The Highly Conserved Stem-Loop II Motif Is Dispensable for SARS-CoV-2

Hongbing Jiang,[a,b] Astha Joshi,[c] Tianyu Gan,[b] Andrew B. Janowski,[d] Chika Fujii,[b] Traci L. Bricker,[c] Tamarand L. Darling,[c] Houda H. Harastani,[c] Kuljeet Seehra,[c] Hongwei Chen,[a] Stephen Tahan,[b] Ana Jung,[e] Binita Febles,[e] Joshua A. Blatter,[b] Scott A. Handley,[e] Bijal A. Parikh,[e] David Wang,[b,e] Adrianus C. M. Boon[b,c,e]

aSchool of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China
bDepartment of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, USA
cDepartment of Medicine, Washington University School of Medicine, St. Louis, Missouri, USA
dDepartment of Pediatrics, Washington University School of Medicine, St. Louis, Missouri, USA
eDepartment of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri, USA

ABSTRACT The stem-loop II motif (s2m) is an RNA structural element that is found in the 3′ untranslated region (UTR) of many RNA viruses, including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Though the motif was discovered over 25 years ago, its functional significance is unknown. In order to understand the importance of s2m, we created viruses with deletions or mutations of the s2m by reverse genetics and also evaluated a clinical isolate harboring a unique s2m deletion. Deletion or mutation of the s2m had no effect on growth *in vitro* or on growth and viral fitness in Syrian hamsters *in vivo*. We also compared the secondary structure of the 3′ UTR of wild-type and s2m deletion viruses using selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) and dimethyl sulfate mutational profiling and sequencing (DMS-MaPseq). These experiments demonstrate that the s2m forms an independent structure and that its deletion does not alter the overall remaining 3′-UTR RNA structure. Together, these findings suggest that s2m is dispensable for SARS-CoV-2.

IMPORTANCE RNA viruses, including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), contain functional structures to support virus replication, translation, and evasion of the host antiviral immune response. The 3′ untranslated region of early isolates of SARS-CoV-2 contained a stem-loop II motif (s2m), which is an RNA structural element that is found in many RNA viruses. This motif was discovered over 25 years ago, but its functional significance is unknown. We created SARS-CoV-2 with deletions or mutations of the s2m and determined the effect of these changes on viral growth in tissue culture and in rodent models of infection. Deletion or mutation of the s2m element had no effect on growth *in vitro* or on growth and viral fitness in Syrian hamsters *in vivo*. We also observed no impact of the deletion on other known RNA structures in the same region of the genome. These experiments demonstrate that s2m is dispensable for SARS-CoV-2.

The RNA genome of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is approximately 30,000 nucleotides in length (1). It consists of a 5′ untranslated region (UTR), coding sequences for structural and nonstructural proteins, and a 3′ UTR. The 3′ UTR contains highly structured RNA elements such as stem-loop sequence 1 (SL1), bulged stem-loop (BSL), pseudoknot (PK), and a hypervariable region (HVR), which have been suggested to function in viral genome replication, transcription, and viral protein translation (2, 3). SARS-CoV-2, SARS-CoV-1, and other members of the *Sarbecovirus* lineage in the *Betacoronavirus* genus, as well as some members of the
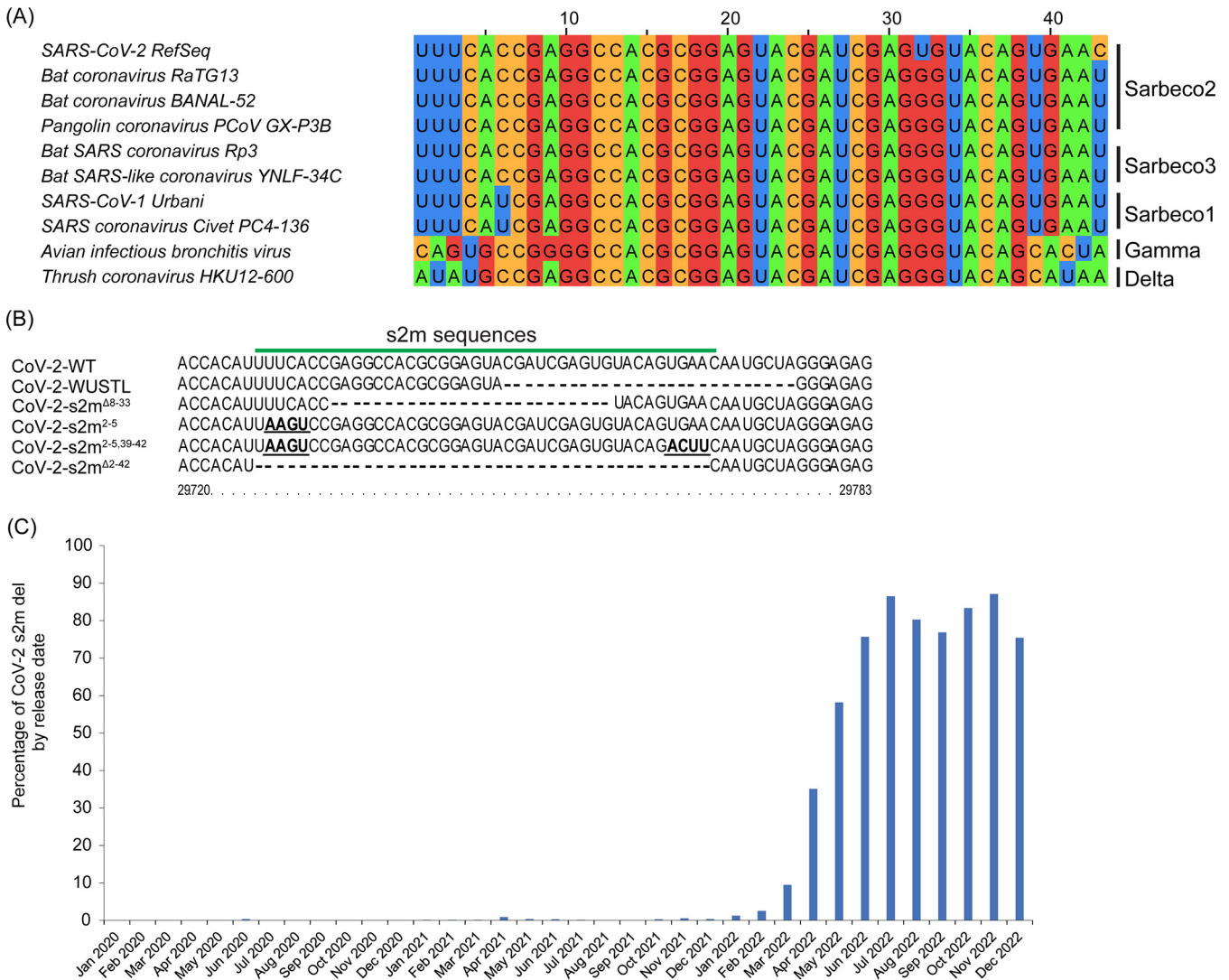
FIG 1 s2m nucleotide sequence alignment. (A) Multiple-sequence alignment of representative coronaviruses in the *Betacoronavirus*, *Gammacoronavirus*, and *Deltacoronavirus* genera that encode an s2m. (B) Sequence alignment of the s2m element found in SARS-CoV-2 isolates. SARS-CoV-2-WT is the original SARS-CoV-2 isolate (GenBank accession no. NC_045512), SARS-CoV-2-WUSTL is the virus isolated from the patient in the St. Louis area (2019-nCoV/WUSTL_000226/2020), SARS-CoV-2-s2m$^{\Delta 8-33}$ is the circulating s2m deletion mutant, SARS-CoV-2-s2m$^{2-5}$ is the engineered s2m stem structure mutant, SARS-CoV-2-s2m$^{2-5,39-42}$ is the engineered s2m stem structure mutation revertant, and SARS-CoV-2-s2m$^{\Delta 2-42}$ is the engineered s2m deletion mutant. The genomic coordinates for the SARS-CoV-2-WT are indicated. (C) SARS-CoV-2 s2m deletion-containing sequences found by release date.

Gammacoronavirus and *Deltacoronavirus* genera, encode a stem-loop II motif (s2m) within the terminal portion of the HVR in the 3′ UTR (4, 5) (Fig. 1A). In contrast, seasonal human coronaviruses (HKU1, 229E, OC43, and NL63) and Middle East respiratory syndrome coronavirus (MERS-CoV) do not contain an s2m in their genomes (4). The s2m element has also been detected in members of the *Astroviridae*, *Caliciviridae*, *Coronaviridae*, *Picornaviridae*, and *Reoviridae* viral families, all with highly conserved nucleotide sequences of 39 to 43 nucleotides in length (4, 6–9). Currently, the function of the s2m for the viral life cycle is poorly understood. Phylogenetic distribution suggests horizontal acquisition of the s2m at different time points, and maintenance of the element suggests that it may confer a fitness advantage (4, 6). The X-ray crystal structure of the s2m element from SARS-CoV-1 demonstrates a stem-loop secondary structure and a tertiary structure that includes a 90° kink in the helix axis, resulting in additional tertiary interactions (7). The secondary structure determination by nuclear magnetic resonance (NMR) and probing methods for the SARS-CoV-2 s2m element revealed two stem structures separated by an internal asymmetric loop (10–15). Antisense oligonucleotides targeting the s2m reduced viral replication for SARS-

CoV-2 and classic human astrovirus 1 (HAstV1) replicons (16). The SARS-CoV-2 s2m was shown to dimerize and interact with host cellular microRNA 1307-3p (17). These results suggest that the secondary structure of the s2m is conserved and potentially important for viral replication or other host-virus interactions.

Interestingly, the SARS-CoV-2 genomes encode a uracil residue at position 32 in the s2m (position 29758 in the virus genome) that is distinct from all known s2m sequences in other viruses and is predicted to perturb the secondary structure (10, 11, 18–20). Additional genetic variants or deletions of the s2m have also been periodically detected in clinical isolates prior to the emergence of the BA.2 Omicron lineage of SARS-CoV-2 (20–23). However, SARS-CoV-2 variants that emerged after December 2021 contain a 26-nucleotide deletion of the s2m element (Fig. 1C) (24). Combined, these data suggest that the s2m element has minimal or no impact on the life cycle of SARS-CoV-2.

Here, we determined the functional significance of s2m *in vitro* and *in vivo* using recombinant SARS-CoV-2 viruses or natural isolates with mutations or deletions in the s2m element in the 3′ UTR of the genome. We also determined the 3′-UTR RNA structure of SARS-CoV-2 in the presence or absence of the s2m element. We show that deletion of s2m in SARS-CoV-2 has no impact on the viral fitness or 3′ UTR structure of SARS-CoV-2.

RESULTS

Natural variation and deletion of s2m found in SARS-CoV-2 circulating strains. The original SARS-CoV-2 virus genome (reference genome, GenBank accession no. NC_045512) contains an s2m element in the 3′-UTR region, which is similar to other sarbecoviruses in the *Betacoronavirus* genus and some members of the *Gammacoronavirus* and *Deltacoronavirus* genera (Fig. 1A). Interestingly, the SARS-CoV-2 s2m encodes a uracil at position 32 of the s2m (position 29758 in the reference genome), while essentially all other s2m sequences in coronaviruses known to date contain guanine (Fig. 1A). In the s2m element of SARS-CoV-1, this position forms a G-C base pair as determined by X-ray crystal structure (7) and by RNAfold prediction. To further examine the variation at this position, we analyzed 1,705,180 complete SARS-CoV-2 genomes uploaded to the NCBI database from January 2020 to December 2022. Of those sequences that contained a complete s2m element, only eight genomes contained guanine at this position. An additional 18 had a cytosine residue, and 1 genome contained an adenine at position 32. We also noticed in the multiple-sequence alignment that position 9 can be variable between viruses within the coronavirus family (Fig. 1A). SARS-CoV-1 and SARS-CoV-2 encode an adenine, while avian infectious bronchitis virus encodes a guanine. This position is not predicted to form any base pair but has been identified to form potential long-distance tertiary interactions with nucleotide 30 in the SARS-CoV-1 s2m crystal structure (7).

Although the s2m element is relatively conserved in the genome of SARS-CoV-2 between the start of the pandemic and early 2022, several genomes were found to have a partial or complete deletion of the s2m element (23). Our analysis of the 1,705,180 complete SARS-CoV-2 genomes also revealed the emergence of SARS-CoV-2 lineages with a 26-nucleotide deletion (positions 8 to 33) in the s2m element (Fig. 1B and C). The genomes that contain this deletion mainly belong to the BA.2 lineage (Pango lineages) of SARS-CoV-2, which includes BA.2.75, BA.4, BA.5, and the recent BQ.1 and XBB.1 variants of SARS-CoV-2.

The s2m element is dispensable for SARS-CoV-2 *in vitro*. To determine the importance of the s2m in the SARS-CoV-2 virus life cycle, we recombinantly generated in the reference backbone wild type (CoV-2-s2m-WT) and three mutant viruses with mutations or deletions in the s2m region, using the reverse genetics system described in Materials and Methods. To remove the s2m element in the 3′ UTR of SARS-CoV-2, we deleted nucleotides 2 to 42 in the s2m region (CoV-2-s2m$^{\Delta2-42}$) (Fig. 1B). We also created a mutant that contained four consecutive nucleotide substitutions in the stem region of the s2m (CoV-2-s2m$^{2-5}$) that is predicted to disrupt the secondary structure, as well as a revertant mutant (CoV-2-s2m$^{2-5,39-42}$) that contained four additional compensatory substitutions that restored the predicted stem region and the SARS-CoV-2 s2m secondary structure
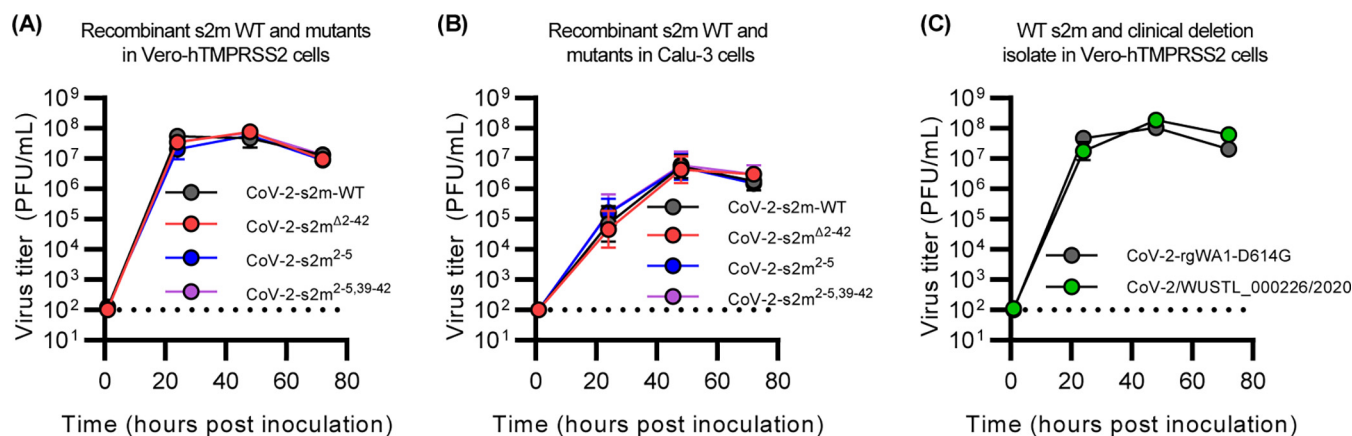
FIG 2 The s2m is dispensable for SARS-CoV-2 *in vitro*. (A and B) Multistep growth curve of CoV-2-s2m-WT and mutants in Vero-hTMPRSS2 (A) and Calu-3 cells (B). Mutant strains include CoV-2-s2m$^{\Delta2-42}$, which contains a deletion of the s2m element; CoV-2-s2m$^{2-5}$, which contains four consecutive substitutions of the stem of the s2m; and CoV-2-s2m$^{2-5,39-42}$, which contains complementary substitutions predicted to restore the secondary structure of the s2m. Infectious virus titer measured in PFU per milliliter at 0, 24, 48, and 72 h postinoculation. No difference in the viral titer was detected by a two-way ANOVA with *post hoc* testing by Dunnett's multiple-comparison test between CoV-2-s2m-WT and all mutants for Vero-hTMPRSS2 [$F_{(3, 20)} = 1.02$, $P = 0.40$] and for Calu-3 cells [$F_{(3, 20)} = 0.48$, $P = 0.7$]. (C) Multistep growth curve of the infectious viral titer using a clinical isolate of SARS-CoV-2 containing a partial deletion of the s2m element (WUSTL_000226/2020) compared to a WA1 strain of SARS-CoV-2 with a D614G mutation. Viral titers were measured at 0, 24, 48, and 72 h postinoculation, and no difference in titers was detected by a two-way ANOVA with *post hoc* testing by Sidak's multiple-comparison test [$F_{(1,20)} = 2.34$, $P = 0.16$]. For all graphs, geometric means $\pm$ geometric standard deviations are depicted. The dotted line is the limit of detection of the assay.

(Fig. 1B). All mutant SARS-CoV-2 genomes yielded infectious viral particles that could be propagated in Vero-hTMPRSS2 cells. We confirmed the presence of the engineered mutations and the absence of spontaneous mutations in the SARS-CoV-2 s2m by sequencing. We next tested whether there were any defects in the growth rate of the mutants by conducting multistep growth curves (Fig. 2A). There was no significant difference between CoV-2-s2m-WT and mutant viruses at any time point in Vero-hTMPRSS2 (African green monkey) cells [two-way analysis of variance (ANOVA) $F_{(3,20)} = 1.02$, $P = 0.4$] or Calu-3 (human) cells [two-way ANOVA $F_{(3,20)} = 0.48$, $P = 0.7$], suggesting that the s2m element is not required for the SARS-CoV-2 life cycle *in vitro* (Fig. 2B).

Growth of a clinical SARS-CoV-2 isolate with a deletion in the s2m. During genomic surveillance for SARS-CoV-2 variants in the St. Louis area, United States, we identified one genome (SARS-CoV-2/WUSTL_000226/2020) that contained a deletion of 27 nucleotides that removes positions 22 to 43 of the s2m and additional nucleotides in the 3′ UTR (Fig. 1B). This mutant belongs to the B1.2 lineage (Pango lineages) (25), which contains the D614G variant in the spike protein and has 99.8% nucleotide identity compared to the reference SARS-CoV-2 genome (GenBank accession no. NC_045512). We were able to culture WUSTL_000226/2020 and verified that the recovered virus maintained the deletion by sequencing. In a multistep growth curve, we did not observe any significant difference in virus titer between WUSTL_000226/2020 and a recombinant WA1-strain of SARS-CoV-2 engineered with the D614G mutation at any time point [two-way ANOVA $F_{(1,10)} = 2.3$, $P = 0.16$] (Fig. 2C).

The s2m element is dispensable for SARS-CoV-2 infection and replication *in vivo*. We next determined if the SARS-CoV-2 s2m was important *in vivo* using the Syrian hamster model. Hamsters were intranasally inoculated with 1,000 PFU of the CoV-2-s2m-WT and mutant viruses, and weights were recorded daily for 6 days. Nasal washes and lungs were collected 3 and 6 days postinfection, and infectious virus titer and viral RNA load were quantified by plaque assay and quantitative reverse transcriptase PCR (RT-qPCR), respectively. No difference in weight loss was observed between the hamsters inoculated with the CoV-2-s2m-WT and deletion or mutant s2m-containing viruses [mixed-effect model with Geisser-Greenhouse correction $F_{(3,44)} = 1.26$, $P = 0.30$]. (Fig. 3A). No significant differences in infectious virus [lung tissue Kruskal-Wallis $H(3) = 0.62$, $P = 0.89$] and viral RNA [lung tissue Kruskal-Wallis $H(3) = 2.0$, $P = 0.57$; nasal wash Kruskal-Wallis $H(3) = 2.2$, $P = 0.54$] were detected for CoV-2-s2m$^{\Delta2-42}$, CoV-2-
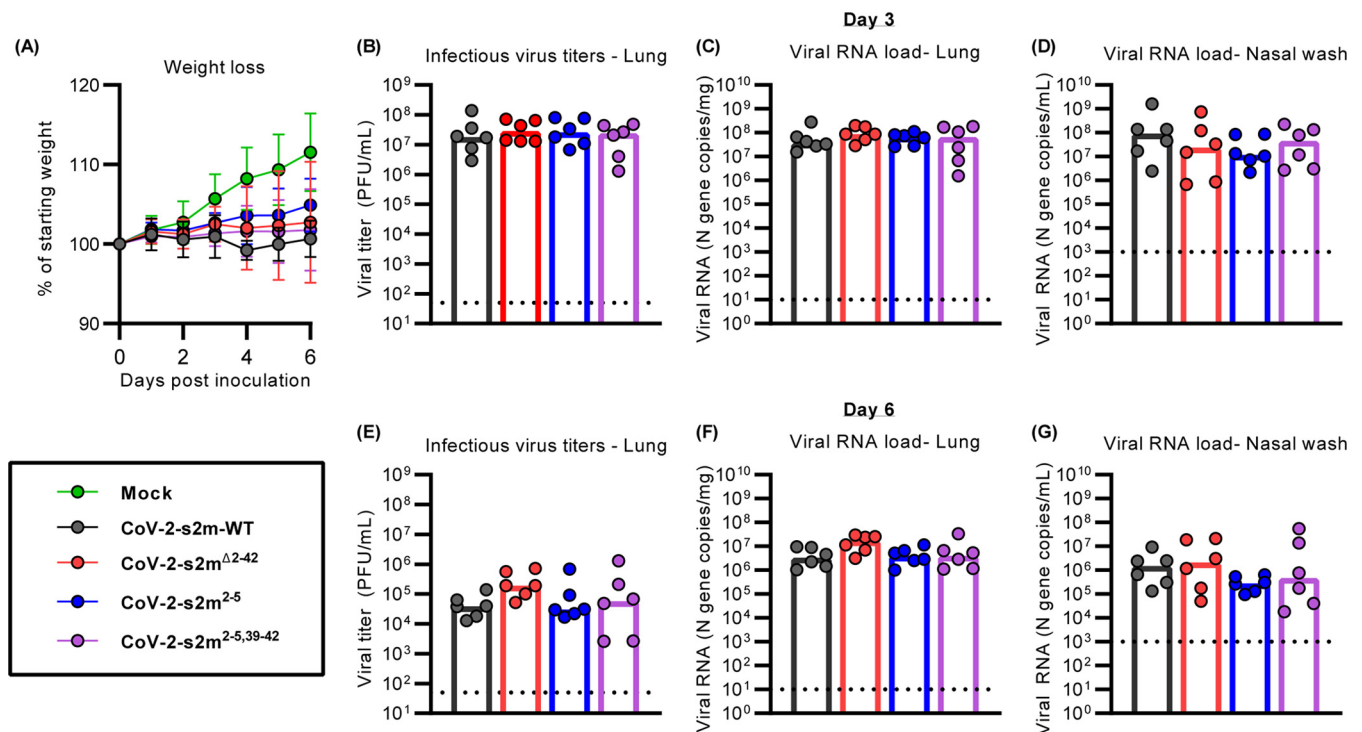
FIG 3 The s2m is dispensable for SARS-CoV-2 *in vivo*. Intranasal inoculation of hamsters with CoV-2-s2m-WT and mutants, including CoV-2-s2m$^{\Delta2-42}$, which contains a deletion of the s2m; CoV-2-s2m$^{2-5}$, which contains four mutations of the stem of the s2m; and CoV-2-s2m$^{2-5,39-42}$, which contains complementary mutations predicted to restore the secondary structure of the s2m. Hamsters were sacrificed on days 3 and 6. (A) Mean hamster weight as percentage of starting weight is graphed, with error bars representing standard deviations. The weights of the hamsters infected with CoV-2-s2m-WT and mutant SARS-CoV-2 viruses were compared, and no difference was detected using a mixed-effect model with Geisser-Greenhouse correction [$F[3,44] = 1.26$, $P = 0.30$]. (B to D) There was no difference in lung infectious viral titer [Kruskal-Wallis $H(3) = 0.62$, $P = 0.89$] (B), lung viral RNA [Kruskal-Wallis $H(3) = 2.0$, $P = 0.57$] (C), and nasal wash viral RNA [Kruskal-Wallis $H(3) = 2.2$ $P = 0.54$] (D) from day 3 were identified comparing CoV-2-s2m-WT and s2m mutant viruses. (E to G) At 6 days postinoculation, there were also no differences in lung infectious viral titer [Kruskal-Wallis $H(3) = 5.5$, $P = 0.14$] (E), lung viral RNA [Kruskal-Wallis $H(3) = 6.2$, $P = 0.10$] (F), and nasal wash viral RNA [Kruskal-Wallis $H(3) = 2.8$, $P = 0.42$] (G) comparing CoV-2-s2m-WT to mutant s2m viruses. For all graphs in panels B to G, each dot is one animal. The dotted line represents the limit of detection of the respective assay.

s2m$^{2-5}$, and CoV-2-s2m$^{2-5,39-42}$ compared to CoV-2-s2m-WT (Fig. 3B to D). At 6 days postinoculation, no difference in infectious virus was detected in the lungs of the mutant SARS-CoV-2-infected hamsters from CoV-2-s2m-WT [Kruskal-Wallis $H(3) = 5.5$, $P = 0.14$] (Fig. 3E). The viral RNA load with CoV-2-s2m$^{\Delta2-42}$, CoV-2-s2m$^{2-5}$, and CoV-2-s2m$^{2-5,39-42}$ also showed no difference in CoV-2-s2m-WT from the lungs [Kruskal-Wallis $H(3) = 6.2$, $P = 0.10$] and nasal washes [Kruskal-Wallis $H(3) = 2.8$, $P = 0.42$] (Fig. 3F and G). Combined, these data suggest that the original SARS-CoV-2 virus does not require the s2m for growth *in vitro* or *in vivo*.

The s2m element is dispensable for SARS-CoV-2 viral fitness *in vivo*. Viral fitness assays are more sensitive than infection and replication assays. Therefore, to further determine if the SARS-CoV-2 s2m element has any effect on viral fitness *in vivo*, we designed a viral competition assay using the CoV-2-s2m-WT and the CoV-2-s2m$^{\Delta2-42}$ virus in Syrian hamsters. CoV-2-s2m-WT and the CoV-2-s2m$^{\Delta2-42}$ mutant virus were mixed at a 1:1 ratio and inoculated intranasally into Syrian hamsters. The ratio of CoV-2-s2m-WT to CoV-2-s2m$^{\Delta2-42}$ mutant virus in the inoculum was determined by RT-PCR on the 3' UTR following RNA extraction of RNase-treated virus (Fig. 4A). Three days postinfection (dpi), the lungs and nasal washes were collected, and the genome copy number of CoV-2-s2m-WT to CoV-2-s2m$^{\Delta2-42}$ was measured by RT-PCR on RNA extracted from these tissues and samples. The relative replicative fitness of CoV-2-s2m$^{\Delta2-42}$ to CoV-2-s2m-WT was calculated for each sample. The mean relative replicative fitness of CoV-2-s2m$^{\Delta2-42}$ to CoV-2-s2m-WT was 1.21 and 1.10 in the lung and nasal washes, respectively (Fig. 4B). This ratio was similar to that of the input, indicating that the CoV-2-s2m-WT virus has no fitness advantage over the CoV-2-s2m$^{\Delta2-42}$ mutant virus *in vivo* in Syrian hamsters.
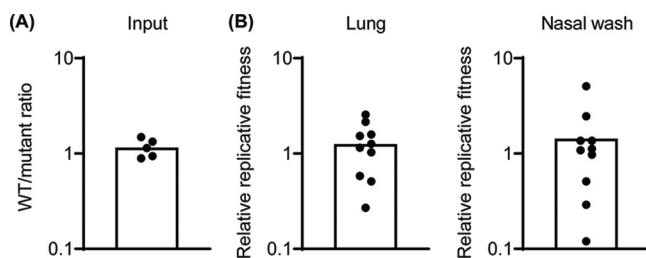
FIG 4 The s2m is dispensable for SARS-CoV-2 viral fitness in Syrian hamster. The fitness of the CoV-2-s2m$^{\Delta 2-42}$ virus was assessed *in vivo* in Syrian hamsters. Hamsters were inoculated intranasally with a 1:1 mixture of CoV-2-s2m-WT and CoV-2-s2m$^{\Delta 2-42}$ viruses, and lungs and nasal washes were collected 3 days postinfection. The genome copy of CoV-2-s2m-WT and CoV-2-s2m$^{\Delta 2-42}$ in the inoculum and infected tissues were measured by RT-PCR. (A) The ratio of CoV-2-s2m-WT to CoV-2-s2m$^{\Delta 2-42}$ in the inoculum was calculated. (B) The replicative fitness of the CoV-2-s2m$^{\Delta 2-42}$ to CoV-2-s2m-WT virus in lungs and nasal washes was calculated. Individual values were shown as dots, and mean values were plotted by box.

Structural analysis of the 3′ UTR of SARS-CoV-2. In order to determine the RNA secondary structure of the 3′ UTR in the SARS-CoV-2 genome and the impact of the s2m deletion on the secondary structure, we performed selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) and dimethyl sulfate mutational profiling and sequencing (DMS-MaPseq) RNA structure probing studies on purified CoV-2-s2m-WT and CoV-2-s2m$^{\Delta 2-42}$ virus using 2-methylnicotinic acid imidazolide (NAI) and DMS, respectively.

RNA structure predictions, using the SHAPE-MaP reactivity data as constraints, identified the bulged stem-loop (BSL), stem-loop 1 (SL1), pseudoknot, and a bulge stem that includes the hypervariable region (HVR), the s2m element, and the octanucleotide motif (ONM) in the 3′ UTR of CoV-2-s2m-WT (Fig. 5A). Interestingly, DMS-MaPseq predicted a similar secondary structure with the BSL, SL1, HVR, s2m element, and the ONM motif (Fig. 5B). Overall, the SHAPE-MaP and DMS-MaPseq reactivities were in good agreement with the predicted structure and with previous findings (11–15). Structure prediction using our SHAPE-MaP data also predicted the formation of a pseudoknot between the base of the BSL and the loop of SL1 (Fig. 5A). This pseudoknot was not predicted using DMS-MaPseq reactivity data (Fig. 5B). Using Superfold, we found that the 3′ UTR is highly structured and all the known structured elements have high base-pairing probabilities, shown by the green color (Fig. 6).

SHAPE-MaP on the 3′ UTR of CoV-2-s2m$^{\Delta 2-42}$ revealed a very similar reactivity profile, and the predicted structure contained the BSL, SL1, HVR, and ONM regions. Comparison between the CoV-2-s2m-WT and CoV-2-s2m$^{\Delta 2-42}$ reactivities showed a high correlation ($R^2 = 0.88$). Similar to CoV-2-s2m-WT, the SL1 region of CoV-2-s2m$^{\Delta 2-42}$ was also predicted to form a pseudoknot with the base of the BSL region. Analogous to the SHAPE-MaP data, the DMS-MaPseq reactivity and predicted structure between CoV-2-s2m-WT and CoV-2-s2m$^{\Delta 2-42}$ showed a high correlation ($R^2 = 0.92$). The only difference in reactivity and structure was observed near the s2m region that was deleted in the CoV-2-s2m$^{\Delta 2-42}$ virus. These data suggest that the s2m region forms an independent structure in the 3′ UTR of SARS-CoV-2 and that deletion of the s2m region did not change the overall 3′-UTR structure.

DISCUSSION

Although the crystal structure of the conserved s2m RNA element was solved for SARS-CoV-1 in 2005 (7) and many recent studies suggested an important role of the s2m structure for SARS-CoV-2, no functional genetic studies on the s2m element in the SARS coronavirus life cycle have been performed (4, 6–9, 26). Here, we demonstrated that the deletion or mutation of the s2m element in the original strain of SARS-CoV-2 did not impact growth *in vitro* or viral fitness *in vivo* and had minimal effect on the predicted RNA secondary structure of the 3′ UTR of SARS-CoV-2. These results suggest that the s2m structure is not essential for SARS-CoV-2.
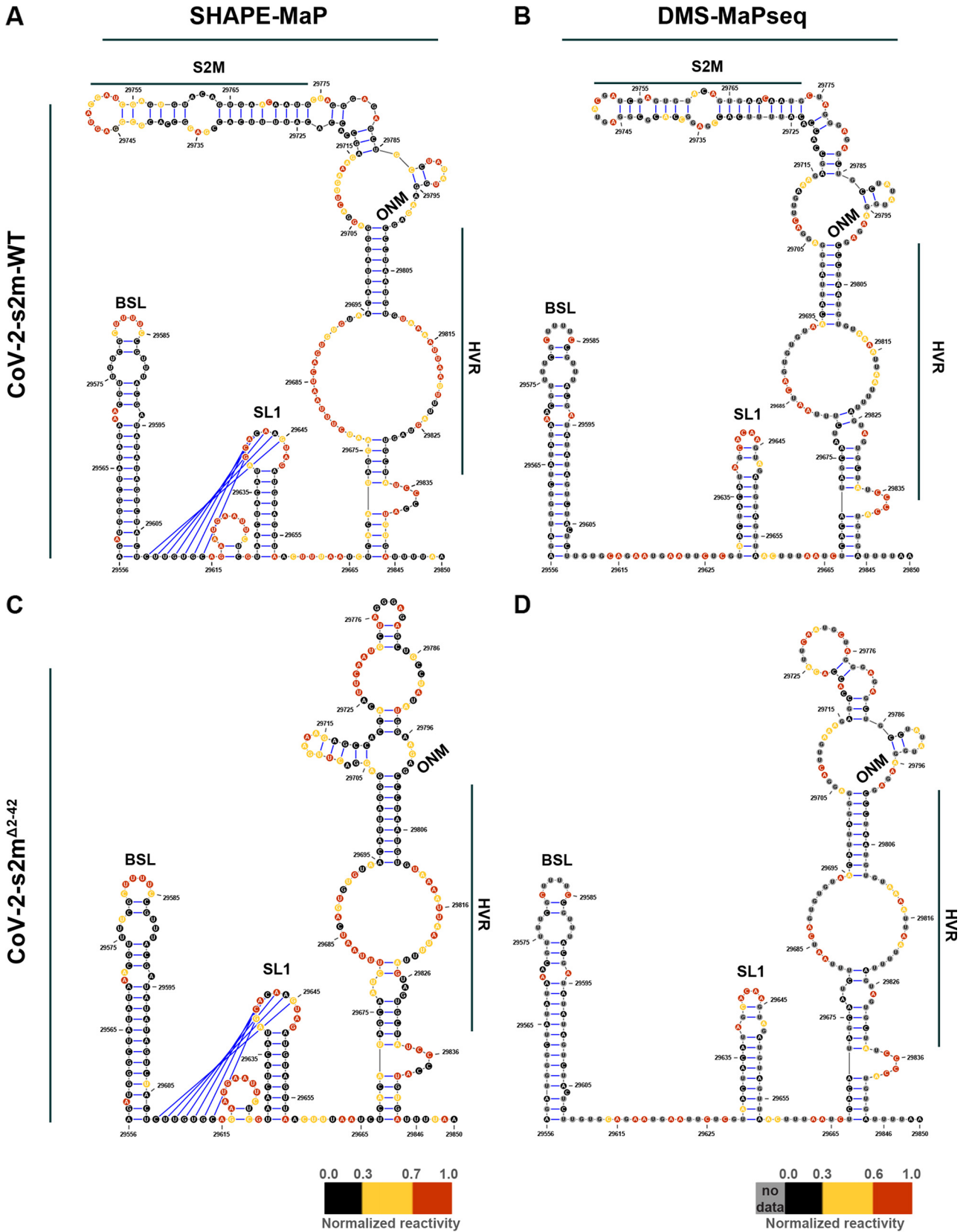
FIG 5 Structure prediction for the 3′ UTR suggests a presence of conserved structural elements. Predicted RNA secondary structure for CoV-2-s2m-WT using SHAPE-MaP (A) and DMS-MaPseq (B) reactivity profiles. Predicted RNA secondary structure for CoV-2-s2m$^{\Delta 2-42}$ using SHAPE-MaP (C) and DMS-MaPseq (D) reactivity profiles. Each nucleotide is colored by its normalized reactivity. BSL, bulged stem-loop; SL1, stem-loop 1; HVR, hypervariable region; ONM, octanucleotide motif. Blue lines indicate the pseudoknot region.
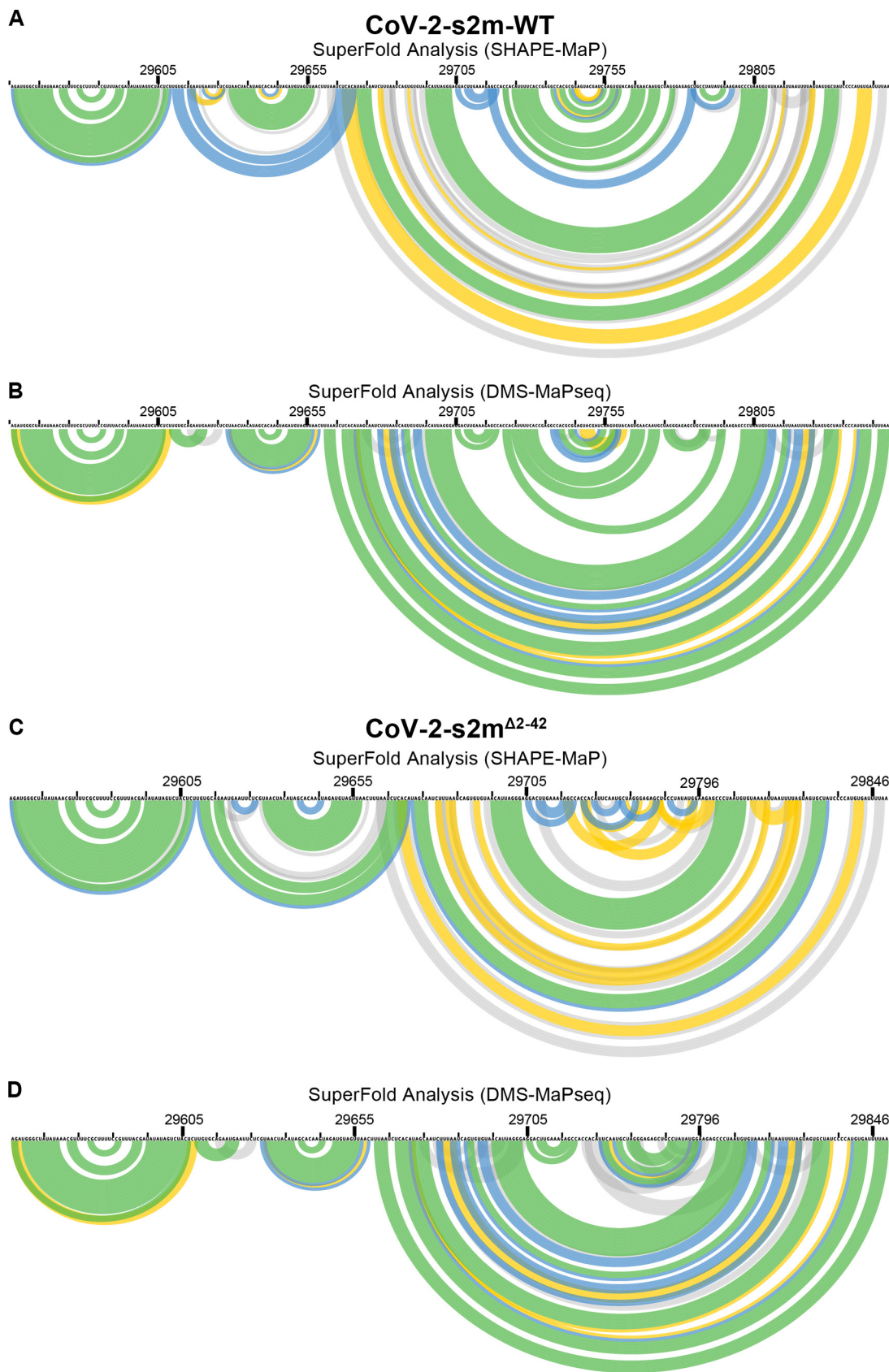
FIG 6 Superfold analysis of the 3′ UTR of wild type and SARS-CoV-2-s2m$^{\Delta 2-42}$. Following the probing of viral genome in purified SARS-CoV-2 virions, the SHAPE-MaP and DMS-MaPseq reactivity profiles were determined as described in Materials

(Continued on next page)

Based on the presence and conservation of the s2m element in different viral families, it has been hypothesized to be beneficial for the virus. However, our results suggest that the function of the s2m element is dispensable in the SARS-CoV-2 genome, perhaps due to redundancy with another uncharacterized virus-derived element, whether a viral protein or RNA motif. Investigating the significance of the s2m element in related sarbecoviruses, including SARS-CoV-1, RaTG13, and pangolin coronaviruses, is needed to better understand its role in coronavirus biology. SARS-CoV-2 is the only coronavirus that contains a uracil at position 32, while all others have a guanine at this site (Fig. 1A) (27). The presence of this uracil may represent a unique and recent evolutionary event for the s2m that is specific to the SARS-CoV-2 lineage, and therefore, the biology of the SARS-CoV-2 s2m may not apply to other coronaviruses. Complete genome analysis of natural isolates of SARS-CoV-2 also found that a small subset contained mutations or deletions in the s2m region (20–23). One of these isolates was found by our group, and we showed no difference in growth potential in two different cell lines from a related SARS-CoV-2 with the s2m element intact. While it is possible that these rare deletions were detected because of the unprecedented amount of genome sequencing that was done during the pandemic, it is also possible that this was an early indication that this region was under neutral or negative selective pressure, facilitating the emergence of the Omicron lineage with a 26-nucleotide deletion in early 2022 (Fig. 1C). Analogous to SARS-CoV-2, the genomes of several seasonal coronaviruses and MERS also do not contain an s2m element (4). Whether this is an adaptation of coronaviruses to the human host remains to be investigated.

Our RNA structure modeling of the 3′ UTR present in virions indicated the presence of several conserved structural elements, including the BSL, SL1, and a bulge stem that includes the HVR, the s2m element, and the ONM. Overall, our model is similar to those identified by others (11–15). The BSL, SL1, and s2m form a stem-loop structure. The bulge in the BSL motif was found to have low SHAPE-MaP reactivity, which can be explained by the binding of a viral or host protein as suggested previously (11). The predicted structure of the HVR was different between the two probing methods and between different studies. It is mostly a single-stranded region with high reactivity bases between two structured regions, which may explain why it tolerates the presence of multiple mutations and is not essential for viral RNA synthesis (21, 28). ONM (5′-GGAAGAGC-3′) is known to be a single-stranded region with a critical biological function (28). In our model, the first two nucleotides (GG) of the ONM form the base of a small hairpin. Currently, it is not known if this hairpin is an artifact of the RNA folding software or whether it is present in the 3′ UTR of SARS-CoV-2. Interestingly, the reactivity of the ONM region was, overall, lower in SHAPE-MaP than DMS-MaPseq, which can be explained by the binding of a viral or host protein to the ONM. The predicted structure of the s2m element was similar between SHAPE-MaP and DMS-MaPseq. Compared to a previously determined crystal structure of s2m in SARS-CoV-1 (7), the main differences are found near the top of the s2m element, with a predicted extended loop in the s2m of SARS-CoV-2. This is potentially caused by the uracil at position 32 resulting in different base pairings. This uracil was found to be moderately reactive by SHAPE-MaP and therefore predicted to interact weakly with the adenosine at position 14 of the s2m element. The weak interaction is supported by DMS-MaPseq, which showed a moderate reactivity of adenosine at position 14 (DMS does not modify uracil residues).

In contrast to DMS-MaPseq, SHAPE-MaP predicted the formation of a tertiary pseudoknot structure between the base of the BSL and the loop of SL1. The confidence of this tertiary structure prediction is relatively low since many of the involved nucleotides are highly reactive by SHAPE-MaP or DMS-MaPseq. As such, it is possible that no

FIG 6 Legend (Continued)
and Methods. SuperFold was used to predict consensus secondary structure prediction with base-pairing probabilities using the SHAPE-MaP (A and C) and DMS-MaPseq (B and D) reactivity profiles for wild-type SARS-CoV-2 (A and B) and SARS-CoV-2-s2m$^{\Delta2-42}$ (C and D). Nucleotides 29729 to 29768 were deleted from the CoV-2-s2m$^{\Delta2-42}$ genome.

pseudoknot is formed, as suggested by others (11). Alternatively, this region of the 3′ UTR switches between a tertiary structure (pseudoknot) and a free SL1 structure in SARS-CoV-2. This model is supported by recent studies on murine hepatitis virus that suggest that both structures contribute to viral replication and may function as molecular switches in different steps of RNA synthesis (2).

The comparison between the predicted RNA structures of the CoV-2-s2m-WT and CoV-2-s2m$^{\Delta2-42}$ 3′ UTRs demonstrated nearly identical structures. Only the region containing the s2m element was significantly different between the two viruses. These data suggest that the s2m hairpin structure does not interact with any other elements in the 3′-UTR region of SARS-CoV-2 as predicted previously by computational analysis (29). Taken together, our RNA structural analysis suggests that there is no impact on the overall 3′-UTR RNA structure upon deletion of the s2m element, as found in the majority of the SARS-CoV-2 viruses isolated since March 2022. However, the lack of interactions between s2m and the rest of the 3′ UTR we observed in SARS-COV-2 by SHAPE-MaP and DMS-MaPseq may differ considerably between different viruses and virus families.

Given the genetic diversity within the coronavirus family, it is possible that the s2m is redundant for SARS-CoV-2, while it is important for other coronaviruses. A recent study investigated the requirement for s2m in avian infectious bronchitis virus (IBV) infection (30). Results showed that deletion of s2m in an infectious clone of IBV impacted viral replication *in vitro*, and RNA modeling suggested that the deletion of the s2m region impacted other structures in the 3′ UTR of IBV. *In vitro* passaging of the s2m deletion virus resulted in a 36-nucleotide insertion, composed of a duplication of the flanking sequences, in place of the deleted s2m region. We did not observe any sequence changes in the s2m region after *in vitro* passage of the CoV-2-s2m$^{\Delta2-42}$ virus, suggesting that there are structural and functional differences in the 3′ UTRs between these two coronaviruses. Additional studies will be needed to elucidate the differences in s2m dependence between different viruses.

Limitations of the study. We note several limitations of our study. (i) The RNA structure analyses were done on purified virus and not on viral genomes inside immunocompetent cells. While our structure predictions of the 3′ UTR are similar to those obtained from infected cells, it is possible that the SHAPE-MaP and DMS-MaPseq reactivities are different in primary human nasal and bronchial epithelial cells. (ii) SHAPE-MaP and DMS-MaPseq are low-resolution probing methods that average out the reactivity of particular nucleotides, potentially obscuring alternative or higher-order structures. (iii) The transmission potential of CoV-2-s2m-WT and s2m mutant viruses was not assessed. Airborne transmission of the CoV-2-s2m-WT viruses is inefficient, and while we did detect airborne transmission of both CoV-2-s2m-WT and mutant virus, the results were inconclusive and therefore not included here. (iv) It is possible that the significance of the s2m element is cell line or host dependent. While we did not test primary differentiated airways' epithelial cell cultures or additional animal models (mouse or nonhuman primates), the emergence and dominance of SARS-CoV-2 viruses lacking the s2m element supports our conclusion that s2m is dispensable for the SARS-CoV-2 life cycle. (v) The impact of the s2m deletion and mutations on lung pathology was not fully assessed. We opted to inoculate the animals with a relatively low dose to allow the detection of replication differences. However, these lower doses do not induce large amounts of weight loss or lung pathology. Future studies could detect transcriptional differences in the lungs of hamsters infected with CoV-2-s2m-WT or s2m deletion viruses in order to identify a role of s2m in modulating host responses. (vi) We used RT-PCR to measure the ratio of CoV-2-s2m-WT to CoV-2-s2m$^{\Delta2-42}$ virus. Since the amplicon of the CoV-2-s2m$^{\Delta2-42}$ is 41 nucleotides shorter (10.5%), it is possible that the outcome is affected by PCR amplification bias of the shorter PCR product (CoV-2-s2m$^{\Delta2-42}$). While we believe this impact to be minimal, we used relative fitness, as determined by the ratio of CoV-2-s2m-WT to CoV-2-s2m$^{\Delta2-42}$ in the tissue divided by the ratio of the inoculum, to determine the effect of the s2m deletion on viral fitness. Therefore, any PCR amplification bias is accounted for.

Overall, we have found that the s2m element is not critical for the SARS-CoV-2 virus life cycle *in vitro* or viral fitness *in vivo*. Further studies are needed to define the mechanistic basis as to why a highly conserved RNA element has no critical functional roles in the viral life cycle.

MATERIALS AND METHODS

Cell culture conditions. BHK-21 cells were maintained in Dulbecco's modified Eagle medium (DMEM) with L-glutamine (Gibco) supplemented with 10% fetal bovine serum (FBS) and 100 units/mL penicillin-streptomycin and incubated at 37°C and 5% CO_2. Vero cells expressing human TMPRSS2 (Vero-hTMPRSS2) (31) or human ACE2 and human TMPRSS2 (Vero-hACE2-hTMPRSS2, a gift from Graham and Creanga at NIH) and BSR cells (a clone of BHK-21) were cultured and maintained in DMEM supplemented with 5% FBS and 100 units/mL of penicillin and streptomycin. Vero-hTMPRSS2 and Vero-hACE2-hTMPRSS2 cells were maintained by selection with 5 μg/mL blasticidin and 10 μg/mL puromycin, respectively.

SARS-CoV-2 reverse genetics system. All work with potentially infectious SARS-CoV-2 particles was conducted under enhanced biosafety level 3 (BSL-3) conditions and approved by the institutional biosafety committee of Washington University in St. Louis. The prototypic SARS-CoV-2 genome (reference genome, GenBank accession no. NC_045512) was split into 7 fragments named A to G, and each DNA fragment was commercially synthesized (GenScript). A T7 promoter sequence was introduced at the 5′ end of fragment A, and a poly(A) sequence of 22 adenosines was introduced at the 3′ end of fragment G. In addition, NotI and SpeI sites were introduced at the 5′ end of fragment A before the T7 promoter and the 3′ end of fragment G after the poly(A) sequence, respectively. Details for the reverse genetics system design are available upon request. To ensure seamless assembly of the full virus DNA genome, the 3′ end of fragment A, both ends of fragments B to F, and the 5′ end of fragment G were appended by class II restriction enzyme recognition sites (BsmBI and BsaI). Fragments A and C to G were cloned into plasmid pUC57 vector and amplified in the *Escherichia coli* DH5α strain. The bacteria-toxic fragment B was cloned into low-copy-inducible BAC vector pCCI and amplified through plasmid induction in EPI300. Low-copy plasmids were extracted by NucleoBond Xtra Midi kit (Macherey-Nagel), and the other plasmids were extracted by plasmid midi kit (Qiagen) according to the manufacturers' protocols. To generate mutant s2m sequences, the SARS-CoV-2 s2m sequence in the G fragment was mutated by site-directed mutagenesis (Fig. 1B). Mutations and deletions in the s2m element were confirmed by Sanger sequencing. To assemble the full-length SARS-CoV-2 genome, each DNA fragment in plasmid was digested by corresponding restriction enzymes, DNA fragments were recovered by gel purification columns (New England Biolabs), and the seven fragments were ligated at equal molar ratios with 10,000 units of T4 ligase (New England Biolabs) in 100 μL at 16°C overnight. The final ligation product was incubated with proteinase K in the presence of 10% SDS for 30 min, extracted twice with equal volume of phenol-chloroform-isoamyl alcohol (25:24:1; Thermo Fisher), isopropanol precipitated, air-dried, and resuspended in RNase/DNase-free water. The DNA was analyzed on a 0.6% agarose gel. Full-length genomic SARS-CoV-2 RNA was *in vitro* transcribed using mMESSAGE mMACHINE T7 Ultra transcription kit (Invitrogen) following the manufacturer's protocol. Four micrograms of DNA template was added to the reaction mixture, supplemented with GTP (7.5 μL per 50-μL reaction mixture). *In vitro* transcription was done overnight at 32°C. Afterward, the template DNA was removed by digestion with Turbo DNase for 30 min at 37°C. The *in vitro* transcript (IVT RNA) mixture was used directly for electroporation. To further enhance rescue of recombinant virus, we also generated SARS-CoV-2 nucleocapsid (N) gene RNA. The SARS-CoV-2 N gene was PCR amplified from plasmid pUC57-SARS-CoV-2-N (GenScript) using forward primers with T7 promoter and reverse primers with poly(T)34 sequences. The N gene PCR product was gel purified and used as the template for *in vitro* transcription using the same mMESSAGE mMACHINE T7 transcription kit with 1 μg of DNA template and 1 μL of supplemental GTP in a 20-μL reaction volume. For SARS-CoV-2 IVT RNA electroporation, low-passage-number BHK-21 cells were trypsinized and resuspended in cold phosphate-buffered saline (PBS) at 0.5×10^7 cells/mL. A total of 20 μg of SARS-CoV-2 IVT RNA and 20 μg of N gene *in vitro* transcript were added to resuspended BHK-21 cells in 2-mm-gap cuvettes and electroporated with settings at 850 V and 25 μF and infinite resistance three times with about 5-s intervals in between pulses. The electroporated cells were allowed to rest for 10 min at room temperature and were then cocultured with Vero-hACE2-hTMPRSS2 cells at a 1:1 ratio in a T75 culture flask. Cell culture medium was changed to DMEM with 2% FBS the next day. Cytopathic effect (CPE) was monitored for 5 days, and cell culture supernatants were harvested for virus titration.

Propagation of a clinical isolate of SARS-CoV-2 containing a deletion of the s2m. As part of ongoing SARS-CoV-2 variant surveillance, a random set of RT-PCR-positive respiratory secretions from the Barnes Jewish Hospital Clinical microbiology laboratory was subjected to whole-genome sequencing using the ARTIC primer amplicon strategy (32). This study was approved by the Washington University Human Research Protection Office (no. 202004259). From the sequences generated, one genome (2019-nCoV/WUSTL_000226/2020; GenBank accession no. OM831956) had a 27-nucleotide deletion that removed 22 nucleotides from the 3′ end of the s2m element (Fig. 1C). The spike protein of this virus harbored L18F, D614G, and E780Q mutations, demonstrating that this virus belonged to the original B.1 lineage of SARS-CoV-2. This virus was expanded twice on Vero-hACE2-hTMPRSS2 cells, and the virus titer was determined by plaque assay. The P2 of 2019-nCoV/WUSTL_000226/2020 was sequenced by next-generation sequencing (NGS) to confirm the presence of the s2m deletion and rule out any tissue culture adaptations in the rest of the genome.

SARS-CoV-2 growth curve and titration assays. Vero-hTMPRSS2 cells were grown to confluence. Cells were inoculated with a multiplicity of infection (MOI) of 0.001 of recombinant CoV-2-s2m-WT or mutant SARS-CoV-2, and culture supernatant was collected at 1, 24, 48, and 72 hpi and saved for viral quantification by plaque assay on Vero-hACE2-hTMPRSS2 cells in 24-well plates as described (33).

SARS-CoV-2 Syrian hamster infection model. Animal studies were carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocols were approved by the Institutional Animal Care and Use Committee at the Washington University School of Medicine (assurance no. A3381-01). Five- to 6-week-old male hamsters were obtained from Charles River Laboratories and housed at Washington University. All animals were housed in the specific-pathogen-free barrier and cared for under U.S. Department of Agriculture (USDA) guidelines for laboratory animals. Hamsters were housed individually and allowed to receive food and water *ad libitum* before and during the experiment. Upon arrival at the facility, the animals were randomly assigned to experimental groups. Next, the animals were sedated with isoflurane and inoculated via the intranasal route with 1,000 PFU of the recombinant CoV-2-s2m-WT or s2m mutant SARS-CoV-2 viruses under enhanced biosafety level 3 (BSL-3) conditions (34). Three mock-infected control animals were included and inoculated with PBS. Animal weights were measured daily for the duration of the experiment. Three and 6 days after the inoculation, the animals were euthanized by carbon dioxide overdose, and lung tissues and nasal washes were collected. The nasal wash was performed with 1.0 mL of PBS containing 1% bovine serum albumin (BSA), clarified by centrifugation for 10 min at $2,000 \times g$, and stored at $-80°C$. The left lung lobe was homogenized in 1.0 mL DMEM, clarified by centrifugation ($1,000 \times g$ for 5 min), and used for viral titer analysis by plaque assay and RT-qPCR using primers and probes targeting the N gene. For viral RNA quantification, RNA was extracted using the RNA isolation kit (Omega Bio-Tek). SARS-CoV-2 RNA levels were measured by one-step quantitative reverse transcriptase PCR (RT-qPCR) TaqMan assay as described previously using a SARS-CoV-2 nucleocapsid (N)-specific primer/probe set from the Centers for Disease Control and Prevention (F primer, GACCCCAAAATCAGCGAAAT; R primer, TCTGGTTACTGCCAGTTGAATCTG; probe, 5′-FAM/ACCCCGCATTACGTTTGGTGGACC/3′-ZEN/IBFQ) (35). Viral RNA was expressed as N gene copy numbers per milligram for lung tissue homogenates or milliliter for nasal swabs, based on a standard included in the assay, which was created via *in vitro* transcription of a synthetic DNA molecule containing the target region of the N gene. The *in vivo* challenge studies were performed twice independently, and a total of six hamsters per virus per day were aggregated and plotted in Fig. 3B to G. No infectious virus was detected in the mock-infected animals, and the amount of viral RNA N gene was below the limit of detection of the assay.

SARS-CoV-2 competition assays in Syrian hamster. Ten golden Syrian hamsters were inoculated intranasally with a total of 1,000 PFU of a 1:1 mixture of CoV-2-s2m-WT and CoV-2-s2m$^{\Delta2–42}$ in 100 μL volume. Three days after inoculation, the hamsters were sacrificed, and the nasal washes, left lobes, and lungs were harvested and homogenized. One hundred microliters of the nasal wash or lung homogenates were added in 300 μL of TRK lysis buffer from E.Z.N.A. Total RNA kit (Omega Bio-Tek) for RNA isolation. For quantitation of the virion-associated RNA levels in the inocula, we prepared five inocula separately and treated them with RNase A (Thermo Scientific) to remove free viral genomic RNAs and subgenomic RNAs. These five RNase-treated inocula were added to TRK lysis buffer for RNA isolation and quantification as described below to get the initial mutant to CoV-2-s2m-WT ratio. cDNA was synthesized from the extracted RNA with random hexamers using SuperScript IV first-strand synthesis kit (Invitrogen) following the manufacturer's protocol. PCR covering the virus s2m region was performed on cDNA samples for 40 cycles with primers HJ551-S2UTRF, 5′-CTCCAAACAATTGCAACAATC-3′, and HJ552-S2UTRR, 5′-GTCATTCTCCTAAGAAGCTATTAAAATC-3′ using the high-fidelity AccuPrime Taq DNA polymerase (Invitrogen) following the manufacturer's protocol. The presence of two different-size amplicons (389 bp for CoV-2-s2m-WT virus, 348 bp for CoV-2-s2m$^{\Delta2–42}$ virus) was verified by gel electrophoresis, and then PCR products were purified with QIAquick PCR purification kit (Qiagen) and quantified with Qubit 4 fluorometer (Invitrogen). Clean PCR products were diluted and subjected to 2100 Bioanalyzer (Agilent) analysis following the manufacturer's protocol to get the molar quantities of the two different-size amplicons. Mutant-to-CoV-2-s2m-WT ratios were calculated based on quantitation readout, and the relative replicative fitness is defined and calculated by dividing the final mutant-to-CoV-2-s2m-WT ratio in hamster samples by the initial mutant to CoV-2-s2m-WT ratio in the inocula.

Structural analysis. Vero-hTMPRSS2 cells were grown to confluence, and the cells were inoculated with an MOI of 0.01 of recombinant CoV-2-s2m-WT or mutant CoV-2-s2m$^{\Delta2–42}$ virus. Supernatant was collected at 24 hours postinfection (hpi), and the virus was purified from the supernatant using polyethylene glycol (PEG) precipitation method (36). The purified virus pellet was resuspended in a buffer (0.05 M HEPES, pH 8, 0.1 M NaCl, and 0.0001 M EDTA for SHAPE-MaP and 0.3 M HEPES, pH 8, and 0.1 M NaCl for DMS-MaPseq).

(i) SHAPE-MaP. For selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP), the resuspended viruses were divided into 3 reactions (modified sample, control sample, and denatured sample). For the modified sample, 2-methylnicotinic acid imidazolide (NAI) from 1 M stock (in dimethyl sulfoxide [DMSO]) was added at a final concentration of 100 mM. For the control sample, the corresponding amount of DMSO was added. Samples were then incubated at 37°C for 15 min, followed by quenching of NAI through the addition of dithiothreitol (DTT) at a final concentration of 0.5 M. For denaturing control, resuspended viruses were set aside without any treatment. TRK lysis buffer was added to the above-described reaction mixtures to lyse the virus. Total RNA was extracted using Zymo RNA Clean and Concentrator-5 kit (Zymo Research). The denatured control RNA sample was incubated at 95°C for 1 min and then treated with 100 mM NAI for 1 min at 95°C, and the reaction was quenched with DTT as described previously. For denatured sample, RNA was again purified using the Zymo RNA clean

and concentrator-5 kit (Zymo Research). Sequencing library preparation was performed according to the amplicon workflow as described previously (37). The primers were designed tiling 3′ UTR across the SARS-CoV-2 genome, 5′-GCAGACCACACAAGGC-3′ (forward) and 5′-CGTCATTCTCCTAAGAAGCTA-3′ (reverse). The RNA was reverse transcribed using the specific primer with SuperScript II (Invitrogen) in MaP buffer (50 mM Tris-HCl [pH 8.0], 75 mM KCl, 6 mM $MnCl_2$, 10 mM DTT, and 0.5 mM deoxynucleoside triphosphate). Amplicons tiling the 3′ UTR SARS-CoV-2 genome were generated using Q5 Hot Start high-fidelity DNA polymerase (catalog no. M0492S), 3′ UTR-specific forward and reverse PCR primers, and 8 μL of purified cDNA. The Nextera XT DNA library preparation kit (Illumina) was used to prepare the sequencing libraries. Final PCR amplification products were size-selected using Agencourt AMPure XP beads (Beckman Coulter). Libraries were quantified using a Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (Thermo Fisher; catalog no. Q32851) to determine the concentration, and quality was assessed with the Agilent high-sensitivity DNA kit (Agilent Technologies) on a Bioanalyzer 2100 system (Agilent Technologies) to determine average library member size and accurate concentration. The libraries were sequenced (2 × 150 bp) on a MiniSeq System (Illumina). Sequencing reads were aligned with the reference sequences, and SHAPE-MaP reactivity profiles for each position were calculated using ShapeMapper 2.15 (38) with default parameters. All SHAPE-MaP reactivities were normalized to an approximate 0 to 2 scale by dividing the SHAPE-MaP reactivity values by the mean reactivity of the 10% most highly reactive nucleotides after excluding outliers (defined as nucleotides with reactivity values that are >1.5 times the interquartile range). High SHAPE-MaP reactivities above 0.7 indicate more flexible (that is, single-stranded) regions of RNA, and low SHAPE-MaP reactivities below 0.3 indicate more structurally constrained (that is, base-paired) regions of RNA.

(ii) DMS-MaPseq. For dimethyl sulfate mutational profiling and sequencing (DMS-MaPseq), the resuspended viruses were divided into 2 reactions (modified and control samples). For modified sample, 2% (vol/vol) dimethyl sulfide (DMS) was added, mixed thoroughly, and incubated immediately at 37°C for 5 min before quenching with 100 μL 30% β-mercaptoethanol in PBS. The control sample was prepared similarly only without addition of DMS. TRK lysis buffer was added to the above-described reaction mixtures to lyse the virus. Total RNA was extracted using Zymo RNA clean and concentrator-5 kit (Zymo Research). For reverse transcription, the 11.5 μL RNA was supplemented with 4 μL 5× first strand buffer (Thermo Fisher Scientific), 1 μL 10 μM reverse primer, 1 μL deoxynucleoside triphosphate (dNTP), 1 μL 0.1 M DTT, 1 μL RNaseOut, and 0.5 μL MarathonRT. The reverse transcription reaction mixture was incubated at 42°C for 3 h. One microliter of RNase H was added to each reaction mixture and incubated at 37°C for 20 min to degrade the RNA. cDNA was purified using QIAquick PCR purification kit (catalog no. 28104). dsDNA was prepared as described above but with a cocktail of 2 forward primers, 5′-GCAGACCACACAAGGC-3′ and 5′-ACGTTTTCGCTTTTCCG-3′. NEBNext Ultra II DNA library prep kit for Illumina (New England Biolabs; catalog no. E7645S) was used to prepare the sequencing libraries as per the manufacturer's instructions. The cleanup, quantification, and sequencing of libraries were performed as described above. DMS-MaPseq reactivity profiles were calculated by aligning the sequencing reads to reference sequences using DREEM Webserver (39). FastQC was used to assess the quality of fastq files, TrimGalore was used to remove adapter sequences, and Bowtie2 was used to align the reads to sequences, which were integrated into DREEM. DREEM mapped the reads and converted them into BitVectors based on the mutation rate (if the mutation rate is >0.5%, it converts to 1; otherwise, matches convert to 0). The BitVector files were then used to count mutations and normalized by sequencing depth in order to provide normalized DMS reactivity.

(iii) Experimentally informed secondary structure modeling. To computationally predict the RNA secondary structure of the 3′ UTR, we have removed the primer binding sequences from our analysis. The SHAPE-MaP and DMS-MaPseq reactivity profiles obtained through ShapeMapper and the DREEM Webserver, respectively, were used as constraints to predict the secondary structure using ShapeKnots (40) with default settings. In order to calculate the correlation between the WT and mutant, the Pearson correlation coefficient was calculated by comparing their nucleotide reactivity. SuperFold (41) with experimental data restraints was used to predict consensus secondary structure prediction with base-pairing probabilities and Shannon entropy.

Sequencing of reverse genetics rescued SARS-CoV-2 virus. For SARS-CoV-2 genome sequencing, 200 μL of supernatant containing virus was collected and added to 600 μL TRK lysis buffer plus beta-mercaptoethanol (BME) according to the E.Z.N.A. Total RNA kit I (Omega Bio-Tek). Total RNA was extracted according to the manufacturer's protocol and eluted into RNase/DNase-free H_2O and used for library construction. rRNA was removed from total RNA by Ribo-Zero depletion (Illumina). Indexed sequencing libraries were prepared using TruSeq RNA library preparation kit (Illumina), pooled, and then sequenced using the Illumina NextSeq system. The raw sequence data were analyzed using the LoFreq pipeline to call the mutations in the entire virus genome (42). In brief, Illumina sequencing fastq data were aligned by BWA with the SARS-CoV-2 reference genome sequence (GenBank accession no. NC_045512) after indexing to generate the aligned SAM and BAM files. The read group was added by Picard after the aligned BAM file sorting and indexing with SAMtools, and then duplicates were removed by Picard MarkDuplicates. Local realignment was achieved by GATK3, and then variants were called by LoFreq to generate the mutant report file.

Multiple-sequence alignment. Representative s2m sequences were aligned in MEGA X using MUSCLE (43). Sequences were visualized using Jalview 2 (44). Representative sequences included SARS-CoV-1 Urbani (GenBank accession no. AY278741), SARS-CoV-2 (GenBank accession no. NC_045512), bat coronavirus RaTG13 (GenBank accession no. MN996532), bat coronavirus BANAL-52 (GenBank accession no. MZ937000), bat SARS coronavirus Rp3 (GenBank accession no. DQ071615), bat SARS-like coronavirus YNLF-34C (GenBank accession no. KP886809), SARS coronavirus Civet PC4-136 (GenBank accession no.

AY613949), pangolin coronavirus PCoV GX-P3B (GenBank accession no. MT072865), avian infectious bronchitis virus (GenBank accession no. NC_001451), and thrush coronavirus HKU12-600 (GenBank accession no. NC_011549).

Analysis of SARS-CoV-2 s2m sequences using the NCBI database. SARS-CoV-2 complete genome sequences deposited between January 2020 and December 2022 were downloaded from NCBI as a single fasta file. The fasta data set was processed with a published SARS-CoV-2 freebayes pipeline to call all the variants along the genome (45). In brief, the single SARS-CoV-2 complete genome fasta file was decomposed into individual FASTA genome sequences. Then, each FASTA genome was aligned individually against the SARS-CoV-2 reference sequence (GenBank accession no. NC_045512) using minimap2 (46). Variant calling was performed on each BAM file using freebayes variant caller to produce the VCF files. Variants in the s2m position 32 and deletion of the s2m Δ8-33 for all the VCF files were extracted and reported.

Statistical analysis. Data were graphed using Prism 9.3.1 (GraphPad). Parametric and nonparametric comparisons were made where appropriate. *Post hoc* testing of Kruskal-Wallis comparisons was completed with correction of multiple comparisons by Dunn's multiple test correction. For multistep growth curves, mutants were compared to CoV-2-s2m-WT by logarithmically transforming the data and comparing using a two-way ANOVA in Prism after excluding the 1-h time point. Hamster weight data were analyzed by mixed-effect modeling with Geisser-Greenhouse correction in Prism. Adjusted P values of ≤ 0.05 were considered significant.

Data availability. The fasta files for SHAPE-MaP and DMS-MaPseq can be found under BioProject accession no. PRJNA936272.

REFERENCES

1. V'Kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 19:155–170. https://doi.org/10.1038/s41579-020-00468-6.

2. Goebel SJ, Hsue B, Dombrowski TF, Masters PS. 2004. Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. J Virol 78:669–682. https://doi.org/10.1128/jvi.78.2.669-682.2004.

3. Yang D, Leibowitz JL. 2015. The structure and functions of coronavirus genomic 3' and 5' ends. Virus Res 206:120–133. https://doi.org/10.1016/j.virusres.2015.02.025.

4. Tengs T, Jonassen CM. 2016. Distribution and evolutionary history of the mobile genetic element s2m in coronaviruses. Diseases 4:27. https://doi.org/10.3390/diseases4030027.

5. Dhama K, Khan S, Tiwari R, Sircar S, Bhat S, Malik YS, Singh KP, Chaicumpa W, Bonilla-Aldana DK, Rodriguez-Morales AJ. 2020. Coronavirus disease 2019–COVID-19. Clin Microbiol Rev 33:e00028-20. https://doi.org/10.1128/CMR.00028-20.

6. Tengs T, Kristoffersen AB, Bachvaroff TR, Jonassen CM. 2013. A mobile genetic element with unknown function found in distantly related viruses. Virol J 10:132. https://doi.org/10.1186/1743-422X-10-132.

7. Robertson MP, Igel H, Baertsch R, Haussler D, Ares M, Scott WG. 2005. The structure of a rigorously conserved RNA element within the SARS virus genome. PLoS Biol 3:e5. https://doi.org/10.1371/journal.pbio.0030005.

8. Kofstad T, Jonassen CM. 2011. Screening of feral and wood pigeons for viruses harbouring a conserved mobile viral element: characterization of novel astroviruses and picornaviruses. PLoS One 6:e25964. https://doi.org/10.1371/journal.pone.0025964.

9. Tengs T, Delwiche CF, Monceyron Jonassen C. 2021. A genetic element in the SARS-CoV-2 genome is shared with multiple insect species. J Gen Virol 102:001551. https://doi.org/10.1099/jgv.0.001551.

10. Wacker A, Weigand JE, Akabayov SR, Altincekic N, Bains JK, Banijamali E, Binas O, Castillo-Martinez J, Cetiner E, Ceylan B, Chiu L-Y, Davila-Calderon J, Dhamotharan K, Duchardt-Ferner E, Ferner J, Frydman L, Fürtig B, Gallego J, Grün JT, Hacker C, Haddad C, Hähnke M, Hengesbach M, Hiller F, Hohmann KF, Hymon D, de Jesus V, Jonker H, Keller H, Knezic B, Landgraf T, Löhr F, Luo L, Mertinkus KR, Muhs C, Novakovic M, Oxenfarth A, Palomino-Schätzlein M, Petzold K, Peter SA, Pyper DJ, Qureshi NS, Riad M, Richter C, Saxena K, Schamber T, Scherf T, Schlagnitweit J, Schlundt A, Schnieders R, et al. 2020. Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. Nucleic Acids Res 48:12415–12435. https://doi.org/10.1093/nar/gkaa1013.

11. Huston NC, Wan H, Strine MS, de Cesaris Araujo Tavares R, Wilen CB, Pyle AM. 2021. Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. Mol Cell 81:584–598.e585. https://doi.org/10.1016/j.molcel.2020.12.041.

12. Lan TCT, Allan MF, Malsick LE, Woo JZ, Zhu C, Zhang F, Khandwala S, Nyeo SSY, Sun Y, Guo JU, Bathe M, Näär A, Griffiths A, Rouskin S. 2022.

Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. Nat Commun 13:1128. https://doi.org/10.1038/s41467-022-28603-2.

13. Yang SL, DeFalco L, Anderson DE, Zhang Y, Aw JGA, Lim SY, Lim XN, Tan KY, Zhang T, Chawla T, Su Y, Lezhava A, Merits A, Wang L-F, Huber RG, Wan Y. 2021. Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. Nat Commun 12:5113. https://doi.org/10.1038/s41467-021-25357-1.

14. Zhao J, Qiu J, Aryal S, Hackett JL, Wang J. 2020. The RNA architecture of the SARS-CoV-2 3'-untranslated region. Viruses 12:1473. https://doi.org/10.3390/v12121473.

15. Manfredonia I, Nithin C, Ponce-Salvatierra A, Ghosh P, Wirecki TK, Marinus T, Ogando NS, Snijder EJ, van Hemert MJ, Bujnicki JM, Incarnato D. 2020. Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. Nucleic Acids Res 48:12436–12452. https://doi.org/10.1093/nar/gkaa1053.

16. Lulla V, Wandel MP, Bandyra KJ, Ulferts R, Wu M, Dendooven T, Yang X, Doyle N, Oerum S, Beale R, O'Rourke SM, Randow F, Maier HJ, Scott W, Ding Y, Firth AE, Bloznelyte K, Luisi BF. 2021. Targeting the conserved stem loop 2 motif in the SARS-CoV-2 genome. J Virol 95:e00663-21. https://doi.org/10.1128/JVI.00663-21.

17. Imperatore JA, Cunningham CL, Pellegrene KA, Brinson RG, Marino JP, Evanseck JD, Mihailescu MR. 2022. Highly conserved s2m element of SARS-CoV-2 dimerizes via a kissing complex and interacts with host miRNA-1307-3p. Nucleic Acids Res 50:1017–1032. https://doi.org/10.1093/nar/gkab1226.

18. Vahed M, Vahed M, Sweeney A, Shirazi FH, Mirsaeidi M. 2020. Mutation in position of 32 (G>U) of S2M differentiate human SARS-CoV2 from bat coronavirus. bioRxiv. https://doi.org/10.1101/2020.09.02.280529.

19. Rangan R, Watkins AM, Chacon J, Kretsch R, Kladwang W, Zheludev IN, Townley J, Rynge M, Thain G, Das R. 2021. De novo 3D models of SARS-CoV-2 RNA elements from consensus experimental secondary structures. Nucleic Acids Res 49:3092–3108. https://doi.org/10.1093/nar/gkab119.

20. Chiara M, Horner DS, Gissi C, Pesole G. 2021. Comparative genomics reveals early emergence and biased spatiotemporal distribution of SARS-CoV-2. Mol Biol Evol 38:2547–2565. https://doi.org/10.1093/molbev/msab049.

21. Ryder SP, Morgan BR, Coskun P, Antkowiak K, Massi F. 2021. Analysis of emerging variants in structured regions of the SARS-CoV-2 genome. Evol Bioinform Online 17:11769343211014167. https://doi.org/10.1177/11769343211014167.

22. Gilbert C, Tengs T. 2021. No species-level losses of s2m suggests critical role in replication of SARS-related coronaviruses. Sci Rep 11:16145. https://doi.org/10.1038/s41598-021-95496-4.

23. Yeh TY, Contreras GP. 2020. Emerging viral mutants in Australia suggest RNA recombination event in the SARS-CoV-2 genome. Med J Aust 213:44.e41–44.e1. https://doi.org/10.5694/mja2.50657.

24. Frye CJ, Shine M, Makowski JA, Kensinger AH, Cunningham CL, Milback EJ, Evanseck JD, Lackey PE, Mihailescu MR. 2023. Bioinformatics analysis of the s2m mutations within the SARS-CoV-2 Omicron lineages. J Med Virol 95:2022. https://doi.org/10.1002/jmv.28141.

25. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, Du Plessis L, Pybus OG. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 5:1403–1407. https://doi.org/10.1038/s41564-020-0770-5.

26. Monceyron C, Grinde B, Jonassen TO. 1997. Molecular characterisation of the 3'-end of the astrovirus genome. Arch Virol 142:699–706. https://doi.org/10.1007/s007050050112.

27. Temmam S, Vongphayloth K, Baquero E, Munier S, Bonomi M, Regnault B, Douangboubpha B, Karami Y, Chrétien D, Sanamxay D, Xayaphet V, Paphaphanh P, Lacoste V, Somlor S, Lakeomany K, Phommavanh N, Pérot P, Dehan O, Amara F, Donati F, Bigot T, Nilges M, Rey FA, van der Werf S, Brey PT, Eloit M. 2022. Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. Nature 604:330–336. https://doi.org/10.1038/s41586-022-04532-4.

28. Goebel SJ, Miller TB, Bennett CJ, Bernard KA, Masters PS. 2007. A hypervariable region within the 3' cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. J Virol 81:1274–1287. https://doi.org/10.1128/JVI.00803-06.

29. Gumna J, Antczak M, Adamiak RW, Bujnicki JM, Chen SJ, Ding F, Ghosh P, Li J, Mukherjee S, Nithin C, Pachulska-Wieczorek K, Ponce-Salvatierra A, Popenda M, Sarzynska J, Wirecki T, Zhang D, Zhang S, Zok T, Westhof E, Miao Z, Szachniuk M, Rybarczyk A. 2022. Computational pipeline for reference-free comparative analysis of RNA 3D structures applied to SARS-CoV-2 UTR models. Int J Mol Sci 23:2022. https://doi.org/10.3390/ijms23179630.

30. Keep S, Dowgier G, Lulla V, Britton P, Oade M, Freimanis G, Tennakoon C, Jonassen CM, Tengs T, Bickerton E. 2023. Deletion of the s2m RNA structure in the avian coronavirus infectious bronchitis virus and human astrovirus results in sequence insertions. J Virol 97:e00038-23. https://doi.org/10.1128/jvi.00038-23.

31. Case JB, Rothlauf PW, Chen RE, Liu Z, Zhao H, Kim AS, Bloyet L-M, Zeng Q, Tahan S, Droit L, Ilagan MXG, Tartell MA, Amarasinghe G, Henderson JP, Miersch S, Ustav M, Sidhu S, Virgin HW, Wang D, Ding S, Corti D, Theel ES, Fremont DH, Diamond MS, Whelan SPJ. 2020. Neutralizing antibody and soluble ACE2 inhibition of a replication-competent VSV-SARS-CoV-2 and a clinical isolate of SARS-CoV-2. Cell Host Microbe 28:475–485.e475. https://doi.org/10.1016/j.chom.2020.06.021.

32. Tyson JR, James P, Stoddart D, Sparks N, Wickenhagen A, Hall G, Choi JH, Lapointe H, Kamelian K, Smith AD, Prystajecky N, Goodfellow I, Wilson SJ, Harrigan R, Snutch TP, Loman NJ, Quick J. 2020. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. bioRxiv. https://doi.org/10.1101/2020.09.04.283077.

33. Bricker TL, Darling TL, Hassan AO, Harastani HH, Soung A, Jiang X, Dai Y-N, Zhao H, Adams LJ, Holtzman MJ, Bailey AL, Case JB, Fremont DH, Klein R, Diamond MS, Boon ACM. 2021. A single intranasal or intramuscular immunization with chimpanzee adenovirus-vectored SARS-CoV-2 vaccine protects against pneumonia in hamsters. Cell Rep 36:109400. https://doi.org/10.1016/j.celrep.2021.109400.

34. Su W, Sia SF, Schmitz AJ, Bricker TL, Starr TN, Greaney AJ, Turner JS, Mohammed BM, Liu Z, Choy KT, Darling TL, Joshi A, Cheng KM, Wong AYL, Harastani HH, Nicholls JM, Whelan SPJ, Bloom JD, Yen H-L, Ellebedy AH, Boon ACM. 2021. Neutralizing monoclonal antibodies that target the spike receptor binding domain confer fc receptor-independent protection against SARS-CoV-2 infection in Syrian hamsters. mBio 12:e02395-21. https://doi.org/10.1128/mBio.02395-21.

35. U.S. Centers for Disease Control and Prevention. 2022. 2019-Novel coronavirus (2019-nCoV) real-time rRT-PCR panel primers and probes. https://www.who.int/docs/default-source/coronaviruse/uscdcrt-pcr-panel-primer-probes.pdf?sfvrsn=fa29cb4b_2. Accessed May, 2020.

36. Heyward JT, Klimas RA, Stapp MD, Obijeski JF. 1977. The rapid concentration and purification of influenza virus from allantoic fluid. Arch Virol 55:107–119. https://doi.org/10.1007/BF01314484.

37. Smola MJ, Rice GM, Busan S, Siegfried NA, Weeks KM. 2015. Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. Nat Protoc 10:1643–1669. https://doi.org/10.1038/nprot.2015.103.

38. Busan S, Weeks KM. 2018. Accurate detection of chemical modifications in RNA by mutational profiling (MaP) with ShapeMapper 2. RNA 24:143–148. https://doi.org/10.1261/rna.061945.117.

39. Jurich CP, Brivanlou A, Rouskin S, Yesselman JD. 2022. Web-based platform for analysis of RNA folding from high throughput chemical probing data. Nucleic Acids Res 50:W266–W271. https://doi.org/10.1093/nar/gkac435.

40. Hajdin CE, Bellaousov S, Huggins W, Leonard CW, Mathews DH, Weeks KM. 2013. Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. Proc Natl Acad Sci U S A 110:5498–5503. https://doi.org/10.1073/pnas.1219988110.

41. Siegfried NA, Busan S, Rice GM, Nelson JA, Weeks KM. 2014. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat Methods 11:959–965. https://doi.org/10.1038/nmeth.3029.

42. Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N. 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 40:11189–11201. https://doi.org/10.1093/nar/gks918.

43. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096.

44. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

45. Farkas C, Mella A, Turgeon M, Haigh JJ. 2021. A novel SARS-CoV-2 viral sequence bioinformatic pipeline has found genetic evidence that the viral 3' untranslated region (UTR) is evolving and generating increased viral diversity. Front Microbiol 12:665041. https://doi.org/10.3389/fmicb.2021.665041.

46. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.