Taylor & Francis
Taylor & Francis Group

BRIEF COMMUNICATION

🔓 OPEN ACCESS   Check for updates

Secondary structure of the SARS-CoV-2 5′-UTR

Zhichao Miao ![ID]a,b,c, Antonin Tidud, Gilbert Eriani ![ID]d, and Franck Martin ![ID]d*

aEuropean Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, UK; bTranslational Research Institute of Brain and Brain-Like Intelligence and Department of Anesthesiology, Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine, Shanghai, China; cNewcastle Fibrosis Research Group, Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; dArchitecture Et Réactivité De l'ARN, Université De Strasbourg, Institut De Biologie Moléculaire Et Cellulaire Du CNRS, Strasbourg, France

ABSTRACT

The SARS-CoV-2, a positive-sense single-stranded RNA Coronavirus, is a global threat to human health. Thus, understanding its life cycle mechanistically would be important to facilitate the design of antiviral drugs. A key aspect of viral progression is the synthesis of viral proteins by the ribosome of the human host. In Coronaviruses, this process is regulated by the viral 5′ and 3′ untranslated regions (UTRs), but the precise regulatory mechanism has not yet been well understood. In particular, the 5′-UTR of the viral genome is most likely involved in translation initiation of viral proteins. Here, we performed inline probing and RNase V1 probing to establish a model of the secondary structure of SARS-CoV-2 5′-UTR. We found that the 5′-UTR contains stable structures including a very stable four-way junction close to the AUG start codon. Sequence alignment analysis of SARS-CoV-2 variants 5′-UTRs revealed a highly conserved structure with few co-variations that confirmed our secondary structure model based on probing experiments.

Introduction

Coronaviruses are found to infect a large variety of animals and humans. Besides enteric diseases, they mainly cause severe respiratory defects sometimes leading to death [1]. The recently emerged SARS-CoV-2 belongs to the betacoronavirus genome, subgenus Sarbecovirus [2]. Its genome is a positive single-stranded RNA molecule (+)ssRNA (Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020). The genomic sequence of SARS-CoV-2 was determined at the end of 2019 [3]. The RNA genome is capped at the 5′ end and polyadenylated at the 3′ end. The genome encodes two long open reading frames (ORF1a and ORF1b) at the 5′ end and several ORFs that are expressed in the late phase of infection from subgenomic RNAs (sgRNAs) [4]. After cell entry, the translation of ORF1a and ORF1b from the whole ss(+) RNA are the first events of the infectious process. The translation of ORF1b requires a − 1 frameshifting event [5,6]. The polyprotein synthesized from ORF1a is processed into eleven non-structural proteins (NSP1-NSP11). The first one, NSP1 binds to the host small ribosomal subunit 40S and recruits a yet unidentified cellular nuclease that triggers the degradation of the host mRNAs, while viral RNA is being translated [7,8]. Thus, the virus specifically degrades the cellular mRNAs that are translated by the canonical cap-dependent translation mechanism. Other studies have shown that NSP1 is also able to prevent 48S ribosomal complex formation by another so far uncharacterized

mechanism [8,9]. Interestingly, it has been shown that IRES (Internal Ribosome Entry Site)-mediated translation initiation of class III- and IV–IRES (e.g., in Hepatitis C Virus (HCV) and Cricket paralysis virus (CrPV)) respectively, is immune to the NSP1 inhibition. Instead, the translation initiation driven by the encephalomyocarditis virus (EMCV) class II–IRES is efficiently inhibited by NSP1 [8]. The NSP1 binding site to the ribosomal 40S subunit has not yet been determined.

Since the SARS-CoV-2 genomic RNA is capped at the 5′ end, it is generally believed that its translation initiation is canonical and cap-dependent. However, two major observations provide hints that the translation mechanism of SARS-CoV-2 RNA is in fact mediated by an unconventional translation initiation mechanism rather than a canonical one. First, the secondary structure of the SARS-CoV-2 5′-UTR is likely to be complex in the proximity of the 5′ cap, based on the experimental SHAPE structure of the 5′-UTR of Mouse Hepatitis Virus (MHV), a Coronavirus belonging to the Embecovirus subgroup, and on related structural predictions of the 5′-UTR of the SARS-CoV Sarbecovirus [10]. RNA structures proximal to the cap are known to inhibit the recruitment of cap-binding translation factors (eIF4E and consequently eIF4F), thus indicating canonical cap-dependent translation improbable [11,12]. Second, after the translation of ORF1a, the rapidly produced NSP1 protein would shut down the canonical cap-dependent translation of cellular mRNAs. Yet, the

translation of SARS-CoV-2 proteins is not inhibited by NSP1. Indeed, class III- and class IV–IRES are immune to this NSP1-mediated inhibitory mechanism [8,13]. This suggests that translation initiation of SARS-CoV-2 may be less dependent on eIF4F. Previous structural studies from other coronavirus 5′-UTRs have shown that stable hairpin structures are found in the proximity of the cap structure. When these structures are present in the vicinity of the cap,
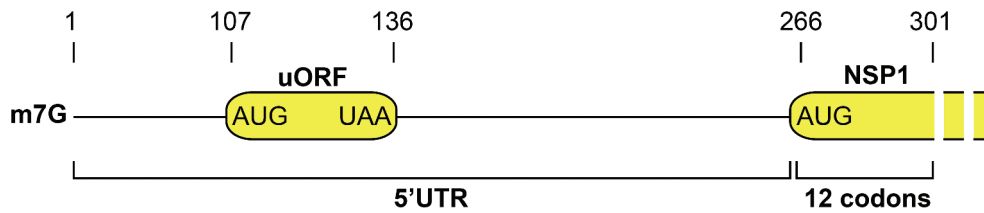

Figure 1. SARS-CoV-2 5′-UTR.
An RNA transcript containing the 301 5′ proximal nucleotides from SARS-CoV-2 is represented. The positions of the uORF and the N-terminal coding sequence of NSP1 are shown in green.
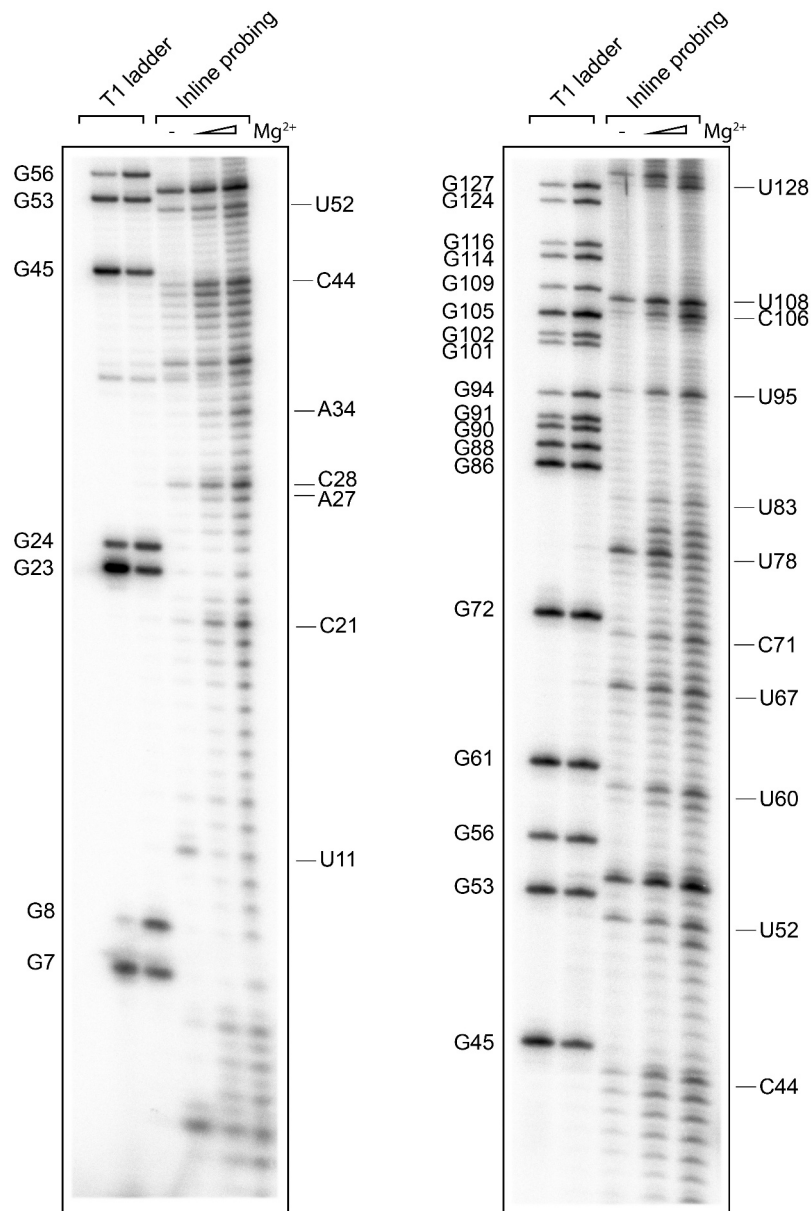

Figure 2. Inline probing of the SARS-CoV-2 5′-UTR from nucleotides 1 to 128.
Inline probing of the SARS-CoV-2 without Mg^{2+}, with 1 and 10 mM Mg^{2+}. The cuts are mapped using an RNase T1 denaturing ladder that cuts after G residues. The left panel shows the reactivity of nucleotides 1–56 and the right panel shows the reactivity of nucleotides 30–128.

translation initiation efficiency is significantly modulated by these secondary structures [11]. In both cap-dependent and cap-independent mechanisms, the secondary structure of the 5′-UTR is critical for translation initiation efficiency.

In order to better understand the translation initiation mechanism of viral translation during SARS-CoV-2 infection, the first step is to determine the secondary structure of the 5′-UTR. Here, we report the first experimental determination of SARS-CoV-2 5′-UTR structure using inline probing and RNAse V1 enzymatic probing.

Material and methods

Coronavirus complete genome sequence

The complete genome sequence of SARS-CoV-2 was downloaded from NCBI nucleotide database Genbank [14] MN908947.3.

Secondary structure probing

The RNA transcripts have been synthesized by *in vitro* transcription. The RNA transcripts were then separated on PAGE containing 8 M urea and purified by electro-elution using Bio Trap apparatus and Schleicher & Schuell membranes. The purified transcript was then ^{32}P-labelled by 5′ capping using the ScriptCap m^{7}G Capping System kit from CELLSCRIPT™. The RNA transcripts were probed directly after purification without any denaturation-renaturation step. Briefly, for inline probing, 50 000 cpm of radiolabelled RNA was incubated in 50 mM Tris-HCl pH 8.8, 100 mM KCl without MgCl$_2$ or with 1 or 10 mM MgCl$_2$ for 72 h at room temperature. The cuts in the RNA backbone were analysed on denaturing PAGE containing 8 M urea. For V1 probing, the RNA was incubated with serial dilutions of RNase V1 in order to have statistically one digestion cut per molecule for 10 min at room temperature as previously described [15]. The cuts were
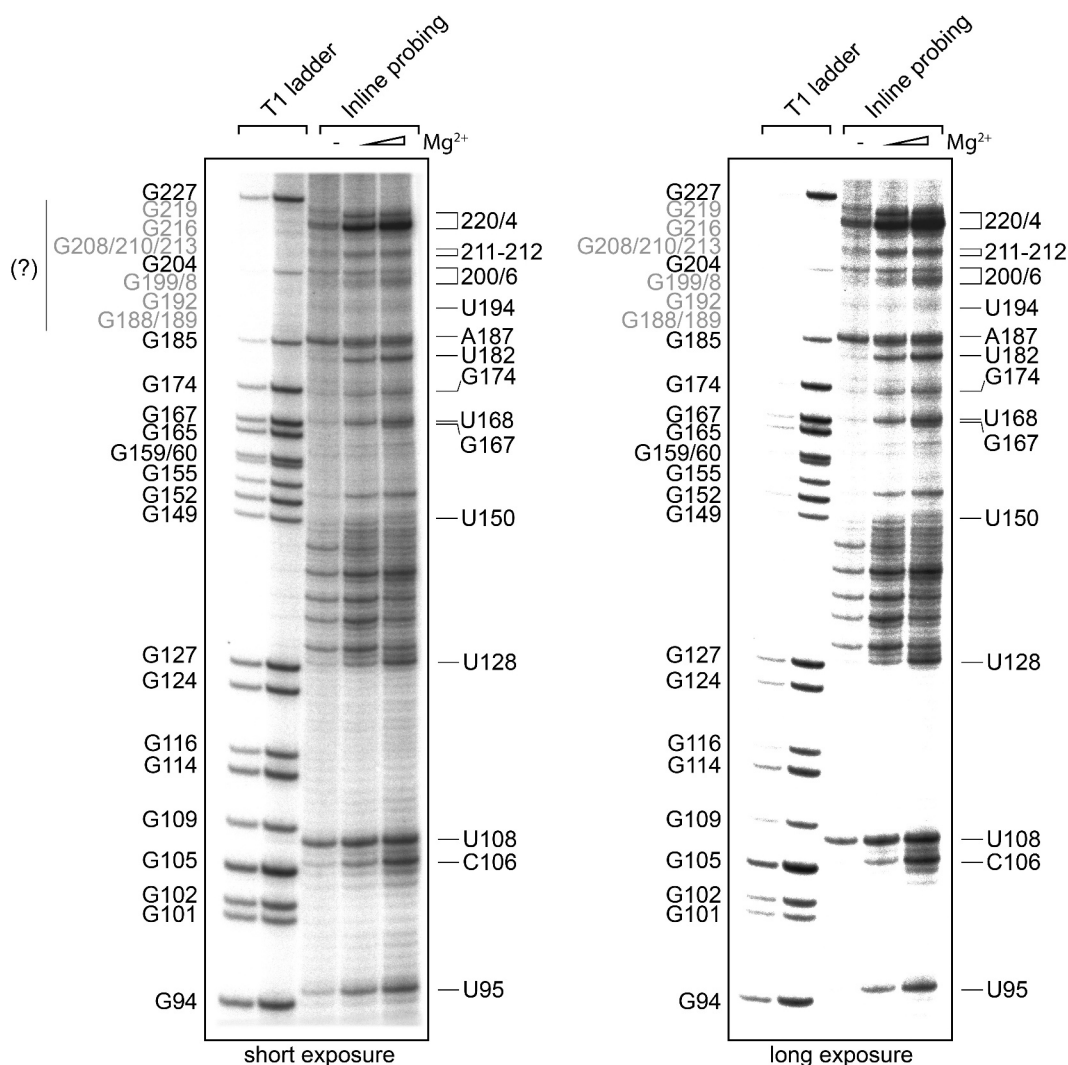

Figure 3. Inline probing of the SARS-CoV-2 5′-UTR from nucleotides 94 to 295.

(A) Inline probing of the SARS-CoV-2 without Mg^{2+}, with 1 and 10 mM Mg^{2+}. The cuts are mapped using an RNAse T1 denaturing ladder that cuts after G residues. The left panel shows the reactivity of nucleotides 94–227 and the right panel shows the same gel with a longer exposure. **(B)** Inline probing of the SARS-CoV-2 without Mg $^{2+}$, with 1 and 10 mM Mg^{2+}. The upper panel shows the reactivity of nucleotides 165–262, and the lower panels show the reactivity of nucleotides 220–295 (short exposure on the left and long exposure on the right).
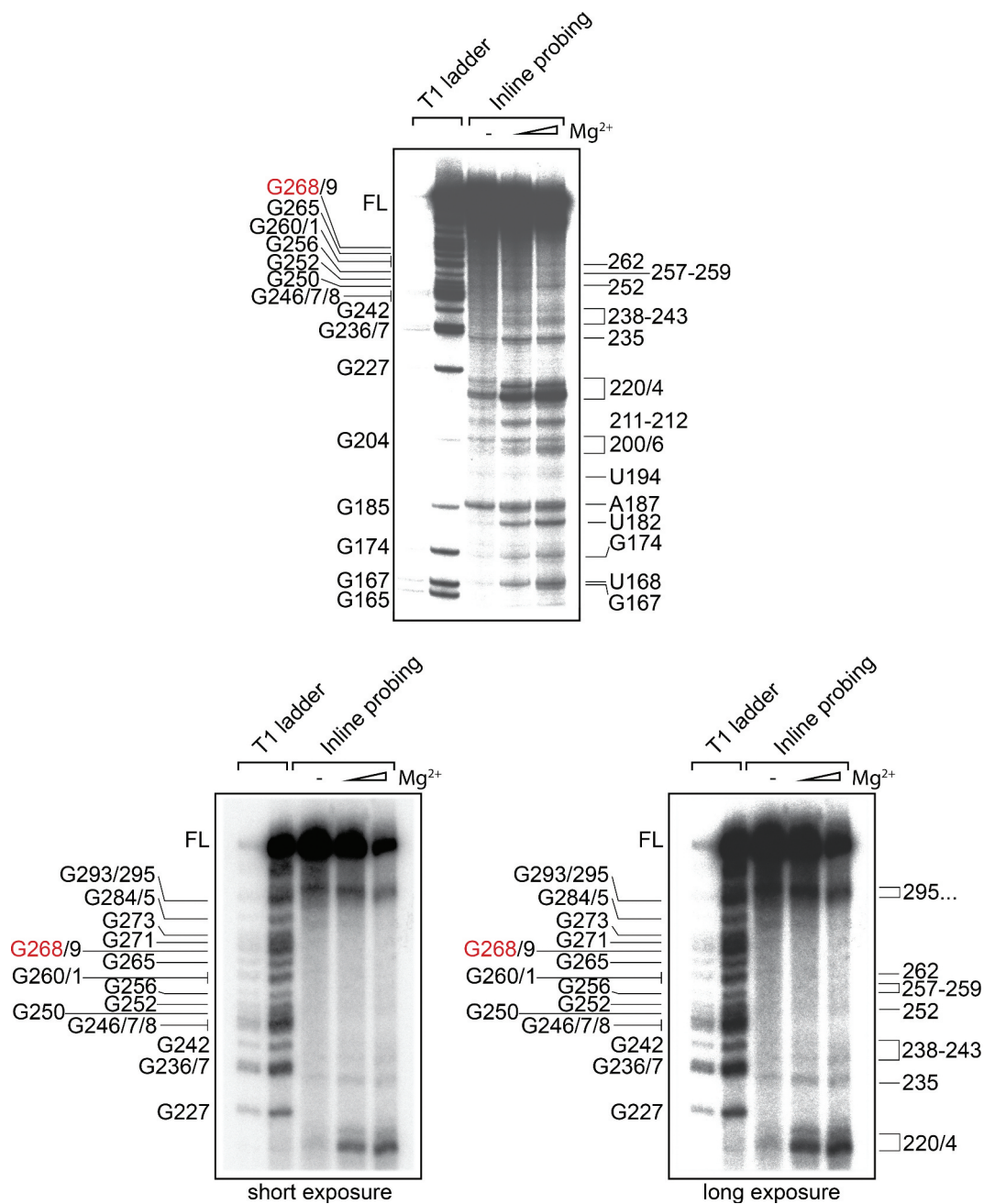
Figure 3. (continued).

mapped by an RNase T1 ladder performed in a denaturing buffer according to previously established protocol [16]. Each segment of the 5′-UTR has been probed at least twice and representative gels for all the parts of the 5′-UTR are shown in the figures and supplemental figures. The inline reactivities have been classified as 'accessible' or 'not accessible' for inline probing. For V1 probing, the reactivities are shown as 'weak' or 'strong' according to the band intensities.

Sequence alignments

Homologous sequences, most of which were from SARS-CoV -2, were retrieved by BLAST [17] search. Sequences were aligned with ClustalW [18] before the alignment-based prediction of RNAalifold [19]. Forna [20] and R2R [21] were used to visualize the secondary structures.

Results and discussion

SARS-CoV-2 5′-UTR inline probing

Using *in vitro* transcription, we synthesized a transcript encompassing nucleotides 1 to 301 from the SARS-CoV-2 variant (GenBank: MN908947.3). It contains the whole 5′-UTR and the sequence coding for the 12 N-terminal codons of NSP1 Fig. 1. Since the viral genomic RNA is capped at the 5′ end, we labelled the transcript at the 5′ end

with a radioactive m^{7}G cap. In order to determine the secondary structure of this transcript, we first performed inline probing in the absence and presence of 1 or 10 mM Mg^{2+}. We then analysed the cuts in the RNA backbone by migration on denaturing polyacrylamide gel electrophoresis. The cuts were mapped using an RNase T1 ladder performed in denaturing conditions. We analysed regions from nt 1 to 128 and 94 to 295 Figs. 2 and 3, respectively.
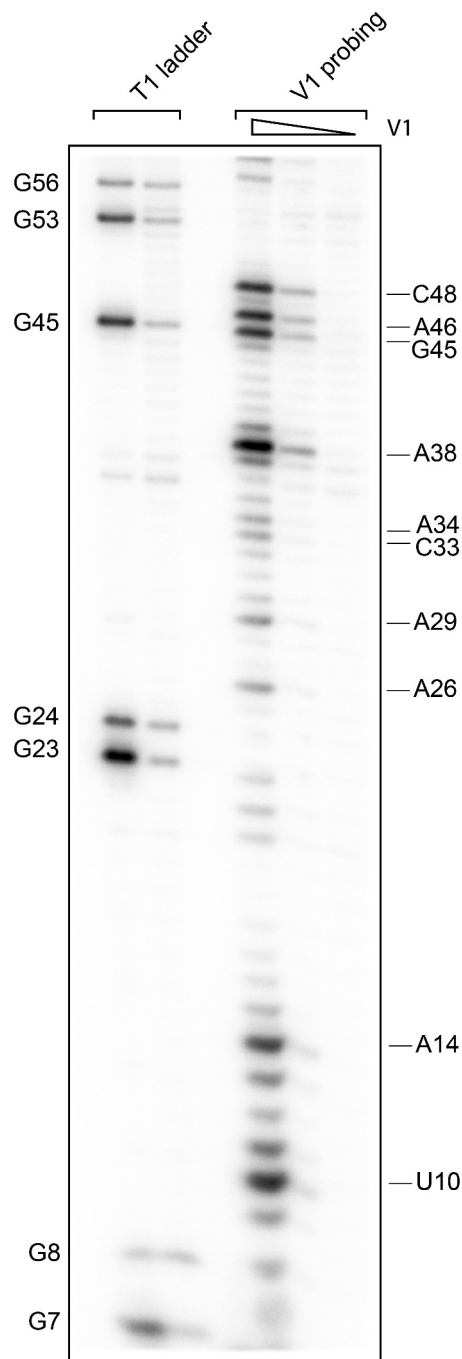
Interestingly, RNase T1 digestion of G residues from 188 to 219 was inefficient even though digestion was performed in denaturing conditions (Fig. 3, right panel). This indicates that these residues are embedded in a highly stable structural region that is still efficiently folded in denaturing conditions thereby preventing the access of RNase T1. Inline probing with and without Mg^{2+} allowed us to map the structurally accessible regions of the RNA that generally correspond to single-stranded regions. Using this method, we could also detect inaccessible areas that are potentially forming base pairs.

In order to confirm these putative stems, we performed RNase V1 probing. The V1 enzyme specifically cuts in stem regions Figs. 4 and 5. We confirmed that most of the inaccessible regions by inline probing do actually correspond to areas containing base pairs. Altogether, these data allowed us to establish a solid model of the 2D
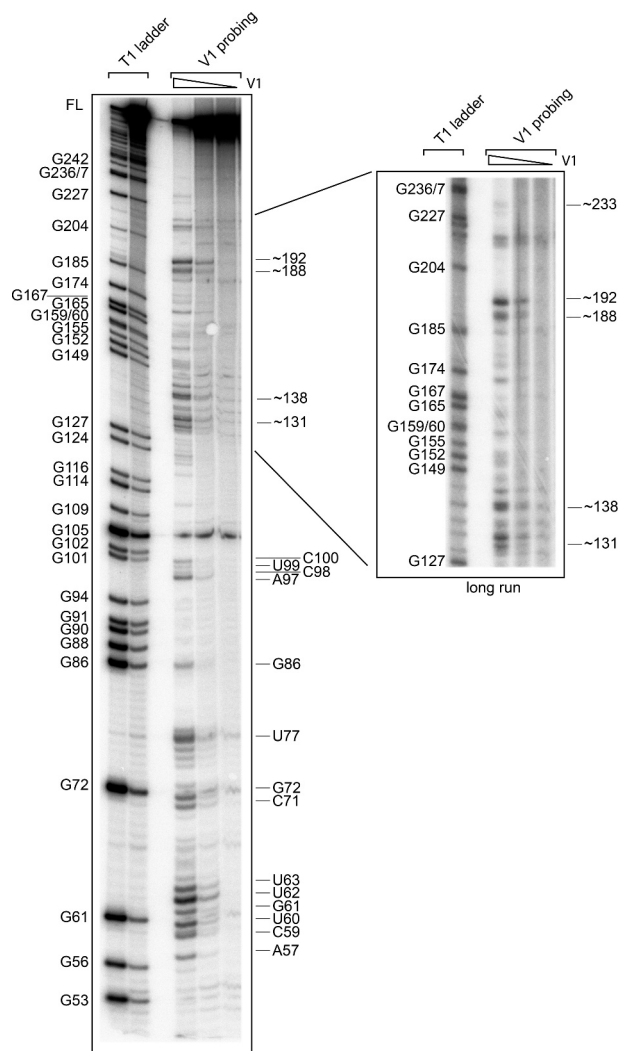

Figure 4. Enzymatic probing by RNase V1 of the SARS-CoV-2 5′-UTR from nucleotides 1 to 56.
Enzymatic probing of the SARS-CoV-2 which are performed by decreasing amount of RNase V1. The cuts are mapped using an RNase T1 denaturing ladder that cuts after G residues.

Figure 5. Enzymatic probing by RNase V1 of the SARS-CoV-2 5′-UTR from nucleotides 105 to 295.
Enzymatic probing of the SARS-CoV-2 which are performed by decreasing amount of RNase V1. The cuts are mapped using an RNase T1 denaturing ladder that cuts after G residues. The panel shows the reactivity of nucleotides 53–295. The right panel shows a long run of the same samples to improve the resolution of the T1 ladder 127 to 236.
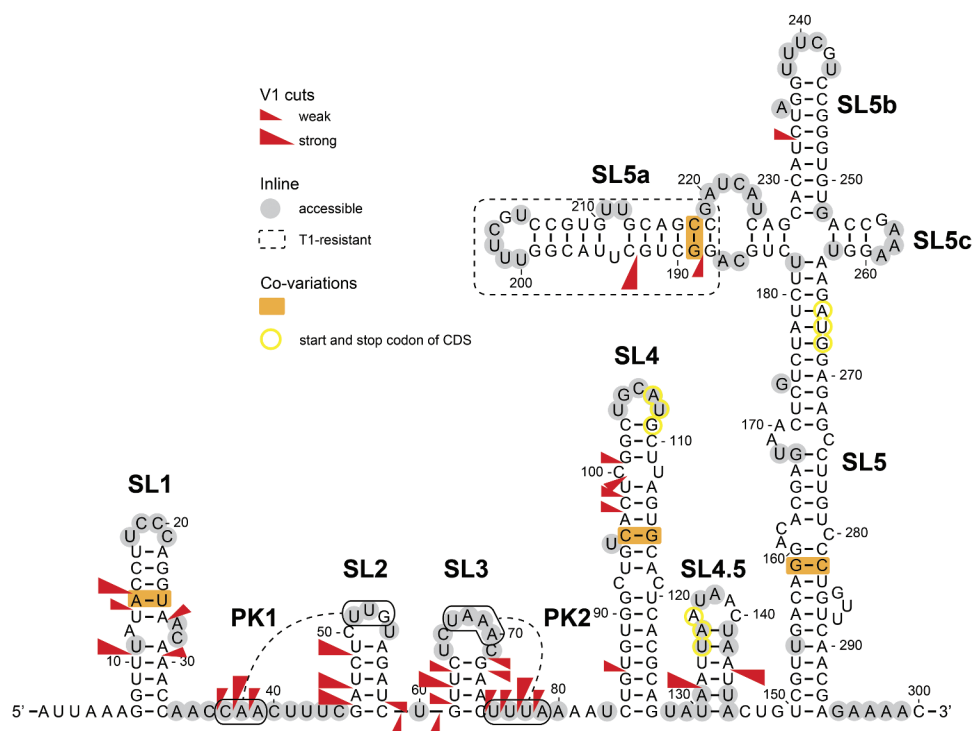
Figure 6. Secondary structural model of SARS-CoV-2 5′-UTR.
The model of the 5′-UTR is represented based on Inline probing and enzymatic probing by RNase V1. Reactivity determined by inline probing experiments is shown in grey and V1 cuts are shown by red triangles. A dashed line boxes the area that is resistant to RNase T1 digestion in denaturing conditions. The positions of the uAUG and the AUG codons are highlighted in yellow.

structure of the whole SARS-CoV-2 5′-UTR Fig. 6. The 5′-UTR is highly structured with a few accessible bulges and loops. It contains five simple hairpin structures that were named SL1, SL2, SL3, SL4 and SL5 in good agreement with bioinformatic secondary structure predictions for the SARS-CoV-2 [22] and also from other coronaviruses [23]. Our model is also highly similar to the models obtained by probing of the whole SARS-CoV-2 genome *in vitro* [24] and *in vivo* [25,26]. However, we found an additional hairpin located between SL4 and SL5 that we named SL4.5. As predicted, SL1 is located close to the 5′ extremity. It has been proposed that the low overall stability of SL1, due to a high proportion of A-U and U-A base pairs, is important for replication in Mouse Hepatitis Virus MHV [27]. The loop of SL1 is not conserved in SARS-CoV-2 variants and the two bulged nucleotides in the middle of SL1 can be involved in base pairs as observed in some variants suggesting that the loop and the bulge of SL1 are not required for efficient viral propagation Fig. 7. The SL2 hairpin domain is comparable to the SL3 domain in Bovine coronavirus (BCoV), which is known to form a hairpin structure according to NMR spectroscopy [28] as well as enzymatic probing [29,30]. This domain is expected to be involved in the replication complex formation Fig. 8. The loop (CUUGY) of SL2 is important for Mouse Hepatitis Virus replication [28]. It contains a U-turn like motif. In contrast, the SARS-CoV SL2 rather adopts a typical tetraloop structure [31]. This is in good agreement with our inline

reactivity profile of the SL2 loop since C50 and G53 are not accessible in the loop but the U54 is highly accessible, a well-characterized feature of a tetraloop structure [31]. Hairpin SL3 is known to encompass the leader TRS (TRS-L) sequence as previously observed for group IIb coronavirus [23] Fig. 8. The SL4 structure is a relatively stable and long hairpin and three base pairs in the stem region have been described to contain covariations (R_{90}-Y_{121}, R_{97}-U_{115} and G_{101}-Y_{111}), indicating structural conservation. SL4 contains the start codon of a uORF that is present in all coronaviruses although there is no evidence of translation of this uORF so far [23]. The integrity of the upper part of SL4 is important for replication of BCoV [29] Fig. 9. It has also been proposed that SL4 is involved in the synthesis of subgenomic RNA fragments [23]. In the 3′ part, the 5′-UTR contains a more complex structure named SL5 that comprises a four-way junction formed by SL5a, SL5b and SL5c. This four-way junction is found in all the coronavirus 5′-UTRs that have been probed so far [23,32–35] Fig. 9. Importantly, in the SARS-CoV-2, the NSP1 AUG start codon is part of the SL5. We observed the same reactivity pattern in the identical loops from SL5a and SL5b, the inline reactivities are in good agreement with classical U-turns [36]. In contrast, SL5c contains a typical GNRA loop as previously observed in group IIb coronaviruses [23]. The loops of SL5 are most likely involved in RNA packaging as previously suggested [37]. These stems are separated by short single-stranded regions that are generally
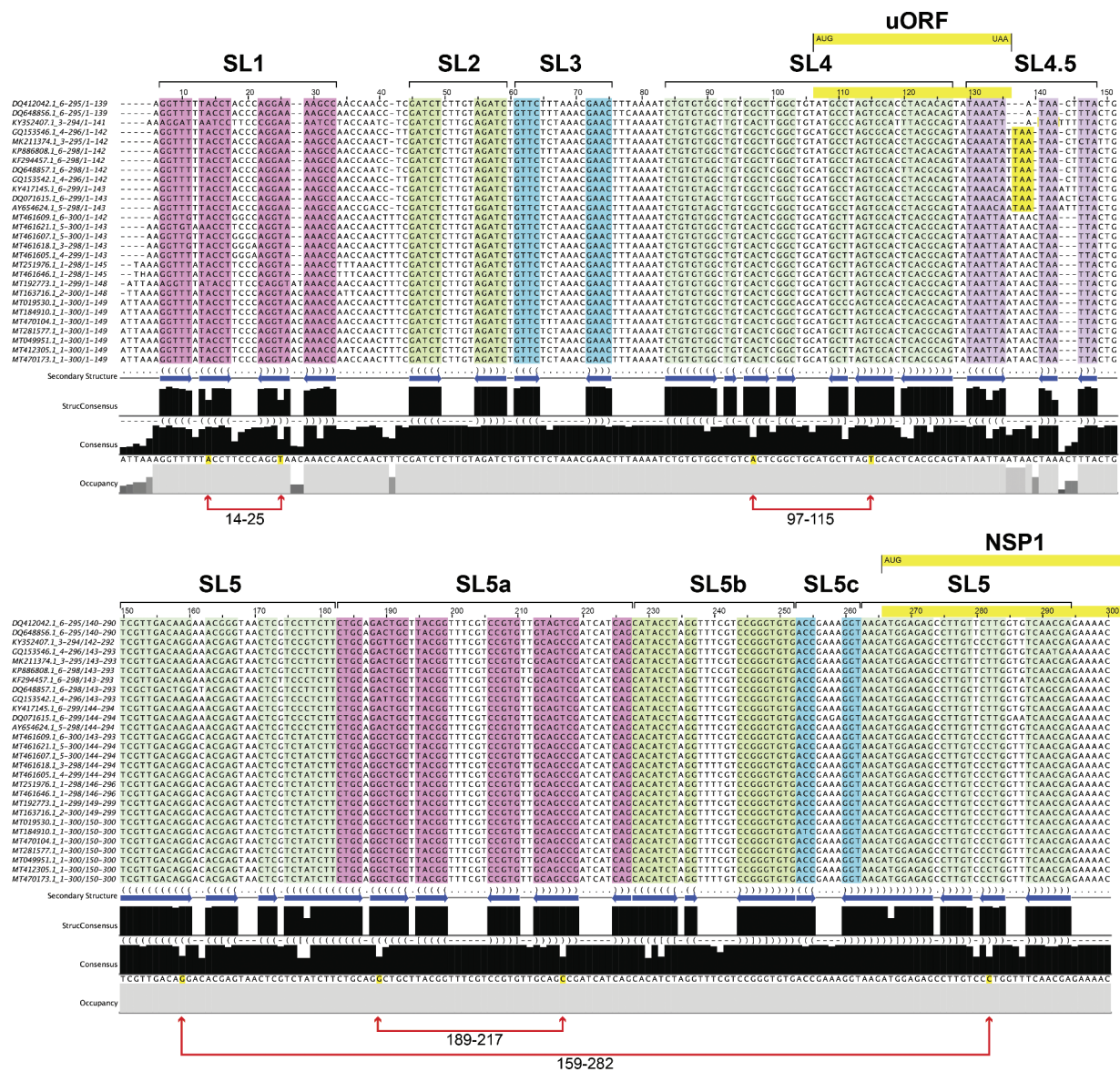
Figure 7. Sequence alignments of 5′-UTRs from SARS-CoV-2 variants.

The sequence alignments of 5′-UTRs from 28 SARS-CoV-2 variants are presented in two panels from nt #1 to 151 (top) and from nt 1#50 to 300 (bottom). The stems of SL1 to SL5 are underlined by distinct colours. The positions of uORF and the beginning of NSP1 ORF are indicated in light green on top of the sequences. In several variants, the stop codon of the uORF is shifted from one codon (also shown in green). Covariations in stems are indicated by red arrows and the corresponding nucleotides are shown in yellow on the consensus sequence.

accessible except for two short regions (A_{34}-C_{40}) and (C_{75} -A_{80}) that are accessible in our inline experiments but also contain V1 cuts indicating the potential forming of base pairs. These apparently contradictory results can be explained as follows: these regions are dynamic and can be considered as single-stranded regions in a so-called 'open' state or double-stranded in a 'closed' state. Interestingly, the sequences can fold into putative pseudo-knots structures with loop of SL2 and SL3 that we named, respectively, PK1 and PK2. The position of SL1 close to the m^7G cap might interfere with the binding of eIF4F complex thereby preventing canonical cap-dependent translation [12]. This observation is important for some SARS-CoV-2 variants that are shorter in which the 5′ extremity is

located just one nucleotide upstream of SL1 Fig. 7. We also probed a truncated version lacking the 5′ proximal half of the 5-UTR (Δ105). Interestingly, the folding of SL4 and SL5 in this shorter version is identical to the folding found in the full-length 5′-UTR (Supplemental Figure 1). This also allowed us to confirm the highly stable structure of SL5abc three-way junction. Moreover, these experiments demonstrate that SL5 can fold without nt 1–105 which suggests the existence of independent motifs in the 5′-UTR that do not interact with each other. Importantly, sequence alignments of recently sequenced SARS-CoV-2 variants also enabled the discovery of co-variations in SL1, SL4 SL5 and SL5a thereby validating our secondary structure model for these structural regions Fig. 7. SL5a is
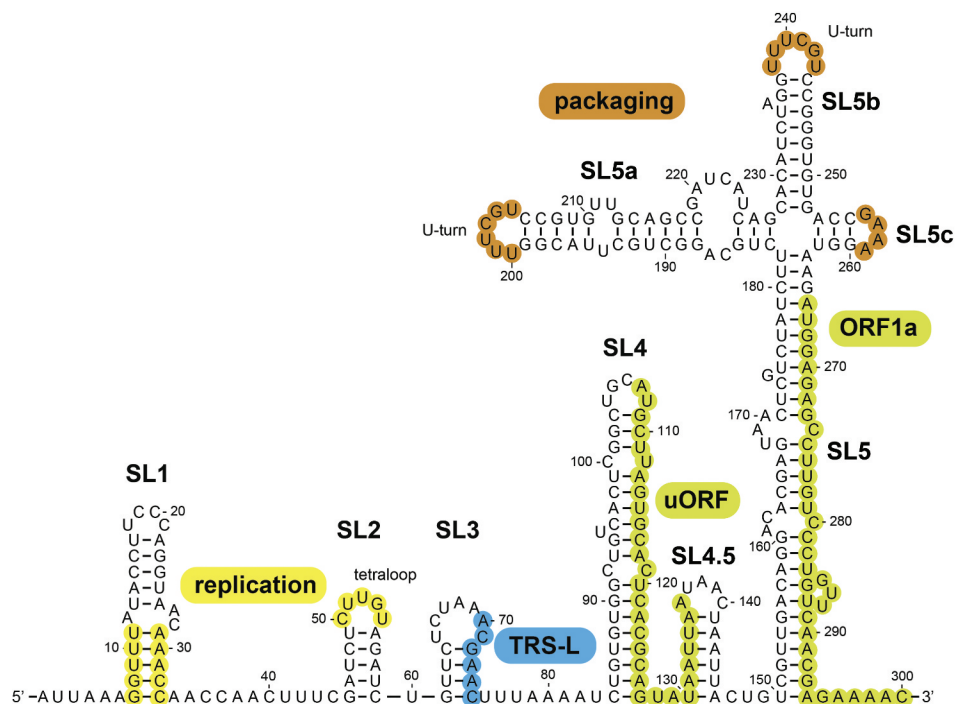
Figure 8. Functional motifs in the 5′-UTR of SARS-CoV-2.

The previously described functional motifs for replication, transcription, viral packaging and translation are shown on our secondary structural model. These motifs have been characterized in the coronavirus family.

a structurally stable region that is not cut by RNase T1 in denaturing conditions. This hairpin contains 10 G-C base pairs and its minimal free energy is predicted to be −17.30 kcal.mol^{-1}, probably explaining its resistance to RNase T1 digestion even in denaturing conditions. The 5′-UTR also contains an upstream opened reading frame (uORF) that is overlapping with SL4 and SL4.5. The location and the sequence of uAUG are absolutely conserved. On the contrary, the UAA stop codon is mutated in a few variants but another in-frame UAA stop is present immediately downstream, implying that the uORF is present and conserved in all variants Fig. 7. The NSP1 AUG start codon is embedded in the four-way junction structure. For an efficient translation initiation, the sequences surrounding the AUG start codon have to be unfolded. The mechanism used by the virus is still unknown and an important issue to investigate is the role and the putative function of the stable SL5 structure (SL5a, SL5b, SL5c) in the translation of the SARS-CoV-2 polyprotein. Although the viral genome is capped at its 5′ end, the translation initiation mechanism used to locate the AUG start codon in SL5 remains elusive. The presence of the 5′ m^{7}G cap and hairpins SL1, SL2, SL3, SL4 and SL5 suggest that a canonical cap-dependent scanning mechanism would require the eIF4A helicase [38,39]. On the other hand, the fact that the AUG start codon is located in the vicinity and downstream of a four-way junction structure is reminiscent of similar structures found in the HCV IRES [40] (and references therein). Indeed, IRES elements are typically highly structured RNA motifs [40].

In addition, the 5′-UTR contains a uORF that is conserved in SARS-CoV-2 variants, which indicates that it can be translated by the host ribosome. The use of uORF is another way of translation regulation [41,42]. It is possible that the viral translation may use all these mechanisms at distinct stages of the infectious process. We hope that our investigations of the 5′-UTR structure in SARS-CoV-2 translation will pave the way to further studies in understanding its functions in the viral infection life cycle.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the authors.
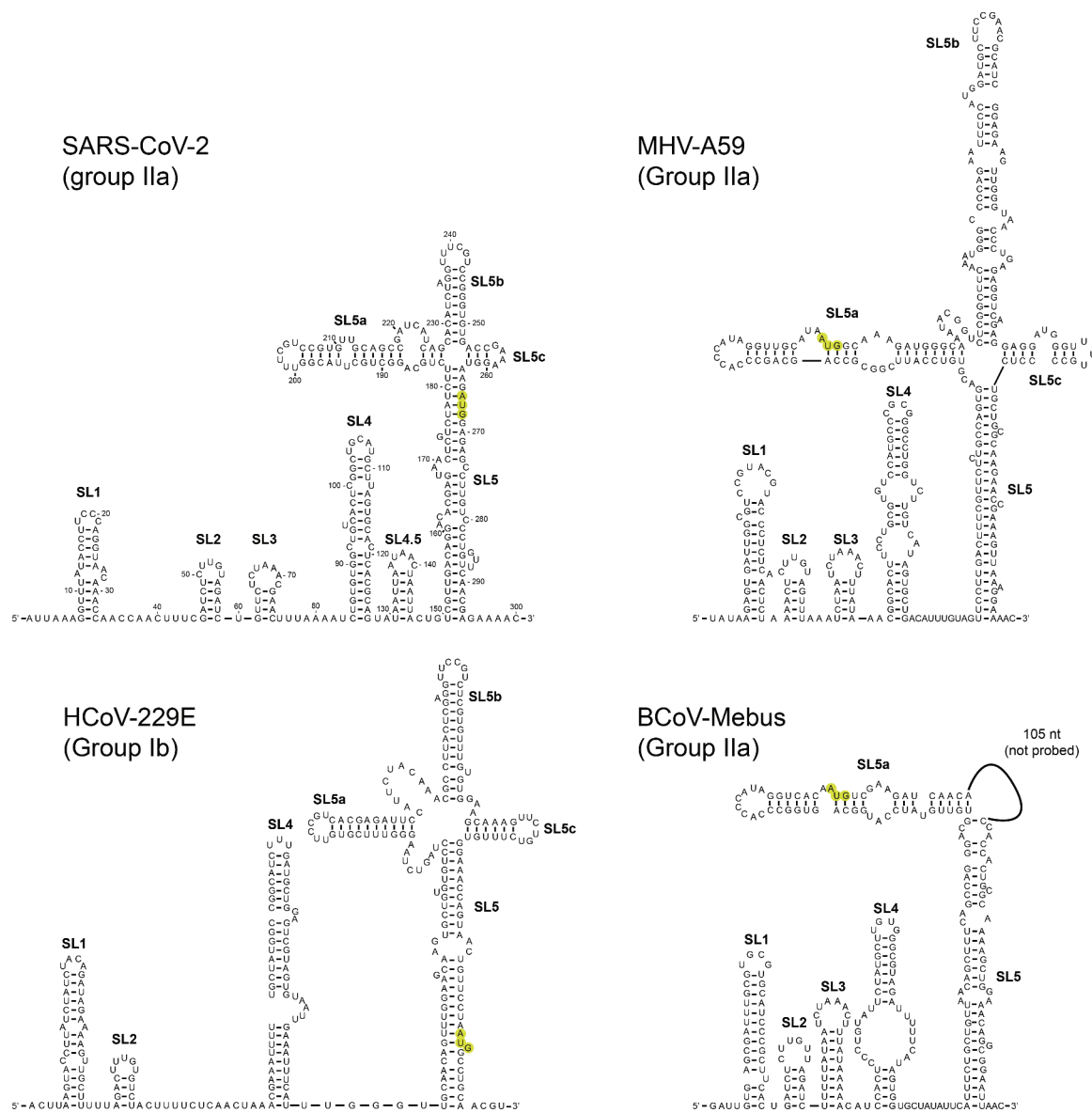
Funding

Figure 9. Comparison of the 2D structure of SARS-CoV-2 5'UTR with other probed coronavirus 5'-UTRs.

Our model for the SARS-CoV-2 5'-UTR is compared with other coronavirus 5'-UTR that have been determined by experimental probing. The 5'-UTR secondary structures from Mouse Hepatitis Virus (MHV) [34], Human coronavirus (HCoV) [35] and Bovine coronavirus (B-CoV) [23,32,33] are shown. The position of the NSP1 AUG start codon is highlighted in green.

ORCID

Zhichao Miao http://orcid.org/0000-0002-5777-9815
Gilbert Eriani http://orcid.org/0000-0002-8518-4675
Franck Martin http://orcid.org/0000-0001-9724-4025

References

[1] Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev. 2005;4(4):635–664.

[2] Andersen KG, Rambaut A, Lipkin WI, et al. The proximal origin of SARS-CoV-2. Nat Med. 2020;26(4):450–452.

[3] Wu A, Peng Y, Huang B. et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell Host Microbe. 2020;27(3):325–328. .

[4] Brian DA, Baric RS. Coronavirus genome structure and replication. Curr Top Microbiol Immunol. 2005;287:1–30.

[5] Plant EP, Rakauskaite R, Taylor DR, et al. Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. J Virol. 2010;84(9):4330–4340.

[6] Plant EP, Sims AC, Baric RS, et al. Altering SARS coronavirus frameshift efficiency affects genomic and subgenomic RNA production. Viruses. 2013;5(1):279–294.

[7] Kamitani W, Narayanan K, Huang C, et al. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc Natl Acad Sci U S A. 2006;103(34):12885–12890.

[8] Lokugamage KG, Narayanan K, Huang C, et al. Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. J Virol. 2012;86(24):13598–13608.

[9] Narayanan K, Ramirez SI, Lokugamage KG, et al. Coronavirus nonstructural protein 1: common and distinct functions in the regulation of host and viral gene expression. Virus Res. 2015;202:89–100.

[10] Yang D, Leibowitz JL. The structure and functions of coronavirus genomic 3′ and 5′ ends. Virus Res. 2015;206:120–133.

[11] Babendure JR, Babendure JL, Ding JH, et al. Control of mammalian translation by mRNA structure near caps. RNA. 2006;5 (5):851–861.

[12] Pestova TV, Kolupaeva VG. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. Genes Dev. 2002;16(22):2906–2922.

[13] Kamitani W, Huang C, Narayanan K, et al. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. Nat Struct Mol Biol. 2009;16 (11):1134–1140.

[14] Clark K, Karsch-Mizrachi I, Lipman DJ, et al. Genbank. Nucleic Acids Res. 2016;44(D1):D67–72.

[15] Jaeger S, Martin F, Rudinger-Thirion J, et al. Binding of human SLBP on the 3′-UTR of histone precursor H4-12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. Nucleic Acids Res. 2006;34(17):4987–4995.

[16] Peattie DA, Gilbert W. Chemical probes for higher-order structure in RNA. Proc Natl Acad Sci U S A. 1980;77(8):4679–4682.

[17] Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–410.

[18] Thompson JD, Higgins DG, Gibson TJ, et al. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–4680.

[19] Lorenz R, Bernhart SH, Höner Zu Siederdissen C, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011;6(1):26.

[20] Kerpedjiev P, Hammer S, Hofacker IL. Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. Bioinformatics. 2015;31(20):3377–3379.

[21] Weinberg Z, Breaker RR. R2R - software to speed the depiction of aesthetic consensus RNA secondary structures. BMC Bioinformatics. 2011;12(1):3.

[22] Rangan R, Zheludev IN, Das R. RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. RNA. 2020;26(8):937–959.

[23] Chen SC, Olsthoorn RCL. Group-specific structural features of the 5′-proximal sequences of coronavirus genomic RNAs. Virology. 2010;401(1):29–41.

[24] Manfredonia I, Nithin C, Ponce-Salvatierra A. et al. Genome-wide mapping of therapeutically-relevant SARS-CoV-2 RNA structures. bioRxiv 2020. 2020;06(15):151647.

[25] Lan TCT, Allan MF, Malsick L, et al. Structure of the full SARS-CoV-2 RNA genome in infected cells. bioRxiv 2020. 2020;2020(6):29.178343.

[26] Huston NC, Wan H, Tavares R de CA, et al. Comprehensive in-vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. bioRxiv 2020. 2020;7 (10):197079.

[27] Li L, Kang H, Liu P, et al. Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. J Mol Biol. 2008;377(3):790–803.

[28] Liu P, Li L, Millership JJ, et al. A U-turn motif-containing stem-loop in the coronavirus 5′ untranslated region plays a functional role in replication. RNA. 2007;13(5):763–780.

[29] Raman S, Bouma P, Williams GD, et al. Stem-loop III in the 5′ untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J Virol. 2003;77 (12):6720–6730.

[30] Raman S, Brian DA. Stem-loop IV in the 5′ untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J Virol. 2005;79(19):12434–12446.

[31] Liu P, Li L, Keane SC, Yang D, Leibowitz JL, Giedroc DP. Mouse Hepatitis Virus Stem-Loop 2 Adopts a uYNMG(U)a-Like Tetraloop Structure That Is Highly Functionally Tolerant of Base Substitutions. J Virol. 2009;83(23):12084–12093.

[32] Guan B-J, Wu H-Y, Brian DA. An Optimal cis-Replication Stem-Loop IV in the 5′ Untranslated Region of the Mouse Coronavirus Genome Extends 16 Nucleotides into Open Reading Frame 1. J Virol. 2011;85(11):5593–5605.

[33] Guan B-J, Su Y-P, Wu H-Y, Brian DA. Genetic Evidence of a Long-Range RNA-RNA Interaction between the Genomic 5′ Untranslated Region and the Nonstructural Protein 1 Coding Region in Murine and Bovine Coronaviruses. J Virol. 2012;86 (8):4631–4643.

[34] Yang D, Liu P, Wudeck E V, Giedroc DP, Leibowitz JL. Shape analysis of the rna secondary structure of the mouse hepatitis virus 5′ untranslated region and n-terminal nsp1 coding sequences. Virology. 2015;475:15–27.

[35] Madhugiri R, Karl N, Petersen D, Lamkiewicz K, Fricke M, Wend U, Scheuer R, Marz M, Ziebuhr J. Structural and functional conservation of cis-acting RNA elements in coronavirus 5′-terminal genome regions. Virology. 2018;517:44–55.

[36] Gutell RR, Cannone JJ, Konings D, et al. Predicting U-turns ribosomal RNA with comparative sequence analysis. J Mol Biol. 2000;300 (4):791–803.

[37] Escors D, Izeta A, Capiscol C, et al. Transmissible gastroenteritis coronavirus packaging signal is located at the 5′ end of the virus genome. J Virol. 2003;77(14):7890–7902.

[38] Boussemart L, Malka-Mahieu H, Girault I. et al. eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies. Nature. 2014;513(7516):105–109. .

[39] Hershey JWB, Sonenberg N, Mathews MB. Principles of translational control. Cold Spring Harb Perspect Biol. 2019;11(9): a032607.

[40] Mailliot J, Martin F. Viral internal ribosomal entry sites: four classes for one goal. Wiley Interdiscip. Rev. 9. RNA2018. e1458. doi: 10.1002/wrna.1458

[41] Leppek K, Das R, Barna M. Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. Nat Rev Mol Cell Biol. 2018;19:158–174.

[42] Orr MW, Mao Y, Storz G, et al. Fs and small ORFs: shedding light on the dark proteome. Nucleic Acids Res. 2020;48(3):1029–1042.