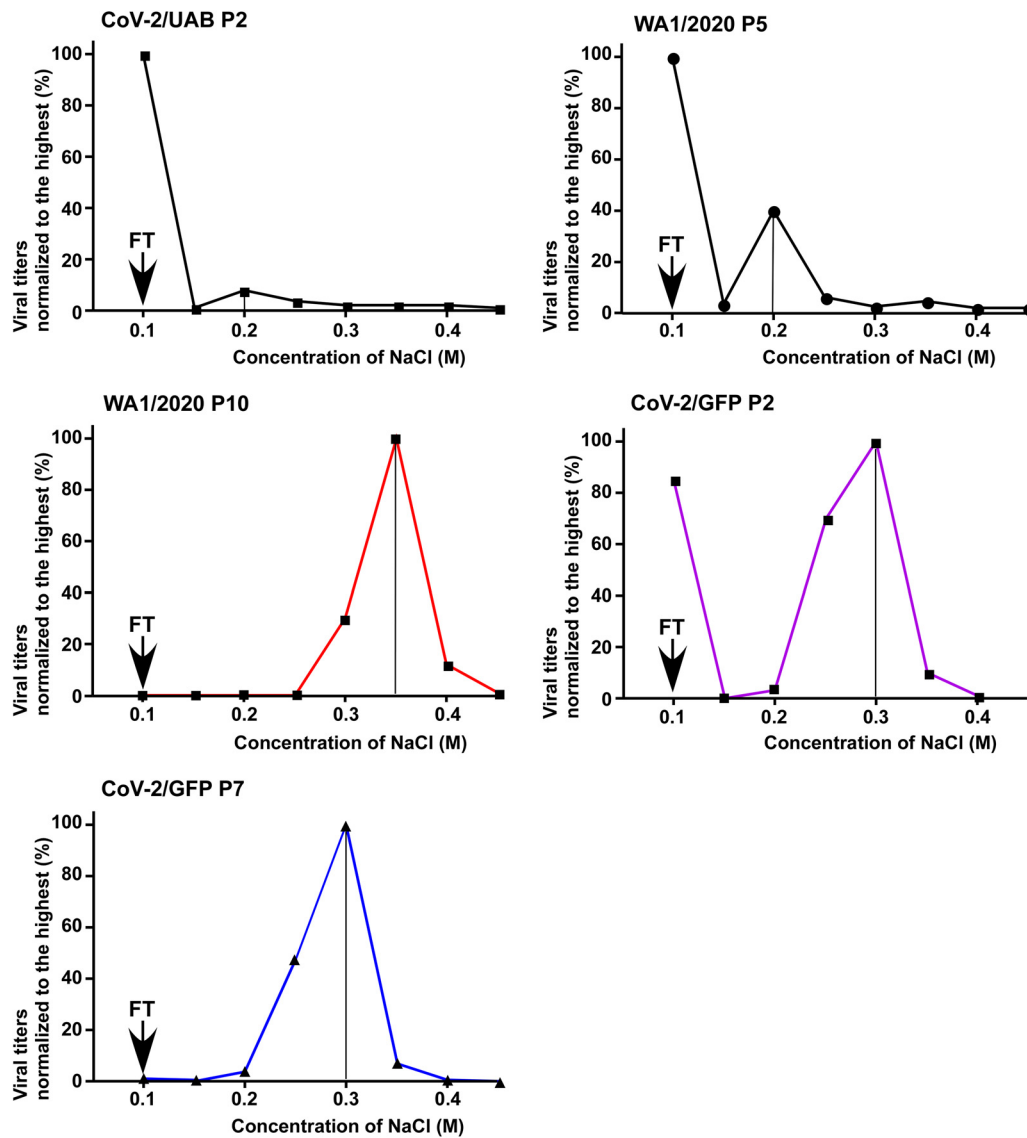
**FIG 6** After passaging on Vero cells, SARS-CoV-2 acquires higher affinity to heparin Sepharose. Samples of the indicated viruses were prepared in the serum-free media, diluted to 0.1 M NaCl, and loaded to a column of heparin Sepharose. After washing with $0.66\times$ PBS, which contains 0.1 M NaCl, viruses were eluted with PBS containing increasing concentrations of NaCl. Viral titers in each fraction, including FT, were determined by plaque assay on Vero/ACE2 cells. Titers were normalized to those in the fraction with the highest viral concentration. The experiments were repeated twice with reproducible results. The data from one of the experiments are presented.

of NaCl (see Materials and Methods for details). The results presented in Fig. 6 demonstrate that the low-passage-number samples of natural isolates WA1/2020 P5 and CoV-2/UAB had low affinities to heparin and mostly remained in the flowthrough (FT) fraction containing 0.1 M NaCl. Smaller peaks of infectious viruses were also eluted by 0.2 M NaCl. CoV-2/GFP P2 demonstrated noticeable heterogeneity in terms of binding to heparin. Some of the virus was still present in the FT fraction, and the rest was eluted with 0.25 to 0.3 M NaCl. Passaged viruses WA1/2020 P10 and CoV-2/GFP P7 bound to heparin Sepharose more efficiently, and less than 1% of the infectious viruses were found in the FT fraction. WA1/2020 P10 and CoV-2/GFP P7 were eluted by 0.35 M and 0.3 M NaCl, respectively. Thus, the mutations accumulated in the S protein of the high-passage-number viruses, which became more infectious for Vero cells, also increased their affinities to heparin.

The pools of the high-passage-number viruses still contained variants with a variety of mutations in the S protein-coding sequence. Therefore, it was difficult to make definitive

conclusions about the effects of GLTSKRN insertion and S686G mutation on virions' affinity to heparin. Therefore, we also evaluated binding of the above-described designed recombinant viruses to heparin Sepharose. Based on sequencing data, P1 virions had no additional aa substitutions in the S proteins. Their samples were also prepared in serum-free media. Binding and elution patterns of CoV-2/GFP/ins P1 and CoV-2/GFP/G P1 (Fig. 7) correlated with those of the high-passage-number WA1/2020 and CoV-2/GFP (Fig. 6), respectively. In the FT fraction, both recombinant viruses were present at very low levels. The peak of CoV-2/GFP/ins variant elution was at 0.35 M NaCl, and CoV-2/GFP/G was eluted by 0.3 M NaCl. Thus, both the insertion and S686G mutation increased affinities of the viruses to heparin Sepharose. The S686G mutation did not increase the positive charge of the S protein surface, but it likely transformed the unprocessed furin cleavage site into the heparin-binding site. Interestingly, the double-mutant CoV-2/GFP/ins/G, which contained both the insertion in the NTD and S686G mutation, was also eluted by 0.35 M NaCl. Thus, the combination of both modifications in the spikes did not increase viral affinity above that of the high-passage-number natural isolates. This was suggestive that, although binding to heparin was likely determined by two separate sites, the higher affinity was mostly determined by the insertion in the NTD and played the dominant role.

Taken together, the results of these experiments demonstrated that passaging of SARS-CoV-2 leads to its rapid evolution toward a more efficient binding to heparin. These adaptive mutations include insertion of the short positively charged peptide after aa 214 in the S protein and likely transformation of the furin cleavage site into a heparin-binding aa sequence. However, additional contributions of other mutations in the NTD of the spike protein, such as S68R or N74K, which may additionally increase its affinity to HS, are also possible.

To confirm biological significance of the above data, we evaluated inhibition of infectivity of different viral variants by heparin (see Materials and Methods for details). The infectivity of the low-passage-number CoV-2/UAB P2 was almost unaffected by heparin at any concentration used (Fig. 8). In contrast, the high-passage-number WA1/2020 P9 and CoV-2/GFP P5 viruses and those having insertion in the NTD of S protein (CoV-2/GFP/ins and CoV-2/GFP/ins/G) demonstrated high sensitivities to the presence of heparin. Their infectivities were reduced by~10- to 20-fold after incubation, even at the lowest concentration of heparin. The recombinant virus with S686G substitution demonstrated intermediate sensitivity. Thus, its partially unprocessed FCS was able to function as a heparin-binding site, but other mutations, which were detected in CoV-2/GFP P5 variant, were likely required.

We have also performed NGS for P2 of all mutant viruses. We found that CoV-2/GFP/G has accumulated several low-frequency mutations in the NTD (N74K, 6%; S274R, 8%; and T259K, 6%). They are predicted to increase the positive charge of the NTD surface. CoV-2/GFP/ins P2 sample did not acquire additional mutations in NTD but had S686G mutation in the FSC in 4% of the genomes. The double-mutant CoV-2/GFP/ins/G did not have any additional mutations. Thus, both introduction of additional positive amino acids into NTD and mutation of FCS are essential for making viruses stable during tissue culture propagation.

Variants having the S686G mutation exhibited one more interesting characteristic. They lost the ability to form syncytia on Vero/ACE2 cells (Fig. 9). In our experimental conditions, CoV-2/UAB, WA1/2020, and CoV-2/GFP/ins formed very large syncytia. The early-passage CoV-2/GFP P2, which had a mixed population of viral variants, was still capable of forming at least small syncytia, but by passage 7, it completely lost this ability (Fig. 9). Both CoV-2/GFP/G and double-mutant CoV-2/GFP/ins/G were unable to produce syncytia at all. These data correlated with the previous study on mouse hepatitis virus (MHV), which demonstrated that processing of S protein plays a critical role in cell-cell but not virus-cell fusion (19).

## DISCUSSION

In this report, we demonstrate that SARS-CoV-2 rapidly adapts to replication in Vero E6 cells, which are most commonly used for virus propagation. This adaptation results in higher infectious titers *in vitro*, which are determined by higher infectivity of the viral particles rather
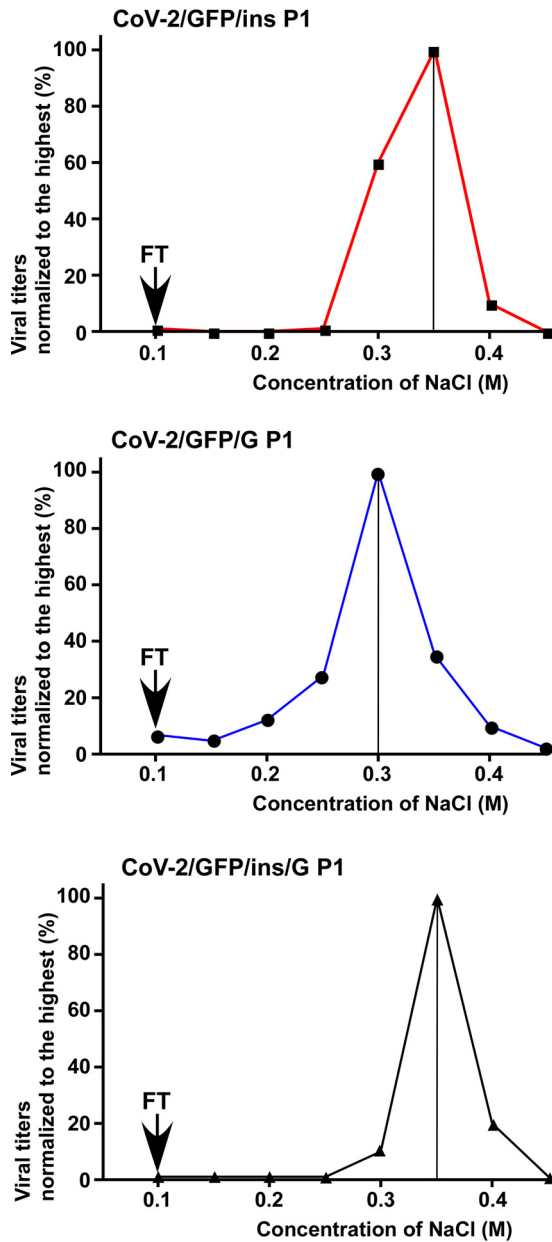
**FIG 7** Recombinant viruses with the indicated mutations in the S protein exhibit high affinity to heparin Sepharose. Viral samples were prepared in serum-free media and analyzed as described in the legend to Fig. 6. Viral titers in each fraction, including FT, were determined by plaque assay on Vero/ACE2 cells. Titers were normalized to those in the fraction with the highest concentration of infectious virus. The experiments were repeated twice with reproducible results. Data from one of the representative experiments are presented.

than by more efficient particle release. The evolved variants form dramatically larger plaques on Vero cells and demonstrate more rapid infection spread. Similar evolution of SARS-CoV-2 to the large plaque-forming phenotype has been detected in previous studies, suggesting that this is a common effect and likely independent of experimental conditions (17).

Viral evolution developed by two mechanisms. The first means of adaptation was determined by the insertion of a 7-aa-long peptide into the S protein of the WA1/2020 isolate (Fig. 2). The variant containing insertion was already present in the provided heterogeneous viral sample and rapidly became dominant during further passaging. This GLTSKRN insertion was located in the NTD (Fig. 2, 9, and 10), suggesting an important role of the latter domain in primary attachment of the virus to cells prior to interaction with
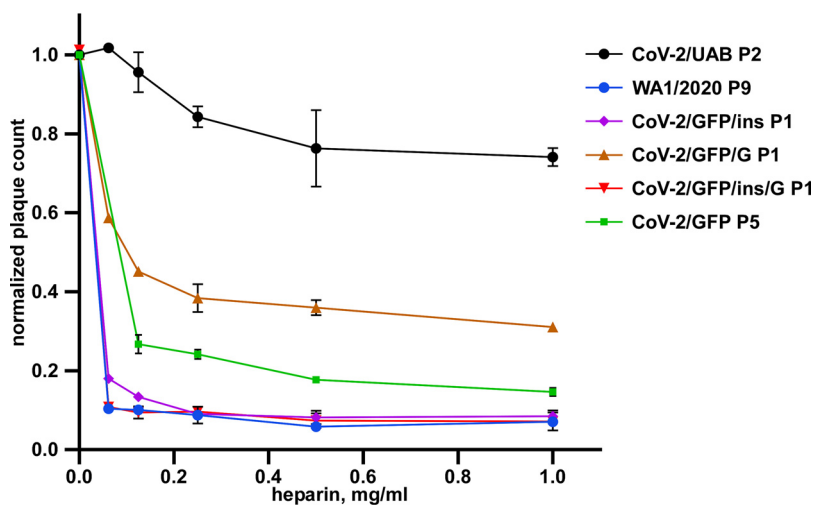
**FIG 8** Heparin efficiently inhibits infectivity of high-passage-number but not low-passage-number viruses. The indicated viruses were incubated with heparin at the indicated concentrations for 1 h and then used for infection of Vero/ACE2 cells (see Materials and Methods for details). The data were normalized to those generated on the samples incubated without heparin. The experiment was repeated 3 times; means and SDs are indicated. Significances of differences between CoV-2/UAB P2 and other variants at each concentration of heparin were determined by one-way analysis of variance (ANOVA) with Dunnett's test. For all concentrations, the differences were highly significant with $P < 0.0001$.

the ACE2 receptor. The inserted peptide contained two positively charged aa and made both the evolved natural variant and its cDNA-derived analog more infectious *in vitro*. It strongly increased affinity of both viruses to heparin Sepharose (Fig. 6 and 7) and decreased their GE/PFU ratios by 2 orders of magnitude compared to those of the low-passage-number natural isolates. Consequently, it had also positive effects on the plaque size and rates of infection spread. Infectivity of the viruses containing the above insertion was also strongly inhibited by heparin (Fig. 8). The identified insertion increased positive charge on the surface of the NTD domain (Fig. 10 and 11), and this provided a plausible explanation for the higher affinity of the mutant virus to heparin. A small fraction of passaged WA1/2020 contained another mutation, S247R, which also contributes to an increase in the positive charge of the NTD surface. Since presence of the latter mutation did not increase during passaging, its contribution to infection spread was likely lower than that of the above-described insertion. In the previous studies, the S protein of MHV also acquired the insertion of a 7-aa-long peptide after passaging on nonmurine cells (20). This R/K-enriched insertion, in combination with additional, uncharacterized point mutations, made the latter virus use HS as a new receptor (21–23). Interestingly, mutation in the FCS site that abolished its cleavage additionally enhanced heparin-dependent, receptor-independent entry of mutant MHV. However, in our case, infection of WA1/2020 P10 remained fully dependent on its natural ACE2 receptor, and binding to heparin facilitated primary virus binding to the cells.

The second mechanism of SARS-CoV-2 adaptation was based on the appearance of the S686G substitution in P1′ position of the S1/S2 cleavage site, FCS. It was rapidly selected in the S protein of recombinant CoV-2/GFP. The previously published data showed that the peptide representing FCS is often entirely or partially deleted during virus passaging (17, 18, 24), and the same S686G mutation was also detected after passaging of other SARS-CoV-2 isolates (17, 18). The identified mutation strongly affected processing of the S protein (Fig. 4C and Fig. 5D) and formation of the syncytia (Fig. 9) during viral replication. It also increased infectivity of harvested viruses, their plaque sizes, and infectious titers (Fig. 3C and Fig. 5C). In prior studies, the FCS could be inactivated by either its deletion or accumulation of point mutations eliminating the arginines essential for furin cleavage (18). Viral mutants with deleted FCS and R682L, S686G, and H655L mutations have been detected in the stock of the WA1/2020 isolate from the CDC (25). The effect of the FCS deletion was experimentally evaluated in the
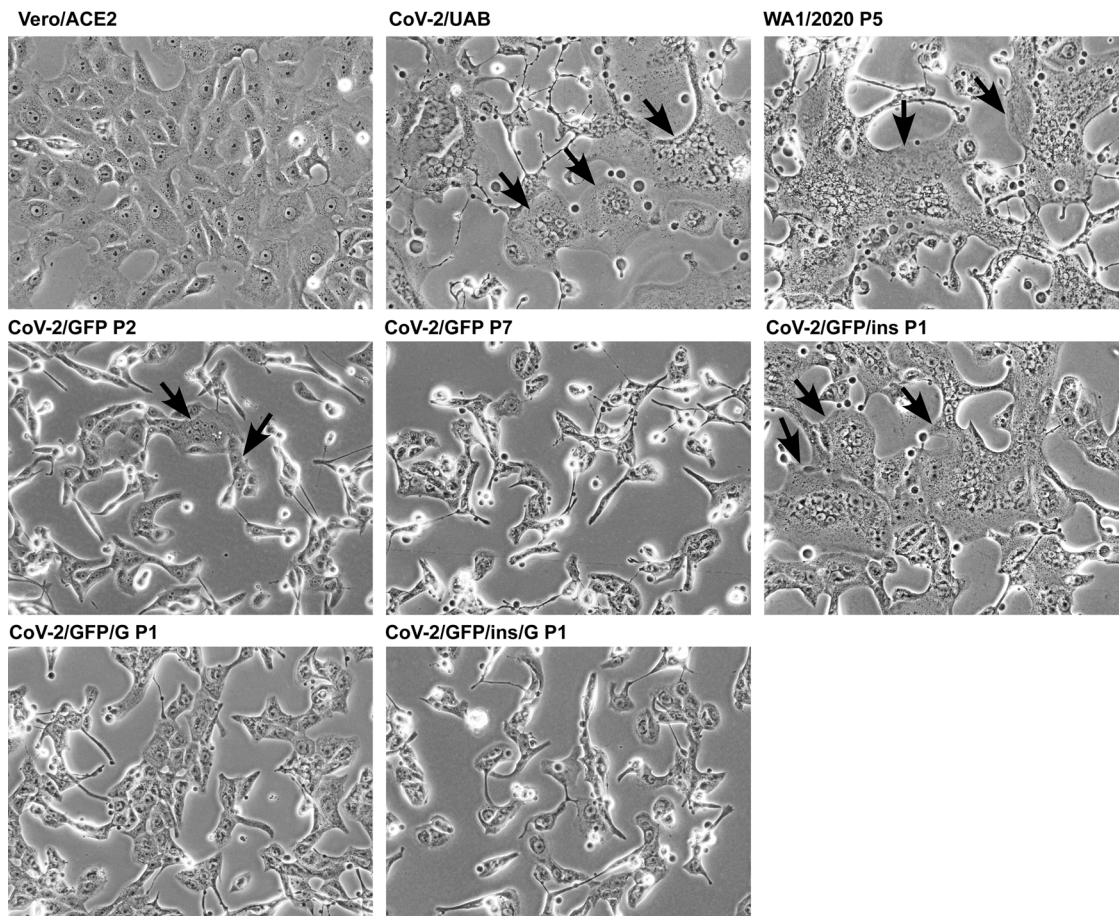
**Vero/ACE2**　　　　　　**CoV-2/UAB**　　　　　　**WA1/2020 P5**

**CoV-2/GFP P2**　　　　　**CoV-2/GFP P7**　　　　　**CoV-2/GFP/ins P1**

**CoV-2/GFP/G P1**　　　**CoV-2/GFP/ins/G P1**



**FIG 9** The S686G mutation has negative effect on the ability of SARS-CoV-2 virus to form syncytia. Vero/ACE2 cells in 6-well Costar plates were infected with the indicated variants at an MOI of 2 PFU/cell and incubated in serum-free media for 16 h at 37°C. Then they were fixed by 4% PFA, and images were acquired on EVOS microscope. Syncytia are indicated by black arrows.

context of pseudoviruses or recombinant SARS-CoV-2 (26, 27). It resulted in higher infectious titers of the viruses harvested from Vero cells, but the detected increase was small. It has also been demonstrated that the deletion of FCS affected viral infectivity in human cells, including Calu-3, Caco2, and primary lung epithelia cells.

In our study, the selected recombinant viruses with the S686G mutation replicated to 2 orders of magnitude higher infectious titers in Vero cells than did the natural isolates and previously designed deletion mutants. These results suggest that, in contrast to FCS deletion, inhibition of the S1/S2 processing by S686G is likely not the only advantage of the new variants. Viral spikes with S686G substitutions retained the entire polybasic motif, which was originally required for cleavage by furin protease (Fig. 4A), and such viruses attained high affinity to heparin Sepharose (Fig. 6 and 7) and higher infectivity and more efficient spread on Vero cells and made the virus more sensitive to the inhibitory effect of heparin. Similar evolution of the FCS into the heparin-binding motif has previously been described for some other coronaviruses (28). The resultant polybasic motif of the S protein-specific FCS was functioning in HS binding and, together with additional R/K substitutions, made MHV capable of utilizing HS as an alternative receptor. Interestingly, mutations in FCS were capable of increasing CoV pathogenicity. In the feline enteric coronavirus (FECV), FCS mutations abolishing S1/S2 cleavage strongly correlated with transformation of the virus into another more pathogenic biotype, feline infectious peritonitis virus (FIPV) (29, 30). However, an additional mutation in the N-terminal fragment of S2 was also required (21, 31). In our study, at the early passages, we were able to detect another CoV-2/GFP variant with mutated positively charged aa in the furin cleavage site (R685H mutation). However,
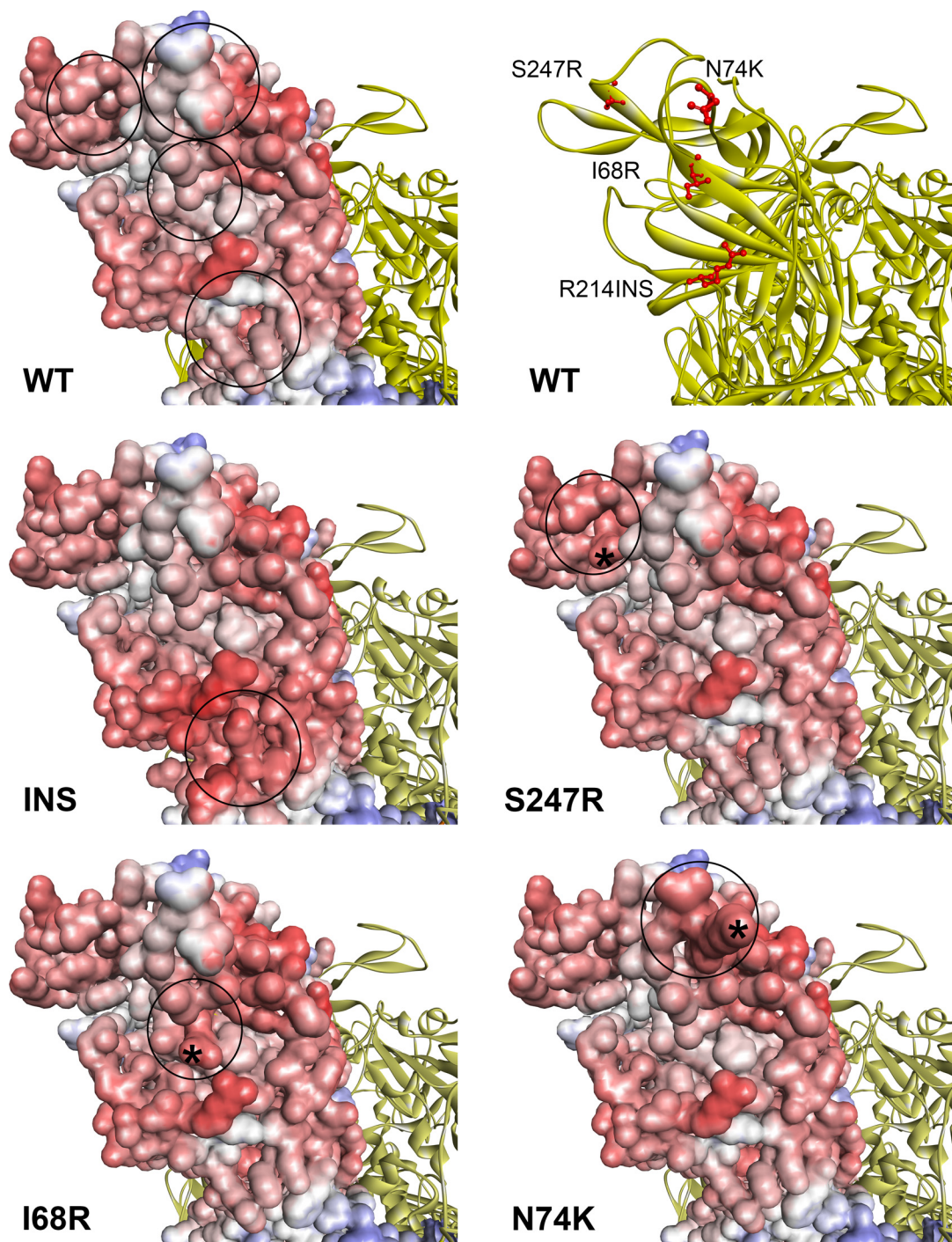
**FIG 10** Mutations acquired by the NTD during virus passaging in Vero cells increase the positive charge of the protein surface. Electrostatic surface rendering is presented. Blue and red indicate electronegative and electropositive surfaces, respectively. Positions of the indicated aa substitutions and peptide insertion GLTSKRN between amino acid residues 214 and 215 are shown by black circles and asterisks. Images were generated with BIOVIA Discovery Studio Visualizer.

it was likely less competitive than the S686G mutant and was eliminated from the viral pool within a few subsequent passages. Thus, the S686G mutant apparently produced higher infectious titers than other variants with mutated arginine-rich motifs. The latter RRAR remains in the S1 subunit of the original isolates after furin-mediated processing. However, in the postcleavage conformation, it likely interacts with heparin less efficiency, if at all, than that in the uncleaved S protein of the selected mutant. Interestingly,
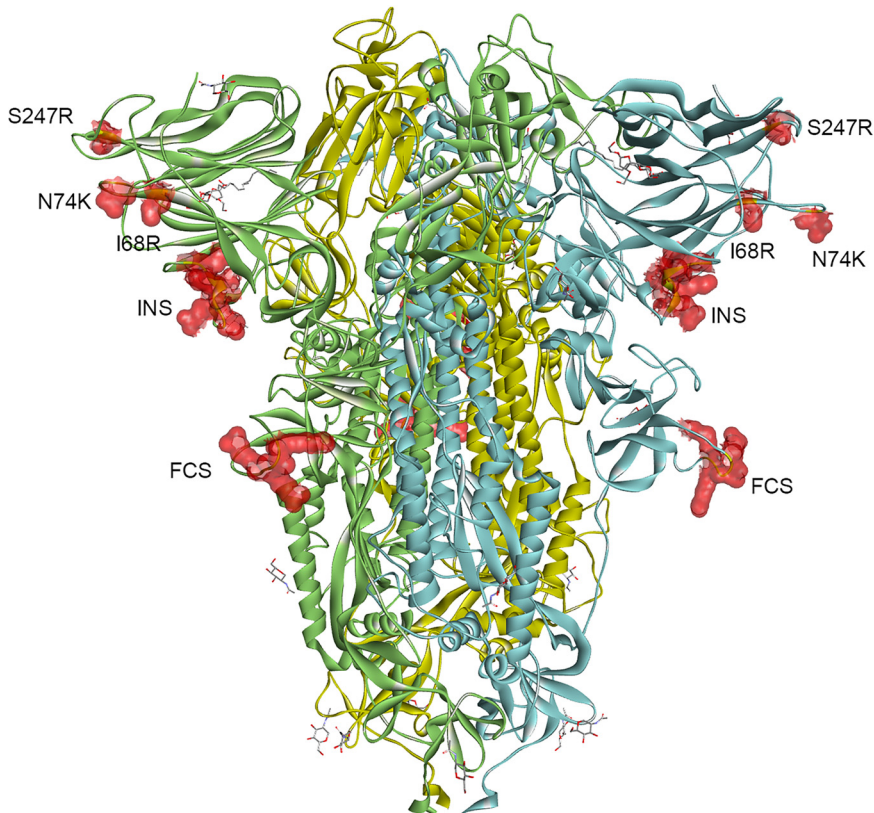
**FIG 11** Three-dimensional model of the trimeric SARS-CoV-2 spike ectodomain is shown. The individual spike subunits are shown in green, yellow, and blue. Locations of the single mutations, the insertion in the NTD, and uncleaved FCS are highlighted with red spheres.

downregulation of the S protein processing had a negative effect on the ability of the recombinant viruses to induce syncytium formation at least in Vero cells (Fig. 9), but it remains unclear whether the mutants became less efficient in forming syncytia *in vivo*.

Our experimental data suggest that the virus with unprocessed FCS remains less efficient in HS binding than the variant with the NTD insertion. First, viruses with S686G are eluted from heparin Sepharose by lower concentrations of NaCl (Fig. 6 and 7). Second, in contrast to the insertion mutant, the originally selected CoV-2/GFP variant with the S686G mutation continued to evolve and attain additional mutations in the NTD, such as I68R and N74K (Fig. 4, 10, and 11). Their accumulation was relatively slow but clearly detectable. Third, the recombinant CoV-2/GFP/G virus demonstrated lower sensitivity to the inhibitory effect of heparin on infectivity. Taken together, these data suggest that I68R and N74K additionally contribute to HS binding. Analysis of sequences submitted to GISAID's EpiCoV database revealed that accumulation of mutations that increase the positive charge of NTD upon passaging in Vero cells is quite common: N74K was detected in EPI_ISL_1039208, EPI_ISL_1190402, EPI_ISL_1718321, EPI_ISL_1707039, EPI_ISL_1785073, and EPI_ISL_1785076, and I68R/K was found in EPI_ISL_493139 and EPI_ISL_2226226. The WA1/2020 from the CDC was reported to contain 9 low-frequency R/K substitutions in NTD, including a 4-aa-long insertion after D215 (25). The indicated I68R and N74K mutations, which additionally increase the positive charge of the NTD surface (Fig. 10), are also located relatively close to the above-described S247R (Fig. 11). Currently, there is not enough information to speculate whether the increase of positive charge on the NTD surface attenuates the virus, makes it more pathogenic, or has no effect on pathogenicity. Importantly, since R/K mutations in NTD are highly beneficial for virus propagation *in vitro*, viruses used for animal studies should be thoroughly analyzed by NGS.

Combining the positively charged insertion and S686G mutation in the same S protein did not detectably increase the affinity of infectious viral particles to heparin Sepharose

above the level determined for CoV-2/GFP/ins (Fig. 7). However, the recombinant double mutant demonstrated significantly higher infection rates than did the single-site mutants (Fig. 5). CoV-2/GFP/ins/G was reaching the peak titers above $10^8$ PFU/ml by 24 hpi, even at low multiplicity of infection (MOI), such as 0.01 PFU/cell. This was an indication that during viral infection, the introduced mutations, GLTSKRN insertion and S686G substitution, were able to function in either synergistic or additive modes. Importantly, the recombinant single mutants continued to accumulate the second-site mutations, either R/K substitutions in NTD for CoV-2/GFP/G or the S686G mutation for CoV-2/GFP/ins. The double mutant containing both mutations CoV-2/GFP/ins/G was stable and did not acquire additional changes. The structural model of the S protein supports the possibility of a combined effect. The above insertion and furin cleavage site are closely located on the surface of the S protein (Fig. 10 and 11) and may have a stimulatory effect on primary binding of the virus to GAGs at the plasma membrane. An important characteristic of the double mutant is that its further evolution in cultured cells appears to be unlikely. The latter virus demonstrated the lowest GE/PFU ratio, and its further decrease, at least in Vero cells, appears to be an improbable event.

Similar rapid evolution to more efficient spread *in vitro* was previously described for alphaviruses. As with SARS-CoV-2 in this study, chikungunya, Venezuelan equine encephalitis, Ross River, and Sindbis viruses acquired 1 or 2 additional basic aa in the E2 glycoprotein, which had strong positive effects on virus interaction with HS (12–14, 32, 33). The mutations increased plaque sizes, stimulated spread of infection, and dramatically decreased the GE/PFU ratio of the evolved alphaviruses. These mutant viruses were usually less pathogenic in mice. Therefore, it is tempting to speculate that, similar to other RNA-positive ($RNA^+$) viruses that adapted to binding to HS, the designed single mutants and the double mutant, in particular, became attenuated *in vivo*, but this hypothesis needs experimental support. Of note, the development of the recombinant live attenuated vaccine for SARS-CoV-2 will also require its large-scale production and passaging in cell culture, most likely in Vero cells. Therefore, viral evolution to a more stable HS-binding phenotype through the acquisition of mutations in the NTD is likely an unavoidable event and needs to be considered. The original isolates form very small plaques and have a very high GE/PFU ratio. Their passaging results in rapid evolution of viral pool and selection of variants with higher infectivities *in vitro*. However, it is unlikely that they fully reproduce SARS-CoV-2 infection and pathogenesis *in vivo*.

Interestingly, previously published data, which were generated on the receptor-binding domain (RBD) and the isolated S protein, suggested that the RBD is capable of efficient interaction with heparin Sepharose and HS (10). This study has also demonstrated inhibition of SARS-CoV-2 infection by heparin. However, the used USA-WA1/2020 isolate likely already had R/K mutations in NTD. In our experiments, the early-passage natural isolates demonstrated relatively inefficient binding to heparin Sepharose. The main fraction of infectious virus remained in the FT fraction (0.1 M NaCl), and some virus was also eluted as a small peak at NaCl concentration of 0.2 M (Fig. 6). CoV-2/UAB P2 was also resistant to the inhibitory effect of heparin on viral infectivity (Fig. 8). However, the insertion- and/or S686G mutation-containing variants, which became adapted to cell culture, were eluted by NaCl at 0.35 and 0.3 concentrations. They had higher affinities to heparin, which strongly correlated with viral infectivities and spread in cultured cells.

Importantly, we observed that rapid adaptation of the recombinant, cDNA-derived virus, and the S686G mutation became dominant by passage 2 because it increased viral infectivity by 100-fold. The SP6 RNA polymerase used in our study and the T7 polymerase used by other groups have mutation rates of $1 \times 10^{-4}$/nt to $2 \times 10^{-4}$/nt (34). Thus, the *in vitro*-synthesized viral genomes already contain a wide range of mutations, which represent a ground for further selection of the most efficiently spreading variants. Unfortunately, no NGS sequencing is being routinely performed on the rescued recombinant viruses, and thus, the results of the studies with recombinant viruses should consider a possibility that the harvested pools rapidly accumulate mutations similar to ones detected in our study.

In conclusion, previously published data and the results of this study suggest that both natural isolates of SARS-CoV-2 and the cDNA-derived recombinant variants rapidly adapt to cell culture. They evolve to higher infectivity, which results in lower GE/PFU ratios and more efficient infection spread by either acquiring mutations in the NTD, which suggests an important role of this domain in viral binding to the cells, or in the FCS. These mutations increase virus affinity to HS during its primary attachment to cells before interaction with the ACE2 receptor. The mutations leading to more efficient spread include (i) substitutions of single aa in the NTD by those positively charged, (ii) insertions of short peptides that contain basic aa, and (iii) an S686G substitution that likely transforms FCS into the HS-interacting peptide. Such viral variants rapidly become dominant, and their further evolution and adaptations are unlikely. These adapted variants appear to be more convenient for CPE-based screening of antiviral drugs. As with the HS-binding mutants of other RNA$^+$ viruses, the evolved SARS-CoV-2 may also be attenuated *in vivo*, particularly the double mutant that demonstrates the most adapted phenotype and alterations in syncytium formation. Thus, they may also be used as a basis for development of stable live attenuated vaccines for COVID-19.

## MATERIALS AND METHODS

**Cell cultures.** The BHK-21 cells were kindly provided by Paul Olivo (Washington University, St. Louis, MO). The NIH 3T3 and Vero E6 cells were obtained from the American Type Culture Collection (Manassas, VA). Huh7.5 cells were kindly provided by Charles Rice (Rockefeller University, New York, NY). BHK-21, Vero E6, NIH 3T3 cells, and their derivatives were maintained in alpha minimum essential medium supplemented with 10% fetal bovine serum (FBS) and vitamins. Huh7.5 cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% FBS.

**SARS-CoV-2 isolates.** The stock of SARS-CoV-2 strain 2019-nCoV/USA_WA1/2020 was derived from the first patient diagnosed in the United States. This virus isolate was originally provided by Natalie Thornburg from the Centers for Disease Control and Prevention in Atlanta, Georgia (35), and amplified on Vero E6 cells at the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA) at the University of Texas Medical Branch at Galveston (UTMB). Passage 4 was received from WRCEVA and further passaged in this study. In the following sections, this isolate is referred to as WA1/2020.

The second isolate (CoV-2/UAB) was kindly provided by William Britt (The University of Alabama at Birmingham, USA), and passage 1 was used in this study. The S protein of this isolate contained a single-mutation D614G, which is relatively common for SARS-CoV-2 isolates (36).

**Plasmid constructs.** The cDNA of SARS-CoV-2 (GenBank accession no. NC_045512.2) genome was assembled in two fragments (5′ and 3′ fragments), which were cloned into low-copy-number plasmids pACNR1180 (37) and Proteus1 (38) and propagated in *Escherichia coli* WM1100. Fragments were assembled from gene blocks synthesized by Integrated DNA Technologies (IDT). The 7a gene in the 3′ fragment was replaced by the GFP-coding sequence. The 5′ fragment was positioned under the control of the SP6 promoter. Sequences of both cDNA fragments were verified by Sanger sequencing. Mutations were introduced into the S protein-coding sequence by cloning corresponding gene blocks to replace the original fragments. Human ACE2 gene (hACE2) was kindly provided by Luis Martinez-Sobrido (Texas Biomedical Research Institute, San Antonio, TX). It was cloned into a modified PiggyBac plasmid under the control of the CMV promoter (PiggyBAC/ACE2). The N protein-coding sequence was cloned into a noncytopathic replicon of Venezuelan equine encephalitis virus under the control of the subgenomic promoter (VEErep/N/Pac) (39). The second promoter was driving expression of the selectable marker, puromycin acetyltransferase (Pac).

**Stable cell lines.** Stable ACE2-expressing cell lines were generated by cotransfecting BHK-21, NIH 3T3, Vero E6, and Huh7.5 cells with PiggyBAC/ACE2 and integrase-encoding plasmids (System Biosciences, Inc.) using Lipofectamine 3000 according to the protocol recommended by the manufacturer (Invitrogen). Blasticidin selection was used to select stable ACE2-expressing cell lines. Stable N protein-expressing cell lines were generated by electroporation of the *in vitro*-synthesized VEErep/N/Pac RNA into BHK-21 or into BHK/ACE2 cells. Electroporated cells were further passaged in the presence of 5 $\mu$g/ml of puromycin in the media.

Infectious titers of SARS-CoV-2 were determined by plaque assay on Vero E6 or Vero/ACE2 cells. Cells were seeded in 6-well Costar plates ($2.5 \times 10^5$ cells/well). In 4 h, they were infected by serial dilutions of the viruses. After 30 min of incubation at 37°C, cells were covered with 2 ml of 0.5% agarose supplemented with DMEM and 3% FBS. After 60 to 70 h of incubation at 37°C, cells were fixed by 10% formaldehyde and then stained with crystal violet.

**Rescuing of the recombinant viruses.** Plasmids encoding cDNAs of viral genomes were purified by ultracentrifugation in CsCl gradients. The 5′ fragment-containing plasmid was digested by SacII, located in the plasmid backbone, dephosphorylated by alkaline phosphatase, and, after its inactivation, digested by SacI. The fragment encoding the SP6 promoter and ~16,000 nt of the 5′ end of viral genome was isolated by agarose gel electrophoresis. The plasmid containing the 3′ fragment of the genome was digested by SacI and SfiI, whose site is located downstream of the poly(A) tail, and the 3′ fragment of the viral sequence was also isolated by agarose gel electrophoresis. The 5′ and 3′ fragments (2 $\mu$g of

each) were ligated for 1 h at 16°C by T4 DNA ligase (Invitrogen) and purified by phenol-chloroform extraction. RNAs were synthesized *in vitro* by SP6 RNA polymerase (New England Biolabs) in the presence of a cap analog (New England Biolabs) under the conditions recommended by the manufacturer. The qualities of the RNAs were evaluated by electrophoresis in nondenaturing agarose gels. RNAs were used for electroporation of cells without additional purification.

The *in vitro*-synthesized RNAs (~1 $\mu$g) were electroporated into BHK-21 or BHK/ACE2 cells containing VEErep/N/Pac under previously described conditions (40, 41). Electroporated cells were seeded onto subconfluent monolayers of Vero/ACE2 cells and incubated at 37°C. Depending on the construct, viruses were harvested at 48 or 72 h postelectroporation, and titers were determined by a plaque assay on Vero/ACE2 cells.

**Analysis of virus binding to heparin Sepharose.** For this analysis, virus stocks were prepared in virus production serum-free (VP-SF) media (Invitrogen). Viruses were harvested before CPE development and additionally clarified by centrifugation at 16,000 $\times$ *g*, and supernatants were used for further analysis. A column of heparin Sepharose (0.4 ml) was equilibrated with 0.66$\times$ phosphate-buffered saline (PBS) (NaCl concentration of 0.1 M). Viral samples (1 ml) were diluted by 0.5 volume of water to reduce concentration of NaCl to 0.1 M and loaded onto the column in the biosafety cabinet. Then, the column was washed by a step gradient of PBS containing NaCl at concentrations increasing from 0.1 to 0.5 M, and viral titers were determined in each fraction, including flowthrough (FT), by plaque assay on Vero/ACE2 cells.

**Analysis of spike protein processing in the released virions.** Viral samples were prepared in VP-SF media. To avoid contamination with cell debris, they were harvested before CPE development. The virus was inactivated by adding 4$\times$ protein gel loading buffer and analyzed by Western blotting using S2- and N protein-specific Abs (Sinobiological and ProSci, respectively) and corresponding secondary Abs labeled with infrared dyes. The membranes were scanned on an Odyssey imaging system (LI-COR Biosciences).

**NGS.** G RNAs were isolated from the viral samples using Direct-zol RNA miniprep kit (Zymo Research) according to the manufacturer's protocol. NGS was performed on the Illumina MiSeq in the UAB Genomics Core Facility as described elsewhere (42). To identify mutational hot spots, the raw sequence FASTQ files were aligned to the reference genomes using BWA version 0.7.17-r1188. Aligned reads were then sorted with Picard version 2.9.2 SortSam. Naïve Variant Caller (Galaxy version 0.0.4) called the variants, and VariantAnnotator (Galaxy version 1.3.2) counted the bases at each position.

**Reverse transcriptase quantitative PCR.** Total RNA was extracted from 0.1 ml viral sample using Direct-zol RNA miniprep kit according to the manufacturer's protocol. cDNA was synthesized using random primers and QuantiTect reverse transcription kit (Qiagen). Quantitative PCR (qPCR) reactions were performed in a final volume of 20 $\mu$l containing SsoFast EvaGreen supermix (Bio-Rad) using CFX96 real-time PCR detection system (Bio-Rad) and 500 nM SARS-Cov-2 nsp1-specific CCTCAACTTGAACAGC CCTATG forward and GAATGCCTTCGAGTTCTGCTAC reverse primers.

**Inhibition of viral infection by heparin.** Viruses were diluted to a concentration of ~500 PFU/ml in PBS supplemented with 1% FBS and incubated in the presence of heparin at concentrations ranging between 0.0625 to 1 mg/ml at 37°C for 1 h. Then, Vero/ACE2 cells were infected with 0.2 ml of the samples and the control, which was incubated without heparin. After 30 min at 37°C, cells were covered with 2 ml of 0.5% agarose supplemented with DMEM and 3% FBS. After 2 to 3 days of incubation at 37°C, cells were fixed and then stained with crystal violet.

**Molecular modeling.** Models for the complete SARS-CoV-2 S ectodomain trimer and trimer with the GLTSKRN insertions after amino acid residue 214 were generated with SWISS-MODEL (43) based on the coordinates of PDB ID 7JJI (44). Additional models of S trimers individually harboring I68R, N74K, or S247R mutations were generated using Coot (45). The structures were visualized in BIOVIA Discovery Studio Visualizer.

**Biosafety.** Virus rescue and analyses, which required infectious virus, were performed in biosafety level 3 (BSL3) Southeastern Biosafety Laboratory Alabama Birmingham (SEBLAB) facility of the University of Alabama at Birmingham according to the protocols approved by the Institutional Biosafety Committee.

## REFERENCES

1. Dong E, Du H, Gardner L. 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20:533–534. https://doi.org/10.1016/S1473-3099(20)30120-1.

2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel Coronavirus Investigating and Research Team. 2020. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 382:727–733. https://doi.org/10.1056/NEJMoa2001017.

3. Sola I, Almazan F, Zuniga S, Enjuanes L. 2015. Continuous and discontinuous RNA synthesis in coronaviruses. Annu Rev Virol 2:265–288. https://doi.org/10.1146/annurev-virology-100114-055218.

4. Xie X, Muruato A, Lokugamage KG, Narayanan K, Zhang X, Zou J, Liu J, Schindewolf C, Bopp NE, Aguilar PV, Plante KS, Weaver SC, Makino S, LeDuc JW, Menachery VD, Shi PY. 2020. An infectious cDNA clone of SARS-CoV-2. Cell Host Microbe 27:841–848.e3. https://doi.org/10.1016/j.chom.2020.04.004.

5. Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. 2020. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. Antiviral Res 178:104792. https://doi.org/10.1016/j.antiviral.2020.104792.

6. Whittaker GR, Daniel S, Millet JK. 2021. Coronavirus entry: how we arrived at SARS-CoV-2. Curr Opin Virol 47:113–120. https://doi.org/10.1016/j.coviro.2021.02.006.

7. Hoffmann M, Kleine-Weber H, Pohlmann S. 2020. 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.e5. https://doi.org/10.1016/j.molcel.2020.04.022.

8. Marsh M, Helenius A. 2006. Virus entry: open sesame. Cell 124:729–740. https://doi.org/10.1016/j.cell.2006.02.007.

9. Lindahl U, Couchman J, Kimata K, Esko JD. 2015. Proteoglycans and sulfated glycosaminoglycans, p 207–221. *In* Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH (ed), Essentials of glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. https://doi.org/10.1101/glycobiology.3e.017.

10. Clausen TM, Sandoval DR, Spliid CB, Pihl J, Perrett HR, Painter CD, Narayanan A, Majowicz SA, Kwong EM, McVicar RN, Thacker BE, Glass CA, Yang Z, Torres JL, Golden GJ, Bartels PL, Porell RN, Garretson AF, Laubach L, Feldman J, Yin X, Pu Y, Hauser BM, Caradonna TM, Kellman BP, Martino C, Gordts P, Chanda SK, Schmidt AG, Godula K, Leibel SL, Jose J, Corbett KD, Ward AB, Carlin AF, Esko JD. 2020. SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. Cell 183:1043–1057.e15. https://doi.org/10.1016/j.cell.2020.09.033.

11. Mandl CW, Kroschewski H, Allison SL, Kofler R, Holzmann H, Meixner T, Heinz FX. 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. https://doi.org/10.1128/JVI.75.12.5627-5637.2001.

12. Bernard KA, Klimstra WB, Johnston RE. 2000. Mutations in the E2 glycoprotein of Venezuelan equine encephalitis virus confer heparan sulfate interaction, low morbidity, and rapid clearance from blood of mice. Virology 276:93–103. https://doi.org/10.1006/viro.2000.0546.

13. Klimstra WB, Ryman KD, Johnston RE. 1998. Adaptation of Sindbis virus to BHK cells selects for use of heparan sulfate as an attachment receptor. J Virol 72:7357–7366. https://doi.org/10.1128/JVI.72.9.7357-7366.1998.

14. Gorchakov R, Wang E, Leal G, Forrester NL, Plante K, Rossi SL, Partidos CD, Adams AP, Seymour RL, Weger J, Borland EM, Sherman MB, Powers AM, Osorio JE, Weaver SC. 2012. Attenuation of Chikungunya virus vaccine strain 181/clone 25 is determined by two amino acid substitutions in the E2 envelope glycoprotein. J Virol 86:6084–6096. https://doi.org/10.1128/JVI.06449-11.

15. Gardner CL, Choi-Nurvitadhi J, Sun C, Bayer A, Hritz J, Ryman KD, Klimstra WB. 2013. Natural variation in the heparan sulfate binding domain of the eastern equine encephalitis virus E2 glycoprotein alters interactions with cell surfaces and virulence in mice. J Virol 87:8582–8590. https://doi.org/10.1128/JVI.00937-13.

16. Ryman KD, Gardner CL, Burke CW, Meier KC, Thompson JM, Klimstra WB. 2007. Heparan sulfate binding can contribute to the neurovirulence of neuroadapted and nonneuroadapted Sindbis viruses. J Virol 81:3563–3573. https://doi.org/10.1128/JVI.02494-06.

17. Klimstra WB, Tilston-Lunel NL, Nambulli S, Boslett J, McMillen CM, Gilliland T, Dunn MD, Sun C, Wheeler SE, Wells A, Hartman AL, McElroy AK, Reed DS, Rennick LJ, Duprex WP. 2020. SARS-CoV-2 growth, furin-cleavage-site adaptation and neutralization using serum from acutely infected hospitalized COVID-19 patients. J Gen Virol 101:1156–1169. https://doi.org/10.1099/jgv.0.001481.

18. Ogando NS, Dalebout TJ, Zevenhoven-Dobbe JC, Limpens R, van der Meer Y, Caly L, Druce J, de Vries JJC, Kikkert M, Barcena M, Sidorov I, Snijder EJ. 2020. SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology. J Gen Virol 101:925–940. https://doi.org/10.1099/jgv.0.001453.

19. de Haan CA, Stadler K, Godeke GJ, Bosch BJ, Rottier PJ. 2004. Cleavage inhibition of the murine coronavirus spike protein by a furin-like enzyme affects cell-cell but not virus-cell fusion. J Virol 78:6048–6054. https://doi.org/10.1128/JVI.78.11.6048-6054.2004.

20. Schickli JH, Thackray LB, Sawicki SG, Holmes KV. 2004. The N-terminal region of the murine coronavirus spike glycoprotein is associated with the extended host range of viruses from persistently infected murine cells. J Virol 78:9073–9083. https://doi.org/10.1128/JVI.78.17.9073-9083.2004.

21. Thackray LB, Holmes KV. 2004. Amino acid substitutions and an insertion in the spike glycoprotein extend the host range of the murine coronavirus MHV-A59. Virology 324:510–524. https://doi.org/10.1016/j.virol.2004.04.005.

22. de Haan CA, Te Lintelo E, Li Z, Raaben M, Wurdinger T, Bosch BJ, Rottier PJ. 2006. Cooperative involvement of the S1 and S2 subunits of the murine coronavirus spike protein in receptor binding and extended host range. J Virol 80:10909–10918. https://doi.org/10.1128/JVI.00950-06.

23. de Haan CA, Li Z, Te Lintelo E, Bosch BJ, Haijema BJ, Rottier PJ. 2005. Murine coronavirus with an extended host range uses heparan sulfate as an entry receptor. J Virol 79:14451–14456. https://doi.org/10.1128/JVI.79.22.14451-14456.2005.

24. Sasaki M, Uemura K, Sato A, Toba S, Sanaki T, Maenaka K, Hall WW, Orba Y, Sawa H. 2021. SARS-CoV-2 variants with mutations at the S1/S2 cleavage site are generated in vitro during propagation in TMPRSS2-deficient cells. PLoS Pathog 17:e1009233. https://doi.org/10.1371/journal.ppat.1009233.

25. Ko SH, Bayat Mokhtari E, Mudvari P, Stein S, Stringham CD, Wagner D, Ramelli S, Ramos-Benitez MJ, Strich JR, Davey RT, Jr., Zhou T, Misasi J, Kwong PD, Chertow DS, Sullivan NJ, Boritz EA. 2021. High-throughput, single-copy sequencing reveals SARS-CoV-2 spike variants coincident with mounting humoral immunity during acute COVID-19. PLoS Pathog 17:e1009431. https://doi.org/10.1371/journal.ppat.1009431.

26. Peacock TP, Goldhill DH, Zhou J, Baillon L, Frise R, Swann OC, Kugathasan R, Penn R, Brown JC, Sanchez-David RY, Braga L, Williamson MK, Hassard JA, Staller E, Hanley B, Osborn M, Giacca M, Davidson AD, Matthews DA, Barclay WS. 2021. The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. Nat Microbiol 6:899–909. https://doi.org/10.1038/s41564-021-00908-w.

27. Johnson BA, Xie X, Bailey AL, Kalveram B, Lokugamage KG, Muruato A, Zou J, Zhang X, Juelich T, Smith JK, Zhang L, Bopp N, Schindewolf C, Vu M, Vanderheiden A, Winkler ES, Swetnam D, Plante JA, Aguilar P, Plante KS, Popov V, Lee B, Weaver SC, Suthar MS, Routh AL, Ren P, Ku Z, An Z, Debbink K, Diamond MS, Shi PY, Freiberg AN, Menachery VD. 2021. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. Nature 591:293–299. https://doi.org/10.1038/s41586-021-03237-4.

28. de Haan CA, Haijema BJ, Schellen P, Wichgers Schreur P, Te Lintelo E, Vennema H, Rottier PJ. 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. https://doi.org/10.1128/JVI.00074-08.

29. Licitra BN, Millet JK, Regan AD, Hamilton BS, Rinaldi VD, Duhamel GE, Whittaker GR. 2013. Mutation in spike protein cleavage site and pathogenesis of feline coronavirus. Emerg Infect Dis 19:1066–1073. https://doi.org/10.3201/eid1907.121094.

30. Andre NM, Cossic B, Davies E, Miller AD, Whittaker GR. 2019. Distinct mutation in the feline coronavirus spike protein cleavage activation site in a cat with feline infectious peritonitis-associated meningoencephalomyelitis. JFMS Open Rep 5:2055116919856103. https://doi.org/10.1177/2055116919856103.

31. Chang HW, Egberink HF, Halpin R, Spiro DJ, Rottier PJ. 2012. Spike protein fusion peptide and feline coronavirus virulence. Emerg Infect Dis 18:1089–1095. https://doi.org/10.3201/eid1807.120143.

32. Smit JM, Waarts BL, Kimata K, Klimstra WB, Bittman R, Wilschut J. 2002. Adaptation of alphaviruses to heparan sulfate: interaction of Sindbis and Semliki forest viruses with liposomes containing lipid-conjugated heparin. J Virol 76:10128–10137. https://doi.org/10.1128/jvi.76.20.10128-10137.2002.

33. Heil ML, Albee A, Strauss JH, Kuhn RJ. 2001. An amino acid substitution in the coding region of the E2 glycoprotein adapts Ross River virus to utilize heparan sulfate as an attachment moiety. J Virol 75:6303–6309. https://doi.org/10.1128/JVI.75.14.6303-6309.2001.

34. Huang J, Brieba LG, Sousa R. 2000. Misincorporation by wild-type and mutant T7 RNA polymerases: identification of interactions that reduce misincorporation rates by stabilizing the catalytically incompetent open conformation. Biochemistry 39:11571–11580. https://doi.org/10.1021/bi000579d.

35. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, Spitters C, Ericson K, Wilkerson S, Tural A, Diaz G, Cohn A, Fox L, Patel A, Gerber SI, Kim L, Tong S, Lu X, Lindstrom S, Pallansch MA, Weldon WC, Biggs HM, Uyeki TM, Pillai SK, Washington State 2019-nCoV Case Investigation Team. 2020. First case of 2019 novel coronavirus in the United States. N Engl J Med 382:929–936. https://doi.org/10.1056/NEJMoa2001191.

36. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, Zhang X, Muruato AE, Zou J, Fontes-Garfias CR, Mirchandani D, Scharton D, Bilello JP, Ku Z, An Z, Kalveram B, Freiberg AN, Menachery VD, Xie X, Plante KS, Weaver SC, Shi PY. 2021. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 592:116–121. https://doi.org/10.1038/s41586-020-2895-3.

37. Ruggli N, Tratschin JD, Mittelholzer C, Hofmann MA. 1996. Nucleotide sequence of classical swine fever virus strain Alfort/187 and transcription of infectious RNA from stably cloned full-length cDNA. J Virol 70:3478–3487. https://doi.org/10.1128/JVI.70.6.3478-3487.1996.

38. Rice CM, Levis R, Strauss JH, Huang HV. 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. J Virol 61:3809–3819. https://doi.org/10.1128/JVI.61.12.3809-3819.1987.

39. Petrakova O, Volkova E, Gorchakov R, Paessler S, Kinney RM, Frolov I. 2005. Noncytopathic replication of Venezuelan equine encephalitis virus and eastern equine encephalitis virus replicons in mammalian cells. J Virol 79:7597–7608. https://doi.org/10.1128/JVI.79.12.7597-7608.2005.

40. Liljeström P, Lusa S, Huylebroeck D, Garoff H. 1991. *In vitro* mutagenesis of a full-length cDNA clone of Semliki Forest virus: the small 6,000-molecular-weight membrane protein modulates virus release. J Virol 65: 4107–4113. https://doi.org/10.1128/JVI.65.8.4107-4113.1991.

41. Gorchakov R, Hardy R, Rice CM, Frolov I. 2004. Selection of functional 5' *cis*-acting elements promoting efficient Sindbis virus genome replication. J Virol 78:61–75. https://doi.org/10.1128/jvi.78.1.61-75.2004.

42. Urakova N, Kuznetsova V, Crossman DK, Sokratian A, Guthrie DB, Kolykhalov AA, Lockwood MA, Natchus MG, Crowley MR, Painter GR, Frolova EI, Frolov I. 2017. beta-D-N(4)-hydroxycytidine is a potent anti-alphavirus compound that induces high level of mutations in viral genome. J Virol:e01965-17. https://doi.org/10.1128/JVI.01965-17.

43. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–W303. https://doi.org/10.1093/nar/gky427.

44. Bangaru S, Ozorowski G, Turner HL, Antanasijevic A, Huang D, Wang X, Torres JL, Diedrich JK, Tian JH, Portnoff AD, Patel N, Massare MJ, Yates JR, III, Nemazee D, Paulson JC, Glenn G, Smith G, Ward AB. 2020. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. Science 370:1089–1094. https://doi.org/10.1126/science.abe1502.

45. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501. https://doi.org/10.1107/S0907444910007493.