

NUCLEIC ACID CONSTRUCT AND USE THEREOF

The present disclosure claims priority to Chinese Patent Application No. CN202210456261.9, filed on April 27, 2022 and entitled “NUCLEIC ACID
5 CONSTRUCT AND USE THEREOF”, which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

The present disclosure belongs to the field of nucleic acid drugs and particularly relates
10 to a nucleic acid construct comprising engineered UTRs and use thereof for preventing or treating a disease (e.g., viral infection).

BACKGROUND

Vaccines are an important area of medical research. Existing vaccines include
15 inactivated vaccines, adenoviral vector vaccines, attenuated influenza virus vector vaccines, recombinant protein vaccines, and nucleic acid vaccines (including RNA vaccines and DNA vaccines). In RNA vaccines, mRNA is the active ingredient, and it mainly contains a cap structure (Cap), a 5' untranslated region (UTR), an open reading frame (ORF) that encodes an antigen protein, a 3'UTR, and a poly(A) tail structure.
20 The 5'UTR is essential for recruiting ribosomes to the mRNA and initiating codon selection, and it plays an important role in regulating translation efficiency and shaping the cellular proteome (Ivanov, et al. Science. 2016, 352(6292): 1413-1416). Eukaryotic 3'UTRs have various regulatory motifs that can be recognized by microRNAs (miRNAs) and rbp to control mRNA stability, localization, and translation (Mazumder
25 et al., 2003; Mayr, 2017). The polyA tail (located at the 3' end of the mRNA) is another element that determines mRNA stability and protein levels. Optional 3'UTRs not only affect mRNA stability and translation but also control mRNA localization (Tushev et al., 2018). The design and optimization of mRNA, particularly the selection and use of its UTRs, are critical steps in the entire preparation process of mRNA vaccine
30 products. Therefore, developing efficient UTRs remains an urgent need in the field of mRNA drugs.

The present disclosure provides mRNA vaccines with new UTR structures that offer the advantage of enabling stable and efficient expression of the target gene. The vaccines can be generally applied to mRNA vaccines (e.g., vaccines for influenza or
35 COVID-19) to regulate the expression of the target gene, showcasing excellent prospects for clinical pharmaceutical applications.

SUMMARY

The present disclosure provides a nucleic acid element capable of regulating the expression of a target gene, and a nucleic acid construct. The nucleotide construct comprises at least one nucleic acid element capable of regulating the expression of a target gene.

Nucleic Acid Element

In some embodiments, the nucleic acid element is a 5' untranslated region element (5'UTR).

In some embodiments, the 5'UTR is selected from the group consisting of or is a 5'UTR derived from any one of the genes Rho GTPase activating protein (ARHGAP), heat shock 27kDa protein 1 (HSPB1), hemoglobin subunit beta (HBB), C-C motif chemokine ligand 13 (CCL13), and the like, or a derivative sequence thereof.

In some embodiments, the 5'UTR in the nucleic acid construct is derived from or is a 5'UTR sequence of an ARHGAP gene or a derivative sequence thereof. In some embodiments, the ARHGAP includes ARHGAP1, ARHGAP2, ARHGAP3, ARHGAP4, ARHGAP5, ARHGAP6, ARHGAP7 (DLC1), ARHGAP8, ARHGAP9, ARHGAP10, ARHGAP12, ARHGAP13 (SRGAP1), ARHGAP14 (SRGAP2), ARHGAP15, ARHGAP17 (RICH1), ARHGAP18, ARHGAP19, ARHGAP20, ARHGAP21, ARHGAP22, ARHGAP23, ARHGAP24, ARHGAP25, and ARHGAP26.

In some embodiments, the 5'UTR is derived from or is a 5'UTR sequence of an ARHGAP15 gene or a derivative sequence thereof. In some embodiments, the ARHGAP15 gene is derived from any species, such as human ARHGAP15, baboon ARHGAP15, monkey ARHGAP15, or mouse ARHGAP15.

In some embodiments, the 5'UTR is derived from or is a 5'UTR sequence of the human ARHGAP15 gene or a derivative sequence thereof. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 1 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the 5'UTR comprises at least one point mutation capable of being used to inhibit an ATG within the UTR from initiating translation. In some embodiments, the point mutation is selected from the group consisting of a mutation at any one or more sites of A, T, or G of the ATG sequence in the 5'UTR. In some embodiments, the point mutation is a mutation at site A of the ATG sequence in the 5'UTR. In some embodiments, the mutation is to mutate A into G, C, or T. In some embodiments, the mutation is to mutate the ATG into a GTG, CTG, or TTG. Further, the mutation is capable of being used to inhibit the ATG within the UTR from initiating translation.

In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises at least one point mutation capable of being used to inhibit an ATG within the UTR from initiating translation. In some embodiments, the point

mutation is selected from the group consisting of a mutation at any one or more sites of A, T, or G of the ATG sequence in the 5'UTR. In some embodiments, the point mutation is a mutation at site A of the ATG sequence in the 5'UTR. In some embodiments, the mutation is to mutate A into G, C, or T.

- 5 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises the nucleotide sequence set forth in SEQ ID NO: 44, or, the 5'UTR comprises a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 44, wherein N₁ is selected from the group consisting of A, G, C, and T.
- 10 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an ATG is mutated into a GTG. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 2 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 15 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an ATG is mutated into a CTG. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 42 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 20 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an ATG is mutated into a TTG. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 43 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 25 In some embodiments, the 5'UTR is a 5'UTR truncation. In some embodiments, the 5'UTR truncation still maintains active functions similar to those of a natural 5'UTR, e.g., still maintains a function of regulating expression of a protein from an encoding target gene of the ORF. In some embodiments, the 5'UTR truncation has enhanced active functions, e.g., an enhanced function of regulating expression of a protein from an encoding target gene of the ORF, as compared to a natural 5'UTR.
- 30 In some embodiments, a truncation mode for the 5'UTR truncation comprises: in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end. In some
- 35 embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end, and keeping a nucleotide sequence of the 3' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end, and keeping a nucleotide sequence of the 5' end. In some
- 40 embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 5' end to

the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and the truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end. In some embodiments, the 5'UTR truncation also comprises the point mutation according to any of the preceding embodiments. In some embodiments, the 5'UTR truncation does not comprise the point mutation according to any of the preceding embodiments. In some embodiments, the 5'UTR truncation retains a sequence the length of which is at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 5%, or 1% of the length of a natural 5'UTR sequence.

In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises or is an ARHGAP15 5'UTR truncation. In some embodiments, a truncation mode for the truncation comprises, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end.

In some embodiments, a truncation mode for the truncation comprises, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 5' end, and keeping a nucleotide sequence of the 3' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the direction from the 3' end to the 5' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 3' end, and keeping a nucleotide sequence of the 5' end. In some embodiments, a truncation mode for the 5'UTR truncation is, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, to delete a sequence of contiguous nucleotides from the 5' end; and, in the sequence direction from the 3' end to the 5' end, to delete a sequence of contiguous nucleotides from the 3' end.

In some embodiments, the truncation retains a sequence the length of which is at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 5%, or 1% of the length of the natural ARHGAP15 5'UTR (SEQ ID NO: 1).

In some embodiments, the 5'UTR truncation comprises a sequence of at least 5 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 or 44; in some embodiments, the 5'UTR truncation comprises a sequence of 5-62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 or 44; in some embodiments, the 5'UTR truncation comprises a sequence of 7-62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 or 44; in some embodiments, the 5'UTR truncation comprises a sequence of 7-59 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2; illustratively, the 5'UTR truncation comprises a sequence of 5, 7, 8, 9, 10, 11, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 or 44.

In some embodiments, the nucleotide at the 5' end of the 5'UTR truncation (i.e., the

starting nucleotide of the 5'UTR truncation) is the nucleotide at any one of positions 1-57, according to natural counting, of the sequence set forth in SEQ ID NO: 2 or 44. In some embodiments, the nucleotide at the 5' end of the 5'UTR truncation is the nucleotide at any one of positions 1-13 or 17-29, according to natural counting, of the sequence set forth in SEQ ID NO: 2 or 44, for example, the nucleotide at any one of positions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, and 57.

Illustratively, the nucleotide at the 5' end of the 5'UTR truncation, and the length of the sequence of contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 that the 5'UTR truncation comprises are shown in the table below:

Starting nucleotide	The length of the sequence of contiguous nucleotides from the sequence set forth in SEQ ID NO: 2
Position 1, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62
Position 2, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62
Position 3, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, or 61
Position 4, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60
Position 5, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59
Position 9, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55
Position 13, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51
Position 17, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47
Position 21, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
Position	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,

23, T	27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41
Position 25, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39
Position 29, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35
Position 33, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
Position 37, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27
Position 41, C	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23
Position 45, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19
Position 49, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Position 53, A	5, 6, 7, 8, 9, 10, 11
Position 57, C	5, 6, 7

In some embodiments, the 5'UTR truncation is 59 bp, 55 bp, 51 bp, 47 bp, 43 bp, 39 bp, 35 bp, 31 bp, 27 bp, 23 bp, 19 bp, 15 bp, 11 bp, or 7 bp in length.

In some embodiments, the ARHGAP15 5'UTR truncation also comprises the point mutation according to any of the preceding embodiments. In some embodiments, the ARHGAP15 5'UTR truncation also comprises a nucleotide sequence in which an ATG is mutated into a GTG. In some embodiments, the aforementioned 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 28-30 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ARHGAP15 5'UTR truncation does not comprise the point mutation according to any of the preceding embodiments. In some embodiments, the aforementioned 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequences set forth in SEQ ID NOs: 31-40 and comprises CTATAAT; or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to any one of the aforementioned sequences.

In some embodiments, the 5'UTR truncation comprises at least CTATAAT or the nucleotide sequence set forth in SEQ ID NO: 193.

In some embodiments, the ARHGAP15 5'UTR truncation comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 41, 76, and 137-196 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity

thereto.

In some embodiments, the aforementioned 5'UTR derived from the HSPB1 gene comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 3-5 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to any one of SEQ ID NOs: 3-5.

In some embodiments, the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 3-5, 28-43, 76, and 137-196, or the following sequence: CTATAAT, or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to any one of the aforementioned sequences.

Nucleic Acid Construct

The present disclosure provides a nucleic acid construct comprising:

- (a) an open reading frame (ORF), and
- (b) a 5' untranslated region element (5'UTR).

In some embodiments, the ORF is a polynucleotide sequence encoding a target gene protein.

In some embodiments, the target gene is heterologous. In some other embodiments, the target gene is endogenous.

In some embodiments, the 5'UTR in the nucleic acid construct is located upstream of the open reading frame. In some embodiments, the 5'UTR in the nucleic acid construct is located at the 5' end of the open reading frame.

In some embodiments, the 5'UTR in the nucleic acid construct is selected from the group consisting of or is a 5'UTR derived from any one of the genes Rho GTPase activating protein (ARHGAP), heat shock 27kDa protein 1 (HSPB1), hemoglobin subunit beta (HBB), C-C motif chemokine ligand 13 (CCL13), and the like, or a derivative sequence thereof.

In some embodiments, the 5'UTR in the nucleic acid construct is derived from or is a 5'UTR sequence of an ARHGAP gene or a derivative sequence thereof. In some embodiments, the ARHGAP includes ARHGAP1, ARHGAP2, ARHGAP3, ARHGAP4, ARHGAP5, ARHGAP6, ARHGAP7 (DLC1), ARHGAP8, ARHGAP9, ARHGAP10, ARHGAP12, ARHGAP13 (SRGAP1), ARHGAP14 (SRGAP2), ARHGAP15, ARHGAP17 (RICH1), ARHGAP18, ARHGAP19, ARHGAP20, ARHGAP21, ARHGAP22, ARHGAP23, ARHGAP24, ARHGAP25, and ARHGAP26.

In some embodiments, the 5'UTR in the nucleic acid construct is derived from or is a 5'UTR sequence of an ARHGAP15 gene or a derivative sequence thereof. In some embodiments, the ARHGAP15 gene is derived from any species, such as human ARHGAP15, baboon ARHGAP15, monkey ARHGAP15, or mouse ARHGAP15.

In some embodiments, the 5'UTR in the nucleic acid construct is derived from or is a 5'UTR sequence of the human ARHGAP15 gene or a derivative sequence thereof. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 1 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the 5'UTR in the nucleic acid construct comprises at least one point mutation capable of being used to inhibit an ATG within the UTR from initiating translation. In some embodiments, the point mutation is selected from the group consisting of a mutation at any one or more sites of A, T, or G of the ATG sequence in the 5'UTR. In some embodiments, the point mutation is a mutation at site A of the ATG sequence in the 5'UTR. In some embodiments, the mutation is to mutate A into G, C, or T. In some embodiments, the mutation is to mutate the ATG into a GTG, CTG, or TTG. Further, the mutation is capable of being used to inhibit the ATG within the UTR from initiating translation.

10 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises at least one point mutation capable of being used to inhibit an ATG within the UTR from initiating translation. In some embodiments, the point mutation is selected from the group consisting of a mutation at any one or more sites of A, T, or G of the ATG sequence in the 5'UTR. In some embodiments, the point mutation is a mutation at site A of the ATG sequence in the 5'UTR. In some embodiments, the mutation is to mutate A into G, C, or T.

15 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises the nucleotide sequence set forth in SEQ ID NO: 44, or, the 5'UTR comprises a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 44, wherein N₁ is selected from the group consisting of A, G, C, and T.

20 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an ATG is mutated into a GTG. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 2 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

25 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an ATG is mutated into a CTG. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 42 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

30 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an ATG is mutated into a TTG. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 43 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

35 In some embodiments, the 5'UTR in the nucleic acid construct of the present disclosure is a 5'UTR truncation. In some embodiments, the 5'UTR truncation still maintains active functions similar to those of a natural 5'UTR, e.g., still maintains a function of regulating expression of a protein from an encoding target gene of the ORF. In some

40 In some embodiments, the 5'UTR truncation has enhanced active functions, e.g., an enhanced

function of regulating expression of a protein from an encoding target gene of the ORF, as compared to a natural 5'UTR.

In some embodiments, a truncation mode for the 5'UTR truncation comprises: in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end, and keeping a nucleotide sequence of the 3' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end, and keeping a nucleotide sequence of the 5' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and the truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end.

In some embodiments, the 5'UTR truncation also comprises the point mutation according to any of the preceding embodiments. In some embodiments, the 5'UTR truncation does not comprise the point mutation according to any of the preceding embodiments. In some embodiments, the 5'UTR truncation retains a sequence the length of which is at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 5%, or 1% of the length of a natural 5'UTR sequence.

In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises or is an ARHGAP15 5'UTR truncation. In some embodiments, a truncation mode for the truncation comprises, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end.

In some embodiments, a truncation mode for the truncation comprises, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 5' end, and keeping a nucleotide sequence of the 3' end.

In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the direction from the 3' end to the 5' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 3' end, and keeping a nucleotide sequence of the 5' end.

In some embodiments, a truncation mode for the 5'UTR truncation is, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, to delete a sequence of contiguous nucleotides from the 5' end; and, in the sequence direction from the 3' end to the 5' end, to delete a sequence of contiguous nucleotides from the 3' end.

In some embodiments, the truncation retains a sequence the length of which is at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 5%, or 1% of the length of the natural ARHGAP15 5'UTR (SEQ ID NO: 1).

In some embodiments, the 5'UTR truncation comprises a sequence of at least 5 contiguous nucleotides from the nucleotide sequence set forth in SEQ ID NO: 2 or 44; in some embodiments, the 5'UTR truncation comprises a sequence of 5-62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 or 44; in some embodiments, the 5'UTR truncation comprises a sequence of 7-62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 or 44; in some embodiments, the 5'UTR truncation comprises a sequence of 7-59 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2; illustratively, the 5'UTR truncation comprises a sequence of 5, 7, 8, 9, 10, 11, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2.

In some embodiments, the nucleotide at the 5' end of the 5'UTR truncation (i.e., the starting nucleotide of the 5'UTR truncation) is the nucleotide at any one of positions 1-57, according to natural counting, of the sequence set forth in SEQ ID NO: 2 or 44. In some embodiments, the nucleotide at the 5' end of the 5'UTR truncation is the nucleotide at any one of positions 1-13 or 17-29, according to natural counting, of the sequence set forth in SEQ ID NO: 2 or 44, for example, the nucleotide at any one of positions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, and 57.

Illustratively, the nucleotide at the 5' end of the 5'UTR truncation, and the length of the sequence of contiguous nucleotides from the sequence set forth in SEQ ID NO: 2 that the 5'UTR truncation comprises are shown in the table below:

Starting nucleotide	The length of the sequence of contiguous nucleotides from the sequence set forth in SEQ ID NO: 2
Position 1, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62
Position 2, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62
Position 3, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, or 61
Position 4, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60
Position 5, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59

Position 9, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55
Position 13, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51
Position 17, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47
Position 21, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
Position 23, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41
Position 25, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39
Position 29, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35
Position 33, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
Position 37, T	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27
Position 41, C	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23
Position 45, A	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19
Position 49, G	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Position 53, A	5, 6, 7, 8, 9, 10, 11
Position 57, C	5, 6, 7

In some embodiments, the 5'UTR truncation is 59 bp, 55 bp, 51 bp, 47 bp, 43 bp, 39 bp, 35 bp, 31 bp, 27 bp, 23 bp, 19 bp, 15 bp, 11 bp, or 7 bp in length.

- 5 In some embodiments, the ARHGAP15 5'UTR truncation also comprises the point mutation according to any of the preceding embodiments. In some embodiments, the ARHGAP15 5'UTR truncation also comprises a nucleotide sequence in which an ATG is mutated into a GTG. In some embodiments, the aforementioned 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 28-30 or a nucleotide sequence having at least 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ARHGAP15 5'UTR truncation does not comprise the point mutation according to any of the preceding embodiments. In some embodiments, the aforementioned 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequences set forth in SEQ ID NOs: 31-40, CTATAAT, or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the 5'UTR truncation comprises at least CTATAAT or the nucleotide sequence set forth in SEQ ID NO: 193.

In some embodiments, the ARHGAP15 5'UTR truncation comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 41, 76, and 137-196 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 5'UTR derived from the HSPB1 gene comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 3-5 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to any one of SEQ ID NOs: 3-5.

In some embodiments, the 5'UTR in the nucleic acid constructs of the present disclosure comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 3-5, 28-43, 76, and 137-196, or the following sequence: CTATAAT; or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some specific embodiments, the 5'UTR in the nucleic acid construct according to any one of the above comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-5, 28-43, 76, and 137-196, or the following sequence: CTATAAT.

In some embodiments, the nucleic acid construct of the present disclosure also further comprises: (c) a 3' untranslated region element (3'UTR).

In some embodiments, the open reading frame of the present disclosure is derived from a different gene than the 5'UTR and/or the 3'UTR. In some embodiments, the nucleic acid construct of the present disclosure comprises at least one open reading frame, at least one 5'UTR, or at least one 3'UTR. In some embodiments, the 5'UTR and 3'UTR in the nucleic acid construct of the present disclosure are from the same source or different sources, e.g., derived from the same gene or different genes. In some embodiments, the 5'UTR and 3'UTR in the nucleic acid construct of the present disclosure are derived from the same species or different species.

In some embodiments, the 3'UTR in the nucleic acid construct of the present disclosure is located downstream of the open reading frame. In some embodiments, the 3'UTR in the nucleic acid construct is located at the 3' end of the open reading frame.

In some embodiments, the 3'UTR in the nucleic acid construct of the present disclosure is selected from the group consisting of or is a 3'UTR derived from any one of the genes hemoglobin subunit beta (HBB), ARHGAP15, coronin 1A (CORO1A), hemopexin (HPX), and the like, or a derivative sequence thereof.

In some embodiments, the 3'UTR is derived from or is a 3'UTR sequence of the HBB or

ARHGAP15 gene or a derivative sequence thereof.

In some embodiments, the aforementioned 3'UTR derived from the HBB gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 7 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 3'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 8 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 3'UTR derived from the CORO1A gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 9 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 3'UTR derived from the HPX gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 10 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the 3'UTR in the nucleic acid construct of the present disclosure comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto. In some specific embodiments, the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 7 or 8.

In some embodiments, the nucleic acid construct comprises a 5'UTR and a 3'UTR, wherein:

the 5'UTR is selected from the group consisting of or is a 5'UTR derived from any one of the genes ARHGAP (e.g., ARHGAP15), HSPB1, HBB, CCL13, and the like, or a derivative sequence thereof, and the 3'UTR is selected from the group consisting of or is a 3'UTR derived from any one of the genes HBB, ARHGAP15, CORO1A, HPX, and the like, or a derivative sequence thereof.

In some embodiments, the nucleic acid construct comprises a 5'UTR and a 3'UTR, the 5'UTR and the 3'UTR being selected from the group consisting of any one of the following:

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof;

- the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof; or
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof; or
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof.
- In some embodiments, the nucleic acid construct of the present disclosure comprises a 5'UTR and a 3'UTR, the 5'UTR and the 3'UTR being selected from the group consisting

of any one of the following:

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 28-43, 76, and 137-196 and CTATAAT or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR

5 comprises or is the nucleotide sequence set forth in SEQ ID NO: 7 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 28-43, 76, and 137-196 and CTATAAT or a nucleotide sequence having at least

10 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 8 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs:

15 1-2, 28-43, 76, and 137-196 and CTATAAT or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 9 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

20 the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 28-43, 76, and 137-196 and CTATAAT or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR

comprises or is the nucleotide sequence set forth in SEQ ID NO: 10 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity

25 thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 3 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 7 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%,

30 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 3 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 8 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%,

35 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 3 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 9 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%,

40 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 3 or a

nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 10 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

- [illegible]

100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 10 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the nucleic acid construct comprises a 5'UTR and a 3'UTR, wherein:

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 28-43, 76, and 137-196 and CTATAAT, and the 3'UTR is derived from or is a 3'UTR derived from the HBB, ARHGAP15, CORO1A, or HPX gene or a derivative sequence thereof.

In some specific embodiments, the nucleic acid construct comprises a 5'UTR and a 3'UTR, wherein:

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-2, 28-43, 76, and 137-196 and CTATAAT, and the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10.

In some embodiments, the nucleic acid construct of the present disclosure also further comprises: (d) a polyadenylic acid (poly-A) tail.

In some embodiments, the poly-A tail in the nucleic acid construct is located downstream of the 3'UTR. In some embodiments, the poly-A tail in the nucleic acid construct is located at the 3' end of the 3'UTR. In some embodiments, the poly-A tail is at the 3' end of the nucleic acid construct. In some embodiments, the poly-A tail is at least about 50, 100, 150, 200, 300, 400, or 500 nucleotides in length.

In some embodiments, the poly-A tail includes, but is not limited to, HGH polyA, SV40polyA, BGH polyA, rbGlob polyA, or SV40late polyA.

In some specific embodiments, the poly-A tail comprises or is the nucleotide sequence set forth in SEQ ID NO: 16 or 135.

In some embodiments, in the 5'-to-3' direction, the nucleic acid construct according to any one of the above comprises any one of i)-iv):

i) a 5'UTR, and an open reading frame (ORF);

ii) an open reading frame (ORF), and a 3'UTR;

iii) a 5'UTR, an open reading frame (ORF), and a 3'UTR; and

iv) a 5'UTR, an open reading frame (ORF), a 3'UTR, and a poly-A tail;

wherein the ORF is derived from a different gene than the 5'UTR and/or the 3'UTR.

In some embodiments, the 5'UTR in i) and iii) to iv) is selected from the group consisting of a 5'UTR derived from any one of the genes ARHGAP, HSPB1, HBB, CCL13, and the like, or a derivative sequence thereof (e.g., a 5'UTR of ARHGAP15 or a derivative sequence thereof).

In some embodiments, the 3'UTR in ii) to iv) is selected from the group consisting of a 3'UTR derived from any one of the genes HBB, ARHGAP15, CORO1A, HPX, and the like, or a derivative sequence thereof (e.g., a 3'UTR of HBB or ARHGAP15 or a derivative sequence thereof).

In some embodiments, the poly-A tail in iv) to iv) comprises, for example, the

nucleotide sequence set forth in SEQ ID NO: 16 or 135.

In some embodiments, the 5'UTR in i) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-5, 28-43, 76, and 137-196 and CTATAAT.

5 In some embodiments, the 3'UTR in ii) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10.

In some embodiments, the 5'UTR in iii) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-5, 28-43, 76, and 137-196 and CTATAAT, and the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10.

10 In some embodiments, the 5'UTR in iv) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 1-5, 28-43, 76, and 137-196 and CTATAAT, the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10, and the poly-A tail comprises or is the nucleotide sequence set forth in SEQ ID NO: 16 or 135.

15 In the present disclosure, in the nucleic acid construct according to any one of the above, the ORF comprises a nucleotide sequence encoding at least one polypeptide or protein. In some embodiments, the nucleotide sequence may be a codon-optimized nucleotide sequence.

20 In some embodiments, the polypeptide or protein encoded by the ORF is a fluorescent protein or a luciferase.

In some embodiments, the polypeptide or protein encoded by the ORF is a viral antigen. Illustratively, the viral antigen includes, but is not limited to, antigens of influenza viruses, coronaviruses, respiratory syncytial virus, human immunodeficiency viruses, herpes simplex virus, rabies virus, or EB virus, among others.

25 In some embodiments, the viral antigen is a coronavirus antigen. In some embodiments, the coronavirus is a human-infecting coronavirus, such as SARS-CoV-2 (COVID-19), SARS-CoV, HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, or MERS-CoV. In some embodiments, the coronavirus is SARS-COV-2. In some embodiments, the coronavirus antigen is a structural protein. In some embodiments, the structural protein
30 is selected from the group consisting of a spike protein (S protein), an envelope protein (E protein), a membrane protein (M protein), and a nucleocapsid protein (N protein). In some embodiments, the structural protein is a spike protein. In some embodiments, the spike protein is a SARS-COV-2 spike protein. In some embodiments, the SARS-COV-2 spike protein is selected from the group consisting of a spike protein of any one of the
35 viral strains SARS-COV-2 (e.g., wild-type SARS-COV-2), SARS-COV-2 Alpha(B.1.1.7), SARS-COV-2 Beta(B.1.351), SARS-COV-2 Gamma(P.1), SARS-COV-2 Kappa(B.1.617.1), SARS-COV-2 Delta(B.1.617.2), SARS-COV-2 Omicron(B.1.1.529), SARS-COV-2 Omicron(B.A.4), and the like.

40 In some embodiments, the SARS-COV-2 spike protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 12, 14, 21, 23, 25, 80, and 136 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino

acid identity thereto.

In some embodiments, the ORF comprises a codon-optimized nucleotide sequence comprising a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a wild-type nucleotide sequence encoding a SARS-COV-2 antigen (e.g., wild-type SARS-COV-2, SEQ ID NO: 81).

In some embodiments, the ORF encoding the polypeptide or protein comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 13, 15, 20, 22, 24, 81, and 97 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ORF encodes an influenza virus antigen. In some embodiments, the influenza virus is selected from the group consisting of influenza A virus and influenza B virus; illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza A virus H3N8, influenza A virus H2N2, influenza A virus H5N1, influenza A virus H9N2, influenza A virus H7N7, influenza B virus/Victoria (e.g., influenza B virus/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like. In some embodiments, the influenza virus antigen is a structural protein of an influenza virus, e.g., hemagglutinin (HA), neuraminidase (NA), M2 ion channel, matrix protein M1, or nucleoprotein (NP). In some specific embodiments, the influenza virus antigen is an HA protein of an influenza virus, e.g., an HA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus Yamagata (e.g., influenza B/Phuket/3073/2013). In some specific embodiments, the influenza virus antigen is an NA protein of an influenza virus, e.g., an NA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus Yamagata (e.g., influenza B/Phuket/3073/2013).

In some embodiments, the HA protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 83-86 and 98-101 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid identity thereto.

In some embodiments, the ORF comprises a codon-optimized nucleotide sequence comprising a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a wild-type nucleotide sequence encoding an HA antigen.

In some embodiments, the ORF comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 87-90 and 102-105 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ORF comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 93-96 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the present disclosure provides a nucleic acid construct comprising, in the 5'-to-3' direction, a 5'UTR, an ORF, a 3'UTR, and a poly-A tail in sequence. In some embodiments, the 5'UTR comprises or is the nucleotide sequence set

forth in any one of SEQ ID NOs: 1-5, 28-43, 76, and 137-196 and CTATAAT; the ORF comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 13, 15, 20, 22, 24, 81, 97, 87-90, 93-96, and 102-105, the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10, and the poly-A tail comprises or is the nucleotide sequence set forth in SEQ ID NO: 16 or 135.

Illustratively, the nucleic acid construct comprises the nucleotide sequences set forth in SEQ ID NOs: 18-19.

In the present disclosure, including any embodiment thereof, the nucleic acid construct is a nucleic acid molecule, e.g., a cDNA.

10 RNA Molecule

The present disclosure provides an RNA molecule comprising:

- (a) an open reading frame (ORF), and
- (b) a 5' untranslated region element (5'UTR).

15 In some embodiments, the ORF is a polynucleotide sequence encoding a target gene protein.

In some embodiments, the target gene is heterologous. In some other embodiments, the target gene is endogenous.

20 In some embodiments, the 5'UTR in the RNA molecule is located upstream of the open reading frame. In some embodiments, the 5'UTR in the RNA molecule is located at the 5' end of the open reading frame.

25 In some embodiments, the 5'UTR in the RNA molecule is selected from the group consisting of or is a 5'UTR derived from any one of the genes Rho GTPase activating protein (ARHGAP), heat shock 27kDa protein 1 (HSPB1), hemoglobin subunit beta (HBB), C-C motif chemokine ligand 13 (CCL13), and the like, or a derivative sequence thereof.

30 In some embodiments, the 5'UTR in the RNA molecule is derived from or is a 5'UTR sequence of an ARHGAP gene or a derivative sequence thereof. In some embodiments, the ARHGAP includes ARHGAP1, ARHGAP2, ARHGAP3, ARHGAP4, ARHGAP5, ARHGAP6, ARHGAP7 (DLC1), ARHGAP8, ARHGAP9, ARHGAP10, ARHGAP12, ARHGAP13 (SRGAP1), ARHGAP14 (SRGAP2), ARHGAP15, ARHGAP17 (RICH1), ARHGAP18, ARHGAP19, ARHGAP20, ARHGAP21, ARHGAP22, ARHGAP23, ARHGAP24, ARHGAP25, and ARHGAP26.

35 In some embodiments, the 5'UTR is derived from or is a 5'UTR sequence of an ARHGAP15 gene or a derivative sequence thereof. In some embodiments, the ARHGAP15 gene is derived from any species, such as human ARHGAP15, baboon ARHGAP15, or mouse ARHGAP15.

40 In some embodiments, the 5'UTR is derived from or is a 5'UTR sequence of the human ARHGAP15 gene or a derivative sequence thereof. In some embodiments, the 5'UTR of the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 45 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the 5'UTR in the RNA molecule of the present disclosure comprises at least one point mutation, the point mutation including the point mutation in any embodiment of the aforementioned nucleic acid construct. In some embodiments, the mutation is to mutate A into G, C, or U.

- 5 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises the nucleotide sequence set forth in SEQ ID NO: 79, or, the 5'UTR comprises a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 79, wherein N₂ is selected from the group consisting of A, G, C, and U.
- 10 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an AUG is mutated into a GUG. In some embodiments, the 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 46 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 15 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an AUG is mutated into a CUG. In some embodiments, the 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 77 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 20 In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises a nucleotide sequence in which an AUG is mutated into a UUG. In some embodiments, the 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 78 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 25 In some embodiments, the 5'UTR in the RNA molecule of the present disclosure is a 5'UTR truncation. In some embodiments, the 5'UTR truncation still maintains active functions similar to those of a natural 5'UTR, e.g., still maintains a function of regulating expression of a protein from an encoding target gene of the ORF.
In some embodiments, a truncation mode for the 5'UTR truncation comprises: in the
30 sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides
35 from the 5' end, and keeping a nucleotide sequence of the 3' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end, and keeping a nucleotide sequence of the 5' end. In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and the
40 truncation mode for the 5'UTR truncation comprises, in the sequence direction from the

3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end. In some embodiments, the 5'UTR truncation also comprises the point mutation according to any of the preceding embodiments. In some embodiments, the 5'UTR truncation does not comprise the point mutation according to any of the preceding embodiments. In some embodiments, the 5'UTR truncation retains a sequence the length of which is at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 5%, or 1% of the length of a natural 5'UTR sequence.

In some embodiments, the aforementioned 5'UTR sequence derived from the ARHGAP15 gene comprises or is an ARHGAP15 5'UTR truncation. In some embodiments, a truncation mode for the truncation comprises, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in the sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end.

In some embodiments, a truncation mode for the truncation comprises, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 5' end, and keeping a nucleotide sequence of the 3' end.

In some embodiments, a truncation mode for the 5'UTR truncation comprises, in the direction from the 3' end to the 5' end of the ARHGAP15 5'UTR, deleting a sequence of contiguous nucleotides from the 3' end, and keeping a nucleotide sequence of the 5' end.

In some embodiments, a truncation mode for the 5'UTR truncation is, in the direction from the 5' end to the 3' end of the ARHGAP15 5'UTR, to delete a sequence of contiguous nucleotides from the 5' end; and, in the sequence direction from the 3' end to the 5' end, to delete a sequence of contiguous nucleotides from the 3' end.

In some embodiments, the truncation retains a sequence the length of which is at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 5%, or 1% of the length of the natural ARHGAP15 5'UTR (SEQ ID NO: 45).

In some embodiments, the 5'UTR truncation comprises a sequence of at least 5 contiguous nucleotides from the nucleotide sequence set forth in SEQ ID NO: 46 or 79; in some embodiments, the 5'UTR truncation comprises a sequence of 5-62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 46 or 79; in some embodiments, the 5'UTR truncation comprises a sequence of 7-62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 46 or 79; in some embodiments, the 5'UTR truncation comprises a sequence of 7-59 contiguous nucleotides from the sequence set forth in SEQ ID NO: 46 or 79; illustratively, the 5'UTR truncation comprises a sequence of 5, 7, 8, 9, 10, 11, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2.

In some embodiments, the nucleotide at the 5' end of the 5'UTR truncation (i.e., the starting nucleotide of the 5'UTR truncation) is the nucleotide at any one of positions 1-57, according to natural counting, of the sequence set forth in SEQ ID NO: 46 or 79.

In some embodiments, the nucleotide at the 5' end of the 5'UTR truncation is the nucleotide at any one of positions 1-13 or 17-29, according to natural counting, of the sequence set forth in SEQ ID NO: 46 or 79, for example, the nucleotide at any one of positions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, and 57.

In some embodiments, the 5'UTR truncation is 59 bp, 55 bp, 51 bp, 47 bp, 43 bp, 39 bp, 35 bp, 31 bp, 27 bp, 23 bp, 19 bp, 15 bp, 11 bp, or 7 bp in length.

In some embodiments, the ARHGAP15 5'UTR truncation also comprises the point mutation according to any of the preceding embodiments. In some embodiments, the ARHGAP15 5'UTR truncation also comprises a nucleotide sequence in which an AUG is mutated into a GUG. In some embodiments, the aforementioned 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequences set forth in SEQ ID NOs: 63-65 or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ARHGAP15 5'UTR truncation does not comprise the point mutation according to any of the preceding embodiments. In some embodiments, the aforementioned 5'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequences set forth in SEQ ID NOs: 66-75, CUAUAAU, or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 5'UTR derived from the ARHGAP15 gene comprises or is sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U) or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 5'UTR derived from the HSPB1 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 47 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 5'UTR derived from the HBB gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 48 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 5'UTR derived from the CCL13 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 49 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the 5'UTR in the RNA molecule described in the present disclosure comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto. Alternatively, the

5'UTR in the RNA molecule comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U) or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

5 In some specific embodiments, the 5'UTR described in the RNA molecule according to any one of the above comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU. Alternatively, the 5'UTR in the RNA molecule comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U).

10 In some embodiments, the RNA molecule of the present disclosure also further comprises: (c) a 3' untranslated region element (3'UTR).

In some embodiments, the open reading frame of the present disclosure is derived from a different gene than the 5'UTR and/or the 3'UTR. In some embodiments, the RNA molecule of the present disclosure comprises at least one open reading frame, at least
15 one 5'UTR, or at least one 3'UTR. In some embodiments, the 5'UTR and 3'UTR in the RNA molecule of the present disclosure are from the same source or different sources, e.g., derived from the same gene or different genes. In some embodiments, the 5'UTR and 3'UTR in the RNA molecule of the present disclosure are derived from the same species or different species.

20 In some embodiments, the 3'UTR in the RNA molecule of the present disclosure is located downstream of the open reading frame. In some embodiments, the 3'UTR in the RNA molecule is located at the 3' end of the open reading frame. In some embodiments, the 3'UTR in the RNA molecule of the present disclosure is selected from the group consisting of or is a 3'UTR derived from any one of the genes hemoglobin subunit beta
25 (HBB), ARHGAP15, coronin 1A (CORO1A), hemopexin (HPX), and the like, or a derivative sequence thereof.

In some embodiments, the 3'UTR is derived from or is a 3'UTR sequence of the HBB or ARHGAP15 gene or a derivative sequence thereof.

In some embodiments, the aforementioned 3'UTR derived from the HBB gene
30 comprises or is the nucleotide sequence set forth in SEQ ID NO: 50 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 3'UTR derived from the ARHGAP15 gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 51 or a nucleotide
35 sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 3'UTR derived from the CORO1A gene comprises or is the nucleotide sequence set forth in SEQ ID NO: 52 or a nucleotide
40 sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the aforementioned 3'UTR derived from the HPX gene

comprises or is the nucleotide sequence set forth in SEQ ID NO: 53 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

5 In some embodiments, the 3'UTR in the RNA molecule of the present disclosure comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 50-53 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some specific embodiments, the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 50 or 51.

10 In some embodiments, the RNA molecule comprises a 5'UTR and a 3'UTR, wherein: the 5'UTR is selected from the group consisting of or is a 5'UTR derived from any one of the genes ARHGAP (e.g., ARHGAP15), HSPB1, HBB, CCL13, and the like, or a derivative sequence thereof, and the 3'UTR is selected from the group consisting of or is a 3'UTR derived from any one of the genes HBB, ARHGAP15, CORO1A, HPX, and
15 the like, or a derivative sequence thereof.

In some embodiments, the RNA molecule comprises a 5'UTR and a 3'UTR, the 5'UTR and the 3'UTR being selected from the group consisting of any one of the following:

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a
20 derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative
25 sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the ARHGAP15 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a
derivative sequence thereof;

30 the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a
35 derivative sequence thereof;

the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a
derivative sequence thereof;

40 the 5'UTR is derived from or is a 5'UTR of the HBB gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof;

- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the HSPB1 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof; or
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HBB gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the ARHGAP15 gene or a derivative sequence thereof;
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the CORO1A gene or a derivative sequence thereof; or
- the 5'UTR is derived from or is a 5'UTR of the CCL13 gene or a derivative sequence thereof, and the 3'UTR is derived from or is a 3'UTR of the HPX gene or a derivative sequence thereof.
- In some embodiments, the RNA molecule of the present disclosure comprises a 5'UTR and a 3'UTR, the 5'UTR and the 3'UTR being selected from the group consisting of any one of the following:
- the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-46, 63-75, and 77-78 and CUAUAAU or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U) or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 50 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;
- the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-46, 63-75, and 77-78 and CUAUAAU or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U) or nucleotide sequences having at

least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 51 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

5 the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-46, 63-75, and 77-78 and CUAUAAU or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U) or nucleotide sequences having at

10 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 52 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs:

15 45-46, 63-75, and 77-78 and CUAUAAU or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U) or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and the

20 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 53 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 47 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or

25 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 50 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 47 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or

30 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 51 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 47 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or

35 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 52 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;

the 5'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 47 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or

40 100% identity thereto, and the 3'UTR comprises or is the nucleotide sequence set forth in SEQ ID NO: 53 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%,

In some embodiments, the RNA molecule comprises a 5'UTR and a 3'UTR, wherein:
the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs:
45-46, 63-75, and 77-78 and CUAUAAU, or the 5'UTR comprises sequences
corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196
(replacing T in the sequences with U); and the 3'UTR is derived from or is a 3'UTR of
any one of the genes HBB, ARHGAP15, CORO1A, and HPX gene or a derivative
sequence thereof.

In some specific embodiments, the RNA molecule comprises a 5'UTR and a 3'UTR,
wherein:

the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs:
45-46, 63-75, and 77-78 and CUAUAAU, or the 5'UTR comprises sequences
corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196
(replacing T in the sequences with U); and the 3'UTR comprises or is the nucleotide
sequence set forth in any one of SEQ ID NOs: 50-53.

In some embodiments, the RNA molecule of the present disclosure also further
comprises: (d) a polyadenylic acid (poly-A) tail.

In some embodiments, the poly-A tail in the RNA molecule is located downstream of
the 3'UTR. In some embodiments, the poly-A tail in the RNA molecule is located at the
3' end of the 3'UTR. In some embodiments, the poly-A tail is at the 3' end of the RNA
molecule. In some embodiments, the poly-A tail is at least about 50, 100, 150, 200, 300,
400, or 500 nucleotides in length.

In some embodiments, the poly-A tail includes, but is not limited to, HGH polyA,
SV40polyA, BGH polyA, rbGlob polyA, or SV40late polyA.

In some specific embodiments, the poly-A tail comprises or is the nucleotide sequence
set forth in SEQ ID NO: 56 or 135.

In some embodiments, the RNA molecule of the present disclosure also further
comprises: (e) a 5' cap structure (5'Cap).

In some embodiments, the 5' cap structure in the RNA molecule is located upstream of
the 5'UTR. In some embodiments, the 5' cap structure in the RNA molecule is located at
the 5' end of the 5'UTR. In some embodiments, the 5' cap structure is a cap structure
known to those skilled in the art, such as Cap0 (methylation of the first nucleobase, e.g.,
m7GpppN), Cap1 (additional methylation of the ribose of a nucleotide adjacent to
m7GpppN, e.g., m7G(5')ppp(5')(2'OMeA)pG), Cap2 (additional methylation of the
ribose of the 3rd nucleotide downstream of m7GpppN), Cap3 (additional methylation of
the ribose of the 3rd nucleotide downstream of m7GpppN), Cap4 (additional
methylation of the ribose of the 4th nucleotide downstream of m7GpppN), ARCA
(anti-reverse cap analog), modified ARCA (e.g., phosphorothioate-modified ARCA),
inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine,
8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

In some embodiments, the 5'-cap structure (e.g., Cap0 or Cap1) is formed using
chemical RNA synthesis or RNA *in vitro* transcription (co-transcriptional capping).

In some embodiments, the 5'-cap structure (e.g., Cap0 or Cap1) is formed via enzymatic capping using a capping enzyme (e.g., vaccinia virus capping enzyme and/or cap-dependent 2'-O methyltransferase). In some embodiments, the 5' cap structure (Cap0 or Cap1) is added using an immobilized capping enzyme. The capping method and means in WO2016/193226 are incorporated herein by reference in their entirety.

In some embodiments, the 5' cap structure includes, but is not limited to, ARCA, 3'OMe-m7G(5')ppp(5')G, m7G(5')ppp(5')(2'OMeA)pU, m7Gppp(A2'O-MOE)pG, m7G(5')ppp(5')(2'OMeA)pG, m7G(5')ppp(5')(2'OMeG)pG, m7(3'OMeG)(5')ppp(5')(2'OMeG)pG, or m7(3'OMeG)(5')ppp(5')(2'OMeA)pG.

In some specific embodiments, the 5' cap structure is m7G(5')ppp(5')(2'OMeA)pG. Other 5' cap structures or cap structure analogs may also be used.

In some embodiments, in the 5'-to-3' direction, the RNA molecule according to any one of the above comprises any one of i)-v):

- i) a 5'UTR, and an open reading frame (ORF);
- ii) an open reading frame (ORF), and a 3'UTR;
- iii) a 5'UTR, an open reading frame (ORF), and a 3'UTR; and
- iv) a 5'UTR, an open reading frame (ORF), a 3'UTR, and a poly-A tail;
- v) a 5' cap structure (5'Cap), a 5'UTR, an open reading frame (ORF), a 3'UTR, and a poly-A tail;

wherein the ORF is derived from a different gene than the 5'UTR and/or the 3'UTR.

In some embodiments, the 5'UTR in i) and iii) to v) is selected from the group consisting of a 5'UTR derived from any one of the genes ARHGAP, HSPB1, HBB, CCL13, and the like, or a derivative sequence thereof (e.g., a 5'UTR of ARHGAP15 or a derivative sequence thereof).

In some embodiments, the 3'UTR in ii) to v) is selected from the group consisting of a 3'UTR derived from any one of the genes HBB, ARHGAP15, CORO1A, HPX, and the like, or a derivative sequence thereof (e.g., a 3'UTR of HBB or ARHGAP15 or a derivative sequence thereof).

In some embodiments, the poly-A tail in iv) to v) comprises, for example, the nucleotide sequence set forth in SEQ ID NO: 56.

In some embodiments, the 5' cap structure in v) includes, but is not limited to, Cap0, Cap1 (e.g., m7G(5')ppp(5')(2'OMeA)pG), Cap2, Cap3, Cap4, and ARCA.

In some embodiments, the 5'UTR in i) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU; or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U).

In some embodiments, the 3'UTR in ii) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 50-53.

In some embodiments, the 5'UTR in iii) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs:

41, 76, and 137-196 (replacing T in the sequences with U); and the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 50-53.

In some embodiments, the 5'UTR in iv) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U); the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 50-53, and the poly-A tail comprises or is the nucleotide sequence set forth in SEQ ID NO: 56.

In some embodiments, the 5'UTR in v) comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and 137-196 (replacing T in the sequences with U); the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 50-53, the poly-A tail comprises or is the nucleotide sequence set forth in SEQ ID NO: 56, and the 5' cap structure is m⁷G(5')ppp(5')(2'OMeA)pG.

In the present disclosure, in the RNA molecule according to any one of the above, the ORF comprises a nucleotide sequence encoding at least one polypeptide or protein. In some embodiments, the nucleotide sequence may be a codon-optimized nucleotide sequence.

In some embodiments, the polypeptide or protein encoded by the ORF is a fluorescent protein or a luciferase.

In some embodiments, the polypeptide or protein encoded by the ORF is a viral antigen. Illustratively, the viral antigen includes, but is not limited to, antigens of influenza viruses, coronaviruses, respiratory syncytial virus, human immunodeficiency viruses, herpes simplex virus, rabies virus, or EB virus, among others.

In some embodiments, the viral antigen is a coronavirus antigen. In some embodiments, the coronavirus is a human-infecting coronavirus, such as SARS-CoV-2 (COVID-19), SARS-CoV, HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, or MERS-CoV. In some embodiments, the coronavirus is SARS-COV-2. In some embodiments, the coronavirus antigen is a structural protein. In some embodiments, the structural protein is selected from the group consisting of a spike protein (S protein), an envelope protein (E protein), a membrane protein (M protein), and a nucleocapsid protein (N protein). In some embodiments, the structural protein is a spike protein. In some embodiments, the spike protein is a SARS-COV-2 spike protein. In some embodiments, the SARS-COV-2 spike protein is selected from the group consisting of a spike protein of any one of the viral strains SARS-COV-2 (e.g., wild-type SARS-COV-2), SARS-COV-2 Alpha(B.1.1.7), SARS-COV-2 Beta(B.1.351), SARS-COV-2 Gamma(P.1), SARS-COV-2 Kappa(B.1.617.1), SARS-COV-2 Delta(B.1.617.2), SARS-COV-2 Omicron(B.1.1.529), SARS-COV-2 Omicron(BA.4), and the like. In some embodiments, the SARS-COV-2 spike protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 12, 14, 21, 23, 25, 80, and 136 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%,

98%, 99%, or 100% amino acid identity thereto.

In some embodiments, the ORF comprises a codon-optimized nucleotide sequence comprising a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a wild-type nucleotide sequence encoding a SARS-COV-2 antigen (e.g., wild-type SARS-COV-2, SEQ ID NO: 82).

In some embodiments, the ORF encoding the polypeptide or protein comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 54-55, 59-61, 82, and 130 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ORF encodes an influenza virus antigen. In some embodiments, the influenza virus is selected from the group consisting of influenza A virus and influenza B virus; illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza A virus H3N8, influenza A virus H2N2, influenza A virus H5N1, influenza A virus H9N2, influenza A virus H7N7, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like. In some embodiments, the influenza virus antigen is a structural protein of an influenza virus, e.g., hemagglutinin (HA), neuraminidase (NA), M2 ion channel, matrix protein M1, or nucleoprotein (NP). In some specific embodiments, the influenza virus antigen is an HA protein of an influenza virus, e.g., an HA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013). In some specific embodiments, the influenza virus antigen is an NA protein of an influenza virus, e.g., an NA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus Yamagata (e.g., influenza B/Phuket/3073/2013). In some embodiments, the HA protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 83-86 and 98-101 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid identity thereto.

In some embodiments, the ORF comprises a codon-optimized nucleotide sequence comprising a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a wild-type nucleotide sequence encoding an HA antigen.

In some embodiments, the ORF encoding the polypeptide or protein comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 126-129 and 131-134 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the present disclosure provides an RNA molecule comprising, in the 5'-to-3' direction, a 5'UTR, an ORF, a 3'UTR, and a poly-A tail in sequence. In some embodiments, the 5'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 45-49, 63-75, and 77-78 and CUAUAAU, or the 5'UTR comprises sequences corresponding to the sequences set forth in SEQ ID NOs: 41, 76, and

137-196 (replacing T in the sequences with U); the ORF comprises or is the nucleotide sequences set forth in SEQ ID NOs: 54-55, 59-61, 82, and 126-134, the 3'UTR comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 50-53, and the poly-A tail comprises or is the nucleotide sequence set forth in SEQ ID NO: 56 or 135. In some specific embodiments, the RNA molecule comprises the nucleotide sequences set forth in SEQ ID NOs: 57-58 and 106-125.

In the present disclosure, including any embodiment thereof, the RNA molecule may be an mRNA.

The present disclosure uses a 5'UTR sequence derived from human Rho GTPase activating protein (ARHGAP) for recombinant expression of a target gene (e.g., SARS-CoV-2 viral protein) for the first time, and screens, designs, and successfully verifies that a 5'UTR derived from ARHGAP15 and an engineered and modified 5'UTR thereof can initiate efficient expression of different target genes. Rho GTPase activating protein 15 (ARHGAP15), a Rac1-specific GTPase activating protein (GAP), belongs to the ARHGAP family, and it is a major negative regulator of Rho family GTPase activity. The present disclosure provides more 5'UTR and 3'UTR options and combinations with excellent expression efficiency. Additionally, The present disclosure provides mRNA vaccines with the potential to efficiently prevent infection with SARS-CoV-2 and influenza viruses, particularly against various variants thereof.

Polynucleotide

The present disclosure also provides an isolated polynucleotide encoding one or more polypeptides or proteins, wherein the polypeptides or proteins are coronavirus antigens. In some embodiments, the coronavirus is SARS-COV-2. In some embodiments, the coronavirus antigen is a structural protein. In some embodiments, the structural protein is selected from the group consisting of a spike protein (S protein), an envelope protein (E protein), a membrane protein (M protein), and a nucleocapsid protein (N protein). In some embodiments, the structural protein is a spike protein. In some embodiments, the spike protein is a SARS-COV-2 spike protein. In some embodiments, the SARS-COV-2 spike protein is selected from the group consisting of a spike protein of any one of the viral strains SARS-COV-2 (e.g., wild-type SARS-COV-2), SARS-COV-2 Alpha(B.1.1.7), SARS-COV-2 Beta(B.1.351), SARS-COV-2 Gamma(P.1), SARS-COV-2 Kappa(B.1.617.1), SARS-COV-2 Delta(B.1.617.2), SARS-COV-2 Omicron(B.1.1.529), SARS-COV-2 Omicron(BA.4), and the like. In some embodiments, the SARS-COV-2 spike protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 12, 14, 21, 23, 25, 80, and 136 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid identity thereto.

In some embodiments, the ORF encodes an influenza virus antigen. In some embodiments, the influenza virus is selected from the group consisting of influenza A virus and influenza B virus; illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza A virus H3N8, influenza A virus H2N2, influenza A virus H5N1, influenza A virus H9N2, influenza A virus H7N7, influenza B

virus/Victoria (e.g., influenza B/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like. In some embodiments, the influenza virus antigen is a structural protein of an influenza virus, e.g., hemagglutinin (HA), neuraminidase (NA), M2 ion channel, matrix protein M1, or nucleoprotein (NP). In some specific embodiments, the influenza virus antigen is an HA protein of an influenza virus, e.g., an HA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013). In some specific embodiments, the influenza virus antigen is an NA protein of an influenza virus, e.g., an NA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus Yamagata (e.g., influenza B/Phuket/3073/2013). In some embodiments, the influenza virus HA protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 83-86 and 98-101 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid identity thereto.

In some embodiments, the ORF comprises a codon-optimized nucleotide sequence comprising a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a wild-type nucleotide sequence encoding a SARS-COV-2 antigen (e.g., wild-type SARS-COV-2 mRNA, SEQ ID NO: 82).

In some embodiments, the isolated polynucleotide comprises or is the nucleotide sequences set forth in SEQ ID NOs: 54-55, 59-61, 82, and 126-134 or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the polynucleotide is an RNA, e.g., an mRNA. In some embodiments, the polynucleotide is an ORF region of an mRNA.

In some embodiments, the present disclosure provides use of any one of the aforementioned nucleic acid constructs or RNA molecules in any one of the following: (1) preparing a vaccine; (2) encoding a viral antigen in a subject or *in vitro*; (3) preparing a medicament encoding a viral antigen in a subject or *in vitro*; and (4) preparing a medicament.

In some embodiments, the nucleic acid construct or RNA molecule in the present disclosure encodes a target protein, wherein the open reading frame (ORF) has enhanced expression of the target protein.

35 Modification

To further improve the stability of protein expression of the RNA or polynucleotide molecule of the present disclosure, the RNA or polynucleotide molecule may further comprise one or more modifications (including chemical modifications), such as backbone modifications, sugar modifications, base modifications, and/or lipid modifications. In some embodiments, the RNA or polynucleotide molecule is uniformly

modified to a particular modification (e.g., complete modification throughout the sequence). For example, the RNA may be uniformly modified with pseudouridines such that each U in the sequence is a pseudouridine.

5 Backbone modifications in connection with the present disclosure refer to chemical modifications of the phosphate of the backbone of the nucleotides that the RNA or polynucleotide molecule of the present disclosure comprises. In some embodiments, the backbone modifications include, but are not limited to, complete substitution of unmodified phosphate moieties in the backbone with modified phosphates; for example, backbone phosphate groups may be modified by substituting one or more oxygen atoms
10 with a different substituent. In some embodiments, the modified phosphates include, but are not limited to, phosphorothioates, phosphoroselenates, boranophosphates, boranophosphates, hydrogen phosphonates, phosphoramidates, alkyl or aryl phosphonates, and phosphotriesters.

Sugar modifications in connection with the present disclosure refer to chemical
15 modifications of the sugar of the nucleotides that the RNA or polynucleotide molecule of the present disclosure comprises. In some embodiments, the sugar modifications include, but are not limited to, modifications or replacements of the 2' hydroxy (OH) of the RNA molecule with many different “oxy” or “deoxy” substituents. In some embodiments, the “oxy” modifications include, but are not limited to, substitution
20 modifications with alkoxy, aryloxy, polyethylene glycol (PEG), and the like. In some embodiments, “deoxy” modifications include, but are not limited to, hydrogen and amino (e.g., NH₂, alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroaryl amino, diheteroaryl amino, or amino acid) modifications.

Base modifications in connection with the present disclosure refer to chemical
25 modifications of the base moiety of the nucleotides that the RNA or polynucleotide molecule of the present disclosure comprises. In some embodiments, the base modifications include modifications to adenine, guanine, cytosine, and uracil in nucleotides. For example, the nucleosides and nucleotides described herein may be chemically modified on the surface of the major groove. In some embodiments,
30 chemical modifications to the major groove may include amino, thiol, alkyl, or halogen groups. In some embodiments, the base modifications include, but are not limited to, modifications with pseudouridine, 1-methyl-pseudouridine, 5-azacytidine, 5-methylcytosine-5'-triphosphate, or 2-methoxyadenine. In some embodiments, the base modifications are pseudouridine modifications. In some specific embodiments, the RNA
35 may be uniformly modified with pseudouridines such that each U in the sequence is a pseudouridine.

Lipid modifications in connection with the present disclosure mean that the RNA or polynucleotide molecule of the present disclosure comprises lipid modifications. In some embodiments, the lipid modifications include, but are not limited to, covalent
40 attachment of the RNA or polynucleotide molecule of the present disclosure to at least one adapter, and covalent attachment of the corresponding adapter to at least one lipid.

In some embodiments, the lipid modifications include, but are not limited to, covalent attachment of the RNA or polynucleotide molecule of the present disclosure to at least one lipid (no adapter).

UTRs

5 UTRs (5'UTR and/or 3'UTR) may be provided as a flanking region to the nucleic acid construct, RNA, or polynucleotide molecule of the present disclosure. UTRs may be homologous or heterologous with coding regions in the nucleic acid construct, RNA, or polynucleotide molecule of the present disclosure. The flanking region may comprise one or more 5'UTRs and/or 3'UTRs, and the UTRs may be identical or different
10 sequences. Any portion of the flanking region may be codon-optimized. Before and/or after codon optimization, any portion of the flanking region may independently comprise one or more different structural or chemical modifications.

To alter one or more properties of the nucleic acid construct, RNA, or polynucleotide of the present disclosure, UTRs heterologous with the ORF of the present disclosure are
15 introduced or engineering-synthesized into the nucleic acid construct, RNA, or polynucleotide of the present disclosure. The recombinant nucleic acid construct, RNA, or polynucleotide is then administered to a cell, tissue, or organism, and the results, such as protein levels, localization, and/or half-life, are measured to assess the beneficial effects of the heterologous UTRs on the RNA or polynucleotide of the present
20 disclosure. In some embodiments, the UTRs include a wild UTR or a variant thereof, the UTR variant comprising the addition or removal of one or more nucleotides, including A, T, C, or G, to or from an end. In some embodiments, the UTR variant also comprises codon optimization or modification performed in any way. In some embodiments, the UTR variant also comprises the derivative sequence according to any
25 embodiment of the present disclosure; for example, on the basis of a natural UTR sequence, an ATG is point-mutated into a GTG in order to inhibit the ATG within the UTR from initiating translation.

Vector

The present disclosure also provides a vector comprising the nucleic acid construct,
30 RNA, or polynucleotide according to any one of the foregoing. The nucleic acid construct, RNA, or polynucleotide may be present in the vector and/or may be part of the vector, the vector being, for example, a plasmid, cosmid, YAC, or viral vector. The vector may be an expression vector, i.e., a vector capable of providing for the expression of a polypeptide encoded by the nucleic acid construct, RNA, or
35 polynucleotide. The expression vector generally comprises at least one of the nucleic acids of the present disclosure, which is operatively linked to one or more suitable expression regulatory elements (e.g., promoters and terminators). The selection of the elements and their sequences for expression in a particular host is within the knowledge of those skilled in the art. Regulatory elements and other elements useful or necessary
40 for expression of the encoded polypeptide of the present disclosure include, for example, promoters, terminators, selection labels, leader sequences, and reporter genes.

The nucleic acid construct of the present disclosure may be prepared or obtained by known means (e.g., by automatic DNA synthesis and/or recombinant DNA technology) based on nucleotide sequence information of the present disclosure, and/or may be isolated from a suitable natural source.

- 5 In some embodiments, the vector of the present disclosure also comprises a promoter; for example, the promoter is at the 5' end of the 5'UTR of the nucleic acid construct, and for example, the promoter is a T7 promoter, a T7 lac promoter, a Tac promoter, a Lac promoter, or a Trp promoter.

10 In some specific embodiments, the T7 promoter comprises or is the nucleotide sequence set forth in SEQ ID NO: 17.

Host Cell

- 15 The present disclosure also provides a host cell comprising the nucleic acid construct, RNA, or polynucleotide according to any one of the foregoing. In some embodiments, the cell is capable of expressing one or more polypeptides encoded by the nucleic acid construct, RNA, or polynucleotide of the present disclosure. In some embodiments, the host cell is a bacterial cell, a fungal cell, or a mammalian cell.

20 Bacterial cells include, e.g., cells of gram-negative bacterial strains (e.g., *Escherichia coli* strains, *Proteus* strains, and *Pseudomonas* strains), and gram-positive bacterial strains (e.g., *Bacillus* strains, *Streptomyces* strains, *Staphylococcus* strains, and *Lactococcus* strains).

Fungal cells include, e.g., cells of the species *Trichoderma*, *Neurospora*, and *Aspergillus*; or cells of the species *Saccharomyces* (e.g., *Saccharomyces cerevisiae*), *Schizosaccharomyces* (e.g., *Schizosaccharomyces pombe*), *Pichia* (*Pichia pastoris* and *Pichia methanolica*), and *Hansenula*.

25 Mammalian cells include, for example, HEK293 cells, CHO cells, BHK cells, HeLa cells, COS cells, and the like.

However, amphibian cells, insect cells, plant cells, and any other cells used in the art for expressing heterologous proteins may also be used in the present disclosure.

Production or Preparation Method

- 30 The present disclosure provides a method for preparing the nucleic acid construct, RNA, or polynucleotide of the present disclosure, and a method for preparing the encoded polypeptide thereof.

35 Methods and reagents for preparation and production of the nucleic acid construct, RNA, or polynucleotide and the encoded polypeptide thereof, e.g., specific suitable vectors, transformation or transfection methods, selection labels, methods for inducing protein expression, and culture conditions, are known in the art. Similarly, protein isolation and purification techniques suitable for use in the method for manufacturing the encoded polypeptide of the present disclosure are well known to those skilled in the art.

- 40 In some embodiments, the method for preparing the nucleic acid construct, RNA, or polynucleotide comprises culturing the aforementioned host cell and collecting the

produced nucleic acid construct, RNA, or polynucleotide from the culture. The nucleic acid construct, RNA, or polynucleotide of the present disclosure and the encoded polypeptide thereof may also be obtained by other production methods known in the art, such as chemical synthesis, including solid-phase, or liquid-phase synthesis.

5 In some embodiments, a method for preparing an RNA molecule comprises: preparing a nucleic acid construct or a vector, and then using the nucleic acid construct or vector to perform reverse transcription to obtain the RNA molecule. In some specific embodiments, the method also comprises adding a 5'Cap to the 5' end of the RNA molecule.

10 Vaccine

The present disclosure also provides a vaccine comprising the nucleic acid construct according to any one of the foregoing, the RNA according to any one of the foregoing, and/or the polynucleotide according to any one of the foregoing. In some embodiments, the nucleic acid construct, RNA, or polynucleotide encodes one or more
15 antigens of one or more viral strains. The vaccine provided by the present disclosure may be a monovalent vaccine, a multivalent vaccine, or a combination vaccine.

Monovalent Vaccine

The present disclosure provides a monovalent vaccine comprising encoding an antigen of an organism. In some embodiments, the monovalent vaccine comprises encoding
20 one antigen of one viral strain.

Multivalent/Combination Vaccine

In some embodiments, the vaccine may comprise a nucleic acid construct, RNA, or polynucleotide molecule or multiple nucleic acid constructs, RNAs, or polynucleotide molecules encoding two or more antigens of the same species or different species. In
25 some embodiments, the vaccine comprises an RNA or multiple RNAs encoding two or more antigens of the same viral strain or different viral strains. In some embodiments, the RNA may encode 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or more viral antigens.

In some embodiments, in the monovalent vaccine and the multivalent/combination vaccine described above, the antigen is a coronavirus antigen, the coronavirus being, for
30 example, SARS-CoV-2 (COVID-19), SARS-CoV, HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, or MERS-CoV. In some embodiments, the coronavirus antigen is a structural protein, e.g., selected from the group consisting of a spike protein (S protein), an envelope protein (E protein), a membrane protein (M protein), and a nucleocapsid protein (N protein). In some embodiments, the structural protein is a spike
35 protein, e.g., a SARS-COV-2 spike protein. In some embodiments, the SARS-COV-2 spike protein is selected from the group consisting of a spike protein of any one of the viral strains SARS-COV-2 (e.g., wild-type SARS-COV-2 mRNA), SARS-COV-2 Alpha(B.1.1.7), SARS-COV-2 Beta(B.1.351), SARS-COV-2 Gamma(P.1),

SARS-COV-2 Kappa(B.1.617.1), SARS-COV-2 Delta(B.1.617.2), SARS-COV-2 Omicron(B.1.1.529), SARS-COV-2 Omicron(BA.4), and the like. In some embodiments, the SARS-COV-2 spike protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 12, 14, 21, 23, 25, 80, and 136 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid identity thereto. In some embodiments, a DNA sequence encoding the SARS-COV-2 spike protein comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 13, 15, 20, 22, 24, 81, and 97 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto. In some embodiments, an RNA sequence encoding the SARS-COV-2 spike protein comprises or is the nucleotide sequences set forth in SEQ ID NOs: 54-55, 59-61, 82, and 130 or nucleotide sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the ORF encodes an influenza virus antigen. In some embodiments, the influenza virus is selected from the group consisting of influenza A virus and influenza B virus; illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza A virus H3N8, influenza A virus H2N2, influenza A virus H5N1, influenza A virus H9N2, influenza A virus H7N7, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like. In some embodiments, the influenza virus antigen is a structural protein of an influenza virus, e.g., hemagglutinin (HA), neuraminidase (NA), M2 ion channel, matrix protein M1, or nucleoprotein (NP). In some specific embodiments, the influenza virus antigen is an HA protein of an influenza virus, e.g., an HA protein of influenza A virus H1N1, influenza A virus H3N2, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), or influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013). In some embodiments, the HA protein comprises or is the amino acid sequence set forth in any one of SEQ ID NOs: 83-86 and 98-101 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid identity thereto. In some embodiments, a DNA sequence encoding the HA protein comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 87-90, 93-97, and 102-105 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto. In some embodiments, an RNA sequence encoding the HA protein comprises or is the nucleotide sequence set forth in any one of SEQ ID NOs: 126-129 and 131-134 or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, two or more different RNAs (e.g., mRNAs) may be formulated in the same lipid nanoparticle. In other embodiments, two or more different RNAs may be formulated separately in separate lipid nanoparticles, and then the lipid nanoparticles may be combined and administered as a single vaccine composition (e.g., comprising multiple RNAs encoding multiple antigens), or may be administered separately.

The present disclosure also provides a multivalent/combination vaccine comprising an RNA encoding one or more coronaviruses or antigens of one or more different organisms. That is, the vaccine of the present disclosure may be a multivalent/combination vaccine that targets one or more antigens of the same viral strain/species, or one or more antigens of different viral strains/species.

Delivery System

The present disclosure also provides a delivery system comprising the nucleic acid construct according to any one of the foregoing, or the RNA molecule according to any one of the foregoing, wherein the delivery vehicle is a cationic lipid delivery particle. In some embodiments, the particle is a nanoparticle. In some embodiments, the delivery vehicle is a lipid nanoparticle. The RNA molecule in the present disclosure may be delivered into cells and/or *in vivo* using any type of lipid nanoparticle in the art; illustratively, lipid nanoparticles include, but are not limited to, lipid particles disclosed in WO2017075531, WO2018081480A1, WO2017049245A2, WO2017099823A1, WO2022245888A1, WO2022150717A1, CN101291653A, CN102119217A, WO2011000107A1, and CN107028886A, which are incorporated herein by reference in their entirety.

Pharmaceutical Composition

The present disclosure also provides a pharmaceutical composition comprising the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, and/or the delivery vehicle according to any one of the foregoing, and a pharmaceutically acceptable carrier, diluent, or excipient; specifically, the pharmaceutical composition is a solid formulation, an injection, a topical formulation, a spray, a liquid formulation, or a composite formulation.

Product or Kit

The present disclosure provides a product or kit comprising the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, the delivery vehicle according to any one of the foregoing, and/or the pharmaceutical composition according to any one of the foregoing. The kit may be used to provide relevant detection or diagnostic use.

Method for Treating and/or Preventing Disease and Pharmaceutical Use

The present disclosure also provides use of administering to a subject in need thereof a therapeutically and/or prophylactically effective amount of the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, the delivery vehicle according to any one of the foregoing, the pharmaceutical composition according to any one of the foregoing, and/or the product

or kit according to any one of the foregoing in the preparation of a medicament for treating and/or preventing a disease. In some embodiments, the disease comprises a viral infectious disease or a respiratory disease associated with viral infection. In some embodiments, the virus is a coronavirus. In some embodiments, the coronavirus is a human-infecting coronavirus, such as SARS-CoV-2 (COVID-19), SARS-CoV, HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, or MERS-CoV. In some embodiments, the coronavirus is SARS-CoV-2. In some embodiments, the virus is an influenza virus. In some embodiments, the coronavirus is an influenza virus that infects humans, such as influenza A virus or influenza B virus. Illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like. In some embodiments, the respiratory disease associated with viral infection (e.g., respiratory disease associated with SARS-CoV-2 infection) includes simple infection, such as fever, cough, and sore throat, headache, rhinitis, and the like, pneumonia, acute respiratory infection, severe acute respiratory infection (SARI), hypoxemic respiratory failure and acute respiratory distress syndrome, sepsis and septic shock, severe acute respiratory syndrome (SARS), and the like.

The present disclosure also provides a method for treating and/or preventing a disease, the method comprising administering to a subject in need thereof a therapeutically and/or prophylactically effective amount of the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, the delivery vehicle according to any one of the foregoing, the pharmaceutical composition according to any one of the foregoing, and/or the product or kit according to any one of the foregoing. In some embodiments, the disease comprises a viral infectious disease or a respiratory disease associated with viral infection. In some embodiments, the virus is a coronavirus. In some embodiments, the coronavirus is a human-infecting coronavirus, such as SARS-CoV-2 (COVID-19), SARS-CoV, HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, or MERS-CoV. In some embodiments, the coronavirus is SARS-CoV-2. In some embodiments, the virus is an influenza virus. In some embodiments, the coronavirus is an influenza virus that infects humans, such as influenza A virus or influenza B virus. Illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like. In some embodiments, the respiratory disease associated with viral infection (e.g., respiratory disease associated with SARS-CoV-2 infection) includes simple infection, such as fever, cough, and sore throat, headache, rhinitis, and the like, pneumonia, acute respiratory infection, severe acute respiratory infection (SARI), hypoxemic respiratory failure and acute respiratory distress syndrome, sepsis and septic shock, severe acute respiratory syndrome (SARS),

and the like.

The present disclosure also provides a method comprising administering to a subject in need thereof an effective amount of the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, the delivery vehicle according to any one of the foregoing, the pharmaceutical composition according to any one of the foregoing, and/or the product or kit according to any one of the foregoing, which is capable of inducing in the subject a neutralizing antibody reaction and/or a T cell immune response, e.g., a neutralizing antibody reaction against a viral antigen, e.g., a CD4+ and/or CD8+ T cell immune response against a viral antigen. The antigen-antibody titer of the subject increases after vaccination relative to the antigen-antibody titer of a subject vaccinated with a prophylactically effective dose of a conventional vaccine against the antigen. In some embodiments, the virus is a coronavirus. In some embodiments, the coronavirus is a human-infecting coronavirus, such as SARS-CoV-2 (COVID-19), SARS-CoV, HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, or MERS-CoV. In some embodiments, the coronavirus is SARS-CoV-2. In some embodiments, the virus is an influenza virus. In some embodiments, the coronavirus is an influenza virus that infects humans, such as influenza A virus or influenza B virus. Illustratively, the influenza virus is influenza A virus H1N1, influenza A virus H3N2, influenza B virus/Victoria (e.g., influenza B/Washington/02/2019), influenza B virus/Yamagata (e.g., influenza B/Phuket/3073/2013), or the like.

In some embodiments, the subject is immunized. In some embodiments, the subject has a lung disease. In some embodiments, the subject is 5 years old or younger, or 65 years old or older.

In some embodiments, the method comprises administering to the subject at least one, two, three, four, and more doses of the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, the delivery vehicle according to any one of the foregoing, the pharmaceutical composition according to any one of the foregoing, or the product or kit according to any one of the foregoing.

In some embodiments, a detectable level of a viral antigen (such as a coronavirus antigen) is produced in the serum of the subject 1-72 hours after administration of the nucleic acid construct according to any one of the foregoing, the polynucleotide according to any one of the foregoing, the vaccine according to any one of the foregoing, the vector according to any one of the foregoing, the delivery vehicle according to any one of the foregoing, the pharmaceutical composition according to any one of the foregoing, or the product or kit according to any one of the foregoing. In some embodiments, a neutralizing antibody titer of at least 100 NU/mL, 200 NU/mL, 300 NU/mL, 400 NU/mL, 500 NU/mL, 600 NU/mL, 700 NU/mL, 800 NU/mL, 900

NU/mL, or 1000 NU/mL is produced in the serum of the subject 1-72 hours after administration.

In some embodiments, the antibody titer produced in the subject is increased by at least 1 log relative to a control. For example, the antibody titer produced in the subject may be increased by at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 log relative to the control.

In some embodiments, the antibody titer produced in the subject is increased at least 2-fold relative to the control. For example, the antibody titer produced in the subject is increased at least 3, 4, 5, 6, 7, 8, 9, or 10-fold relative to the control. In some examples, a geometric mean is the n th power of the product of n numbers and is generally used to describe proportional growth. In some embodiments, the geometric mean is used to characterize the antibody titer produced in the subject. In some embodiments, the control may be a subject who has not been vaccinated, or a subject who has been administered a live attenuated virus vaccine, an inactivated virus vaccine, or a protein subunit vaccine.

Vaccine Efficacy

In some embodiments, an antigen-specific immune response is characterized by measuring an anti-(corona)virus antigen antibody titer produced in a subject who has undergone the administration of the preceding embodiments. An antibody titer is a measurement of the amount of antibodies within a subject, e.g., antibodies specific to a particular antigen or epitope of an antigen. Antibody titers are typically expressed as the reciprocal of the maximum dilution that provides a positive result. Enzyme-linked immunosorbent assay (ELISA) is a common assay for determining antibody titers.

In some embodiments, the antibody titer is used to assess whether a subject has been infected or to determine whether the immunization is required. In some embodiments, the antibody titer is used to determine the intensity of the autoimmune response, to determine whether the boost immunization is required, to determine whether a previous vaccine dose is effective, and to identify any recent or previous infections. According to the present disclosure, the antibody titer can be used to determine the intensity of an immune response in a subject induced by an immune composition (e.g., an RNA vaccine).

In some embodiments, the antibody titer against a virus (coronavirus) antigen produced in the subject is increased by at least 1 log relative to the control. For example, the antibody titer produced in the subject may be increased by at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 log relative to the control.

In some embodiments, the antibody titer produced in the subject is increased at least 2-fold relative to the control. For example, the antibody titer produced in the subject is increased at least 3, 4, 5, 6, 7, 8, 9, or 10-fold relative to the control.

In some embodiments, the antigen-specific immune response is measured by the ratio of the neutralizing antibody titer in serum to the geometric mean titer (GMT) of the coronavirus, referred to as the geometric mean ratio (GMR). The geometric mean titer (GMT) is the mean antibody titer in a group of subjects calculated by multiplying all

values and taking the n-th root, where n is the number of subjects with available data.

In some embodiments, the control may be a subject who has not been vaccinated, or a subject who has been administered a live attenuated virus vaccine, an inactivated virus vaccine, or a protein subunit vaccine.

- 5 In some embodiments, the vaccine efficacy can be assessed using standard analytical methods (see, e.g., Weinberg et al., J. Infect. Dis., Jun 1, 2010; 201 (11):1607-10). For example, the vaccine efficacy can be measured in a double-blind, randomized, controlled clinical trial. The vaccine efficacy may be expressed as a proportional reduction in the attack rates (ARs) of unvaccinated study cohort (ARU) and vaccinated study cohort (ARV) and can be calculated from the relative risk (RR) of a disease in the vaccinated group using the following equation.

Efficacy = (ARU - ARV)/ARU × 100; and

Efficacy = (1 - RR) × 100

- 15 In some other embodiments, the vaccine effectiveness can be assessed using standard analytical methods (see, e.g., Weinberg et al., J. Infect. Dis., Jun 1, 2010; 201 (11):1607-10). The vaccine effectiveness is an assessment of how a vaccine (which may have been demonstrated to have high vaccine efficacy) reduces a disease in a population. This measure allows to assess the net balance between advantageous and disadvantageous effects of vaccination programs but not the vaccine in a natural condition rather than in a controlled clinical trial. The vaccine effectiveness is proportional to the vaccine efficacy, but is also affected by the degree of immunity of the target population, as well as other non-vaccine related factors such as hospitalization, outpatient visits, or costs. For example, a retrospective case-control analysis can be conducted, i.e., comparing the vaccination rates in a group of infected subjects with an appropriate control group. The vaccine effectiveness can be expressed as a ratio difference, i.e., the probability of infection after vaccination (odds ratio, OR). Effectiveness = (1 - OR) × 100.

BRIEF DESCRIPTION OF THE DRAWINGS

- 30 FIG. 1 illustrates a schematic representation of a plasmid template containing a T7 promoter, a 5'UTR, a 3'UTR, a polyA tail, and an exogenous gene CDS sequence.
- FIG. 2 illustrates the chip electrophoresis photo of the mRNA synthesized by *in vitro* transcription using a linearized plasmid as the template. L denotes the RNA ladder, and 1 denotes the purified mRNA sample.
- 35 FIG. 3 illustrates the difference in expression level of SAS-Cov-2 B.1.351 spike protein caused by different combinations of 5'UTRs and 3'UTRs as measured by ELISA. The 5'UTRs and 3'UTRs are abbreviated as shown in Table 3.
- FIG. 4 illustrates the difference in expression level of SAS-Cov-2 B.1.1.7 spike protein caused by different combinations of 5'UTRs and 3'UTRs as measured by ELISA. The

5'UTRs and 3'UTRs are abbreviated as shown in Table 3.

FIG. 5 illustrates a schematic representation of the point mutation from 5'UTR I to 5'UTR I'.

FIG. 6 illustrates the difference in expression level of SAS-Cov-2 B.1.617.2 spike protein caused by the 5'UTR I point mutation as measured by Western Blot. 5'UTR I was used for samples 1 and 3, and 5'UTR I' was used for samples 2 and 4. The mRNA synthesis templates for samples 1 and 2 were from plasmid miniprep, and the mRNA synthesis templates for samples 3 and 4 were from plasmid maxiprep; sample 5 was an untransfected cell supernatant control; m denotes the molecular weight markers. FIG. 6A illustrates a Western Blot assay, and FIG. 6B illustrates a sample ratio graph after quantification of the Western Blot assay results.

FIG. 7 illustrates the effects of 5'UTR I' on expression levels of different heterologous genes as measured by Western Blot. FIG. 7A illustrates a Western Blot assay, and FIG. 7B illustrates a sample relative value graph after quantification of the Western Blot assay results.

FIG. 8 illustrates the luciferase expression mediated by 5'UTR I'-3'UTR B and 5'UTR I-3'UTR B compared to the BNT162b2 UTR combination.

FIG. 9 illustrates the luciferase expression regulated by 5'UTR I' truncations and 5'UTR I' mutants.

FIG. 10 illustrates the difference in HA mRNA expression level regulated by 5'UTR I' versus a control 5'UTR.

DETAILED DESCRIPTION

Terminology

In order to facilitate the understanding of the present disclosure, some technical and scientific terms are specifically defined below. Unless otherwise specifically defined herein, all other technical and scientific terms used herein have the meanings generally understood by those of ordinary skill in the art to which the present disclosure pertains. The “2019-nCoV” is known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

The “disease caused by 2019-nCoV” is known as COVID-19.

The “nucleic acid” or “nucleotide” includes RNA, DNA, and cDNA molecules. It will be appreciated that due to the degeneracy of the genetic codes, a large number of nucleotide sequences encoding a given protein may be generated. The term nucleic acid can be used interchangeably with the term “polynucleotide”. The “oligonucleotide” refers to a short nucleic acid molecule. The “primer” is an oligonucleotide, whether naturally occurring in a purified restriction enzymatic digest or produced synthetically, that is capable of acting as a start point of synthesis when placed in a condition inducing the synthesis of a primer extension product complementary to a nucleic acid strand (i.e.,

in the presence of nucleotides and an inducing agent such as DNA polymerase, and at a suitable temperature and pH). To maximize the amplification efficiency, the primer is preferably single-stranded, but may optionally be double-stranded. If double-stranded, the primer is first treated to separate the strands before use in preparing an extension product. Preferably, the primer is a deoxyribonucleotide. The primer must be long enough to initiate the synthesis of the extension product in the presence of an inducer. The exact length of the primer will depend on many factors, including temperature, source of the primer, and the method used.

The “vector” or “expression vector” is meant a replicon, such as a plasmid, bacmid, phage, virus, virion, or cosmid, to which another DNA segment, i.e., an “insert”, may be attached, so as to achieve the replication of the attached segment in a cell. The vector may be a nucleic acid construct designed for the delivery to a host cell or for the transfer among different host cells. As used herein, the vector may be viral or non-viral in its original and/or final form, such as the PUC57 DNA vector used herein. The term “vector” encompasses any genetic element capable of replicating when combined with appropriate control elements and capable of transferring a gene sequence to a cell. In some embodiments, the vector may be an expression vector or a recombinant vector.

The “promoter” refers to any nucleic acid sequence that regulates the expression of another nucleic acid sequence by driving the transcription of the nucleic acid sequence, which may be a heterologous target gene encoding a protein or RNA. The promoter may be constitutive, inducible, repressible, or tissue-specific, or any combination thereof. The promoter is a control region of a nucleic acid sequence where the initiation and rate of the transcription of the remainder of the nucleic acid sequence is controlled.

The “gene” refers to a segment of DNA involved in the production of a polypeptide chain, which may or may not include preceding and succeeding coding regions, e.g., 5' untranslated region (5'UTR) or “leader” sequence and 3'UTR or “non-transcribed tail” sequence, as well as intervening sequences (introns) between the coding segments (exons).

The “recombinant” refers to that a polynucleotide is a product of various combinations of cloning, restriction, or ligation procedures, as well as other procedures resulting in constructs different and/or distinct from naturally occurring polynucleotides.

The “introduction”, in the context of inserting a nucleic acid sequence into a cell, refers to “transfection”, “transformation”, or “transduction” and includes reference to the incorporation of a nucleic acid sequence into a eukaryotic or prokaryotic cell, where the nucleic acid sequence may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastidial, or mitochondrial DNA), transformed to autonomous replication, or transiently expressed (e.g., transfected mRNA).

The “nucleic acid construct” refers to a nucleic acid molecule, e.g., a DNA fragment, either single- or double-stranded, that is modified or synthesized to comprise a nucleic acid segment in a manner that does not naturally occur; the nucleic acid molecule comprises one or more control sequences or regulatory elements. In the context of the

present disclosure, the nucleic acid construct contains a recombinant nucleotide sequence consisting essentially of optionally one, two, three, or more isolated nucleotide sequences including 5'UTR, open reading frame (ORF), and 3'UTR. In embodiments involving constructs comprising two or more sequences, the sequences are operably linked to each other in the construct.

The “derivative sequence” refers to a nucleotide sequence that is highly homologous (e.g., having at least 80%, 85%, 88%, 90%, 93%, 95%, 96%, 97%, 98%, 99%, or 100% identity) to the UTR sequence of the present disclosure and still retains the same or similar functional activity as the UTR sequence of the present disclosure. In some embodiments, the derivative sequence is a nucleotide sequence having one or more substituted, deleted, or added nucleotides based on a natural UTR sequence. In some embodiments, the derivative sequence is a nucleotide sequence that is truncated on the basis of the natural UTR sequence. In some embodiments, the derivative sequence has a point mutation of an ATG into a GTG, CTG, or TTG on the basis of a natural UTR sequence, for inhibiting the ATG within the UTR from initiating translation.

The “operably linked” is defined herein as a structure where a control sequence, i.e., a promoter sequence and/or a 5'UTR sequence, are suitably positioned relative to the coding DNA sequence such that the control sequence directs the transcription of the coding sequence and the translation of the mRNA into a polypeptide sequence encoded by the coding DNA.

The “open reading frame”, abbreviated as “ORF”, refers to a segment or region encoding an mRNA molecule of a polypeptide. The ORF comprises contiguous, non-overlapping, in-frame codons, starting with a start codon and ending with a stop codon, which are translated by the ribosome.

The “endogenous” refers to any substance derived from or produced in an organism, cell, tissue, or system.

The “exogenous” refers to any substance introduced or produced from outside an organism, cell, tissue, or system.

The “sequence identity” refers to sequence identity between genes or proteins at the nucleotide or amino acid level, respectively. The “sequence identity” is a measure of identity between proteins at the amino acid level and between nucleic acids at the nucleotide level. The protein sequence identity can be determined by comparing the amino acid sequences at given positions in the aligned sequences. Similarly, the nucleic acid sequence identity can be determined by comparing the nucleotide sequences at given positions in the aligned sequences. Methods for aligning sequences for comparison are well known in the art, including GAP, BESTFIT, BLAST, FASTA, and TFASTA. The BLAST algorithm calculates the percent sequence identity and performs a statistical analysis of the similarity between two sequences. Software for performing BLAST analysis is publicly available through the National Center for Biotechnology Information (NCBI) website.

The “homologous” or “homology” is defined as the percentage of nucleotide residues in

a target chromosome identical to those in a corresponding sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percentage sequence identity. Alignment to determine percentage nucleotide sequence homology can be accomplished in a variety of ways within the skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ClustalW2, or Megalign (DNASTAR). In some specific embodiments, the present disclosure calculates percentage sequence homology on the basis of BLAST. Those skilled in the art can determine suitable parameters for aligning the sequences, including any algorithms necessary to achieve the maximum alignment over the full length of the sequences being compared. In some embodiments, the sequences are considered “homologous” when, for example, a nucleic acid sequence (e.g., a DNA sequence) of a homology arm has at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or higher identity to a corresponding natural or unedited nucleic acid sequence (e.g., a genomic sequence) of a host cell.

The “substitution” is defined as a change in amino acid or nucleotide sequence as follows: one or more amino acids or nucleotides are substituted with a different amino acid or nucleotide, as compared to the amino acid sequence or nucleotide sequence of a reference polypeptide. If the substitution is conservative, the substituent amino acid has similar structural or chemical properties (e.g., charge, polarity, hydrophobicity, etc.) as the substituted amino acid. In some embodiments, a polypeptide variant may have a “non-conservative” change, where the amino acids before and after the substitution differ in structural and/or chemical properties.

The “deletion” is defined as a change in amino acid or nucleotide sequence as follows: one or more amino acid or nucleotide residues are deleted, as compared to the amino acid sequence or nucleotide sequence of a reference polypeptide. In the case of a polypeptide or polynucleotide sequence, considering the length of the polypeptide or polynucleotide sequence being modified, the deletion may involve deletions of 2, 5, 10, up to 20, up to 30, or up to 50 or more amino acid or nucleotide residues.

The “insertion” or “addition” refers to a change in amino acid or nucleotide sequence as follows: one or more amino acids or nucleotides are added, as compared to the amino acid sequence or nucleotide sequence of a reference polypeptide. The “insertion” generally refers to the addition of one or more amino acid residues within the amino acid sequence of a polypeptide (or nucleotide residues within a polynucleotide), while the “addition” may be an insertion or the addition of amino acid residues at the N- or C-terminus of a polypeptide (or the addition of nucleotide residues at the 5' or 3' terminus of a polynucleotide). In the case of a polypeptide or polynucleotide sequence, the addition or insertion may involve up to 10, up to 20, up to 30, or up to 50 or more amino acids (or nucleotide residues).

The “codon optimization” refers to the replacement of a codon in the target sequence that is rarely seen in highly expressed genes of a given species with a codon commonly

seen in highly expressed genes of the species, with the codons before and after the replacement encoding the same amino acid. Various species may exhibit specific usage biases for certain codons for a particular amino acid. The codon usage bias (the difference in codon usage between organisms) is often correlated with the translation efficiency of messenger RNA (mRNA), which in turn is believed to particularly depend on the characteristics of the codons translated and the availability of specific transfer RNA (tRNA) molecules. The predominance of a selected tRNA in a cell is typically a reflection of the codons most frequently used in peptide synthesis. Thus, on the basis of codon optimization, genes can be modified for the optimal gene expression in a given organism. Thus, the option of optimal codons depends on the codon usage bias of the host genome.

The “antigen antibody” refers to a serum antibody that specifically binds to an antigen. The “cell” or “host cell” includes any cell type susceptible to transformation, transfection, transduction, and the like with the nucleic acid construct or vector of the present disclosure. As a non-limiting example, the host cell may be an isolated primary cell, a pluripotent stem cell, a CD34⁺ cell, an induced pluripotent stem cell, or any one of a number of immortalized cell lines (e.g., a HepG2 cell). Alternatively, the host cell may be an *in situ* or *in vivo* cell in a tissue, organ, or organism.

The “treatment” refers to administering a therapeutic agent, such as a composition comprising any one of the nucleic acid constructs of the present disclosure, either internally or externally to a patient with one or more symptoms of a disease on which the therapeutic agent is known to have a therapeutic effect. Typically, the therapeutic agent is administered in an amount effective to alleviate one or more symptoms of the disease in the patient or population being treated to induce regression of such symptoms or to inhibit the development of such symptoms to any clinically measurable degree. The amount of therapeutic agent effective to alleviate any particular symptom of the disease (also known as a “therapeutically effective amount”) may vary depending on a variety of factors, such as the disease state, age, and weight of the patient, and the ability of the drug to produce a desired therapeutic effect in the patient. Whether a disease symptom has been alleviated can be evaluated by any clinical testing methods commonly used by doctors or other health care professionals to evaluate the severity or progression of the symptom.

The “effective amount” or “pharmaceutically effective amount” comprises an amount sufficient to ameliorate or prevent a symptom or condition of a medical disorder. The effective amount also refers to an amount sufficient to allow or facilitate diagnosis. The effective amount for a particular patient or veterinary subject may vary depending on the factors such as the disorder to be treated, the general health of the patient, the method and route and dose of administration, and the severity of adverse effects. The effective amount may be the maximum dose or administration regimen to avoid significant adverse effects or toxic effects. In some embodiments, the “effective amount” is an effective amount of RNA that can produce an antigen-specific immune

response.

The “prophylactically effective dose” refers to an effective dose to prevent viral infection at a clinically acceptable level. In some embodiments, the effective dose is a dose listed in the vaccine package insert. The conventional vaccine, as used herein,
5 refers to a vaccine other than the mRNA vaccine of the present disclosure. For example, the conventional vaccine includes, but is not limited to, a live microbial vaccine, an inactivated microbial vaccine, a subunit vaccine, a protein antigen vaccine, a DNA vaccine, a virus-like particle (VLP) vaccine, and the like. In exemplary embodiments, the conventional vaccine is one that has been approved by regulatory authorities and/or
10 listed by a national drug administration, such as the Chinese National Medical Products Administration.

The term “pharmaceutically acceptable” refers to those therapeutic agents, materials, compositions, and/or dosage forms that are, within the scope of sound medical judgment, suitable for use in contact with the tissues of patients without excessive
15 toxicity, irritation, allergic reaction, or other problems or complications, and are commensurate with a proper benefit/risk ratio and effective for the intended use.

The “monovalent vaccine” refers to a vaccine preparation for immunization made from an antigen of one serotype from a pathogenic organism. For example, a vaccine containing only one antigen of one SARS-CoV-2 strain of the present disclosure is a
20 monovalent vaccine.

The “multivalent vaccine” refers to a vaccine preparation for immunization made from antigens of more than one serotype from the same pathogenic organism. For example, a vaccine containing an antigen from one SARS-CoV-2 variant mixed with one or more antigens from other SARS-CoV-2 variants of different subtypes is a multivalent
25 vaccine.

The “combined vaccine” refers to a vaccine preparation for immunization made from antigens of more than one serotype from different pathogenic organisms. For example, a vaccine containing an antigen of the SARS-CoV-2 strain of the present disclosure mixed with several pathogenic microorganisms that can prevent and/or treat other viral
30 infections is a combined vaccine. The combined vaccine is generally prepared from monovalent vaccines prepared by the same preparation method, and then is mixed in the same injection for immunization; alternatively, antigens of several pathogenic microorganisms are mixed and then prepared into a vaccine product. In some specific embodiments, antigens of pathogenic microorganisms are first mixed and then prepared
35 into a vaccine preparation. Combined and multivalent vaccines are the future development trends due to their advantages of: (1) cost-efficiency in vaccination; (2) cost-efficiency in packaging in doses, transportation, and assembly; (3) reduced damage to vaccinees, particularly infants; (4) increased vaccination coverage; (5) efficiency in vaccination; (6) compactness for vaccine storage, etc. The combined or multivalent
40 vaccine can save direct and indirect costs including multiple visits, missed vaccination doses, and the like. For example, combined, multivalent vaccines commonly used at

present include vaccines against *Streptococcus pneumoniae*, *Neisseria meningitidis*, poliovirus, rotavirus, influenza, HPV, and the like.

The “immune response” of the vaccine of the present disclosure refers to the development of a humoral and/or cellular immune response in a subject to protein(s) of a virus (e.g., coronavirus) present in the vaccine. For the purposes of the present disclosure, the “humoral” immune response refers to an immune response mediated by antibody molecules, including, for example, secretory IgA or IgG molecules, while the “cellular” immune response is an immune response mediated by T lymphocytes (e.g., CD4⁺ helper and/or CD8⁺ T cells (e.g., CTLs) and/or other leukocytes). An important aspect of the cellular immunity involves the antigen-specific reactions of cytolytic T lymphocytes (CTLs). The CTLs are specific for peptide antigens that are presented in combination with proteins encoded by the major histocompatibility complex (MHC) and expressed on the cell surface, and help induce and promote the destruction of intracellular microorganisms or lyse cells infected with these microorganisms. Another aspect of the cellular immunity involves antigen-specific responses by helper T cells, which help stimulate function and concentrate the activities of non-specific effector cells on cells with peptidic antigens associated with MHC molecules. The cellular immune responses also result in cytokines, chemokines, and other like molecules produced by activated T cells and/or other leukocytes.

Examples

The present disclosure is further described below with reference to examples; however, these examples are not intended to limit the scope of the present disclosure. Experimental procedures without conditions specified in the examples of the present disclosure were generally conducted according to conventional conditions such as cell culture manuals or molecular cloning manuals, or according to conditions recommended by the manufacturers of the raw material or the goods. Reagents without specific origins indicated are commercially available conventional reagents.

The following general experimental procedures 1) to 4) are given by means of examples:

1) mRNA production

To generate *in vitro* transcribed mRNA, the plasmid was linearized downstream of the polyA tail using BspQ 1 (Vazyme, DD4302-PC, China) and purified by a PCR purification kit (QIAGEN, 28106, Germany). *In vitro* transcription was performed using the T7 *in vitro* transcription kit (Invitrogen, AM1333, USA) with the purified linearized plasmid as template. The mRNA synthesized by *in vitro* transcription was purified by MEGAclean™ Kit (Invitrogen, AM1908, USA) to give an mRNA with a relatively higher purity for subsequent *in vitro* cell transfection and other experiments.

2) Chip electrophoresis

To examine the integrity and purity of the mRNA obtained by *in vitro* transcription, the mRNA was analyzed using an on-chip electrophoresis system (Agilent 2100, USA) and

the RNA 6000 Nano kit (Agilent, 5067-1511, USA). The procedures are as follows: the mRNA was denatured (incubated for 2 min at 70 °C and quickly transferred on ice), and the treated samples were sequentially transferred on corresponding wells of the chip, loaded for analysis until the analysis results were generated.

3) Cell culture

Materials: FBS (Gibco), DMEM medium (Gibco), Opti-MEM (Gibco), PBS (Gibco), trypsin-EDTA (Gibco), antibiotics (pen/strep, Gibco), and Lipofectamine™ 2000 (Invitrogen, 52887).

Human embryonic kidney cells (293T, ATCC) were cultured in DMEM medium supplemented with 10% of FBS and 1% of the antibiotics in a humidified environment at 5% CO₂.

4) *In vitro* transfection

Cells were seeded in 12-well cell culture plates 8 h before the transfection at 8×10^5 cells/well. The cells were transfected using a commercially available transfection reagent Lipofectamine™ 2000 in a ratio of 1 µg of mRNA to 2.5 µL of Lipofectamine™ 2000, at 2 µg of mRNA/well. The procedures are as follows: Lipofectamine™ 2000 and mRNA were separately diluted in Opti-MEM medium in a final volume of 100 µL, and the Lipofectamine™ 2000 solution and the mRNA solution were mixed and incubated at room temperature for 10-15 min to give a liposome complex containing mRNA. After the DMEM medium in the cell culture plates was discarded, the liposome complexes containing different mRNAs were added to the corresponding plates. Opti-MEM medium containing 10% of FBS was then added at 400 µL/well, and the cells were cultured at 37 °C (5% CO₂) for 4-6 h. The transfection medium was then discarded. DMEM medium containing 10% of FBS and 1% of antibiotics was added at 1 mL/well, and the cells were cultured at 37 °C (5% CO₂) for 48 h. Meanwhile, part of the cells were transfected with Lipofectamine™ 2000 solution with no mRNA as a negative control in the above manner.

The sequences of the target genes used in the present disclosure are summarized in Table 1.

Table 1. Sequence information of target genes of the present disclosure

Protein encoded by target gene	Polypeptide sequence	DNA sequence	mRNA sequence	Optimized DNA sequence	Optimized mRNA sequence	Mutation position compared to wildtype SARS-CoV-2 (KV-PP)
B.1.1.7 (Alpha) spike protein	SEQ ID NO:23	/	/	SEQ ID NO:22	SEQ ID NO:60	983-984
B.1.351 (Beta) spike protein	SEQ ID NO:21	/	/	SEQ ID NO:20	SEQ ID NO:59	983-984
B.1.617.2 (Delta) spike protein	SEQ ID NO:12	/	/	SEQ ID NO:13	SEQ ID NO:54	984-985

B.1.1.529 (Omicron) spike protein	SEQ ID NO:14	/	/	SEQ ID NO:15	SEQ ID NO:55	983-985
SARS-CoV-2 spike protein	SEQ ID NO:25	/	/	SEQ ID NO:24	SEQ ID NO:61	986-987
Wildtype SARS-CoV-2 spike protein	SEQ ID NO:80	SEQ ID NO:81	SEQ ID NO:82	/	/	/

The mutation from amino acids KV to PP in SARS-CoV-2 spike protein can change the spike protein from an unstable prefusion conformation to a stable postfusion conformation (Structure-based design of prefusion-stabilized SARS-CoV-2 spikes, Science, VOL. 369, NO. 6510), such that the spike protein is in a stable state in conformation, which is favorable for vaccine design and production by taking the spike protein as immunogen.

Example 1. Screening and preparation of 5'UTR sequences

1. Screening of UTRs

In order to give a regulatory element capable of highly expressing a target protein in dendritic cells (DCs), this example was conducted with four-stage screening to give a new UTR.

The procedures are as follows: I, a high-protein-expression gene library was established, including genes with a high protein expression in monocytes, genes with a higher protein expression in DCs than in monocytes, genes with a higher protein level, and genes of a UTR with confirmed capability of increasing the protein expression, thus giving 2140 genes. II, a UTR library was established by downloading all sequences in the high-protein-expression gene library from Genbank database (<https://www.ncbi.nlm.nih.gov/genbank/>), analyzed for the integrity, and then analyzed by methods such as repeat combination, CDS (coding sequence) interception and the like to give UTR data, thus giving 940 UTRs. III, a UTR selection library was established by analyzing the UTR library, screening sequences with a proper sequence length (40 to 70 for 5'UTR sequences and 100 to 150 or 300 to 400 for 3'UTR sequences) and free energy, predicting the free energy of the sequences one by one, selecting sequences with a high 5'UTR free energy, and selecting sequences with a low 3'UTR free energy, thus acquiring 64 UTRs from the selection library. IV, final candidate sequences were determined by a screening strategy with a 5'UTR free energy of -10 or higher and a low 3'UTR free energy, finally giving 20 5'UTRs and 3'UTRs for screening and intracellular verification.

On this basis, we believe that the 5'UTR sequence (designated as 5'UTR I, with the

DNA sequence set forth in SEQ ID NO: 1 and the RNA sequence set forth in SEQ ID NO: 45) derived from human Rho GTPase activating protein 15 (abbreviated as ARHGAP15) has the potential of efficiently expressing target proteins in human cells (particularly DCs).

2. Preparation of UTR

This example illustrates the preparation of 5'UTR I-containing mRNA. First, a sequence containing 5'UTR, 3'UTR, polyA tail, and the CDS of a heterologous target gene was artificially synthesized and inserted into pUC57 vector to give a template vector, as shown in FIG. 1. The sequence of the target gene inserted in the vector was confirmed by sequencing.

As shown in the pUC57 Delta13 vector in FIG. 1, the 5'UTR is set forth in SEQ ID NO: 1; the 3'UTR was a human hemoglobin beta subunit (HBB) 3'UTR sequence (SEQ ID NO: 7, designated as 3'UTR B); the CDS sequence was a base sequence (SEQ ID NO: 13) with optimized codons encoding the B.1.617.2 (Delta) spike protein; the polyA tail was a polyadenylic acid sequence (with the DNA sequence set forth in SEQ ID NO: 16 and the RNA sequence set forth in SEQ ID NO: 56); the T7 promoter is set forth in SEQ ID NO: 17.

To generate *in vitro* transcribed mRNA, the plasmid was linearized downstream of the poly A tail using BspQ I (Vazyme, DD4302-PC, China) and purified by a PCR purification kit (QIAGEN, 28106, Germany). *In vitro* transcription was performed to synthesize the mRNA with the purified linearized plasmid as template according to the following procedures. A 100 μ L mRNA reaction system, as shown in Table 2, was synthesized. In the reaction system, 2 μ g of linearized template was diluted to 100 μ L with the non-nucleoside enzyme water. The reaction system was incubated for 4 h at 37 °C and digested for 30 min by DNase I. Detections showed a small amount of byproduct dsRNA in the synthesized mRNA and thus a good yield. The capping process was achieved by chemical capping in the *in vitro* transcription synthesis, with the cap structure being m7G(5')ppp(5')(2'OMeA)pG·NH₄ (Hongene, ON-134). The synthesized mRNA was purified by MEGAclean™ Kit (Invitrogen, AM1908, USA) to give an mRNA with a relatively higher purity for subsequent *in vitro* cell transfection and other experiments.

Table 2. mRNA reaction system

Component	Cat. No.	Manufacturer	System (μ L)
ATP solution (100 mM)	R1331	Hongene	4.41
CTP solution (100 mM)	R3331	Hongene	5
GTP solution (100 mM)	R2331	Hongene	2.79
10 $\times$ transcription buffer	DD4101R-02	Vazyme	10
T7 RNA polymerase	DD4101-PC-02	Vazyme	6.6
Mouse RNase inhibitor	DD4102-PA	Vazyme	2.75

Pyrophosphatase	DD4103-PC-01	Vazyme	0.66
Pseudo-UTP	R5-022	Hongene	2.79
Cap-GAG Trimer	m7G(5')ppp(5')(2'OMeA)pG	Hongene	1

Through screenings and functional verification, the 5'UTR I with the potential of efficiently expressing the target protein was acquired, with a sequence set forth SEQ ID NO: 1; and the construct containing the 5'UTR I can efficiently transcribe the corresponding mRNA, with a size consistent with the theoretical value as determined by chip electrophoresis (see FIG. 2).

Example 2. Screening of combinations of different 5'UTRs and 3'UTRs for efficiently expressing heterologous proteins

In this example, different combinations of 5 'UTRs and different combinations of 3'UTRs were screened to give UTR combinations capable of efficiently expressing a target gene.

The procedures are as follows: Constructs of plasmids containing the CDS of the same target gene with different UTR combinations were artificially synthesized. The basic plasmid was constructed according to the procedures in Example 1. The target gene CDS sequence was introduced by recombinant construction via artificial synthesis and HindIII/XhoI digestion. The reconstructed plasmid was linearized and purified mRNA was synthesized according to the procedures in Example 1. mRNA was transfected into 293T cells by Lipofectamine™ 2000. 293T cell culture supernatant was collected after 48 h and detected by Western Blot and ELISA to determine the level of secreted spike proteins in the cell supernatant.

The target genes used were: the gene encoding the SARS-CoV-2 B.1.351 spike protein (with a CDS set forth in SEQ ID NO: 20), or the gene encoding the SARS-CoV-2 B.1.1.7 spike protein (with a CDS set forth in SEQ ID NO: 22). The information of different 5'UTRs and 3'UTRs and their combinations are shown in Tables 3 and 4.

Western Blot procedures: To detect the *in vitro* expression level of the constructed mRNAs, the cell supernatant in the plate was collected after the transfection was completed, and the *in vitro* expression level of the constructed mRNAs was finally obtained by polyacrylamide gel electrophoresis, membrane transfer, co-incubation with the primary antibody and secondary antibody, and color development. The primary antibody was a SARS-CoV-2 (2019-nCoV) spike RBD antibody (Sino Biology, 40592-T62), and the secondary antibody was an HRP-linked anti-rabbit IgG (Transgene, HS101).

ELISA procedures: after the transfection was completed, the cell supernatant was collected, diluted in a ratio of 1:50 to 1:100, and detected for the expression level of

the constructed mRNAs *in vitro* using an ELISA kit (SARS-CoV-2 (2019-nCoV) spike detection ELISA KIT, KIT40591, Sino Biology). The absorbance at 450 nm was read and the concentration of the relevant antigen was calculated.

Table 3. Sources and sequences of 5'UTR and 3'UTR

UTR	Abbreviation	Source	GenBank ID	DNA sequence (SEQ ID NO:)	RNA sequence (SEQ ID NO:)
5'UTR	I	Human ARHGAP15	XM_003822828.3	1	45
	G	Human heat shock 27kDa protein 1 (HSPB1)	BC012292.1	3	47
	A	Human DNA, containing HBB gene	LC507571.1	4	48
	J	Human C-C motif chemokine ligand 13 (CCL13)	NM_005408.3	5	49
3'UTR	B	Human HBB	MK476504.1	7	50
	F	Human ARHGAP15	NM_018460.4	8	51
	E	Coronin 1A (CORO1A)	NM_007074.4	9	52
	D	Hemopexin (HPX)	NM_000613.3	10	53

5

Table 4. Different combinations of 5'UTR and 3'UTR

Name	5'UTR	3'UTR
A-B	A	B
A-E	A	E
A-D	A	D
A-F	A	F
G-E	G	E
G-B	G	B
G-D	G	D
I-E	I	E
I-B	I	B
I-F	I	F
I-D	I	D
J-E	J	E

In Table 4, 5'UTRs A, G, I, and J are set forth in SEQ ID NOs: 4, 3, 1, and 5,

respectively; 3'UTRs B, E, D, and F are set forth in SEQ ID NO: 7, 9, 10, and 8, respectively.

The effects of different combinations of 5'UTR with 3'UTR on the expression of SARS-CoV-2 B.1.351 spike protein and SARS-CoV-2 B.1.1.7 spike protein are shown in FIG. 3 and Table 5 and FIG. 4 and Table 6, respectively.

Table 5. Regulatory result of different UTR combinations on B.1.351 spike protein expression level (ELISA)

Sample No.	#1: A-B	#3: A-E	#5: A-D	#8: G-B	#10: G-D	#11: I-E	#13: I-B
Spike protein concentration (mean±SD, ng/mL)	48.2±13.4	6.3±5.4	2.9±0.28	36.7±10.5	26±0.14	43.4±2.97	69.4±10.88

(sample numbering rule is that, for example, #1: A-B denotes sample #1 and a UTR combination of A-B)

Table 6. Regulatory result of different UTR combinations on B.1.1.7 spike protein expression level (ELISA)

Sample No.	UK-#11: I-E	UK-#13: I-B	UK-#14: I-F	UK-#15: I-D	NC (Lipo control)
Spike protein concentration (mean±SD, ng/mL)	118.35±3.61	206.1±19.66	227.6±4.38	153.55±7.99	-9.65±0.64

(sample numbering rule is that, for example, UK-#11: I-E denotes sample UK-#11 and a UTR combination of I-E)

As can be seen from FIG. 3 and Table 5, the comparison of B.1.351 spike protein expression level mediated by different UTR combinations is as follows: I-B > A-B > G-B, and I-B > I-E. For the 5'UTR, I is superior to A and G; for relevant combinations of I, I-B is superior to I-E. As can be seen from FIG. 4 and Table 6, the comparison of B.1.1.7 spike protein expression level mediated by different UTR combinations is as follows: I-F > I-B > I-D > I-E. By analyzing the results of B.1.351 and B.1.1.7 spike protein expression levels, the I-B UTR combination exhibited more significant advantages in mediating the efficient expression of the target gene, while the I-F combination also exhibited advantages in high-level expression of the target gene.

Furthermore, the mRNAs containing I-B or I-F regulated B.1.617.2 spike protein were prepared, transfected into cells, and tested by ELISA. The spike protein expression level results are shown in Table 7.

Table 7. I-B and I-F UTR regulated B.1.617.2 spike protein expression level (ELISA)

Sample	I-B	I-F
Spike protein concentration (ng/mL)	57.1	58.8

Unlike the 5'UTR and 3'UTR in the I-B combination that are from different genes, the 5'UTR and 3'UTR in the I-F combination come from the same gene ARHGAP15. The results in Table 7 show that the I-B (SEQ ID NO: 1 and SEQ ID NO: 7) and I-F (SEQ ID NO: 1 and SEQ ID NO: 8) combinations regulate the expression of the target gene at comparable levels, further illustrating the universality of 5'UTR I for combination with different 3'UTRs to mediate efficient expression of the target gene (e.g., the gene encoding the spike protein).

Example 3. Sequence engineering and preliminary functional validation of 5'UTR I

1) Bioinformatic analysis and mutation design

5'UTR I (SEQ ID NO: 1) is a naturally occurring promoter of the human Rho GTPase activating protein family. The data of Examples 1-2 support that 5'UTR I assists in efficiently expressing the SARS-CoV-2 spike protein. However, 5'UTR I contains an ATG and may express a 27aa short peptide, which may affect normal protein expression of the target gene. In order to avoid the phenomenon that ribosome initiates the translation from the ATG inside the UTR, increase the probability of translation from the ATG in the CDS region, and increase the expression, a point mutation (A→G) was introduced to give a mutant 5'UTR I' (with a DNA sequence set forth in SEQ ID NO: 2 and an RNA sequence set forth in SEQ ID NO: 46). The mutation positions are shown in FIG. 5.

2) Point mutation from 5'UTR I to 5'UTR I' effectively increasing the expression of target genes

Plasmids containing 5'UTR I-B and 5'UTR I'-B to regulate the target gene expression were prepared according to the method in Example 1. The target gene was the gene encoding B.1.617.2 spike protein (SEQ ID NO: 13). The plasmids were designated as Delta 13 and Delta 24, respectively, and the mRNAs were synthesized by PCR and transcription and purified (the cap-free I'-Delta24-B mRNA sequence is set forth in SEQ ID NO: 57). The prepared mRNA was transfected into 293T cells, and the B.1.617.2 spike protein expression was detected using the Western Blot method in Example 2. By gray scanning using Image software, the grayscale value was calculated and normalized by the expression level of sample 1 (Delta 13) to give the ratio in Table

8.

The results are shown in FIGs. 6A and 6B and Table 8. The results show that for mRNAs prepared using either the minipreparation plasmid (with no endotoxin removal) or the maxipreparation plasmid (with endotoxins removed) as the template, the expression level of the spike protein mediated by the construct of 5'UTR I' after point mutation was greater than that mediated by the construct containing 5'UTR I. Thus, the 5'UTR I modified by the point mutation in Example 3 effectively improves the expression of heterologous proteins.

Table 8. Target gene expression levels regulated by I-B and I'-B UTRs (Western Blot assay)

Sample	Plasmid mini-pre		Plasmid maxi-pre	
	Sample 1 (Delta 13)	Sample 2 (Delta 24)	Sample 3 (Delta 13)	Sample 4 (Delta 24)
Ratio to expression of Sample 1	1.00	4.88	3.98	5.44

Example 4. Assay of mRNA expression of different target genes regulated by 5'UTR I

Plasmids containing the gene encoding B.1.617.2 (Delta) spike protein (SEQ ID NO: 13) and the gene encoding B.1.1.529 (Omicron) spike protein (SEQ ID NO: 15) were constructed by replacing the target gene via digestion, with the 5'UTR I' and the 3'UTR B retained. The target gene sequences contained in the plasmids, as set forth in SEQ ID NO: 13 and SEQ ID NO: 15, were codon-optimized sequences encoding target protein sequences set forth in SEQ ID NO: 12 and SEQ ID NO: 14. mRNAs were prepared and purified as in Example 1. 2 µg of mRNA was transfected to 293T cells by Lipofectamine™ 2000, and the cell supernatant was harvested after 2 days. As per the Western Blot method in Example 2, the expression of I-B regulated B.1.617.2 spike protein mRNA (designated as Delta 13), I'-B regulated B.1.617.2 spike protein mRNA (designated as Delta 24), and I'-B regulated B.1.1.529 spike protein mRNA (designated as Omicron 24) were determined. The sequences of cap-free Delta 24 mRNA and Omicron 24 mRNA are set forth in SEQ ID NO: 57 and SEQ ID NO: 58. By gray scanning using Image software, the grayscale value was calculated and normalized by the expression level of Delta 13 to give the ratio in Table 9.

FIGs. 7A and 7B show that under the regulation of I'-B UTR combination, the SARS-CoV-2 B.1.617.2 and B.1.1.529 genes were expressed efficiently. The results suggest that 5'UTR I' has universality and can initiate the efficient expression of

different target genes. Comparing 5'UTRs I and I', the Delta 24 spike protein exhibited an expression level significantly higher than that of the Delta 13 spike protein, indicating that the point mutation modification of 5'UTR I' effectively improves the protein expression efficiency, and the result further confirms the conclusion of Example 3.

Table 9. Target gene expression levels regulated by I-B and I'-B UTR combinations (Western Blot assay)

Sample	Delta 13	Delta 24	Omicron 24
Ratio to expression of sample Delta 13	1.00	1.69	1.48

Example 5. Assay of luciferase expression regulated by I'-B UTR combination

A vector containing the luciferase CDS sequence (SEQ ID NO: 26) and the same T7 promoter (SEQ ID NO: 17) and polyA tail as in the previous vector containing the SARS-CoV-2 gene was constructed. Three vectors Luc, Luc13, and Luc24 with different UTRs were separately constructed, and the UTRs contained in the vectors are shown in Table 10. Among these, the 5'UTR and 3'UTR combination of BNT162b2 in Luc was used as the control, which is from CN113521269A. Capped mRNA was synthesized by *in vitro* transcription as in Example 1 and transfected *in vitro*. The supernatant was collected and luciferase substrate was added for detection. The plate was read, and the results are shown in Table 11 and FIG. 8.

Table 10. Vector structure

Sample No. Component	Luc	Luc13	Luc24
5'UTR	BNT162b2(SEQ ID NO: 6)	I	I'
3'UTR	BNT162b2(SEQ ID NO: 11)	B	B

Table 11. Luciferase expression levels regulated by I'-B and I-B

Sample	Luc	Luc13	Luc24
Run I	774.81	992.36	1619.76
Run II	841.14	923.83	1519.50

Note: The parameters are expressed in relative light unit (RLU).

The results show that the luciferase protein expression regulated by 5'UTR I' after point mutation was higher than that by 5'UTR I. Compared to the 5'UTR and 3'UTR combination of BNT162b2 in Luc control, the I-B combination of the present disclosure results in a higher expression level of the target gene, while the I'-B combination results in a significant increase in the expression level of the target gene.

The assay results using luciferase as the target gene also suggest that the 5'UTR I or 5'UTR I' can regulate the efficient expression of any target protein with universality.

Example 6. Assay of luciferase expression regulated by 5'UTR I' truncations and mutants

6.1 The 5' end of the 5'UTR I' was successively truncated, and the efficiency of the 5'UTR I' truncations of different lengths in regulating protein expression was analyzed by bioinformatic methods (Nat Biotechnol. Jul., 2019; 37(7):803-809). The relative MRL values of the 5'UTR I' truncations to the 5'UTR I' were calculated, and the results are shown in Table 12 below. The results in Table 12 show that successive truncations of the nucleotide sequence at the 5' end of the 5'UTR I' did not cause significant reduction in the MRL value, indicating that the 5'UTR I' truncations can retain the functional activity of regulating the expression of the target protein. The 3' end of the 5'UTR I' truncation set forth in SEQ ID NO: 153 was successively truncated, and the relative MRL values of the 3' end truncations to the nucleotide sequence set forth in SEQ ID NO: 153 were calculated, as shown in Table 13. As can be seen from Table 13, the 3' end truncations did not exhibit a significant effect on the regulatory efficiency of the 5'UTR I' truncations.

Table 12. Relative MRL values of 5'UTR I' truncations and 5'UTR I'

5'UTR I' truncation	Relative MRL value	SEQ ID NO:
AGTTAGGAGTTGGTGGCAGTTTCAATAACA GGTCATTGCCGAGAAAAGGATAGCACTATA AT	1	SEQ ID NO: 137
GTTAGGAGTTGGTGGCAGTTTCAATAACAG GTCATTGCCGAGAAAAGGATAGCACTATAA T	1	SEQ ID NO: 138
TTAGGAGTTGGTGGCAGTTTCAATAACAGG TCATTGCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 139
TAGGAGTTGGTGGCAGTTTCAATAACAGGT CATTGCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 28
AGGAGTTGGTGGCAGTTTCAATAACAGGTC ATTGCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 140
GGAGTTGGTGGCAGTTTCAATAACAGGTCA	1	SEQ ID NO: 141

TTGCCGAGAAAAGGATAGCACTATAAT		
GAGTTGGTGGCAGTTTCAATAACAGGTCAT TGCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 142
AGTTGGTGGCAGTTTCAATAACAGGTCATT GCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 29
GTTGGTGGCAGTTTCAATAACAGGTCATTG CCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 143
TTGGTGGCAGTTTCAATAACAGGTCATTGC CGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 144
TGGTGGCAGTTTCAATAACAGGTCATTGCC GAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 145
GGTGGCAGTTTCAATAACAGGTCATTGCCG AGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 30
GTGGCAGTTTCAATAACAGGTCATTGCCGA GAAAAGGATAGCACTATAAT	1	SEQ ID NO: 146
TGGCAGTTTCAATAACAGGTCATTGCCGAG AAAAGGATAGCACTATAAT	0.8	SEQ ID NO: 147
GGCAGTTTCAATAACAGGTCATTGCCGAGA AAAGGATAGCACTATAAT	1	SEQ ID NO: 148
GCAGTTTCAATAACAGGTCATTGCCGAGAA AAGGATAGCACTATAAT	1.02	SEQ ID NO: 31
CAGTTTCAATAACAGGTCATTGCCGAGAAA AGGATAGCACTATAAT	1.02	SEQ ID NO: 149
AGTTTCAATAACAGGTCATTGCCGAGAAAA GGATAGCACTATAAT	1.02	SEQ ID NO: 150
GTTTCAATAACAGGTCATTGCCGAGAAAAG GATAGCACTATAAT	1	SEQ ID NO: 151
TTTCAATAACAGGTCATTGCCGAGAAAAGG ATAGCACTATAAT	1.01	SEQ ID NO: 32
TTCAATAACAGGTCATTGCCGAGAAAAGGA TAGCACTATAAT	1.01	SEQ ID NO: 152
TCAATAACAGGTCATTGCCGAGAAAAGGAT AGCACTATAAT	1.01	SEQ ID NO: 153
CAATAACAGGTCATTGCCGAGAAAAGGATA GCACTATAAT	1.02	SEQ ID NO: 154

AATAACAGGTCATTGCCGAGAAAAGGATAG CACTATAAT	1.02	SEQ ID NO: 33
ATAACAGGTCATTGCCGAGAAAAGGATAGC ACTATAAT	1.02	SEQ ID NO: 155
TAACAGGTCATTGCCGAGAAAAGGATAGCA CTATAAT	1.01	SEQ ID NO: 156
AACAGGTCATTGCCGAGAAAAGGATAGCA CTATAAT	1.01	SEQ ID NO: 157
ACAGGTCATTGCCGAGAAAAGGATAGCACT ATAAT	1	SEQ ID NO: 34
CAGGTCATTGCCGAGAAAAGGATAGCACTA TAAT	1	SEQ ID NO: 158
AGGTCATTGCCGAGAAAAGGATAGCACTAT AAT	1	SEQ ID NO: 159
GGTCATTGCCGAGAAAAGGATAGCACTATA AT	0.97	SEQ ID NO: 160
GTCATTGCCGAGAAAAGGATAGCACTATAA T	0.98	SEQ ID NO: 35
TCATTGCCGAGAAAAGGATAGCACTATAAT	0.99	SEQ ID NO: 161
CATTGCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 162
ATTGCCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 163
TTGCCGAGAAAAGGATAGCACTATAAT	0.98	SEQ ID NO: 36
TGCCGAGAAAAGGATAGCACTATAAT	0.93	SEQ ID NO: 164
GCCGAGAAAAGGATAGCACTATAAT	0.99	SEQ ID NO: 165
CCGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 166
CGAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 37
GAGAAAAGGATAGCACTATAAT	1	SEQ ID NO: 167
AGAAAAGGATAGCACTATAAT	1.01	SEQ ID NO: 168
GAAAAGGATAGCACTATAAT	1	SEQ ID NO: 169
AAAAGGATAGCACTATAAT	1	SEQ ID NO: 38
AAAGGATAGCACTATAAT	1	SEQ ID NO: 170
AAGGATAGCACTATAAT	0.99	SEQ ID NO: 171
AGGATAGCACTATAAT	0.99	SEQ ID NO: 172
GGATAGCACTATAAT	0.97	SEQ ID NO: 39

Table 13. Relative MRL values of 3' end-truncated 5'UTR I' truncations to untruncated

form

5'UTR I' truncation	Relative MRL value	SEQ NO ID:
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCA	0.99	SEQ ID NO: 173
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC TA	0.98	SEQ ID NO: 174
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC T	1	SEQ ID NO: 175
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC TATA	0.99	SEQ ID NO: 176
TCAATAACAGGTCATTGCCGAGAAAAG	0.97	SEQ ID NO: 177
TCAATAACAGGTCATTGCCGAGAAAAGGATA	0.97	SEQ ID NO: 178
TCAATAACAGGTCATTGCCGAGAAAAGGAT	0.98	SEQ ID NO: 179
TCAATAACAGGTCATTGCCGAGAAAAGGATAGC	0.99	SEQ ID NO: 180
TCAATAACAGGTCATTGCCGAGAAAA	0.95	SEQ ID NO: 181
TCAATAACAGGTCATTGCCGAGAAAAGG	0.95	SEQ ID NO: 182
TCAATAACAGGTCATTGCCGAGAAA	0.96	SEQ ID NO: 183
TCAATAACAGGTCATT	0.93	SEQ ID NO: 184
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC TATAA	1.01	SEQ ID NO: 185
TCAATAACAGGTCATTGCCGAGAAAAGGA	0.97	SEQ ID NO: 186
TCAATAACAGGTCATTGCCGAG	0.94	SEQ ID NO: 187
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC TAT	0.93	SEQ ID NO: 188
TCAATAACAGGTCATTGCCGAGAA	0.96	SEQ ID NO: 189
TCAATAACAGGTCATTG	0.9	SEQ ID NO: 190
TCAATAACAGGTCATTGCCGAGAAAAGGATAG	0.98	SEQ ID NO: 191
TCAATAACAGGTCATTGCCGAGA	0.94	SEQ ID NO: 192
TCAATAACAGGTCAT	0.91	SEQ ID NO: 193
TCAATAACAGGTCATTGCC	0.92	SEQ ID NO: 194
TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC	0.94	SEQ ID NO: 195
TCAATAACAGGTCATTGCCG	0.93	SEQ ID NO: 196
TCAATAACAGGTCATTGCCGA	0.92	SEQ ID NO: 41
TCAATAACAGGTCATTGC	0.85	SEQ ID NO: 76

6.2 16 constructs containing 5'UTR I' truncations with different lengths (SEQ ID NOs: 28-40, CTATAAT) and 2 point mutations A→C and A→T (SEQ ID NOs: 42 and 43) different from the A→G point mutation were synthesized by Genewiz, Suzhou. The 16 synthesized sequences were inserted into Luc 24 in Example 5 by HindIII and ScaI digestion and ligation, thereby given constructs with corresponding luciferase 5'UTR I' truncations and point mutations, designated as LUC1-16, with corresponding sequences of the 5'UTR I' truncations (1-I' to 14-I') and mutants (15-I' and 16-I') set forth in SEQ ID NO: 63-78, respectively. Capped mRNA was synthesized by *in vitro* transcription as in Example 1 and transfected into 293T cells *in vitro* for 24 h. The luciferase substrate was added for detection, and the plate was read. The results are shown in FIG. 9. The luciferase expression levels mediated by Luc24 (I'-B combination) and 5'UTR truncations and mutants thereof were significantly higher than that mediated by Luc (5' and 3'UTRs of BNT162b2). The 5'UTR I' truncations can retain the activity of 5'UTR I' in regulating the expression of the target protein. Part of the 5'UTR I' truncations (1-I', 2-I', 3-I', 4-I', 5-I', 6-I', 7-I', 9-I', 10-I', 11-I', and 13-I') exhibited increased activity in regulating the expression of the target protein. Meanwhile, the mutants 15-I' and 16-I' of the 5'UTR I' mediated higher expression levels of the target gene than that of 5'UTR I', indicating that the functional activity of 5'UTR in regulating the expression of the target gene can be further enhanced by truncating the 5' end of 5'UTR I' or point mutation inside.

Example 7. Differences in influenza HA mRNA expression levels regulated by control 5'UTR and 5'UTR I'

Plasmids of the gene encoding influenza A virus H1N1 (A/Wisconsin/588/2019) HA protein (SEQ ID NO: 83), the gene encoding influenza A virus H3N2 (A/Cambodia/e0826360/2020) HA protein (SEQ ID NO: 84), the gene encoding influenza B virus (Washington/02/2019) HA protein (SEQ ID NO: 85), and the gene encoding influenza B virus (Phuket/3073/2013) HA protein (SEQ ID NO: 86) were constructed by replacing the target gene via digestion, with the 5'UTR I' and the 3'UTR B retained. The target gene sequences contained in the plasmids were SEQ ID NOs: 87-90, which were all codon-optimized sequences. Four HA CDS sequences encoding SEQ ID NOs: 83-86 were synthesized with reference to 5' and 3'UTR sequences (SEQ ID NOs: 91 and 92) in Patent Nos. WO2022/245888A1 and WO2022/150717A1, as set forth in SEQ ID NOs: 93-96, respectively. mRNAs were prepared and purified as in Example 1. 1 µg of mRNA was transfected to 293T cells by Lipofectamine™ 2000, and the lysate was harvested after 1 day. The expression of the four influenza HA

protein mRNAs regulated by I'-B was detected by the Western Blot method in Example 2. The primary antibody was an influenza A virus hemagglutinin/HA antibody (Sino Biology, 86001-RM01) or an influenza B virus hemagglutinin/HA antibody against (Sino Biology, 11053-R004), and the secondary antibody was an HRP-linked anti-rabbit IgG (Transgene, HS101). The results are shown in FIGs. 10 and 11. The expression levels of the four influenza HA regulated by 5'UTR-I' were significantly higher as compared to the control 5'UTR.

Part of the sequences of the present disclosure are shown as follows:

10 >5'UTR I (63bp)
GAGTTAGGAGTTGATGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGAT
AGCACTATAAT

SEQ ID NO: 1

>5'UTR I' (63bp)
15 GAGTTAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGAT
AGCACTATAAT

SEQ ID NO: 2

>5'UTR G
GCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCGCCACC
20

SEQ ID NO: 3

>5'UTR A
ACATTTGCTTCTGACACAACGTGTGTTCACTAGCAACCTCAAACGCCGCCACC
SEQ ID NO: 4

>5'UTR J
25 AGAGAGGCCAAAGAAACATTGTGAAATCTCCAACCTTAAACCTTCAAGCCGCC
ACC
SEQ ID NO: 5

>5'UTR BNT162b2
GAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACC
30

SEQ ID NO: 6

>3'UTR B
GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCC
AACTACTAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCT
AATAAAAAACATTTATTTTCATTGCAA
35

SEQ ID NO: 7

>3'UTR F
CAGACAAGACAAGCTACTGAATACGTTACATCTGTCTTGATGCCTAATATTT
TTACATTTCTGTAAACATATTTCTGAAATATTTTTGCCTTTCAAGCGACAGAT
GCCTCATTTTGTGAAAACCTTAATGATGATTTTGTGTTTAAAGTTCCAAACATTTG
40 AATAAAATAATTGACAATA

SEQ ID NO: 8

>3'UTR E

CCTGAGGCCTCAGAACACAGGTTTCAGATTGATAGGCCTGCAGGTCTCCAGG
CAGCAACCAGCTGAGCGACTAAAGGGCCCAAGGCCAGGGCTCTAGGGATGG
5 GGCTCAGCAGAGGCTGGGGTAAGGGGAGCCAGGGAGGAGCTGGGCCTAATG
CAGCACCGGGTCCCCAGGATGCCGTGTCTCGGCTGGAGGAGGAGATGCGGA
AGCTCCAGGCCACGGTGCAGGAGCTCCAGAAGCGCTTGGACAGGCTGGAGG
AGACAGTCCAGGCCAAGTAGAGCCCCGCAGGGCCTCCAGCAGGGTCAGCCA
TTCACACCCATCCACTCACCTCCCATTCCCAGCCACATGGCAGAGAAAAAAA
10 TCATAATAAAATGGCTTTATTTTCTGGTA

SEQ ID NO: 9

>3'UTR D

GGGGCCTTCTGACATGAGTCTGGCCTGGCCCCACCTCCTAGTTCCTCATAATA
AAGACAGATTGCTTCTTCGCTTCTCACTGAGGGGCCTTCTGACATGAGTCTG
15 GCCTGGCCCCACCTCCCCAGTTTCTCATAATAAAGACAGATTGCTTCTTCACT
TGAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAA
GTCCAACCTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCT
GCCTAATAAAAAACATTTATTTTCATTGCAA

SEQ ID NO: 10

20 >3'UTR BNT162b2

CTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCCCTTCCCGTCCTGG
GTACCCCGAGTCTCCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACC
TGCCCCACTCACCACTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCA
ATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTG
25 ATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCCAGGG
TTGGTCAATTTTCGTGCCAGCCACACCCTGGA

SEQ

ID NO: 11

>B.1.617.2 (Delta) spike protein amino acid sequence

30 MFVFLVLLPLVSSQCVNLRTTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLF
LPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTL
DSKTQSLIVNNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESGVYSSANN
CTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGF
SALEPLVDLPIGINITRFQTLALHRSYLT PGDSSSGWTAGAAAYYVGYLQPRFTL
35 LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNI
TNLCPFGGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL
NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNL
DSKVGGNYNYRYRLFRKSNLKPFERDISTEYIYQAGSKPCNGVEGFNCYFPLQSY
GFQPTNGVGYQP YRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG
40 TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNT

SNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVN
 NSYECDIPIGAGICASYQTQTNSRRRARSVASQSIIAYTMSLGAENSVAYSNNNSIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGI
 AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKV
 5 TLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG
 TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKI
 QDSLSTASALGKLQNVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDPPEA
 EVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFC
 GKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGDKAHFPREGVVFV
 10 SNGTHWVFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEE
 LDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKY
 EQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCCLKGCCSCGSCCKFDEDD
 SEPVLKGVKLHYT

SEQ ID NO: 12

15 >B.1.617.2 (Delta) spike protein CDS
 ATGTTTCGTGTTCTTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACC
 TGAGGACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAG
 GCGTGTACTACCCCGACAAGGTGTTTCTCAGAAGCAGCGTGCTGCACTCCACCCA
 AGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCCAC
 20 GTCAGCGGCACCAACGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTC
 AACGACGGCGTGACTTCGCTAGCACCGAAAAGAGCAACATCATCAGAGGC
 TGGATCTTCGGCACCACTAGACTCCAAGACACAAAGCCTGCTGATCGTTA
 ACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACG
 ACCCCTTCCTGGACGTGTACTACCACAAGAACAACAAGAGCTGGATGGAGA
 25 GCGGGGTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCC
 CTTCTGATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGA
 GTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACC
 CCCATCAACCTGGTGAGAGACCTGCCCCAAGGCTTCAGCGCCCTGGAGCCC
 CTGGTGACCTGCCCATCGGCATCAACATCACAAGATTTACAGACCCTGCTGG
 30 CCCTGCACCGTAGCTACCTAACCCCCGGCGATAGTTCTCTGGCTGGACGGC
 CGGCGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTG
 AAGTACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACC
 CCCTGAGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAAAGGGCA
 TCTATCAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTGAGATT
 35 CCCCCAACATCACCAACCTGTGCCCTTCGGCGAGGTGTTCAACGCCACAAGA
 TTCGCTAGCGTGACGCGTGGAACAGAAAGAGAATCAGCAACTGCGTGGCC
 GACTACAGCGTGCTGTACAACAGCGCTAGCTTCAGCACCTTCAAGTGCTATG
 GCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCCG
 ACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCCGGGCAAA
 40 CCGGCAAGATCGCCGACTACAACCTACAAGCTGCCCGACGACTTCACCGGCTG
 CGTGATCGCGTGGAACAGCAATAACCTGGATTCCAAGGTGGGCGGCAACTAC

AACTACCGGTACAGACTGTTTCAGAAAGAGCAACCTGAAGCCCTTCGAGAGA
 GACATCAGCACCGAGATCTACCAAGCCGGCAGCAAGCCCTGCAACGGCGTG
 GAGGGCTTCAACTGCTACTTCCCCCTGCAGAGCTACGGCTTTCAGCCCACCA
 ACGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGAGCTGCT
 5 GCACGCCCCCGCAACCGTGTGCGGCCCCAAAAAGAGCACCAACCTGGTGAA
 GAACAAGTGCGTGAACCTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTG
 ACCGAGAGCAACAAGAAGTTCCTGCCATTTTCAGCAATTCGGCAGAGACATC
 GCCGACACCACCGACGCGGTACAGAGACCCTCAGACCCTGGAGATCCTGGAC
 ATCACCCCCTGTAGCTTCGGCGGGCGTGAGCGTGATCACCCCGGCACTAACA
 10 CAAGCAACCAAGTGGCCGTGCTGTACCAAGGCGTGAACCTGCACCGAGGTGC
 CCGTGGCCATCCACGCCGATCAGCTGACCCCCACCTGGAGAGTATACAGCAC
 CGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGGCGCCGAGCA
 CGTGAACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATCTGCGCT
 AGCTATCAGACACAGACCAACAGCCGCCGGAGAGCTAGAAGCGTGGCTTCT
 15 CAGAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCT
 ACAGCAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACCAC
 CGAGATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCATGTAC
 ATCTGCGGCGACAGCACCGAGTGACAGCAACCTGCTGCTGCAGTACGGCAGC
 TTCTGCACACAGCTGAACAGAGCCCTGACCGGCATCGCCGTGGAGCAAGAC
 20 AAGAACACCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTACAAGACCCCC
 CCCATCAAGGACTTCGGCGGGCTTCAACTTCAGCCAAATCCTGCCCGACCCTA
 GCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGA
 CCCTGGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACAT
 CGCAGCTAGAGACCTGATCTGCGCTCAGAAGTTCAATGGACTCACCGTGCTC
 25 CCCCCCTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTGCTGG
 CGGGAACCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTGCAAA
 TCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCATCGGCGTGACACA
 GAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAGCGCC
 ATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCAAG
 30 CTGCAAAACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAG
 CAGCTGAGCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGT
 CTAGACTGGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCG
 GCAGACTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGAGCCG
 CCGAGATCAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAGTGCG
 35 TGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCACCTGA
 TGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGACCTA
 CGTGCCCGCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCACGA
 CGGCAAGGCCCACTTCCCTAGAGAGGGCGTGTTTCGTGAGCAACGGCACCCA
 CTGGTTTCGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACCGAC
 40 AACACCTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTGAACAAC
 ACCGTATACGACCCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAGGAGCTG

GACAAGTACTTCAAAAACCATACTAGCCCCGACGTGGACCTGGGAGACATCT
 CTGGCATCAACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCGGCTGAA
 TGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCTGGG
 CAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATC
 5 GCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAA
 GCTGTTGCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTT
 CGACGAGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACAC
 CTGATGA

SEQ ID NO: 13

10 >B.1.1.529 (Omicron) spike protein amino acid sequence
 MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLF
 LPFFSNVTWFHVISGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSK
 TQSLIVNATNVVIKVECFQFCNDPFLDHKNKSWMESEFRVYSSANNCTFEY
 VSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPIIVREPEDLPQGFSAL
 15 EPLVDLPIGINITRFQTLALHRSYLTGDSSTGWTAGAAAYYVGYLQPRTFLK
 YNENGITITDAVDCALDPLSETKCTLKSTVEKGIYQTSNFRVQPTESIVRFPNITN
 LCPFDEVFNATRFASVYAWNRKRISNCVADYSVLYNLAPFFTFKCYGVSPTKLN
 DLCFTNVYADSFVIRGDEVQRQIAPGQTGNIADYNYKLPDDFTGCVIAWNSNKLD
 SKVSGNLYLYRLFRKSNLKPFERDISTEYQAGNKPCNGVAGFNCYFPLRSYSF
 20 RPTYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLKGTG
 VLTESNKKFLPFQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSN
 QVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA EYVNS
 YECDIPGAGICASYQTQTKSHRRARSVASQSIIAYTMSLGAENSVAYSNNNSIAIPT
 NFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLKRALTGIAV
 25 EQDKNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTL
 ADAGFIKQYGDCLGDIAARDLCAQKFKGLTVLPPLTDEMIAQYTSALLAGTIT
 SGWTFGAGAAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDS
 LSSTASALGKLQDVVNHNAAQALNTLVKQLSSKFGAISSVLNDIFSRLDPPEAEVQ
 IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKG
 30 YHLSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPREGV FVSNGT
 HWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDK
 YFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQY
 IKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEP
 VLKGVKLHYT

SEQ ID NO: 14

35 >B.1.1.529 (Omicron) spike protein CDS
 ATGTTCTGTGTTCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACC
 TGACCACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAG
 GCGTGTACTACCCCGACAAGGTGTTTCAAGAAGCAGCGTGCTGCACTCCACCCA
 40 AGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGTGATCAGC
 GGCACCAACGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTCAACGAC

GGCGTGTACTTCGCTAGCATCGAAAAGAGCAACATCATCAGAGGCTGGATCT
TCGGCACCACCCTAGACTCCAAGACACAAAGCCTGCTGATCGTTAACAACGC
GACCAACGTGGTGTATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACCCCTTC
CTGGACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCAGAGTGTAC
5 AGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCTGATGG
ACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTCGTGTTCA
AGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCCATCATCGT
GAGAGAGCCCGAGGACCTGCCCAAGGCTTCAGCGCCCTGGAGCCCTGGT
GGACCTGCCCATCGGCATCAACATCACAAGATTTAGACCCTGCTGGCCCTG
10 CACCGTAGCTACCTAACCCCGGCGATAGTTCCTCTGGCTGGACGGCCGGCG
CTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAGTA
CAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTG
AGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTAT
CAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTGAGATTCCCC
15 AACATCACCAACCTGTGCCCTTCGACGAGGTGTTCAACGCCACCAGATTTG
CTAGCGTGTACGCCTGGAACCGGAAGAGAATCAGCAACTGCGTGGCCGACT
ACAGCGTGCTGTACAACCTGGCCCCCTTCTTCACCTTCAAGTGCTATGGCGTC
AGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCCGACAGC
TTCGTGATCAGAGGCGACGAGGTGAGACAGATCGCCCCCGGGCAGACCGGC
20 AACATCGCCGACTACAACCTACAAGCTGCCCGACGACTTCACCGGCTGCGTGA
TCGCCTGGAACAGCAACAAGCTGGATAGCAAGGTGAGCGGCAACTACAATT
ACCTGTACAGACTGTTTCAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACA
TCAGCACCGAGATCTACCAAGCCGGCAACAAGCCCTGCAACGGCGTGGCGG
GCTTCAACTGCTACTTCCCCCTCAGGAGCTACAGCTTCAGACCCACCTACGG
25 CGTGGGCCATCAGCCCTACAGAGTGGTCGTGCTGAGCTTCGAGCTGCTGCAC
GCCCCTGCCACAGTGTGCGGCCCCAAAAGAGCACCAACCTGGTGAAGAAC
AAGTGCGTGAACCTTCAACTTCAACGGCCTGAAGGGCACCGGCGTGCTGACC
GAGAGCAACAAGAAGTTCCTGCCATTTTCAGCAATTCGGCAGAGACATCGCC
GACACCACCGACGCGGTACAGAGACCCTCAGACCCTGGAGATCCTGGACATC
30 ACCCCCTGTAGCTTCGGCGGCGTGAGCGTGATACCCCCGGCACTAACACAA
GCAACCAAGTGGCCGTGCTGTACCAAGGCGTGAACCTGCACCGAGGTGCCCG
TGGCCATCCACGCCGATCAGCTGACCCCCACCTGGAGAGTGTATAGCACCGG
CAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGGCGCCGAGTACGT
GAACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATCTGCGCTAGC
35 TATCAGACACAGACCAAGAGCCACCGGAGAGCTAGAAGCGTGGCTAGCCAA
AGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTACA
GCAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACCACCGA
GATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCATGTACATC
TGCGGCGACAGCACCGAGTGCAGCAACCTGCTCCTGCAGTACGGCAGCTTC
40 TGCACACAGCTGAAGAGAGCCCTGACCGGCATCGCCGTGGAGCAAGACAAG
AACACCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTACAAGACCCCCCCC

ATCAAGTACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGACCCTAGCA
AGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCT
GGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCA
GCTAGAGACCTGATCTGCGCTCAGAAGTTCAAGGGACTCACCGTGCTCCCCC
5 CCCTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGG
AACCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTGCAAATCCCC
TTCGCCATGCAGATGGCCTACAGATTCAACGGCATCGGCGTGACACAGAACG
TGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAGCGCCATCGG
CAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCAAGCTGCA
10 AGACGTGGTGAACCACAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCT
GAGCAGCAAGTTCGGCGCCATCAGCAGCGTGCTGAATGACATCTTCTCTAGA
CTGGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGA
CTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAG
ATCAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAGTGCGTGCTG
15 GGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCACCTGATGAGC
TTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGACCTACGTGC
CCGCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCACGACGGCA
AGGCCCACTTCCCTAGAGAGGGCGTGTTTCGTGAGCAACGGCACCCACTGGT
TCGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACCGACAACAC
20 CTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTGAACAACACCGTA
TACGACCCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAG
TACTTCAAAAACCATAGCCCCGACGTGGACCTGGGAGACATCTCTGGCA
TCAACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCGGCTGAATGAGG
TGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCTGGGCAAGT
25 ACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGG
CCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTT
GCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACG
AGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGAT
GA

30 SEQ ID NO: 15

>poly A

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAA
AAA
AAAAAAAAAAAAA

35 SEQ ID NO: 16

>T7 promoter

TAATACGACTCACTATAAGG

SEQ ID NO: 17

>I'-Delta24-B

40 GAGTTAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGAT
AGCACTATAATATGTTTCGTGTTCTGGTGCTGCTGCCCCTGGTGAGCTCTCAG

TCGGTGAACCTGAGGACAAGAACACAGCTGCCCCCGCCTACACCAACAGC
 TTCACAAGAGGCGTGTACTACCCCGACAAGGTGTTTCAGAAGCAGCGTGCTG
 CACTCCACCCAAGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCC
 ACGCCATCCACGTCAGCGGCACCAACGGCACCAAGAGATTCGACAACCCCG
 5 TGCTGCCCTTCAACGACGGCGTGTACTTCGCTAGCACCGAAAAGAGCAACAT
 CATCAGAGGCTGGATCTTCGGCACCAACCTAGACTCCAAGACACAAAGCCTG
 CTGATCGTTAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGT
 TCTGCAACGACCCCTTCCTGGACGTGTACTACCACAAGAACAACAAGAGCT
 GGATGGAGAGCGGGGTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACG
 10 TGTCTCAGCCCTTCCTGATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAA
 CCTAAGAGAGTTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGC
 AAGCACACCCCCATCAACCTGGTGAGAGACCTGCCCCAAGGCTTCAGCGCC
 CTGGAGCCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAGATTTCAGA
 CCCTGCTGGCCCTGCACCGTAGCTACCTAACCCCCGGCGATAGTTCCTCTGGC
 15 TGGACGGCCGGCGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCT
 TCCTGCTGAAGTACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGC
 CCTGGACCCCTGAGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGA
 GAAGGGCATCTATCAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCAT
 CGTGAGATTCCCCAACATCACCAACCTGTGCCCTTCGGCGAGGTGTTCAAC
 20 GCCACAAGATTCGCTAGCGTGTACGCGTGGAACAGAAAGAGAATCAGCAAC
 TCGGTGGCCGACTACAGCGTGCTGTACAACAGCGCTAGCTTCAGCACCTTCA
 AGTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGT
 GTACGCCGACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCC
 CGGGCAAACCGGCAAGATCGCCGACTACAACCTACAAGCTGCCCCGACGACTT
 25 CACCGGCTGCGTGATCGCGTGGAACAGCAATAACCTGGATTCCAAGGTGGGC
 GGCAACTACAACCTACCGGTACAGACTGTTTCAGAAAGAGCAACCTGAAGCCC
 TTCGAGAGAGACATCAGCACCGAGATCTACCAAGCCGGCAGCAAGCCCTGC
 AACGGCGTGAGGGCTTCAACTGCTACTTCCCCCTGCAGAGCTACGGCTTTC
 AGCCCACCAACGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTT
 30 CGAGCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCCAAAAAGAGCACCAA
 CCTGGTGAAGAACAAGTGCCTGAACCTTCAACTTTAATGGCCTTACCGGCACG
 GGCGTGCTGACCGAGAGCAACAAGAAGTTCCTGCCATTTTCAGCAATTCGGC
 AGAGACATCGCCGACACCACCGACGCGGTCAGAGACCCTCAGACCCTGGAG
 ATCCTGGACATCACCCCTGTAGCTTCGGCGGGCGTGAGCGTGATCACCCCG
 35 GCACTAACACAAGCAACCAAGTGGCCGTGCTGTACCAAGGCGTGAACCTGCA
 CCGAGGTGCCCGTGGCCATCCACGCCGATCAGCTGACCCCCACCTGGAGAGT
 ATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGG
 CGCCGAGCACGTGAACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGG
 CATCTGCGCTAGCTATCAGACACAGACCAACAGCCGCCGGAGAGCTAGAAG
 40 CGTGGCTTCTCAGAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAAC
 AGCGTGGCCTACAGCAACAACAGCATCGCCATCCCCACCAACTTCACCATCA

GCGTGACCACCGAGATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACT
GCACCATGTACATCTGCGGCGACAGCACCGAGTGCAGCAACCTGCTGCTGCA
GTACGGCAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGGCATCGCCGT
GGAGCAAGACAAGAACACCCAAGAGGTGTTGCCCCAAGTGAAGCAGATCTA
5 CAAGACCCCCCCCCATCAAGGACTTCGGGCGGCTTCAACTTCAGCCAAATCCTG
CCCGACCCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCA
ACAAGGTGACCCTGGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCT
AGGCGACATCGCAGCTAGAGACCTGATCTGCGCTCAGAAGTTCAATGGACTC
ACCGTGCTCCCCCCCCCTGCTGACCGACGAGATGATCGCTCAGTACACCTCTG
10 CCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCG
CACTGCAAATCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCATCGG
CGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTC
AACAGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCC
CTGGGCAAGCTGCAAAACGTGGTGAATCAGAACGCCCAAGCCCTGAACACC
15 CTGGTGAAGCAGCTGAGCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATG
ACATCCTGTCTAGACTGGACCCCCCGAGGCCGAGGTGCAGATCGACAGACT
GATCACCGGCAGACTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTGAT
CAGAGCCGCCGAGATCAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAG
CGAGTGCGTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTA
20 CCACCTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCAC
GTGACCTACGTGCCCCGCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCT
GCCACGACGGCAAGGCCCACTTCCCTAGAGAGGGCGTGTTCTGTGAGCAACG
GCACCCACTGGTTCGTGACACAGAGAACTTCTACGAGCCTCAGATCATCAC
CACCGACAACACCTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTG
25 AACAACACCGTATACGACCCCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAG
GAGCTGGACAAGTACTTCAAAAACCATACTAGCCCCGACGTGGACCTGGGA
GACATCTCTGGCATCAACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACC
GGCTGAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAG
AGCTGGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGG
30 CTTTCATCGCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCA
TGACAAGCTGTTGCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTG
CAAGTTCGACGAGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGC
ACTACACCTGATGACTCGAGGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAG
GTTCCCTTTGTTCCCTAAGTCCAATACTAACTGGGGGATATTATGAAGGGCC
35 TTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAAAAAAA
AA
AA
AAAAAAA

SEQ ID NO: 18

40 >I'-OC24-B
GAGTTAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGAT

AGCACTATAATATGTTTCGTGTTCTGTTGGTGCTGCTGCCCCCTGGTGAGCTCTCAG
 TGCGTGAACCTGACCACAAGAACACAGCTGCCCCCGCCTACACCAACAGC
 TTCACAAGAGGCGTGTACTACCCCGACAAGGTGTTTCAGAAGCAGCGTGCTG
 CACTCCACCCAAGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCC
 5 ACGTGATCAGCGGCACCAACGGCACCAAGAGATTCGACAACCCCGTGCTGC
 CCTTCAACGACGGCGTGTACTTCGCTAGCATCGAAAAGAGCAACATCATCAG
 AGGCTGGATCTTCGGCACCAACCTAGACTCCAAGACACAAAGCCTGCTGATC
 GTTAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCA
 ACGACCCCTTCCTGGACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGT
 10 TCAGAGTGTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCC
 CTTCTGATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGA
 GTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACC
 CCCATCATCGTGAGAGAGCCCGAGGACCTGCCCCAAGGCTTCAGCGCCCTG
 GAGCCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAGATTTAGACCC
 15 TGCTGGCCCTGCACCGTAGCTACCTAACCCCCGGCGATAGTTCCTCTGGCTGG
 ACGGCCGGCGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCC
 TGCTGAAGTACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCT
 GGACCCCTGAGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAA
 GGGCATCTATCAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTG
 20 AGATTCCCCAACATCACCACCTGTGCCCTTCGACGAGGTGTTCAACGCCA
 CCAGATTTGCTAGCGTGTACGCCTGGAACCGGAAGAGAATCAGCAACTGCGT
 GGCCGACTACAGCGTGCTGTACAACCTGGCCCCCTTCTTCACCTTCAAGTGC
 TATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACG
 CCGACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATCGCCCCCGGGC
 25 AGACCGGCAACATCGCCGACTACAACCTACAAGCTGCCCGACGACTTCACCG
 GCTGCGTGATCGCCTGGAACAGCAACAAGCTGGATAGCAAGGTGAGCGGCA
 ACTACAATTACCTGTACAGACTGTTTCAGAAAGAGCAACCTGAAGCCCTTCGA
 GAGAGACATCAGCACCGAGATCTACCAAGCCGGCAACAAGCCCTGCAACGG
 CGTGGCGGGCTTCAACTGCTACTTCCCCCTCAGGAGCTACAGCTTCAGACCC
 30 ACCTACGGCGTGGGCCATCAGCCCTACAGAGTGGTCGTGCTGAGCTTCGAGC
 TGCTGCACGCCCCCTGCCACAGTGTGCGGCCCCAAAAAGAGCACCAACCTGG
 TGAAGAACAAGTGCGTGAACCTTCAACTTCAACGGCCTGAAGGGCACCGGCG
 TGCTGACCGAGAGCAACAAGAAGTTCCTGCCATTTAGCAATTCGGCAGAG
 ACATCGCCGACACCACCGACGCGGTGAGAGACCCTCAGACCCTGGAGATCC
 35 TGGACATCACCCCTGTAGCTTCGGCGGCGTGAGCGTGATCACCCCGGCAC
 TAACACAAGCAACCAAGTGCCGTGCTGTACCAAGGCGTGAACCTGCACCGA
 GGTGCCCCGTGGCCATCCACGCCGATCAGCTGACCCCCACCTGGAGAGTGTAT
 AGCACCGGCAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGGCGCC
 GAGTACGTGAACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATCT
 40 GCGCTAGCTATCAGACACAGACCAAGAGCCACCGGAGAGCTAGAAGCGTGG
 CTAGCCAAAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGT

GGCCTACAGCAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTG
ACCACCGAGATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCA
TGTACATCTGCGGCGACAGCACCGAGTGCAGCAACCTGCTCCTGCAGTACGG
CAGCTTCTGCACACAGCTGAAGAGAGCCCTGACCGGCATCGCCGTGGAGCA
5 AGACAAGAACACCCAAGAGGTGTTGCCCCAAGTGAAGCAGATCTACAAGAC
CCCCCCCATCAAGTACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGAC
CCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAG
GTGACCCTGGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCG
ACATCGCAGCTAGAGACCTGATCTGCGCTCAGAAGTTCAAGGGACTCACCGT
10 GCTCCCCCCCCCTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTG
CTGGCGGGAACCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTG
CAAATCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCATCGGCGTGA
CACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAG
CGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGG
15 CAAGCTGCAAGACGTGGTGAACCACAACGCCCAAGCCCTGAACACCCTGGT
GAAGCAGCTGAGCAGCAAGTTCGGCGCCATCAGCAGCGTGCTGAATGACAT
CTTCTCTAGACTGGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATC
ACCGGCAGACTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGA
GCCGCCGAGATCAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAG
20 TGCGTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCAC
CTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGA
CCTACGTGCCCGCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCA
CGACGGCAAGGCCCACTTCCCTAGAGAGGGCGTGTTTCGTGAGCAACGGCAC
CCACTGGTTTCGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACC
25 GACAACACCTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTGAAC
AACACCGTATACGACCCCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAGGAG
CTGGACAAGTACTTCAAAAACCATACTAGCCCCGACGTGGACCTGGGAGACA
TCTCTGGCATCAACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCGGCT
GAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCT
30 GGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTC
ATCGCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGAC
AAGCTGTTGCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAA
GTTTCGACGAGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTA
CACCTGATGACTCGAGGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTT
35 CTTTGTTCCTAAGTCCAATACTAACTGGGGGATATTATGAAGGGCCTTGA
GCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAA
AAA
AAAA

40

SEQ ID NO: 19

>B.1.351 spike protein CDS

ATGTTTCGTGTTTCCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACTT
 CACGACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAGG
 CGTGTAACACCCCGACAAGGTGTTTCAGAAGCAGCGTGCTGCACTCCACCCAA
 GACCTGTTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCCACGT
 5 GAGCGGCACCAACGGCACCAAGAGATTTCGCCAACCCCGTGCTGCCCTTCAA
 CGACGGCGTGTAACCTTCGCTAGCACCGAAAAGAGCAACATCATCAGAGGCTG
 GATCTTCGGCACCACTTAGACTCCAAGACACAAAGCCTGCTGATCGTTAAC
 AACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGAC
 CCCTTCCTGGGCGTGTAACCTACCAAGAACAACAAGAGCTGGATGGAGAGC
 10 GAGTTCCGTGTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTC
 AGCCCTTCCTGATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAA
 GAGAGTTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCA
 CACCCCATCAACCTGGTGAGAGGCCTGCCCCAAGGCTTCAGCGCCCTGGA
 GCCCCTGGTGACCTGCCCATCGGCATCAACATCACAAGATTTTCAGACCCTG
 15 CACCGTAGCTACCTAACCCCGGCGATAGTTCCCTCTGGCTGGACGGCCGGCG
 CTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAGTA
 CAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTG
 AGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTAT
 CAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTGAGATTCCCC
 20 AACATCACCAACCTGTGCCCTTCGGCGAGGTGTTCAACGCCACAAGATTTCG
 CTAGCGTGTAACGCGTGGAACAGAAAGAGAATCAGCAACTGCGTGCCGACT
 ACAGCGTGCTGTACAACAGCGCTAGCTTCAGCACCTTCAAGTGCTATGGCGT
 CAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCCGACAG
 CTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCCGGGCAAACCGG
 25 CAACATCGCCGACTACAACCTACAAGCTGCCCCGACGACTTCACCGGCTGCGTG
 ATCGCGTGGAACAGCAATAACCTGGATTCCAAGGTGGGCGGCAACTACAAC
 ACCTGTACAGACTGTTTCAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACA
 TCAGCACCGAGATCTACCAAGCCGGCAGCACCCCTGCAACGGCGTGAAAG
 GCTTCAACTGCTACTTCCCCCTGCAGAGCTACGGCTTTCAGCCCACCTACGG
 30 CGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGAGCTGCTGCAC
 GCCCCCGCAACCGTGTCGGGCCCCAAAAAGAGCACCAACCTGGTGAAGAAC
 AAGTGCGTGAACTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTGACCG
 AGAGCAACAAGAAGTTCCTGCCATTTTCAGCAATTCGGCAGAGACATCGCCG
 ACACCACCGACGCGGTCAGAGACCCTCAGACCCTGGAGATCCTGGACATCA
 35 CCCCCTGTAGCTTCGGCGGCGTGAGCGTGATCACCCCGGCACTAACACAAG
 CAACCAAGTGGCCGTGCTGTACCAAGGCGTGAACCTGCACCGAGGTGCCCGT
 GGCCATCCACGCCGATCAGCTGACCCCCACCTGGAGAGTATACAGCACCGGA
 AGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGGCGCCGAGCACGTG
 AACAAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATCTGCGCTAGCT
 40 ATCAGACACAGACCAACAGCCCCCGGAGAGCTAGAAGCGTGGCTTCTCAGA
 GCATCATCGCCTACACCATGAGCCTGGGCGTCGAGAACAGCGTGGCCTACAG

CAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACCACCGAG
 ATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCATGTACATCT
 GCGGCGACAGCACCGAGTGCAGCAACCTGCTGCTGCAGTACGGCAGCTTCT
 GCACACAGCTGAACAGAGCCCTGACCGGCATCGCCGTGGAGCAAGACAAGA
 5 ACACCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTACAAGACCCCCCCCAT
 CAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGACCCTAGCAAG
 CCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCTGG
 CCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGC
 TAGAGACCTGATCTGCGCTCAGAAGTTCAATGGACTCACCGTGCTCCCCCCC
 10 CTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAA
 CCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTGCAAATCCCCTT
 CGCCATGCAGATGGCCTACAGATTCAACGGCATCGGCGTGACACAGAACGTG
 CTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAGCGCCATCGGCA
 AGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCAAGCTGCAAG
 15 ACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCTGA
 GCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGTCTAGACT
 GGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGACT
 GCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAGAT
 CAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAGTGCCTGCTGGG
 20 GCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCACCTGATGAGCTT
 CCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGACCTACGTGCCC
 GCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCACGACGGCAAG
 GCCCCACTTCCCTAGAGAGGGCGTGTTTCGTGAGCAACGGCACCCACTGGTTCG
 TGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACCGACAACACCTT
 25 CGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTGAACAACACCGTATAC
 GACCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTAC
 TTCAAAAACCATACTAGCCCCGACGTGGACCTGGGAGACATCTCTGGCATCA
 ACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCGGCTGAATGAGGTGG
 CCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCTGGGCAAGTACG
 30 AGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCT
 GATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGCA
 GCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGG
 ACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 20

35 >B.1.351 spike protein amino acid sequence
 MFVFLVLLPLVSSQCVNFTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLF
 LPFFSNVTWFHAIHVSGTNGTKRFANPVLPFNDGVYFASTEKSNIIRGWIFGTTL
 DSKTQSLIVN NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSA
 NNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRGLP
 40 QGFSALEPLVDLPIGINITRFQTLHRSYLT PGDSSSGWTAGAAAYYVGYLQPRTF
 LLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPN

ITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASFSTFKCYGVSP
 LNDLCFTNVYADSFVIRGDEVQRQIAPGQTGNIADYNYKLPDDFTGCVIAWNSNN
 LDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVKGFNCYFPLQS
 YGFQPTYGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLT
 5 GTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN
 TSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVN
 NSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGVENSVAYSNNSIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSSTECNLLLQYGSFCTQLNRALTGI
 AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKV
 10 TLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLTDEMIAQYTSALLAG
 TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKI
 QDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDLPPEA
 EVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFC
 GKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGDKAHFPREGVVFV
 15 SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEE
 LDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKY
 EQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCCLKGCCSCGSCCKFDEDD
 SEPVLKGVKLHYT

SEQ ID NO: 21

20 >B.1.1.7 spike protein CDS
 ATGTTTCGTGTTCTGCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACC
 TGACCACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAG
 GCGTGTACTACCCCGACAAGGTGTTTCAAGAGCAGCGTGCTGCACTCCACCCA
 AGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCAGC
 25 GGCACCAACGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTCAACGAC
 GCGTGTACTTCGCTAGCACCGAAAAGAGCAACATCATCAGAGGCTGGATCT
 TCGGCACCACCCTAGACTCCAAGACACAAAGCCTGCTGATCGTTAACAACGC
 GACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACCCCTTC
 CTGGGCGTGTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCCGT
 30 GTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCC
 TGATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTTCG
 TGTTCAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCCAT
 CAACCTGGTGAGAGACCTGCCCCAAGGCTTCAGCGCCCTGGAGCCCCTGGT
 GGACCTGCCCATCGGCATCAACATCACAAGATTTAGACCCCTGCTGGCCCTG
 35 CACCGTAGCTACCTAACCCCGGCGATAGTTCCCTCTGGCTGGACGGCCGGCG
 CTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAGTA
 CAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTG
 AGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTAT
 CAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTGAGATTCCCC
 40 AACATCACCAACCTGTGCCCTTCGGCGAGGTGTTCAACGCCACAAGATTTCG
 CTAGCGTGTACGCGTGGAACAGAAAGAGAATCAGCAACTGCGTGGCCGACT

ACAGCGTGCTGTACAACAGCGCTAGCTTCAGCACCTTCAAGTGCTATGGCGT
 CAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCCGACAG
 CTTCTGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCCGGGCAAACCGG
 CAAGATCGCCGACTACAACCTACAAGCTGCCCCGACGACTTCACCGGCTGCGTG
 5 ATCGCGTGGAACAGCAATAACCTGGATTCCAAGGTGGGCGGCAACTACAAC
 ACCTGTACAGACTGTTTCAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACA
 TCAGCACCGAGATCTACCAAGCCGGCAGCACCCCCTGCAACGGCGTGGAGG
 GCTTCAACTGCTACTTCCCCCTGCAGAGCTACGGCTTTCAGCCCACCTACGG
 CGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGAGCTGCTGCAC
 10 GCCCCCGCAACCGTGTGCGGCCCCCAAAAAGAGCACCAACCTGGTGAAGAAC
 AAGTGCGTGAACTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTGACCG
 AGAGCAACAAGAAGTTCCTGCCATTTTCAGCAATTCGGCAGAGACATCGACG
 ACACCACCGACGCGGTCAGAGACCCTCAGACCCTGGAGATCCTGGACATCA
 CCCCCTGTAGCTTCGGCGGCGTGAGCGTGATACCCCCGGCACTAACACAAG
 15 CAACCAAGTGGCCGTGCTGTACCAAGGCGTGAAGTGCACCGAGGTGCCCCGT
 GGCCATCCACGCCGATCAGCTGACCCCCACCTGGAGAGTATACAGCACCGGA
 AGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGGCGCCGAGCACGTG
 AACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATCTGCGCTAGCT
 ATCAGACACAGACCAACAGCCACCGGAGAGCTAGAAGCGTGGCTTCTCAGA
 20 GCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTACAG
 CAACAACAGCATCGCCATCCCCATCAACTTCACCATCAGCGTGACCACCGAG
 ATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCATGTACATCT
 GCGGCGACAGCACCGAGTGACAGCAACCTGCTGCTGCAGTACGGCAGCTTCT
 GCACACAGCTGAACAGAGCCCTGACCGGCATCGCCGTGGAGCAAGACAAGA
 25 ACACCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTACAAGACCCCCCCCCAT
 CAAGGACTTCGGCGGGCTTCAACTTCAGCCAAATCCTGCCCGACCCTAGCAAG
 CCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCTGG
 CCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGC
 TAGAGACCTGATCTGCGCTCAGAAGTTCAATGGACTCACCGTGCTCCCCCCC
 30 CTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAA
 CCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTGCAAATCCCCTT
 CGCCATGCAGATGGCCTACAGATTCAACGGCATCGGCGTGACACAGAACGTG
 CTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAGCGCCATCGGCA
 AGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCAAGCTGCAAG
 35 ACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCTGA
 GCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGGCTAGACT
 GGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGACT
 GCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAGAT
 CAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAGTGCGTGCTGGG
 40 GCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCACCTGATGAGCTT
 CCCTCAGAGCGCCCCCCCACGGCGTGGTGTTCCTGCACGTGACCTACGTGCCC

GCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCACGACGGCAAG
 GCCCCTTCCCTAGAGAGGGCGTGTTCTGTGAGCAACGGCACCCACTGGTTCG
 TGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACCCACAACACCTT
 CGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTGAACAACACCGTATAC
 5 GACCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTAC
 TTCAAAAACCATAGCCCCGACGTGGACCTGGGAGACATCTCTGGCATCA
 ACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCGGCTGAATGAGGTGG
 CCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCTGGGCAAGTACG
 AGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCT
 10 GATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGCA
 GCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGG
 ACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 22

>B.1.1.7 spike protein amino acid sequence

15 MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLF
 LPFFSNVTWFHAISGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDS
 KTQSLNIVNNATNVVIKVFCEFCNDPFLGVYHKNNKSWMESEFRVYSSANNC
 TFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFS
 ALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYYVGYLQPRFTL
 20 LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNI
 TNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL
 NDLCFTNVYADSFVIRGDEVIRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNL
 DSKVGGNYNYLYRLFRKSNLKPFERDISTEYQAGSTPCNGVEGFNCYFPLQSY
 GFQPTYGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG
 25 TGVLTESNKKFLPFQFGRDIDDTDAVRDPQTLEILDITPCSFGGVSVITPGTNT
 SNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVN
 NSYECDIPIGAGICASYQTQTNSHRRARSVASQSIIAYTMSLGAENSVAYSNNNSIAI
 PINFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGI
 AVEQDKNTQEVEFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKV
 30 TLADAGFIKQYGDCLGDIAARDLCAQKFNGLTVLPLLTDEMIAQYTSALLAG
 TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKI
 QDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILARLDPPE
 AEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDF
 CGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGDKAHFPREGVF
 35 VSNQTHWFVTQRNFYEPQIITHTNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKE
 ELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK
 YEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCLKGCCSCGSCCKFDED
 DSEPVLKGVKLHYT

SEQ ID NO: 23

40 >SARS-CoV-2 spike protein CDS

ATGTTCTGTGTTCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACC

TGACCACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAG
 GCGTGTACTACCCCGACAAGGTGTTTCAAGAAGCAGCGTGCTGCACTCCACCCA
 AGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCCAC
 GTGAGCGGCACCAACGGCACCAAGAGATTTCGACAACCCCGTGCTGCCCTTC
 5 AACGACGGCGTGTACTTCGCTAGCACCGAAAAGAGCAACATCATCAGAGGC
 TGGATCTTCGGCACCAACCCTAGACTCCAAGACACAAAGCCTGCTGATCGTTA
 ACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACG
 ACCCCTTCCTGGGCGTGTAATAACAAGAACAACAAGAGCTGGATGGAGA
 GCGAGTTCGCTGTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTC
 10 TCAGCCCTTCCTGATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTA
 AGAGAGTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGC
 ACACCCCATCAACCTGGTGAGAGACCTGCCCCAAGGCTTCAGCGCCCTGG
 AGCCCTGGTGACCTGCCCATCGGCATCAACATCACAAGATTTTCAGACCCT
 GCTGGCCCTGCACCGTAGCTACCTAACCCCGGCGATAGTTCTCTGGCTGG
 15 ACGGCCGGCGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCC
 TGCTGAAGTACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCT
 GGACCCCTGAGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAA
 GGGCATCTATCAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTG
 AGATTCCCAACATCACCAACCTGTGCCCTTCGGCGAGGTGTTCAACGCCA
 20 CAAGATTCGCTAGCGTGACGCGTGGAACAGAAAGAGAATCAGCAACTGCG
 TGGCCGACTACAGCGTGCTGTACAACAGCGCTAGCTTCAGCACCTTCAAGTG
 CTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTAC
 GCCGACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCGGG
 CAAACCGGCAAGATCGCCGACTACAACCTACAAGCTGCCCCGACGACTTCACC
 25 GGCTGCGTGATCGCGTGGAACAGCAATAACCTGGATTCCAAGGTGGGCGGC
 AACTACAACCTGTACAGACTGTTTCAAGAAAGAGCAACCTGAAGCCCTTCG
 AGAGAGACATCAGCACCGAGATCTACCAAGCCGGCAGCACCCCTGCAACG
 GCGTGGAGGGCTTCAACTGCTACTTCCCCCTGCAGAGCTACGGCTTTCAGCC
 CACCAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGAG
 30 CTGCTGCACGCCCCCGCAACCGTGTGCGGCCCCAAAAAGAGCACCAACCTG
 GTGAAGAACAAGTGCGTGAACCTTCAACTTTAATGGCCTTACCGGCACGGGCG
 TGCTGACCGAGAGCAACAAGAAGTTCTGCCATTTTCAGCAATTCGGCAGAG
 ACATCGCCGACACCACCGACGCGGTGAGAGACCCTCAGACCCTGGAGATCC
 TGGACATCACCCCTGTAGCTTCGGCGGCGTGAGCGTGATCACCCCGGCAC
 35 TAACACAAGCAACCAAGTGGCCGTGCTGTACCAAGACGTGAACTGCACCGA
 GGTGCCCCGTGGCCATCCACGCCGATCAGCTGACCCCACTGGAGAGTATAC
 AGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGGCGCC
 GAGCACGTGAACAACAGCTACGAGTGCAGCATCCCATCGGCGCCGGCATCT
 GCGCTAGCTATCAGACACAGACCAACAGCCCTCGGAGAGCTAGAAGCGTGG
 40 CTTCTCAGAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGT
 GGCCTACAGCAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTG

ACCACCGAGATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCA
 TGTACATCTGCGGCGACAGCACCGAGTGCAGCAACCTGCTGCTGCAGTACGG
 CAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGGCATCGCCGTGGAGCA
 AGACAAGAACACCCAAGAGGTGTTGCCCCAAGTGAAGCAGATCTACAAGAC
 5 CCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGAC
 CCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAG
 GTGACCCTGGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCG
 ACATCGCAGCTAGAGACCTGATCTGCGCTCAGAAGTTCAATGGACTCACCGT
 GCTCCCCCCCCCTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTG
 10 CTGGCGGGAACCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTG
 CAAATCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCATCGGCGTGA
 CACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAG
 CGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGG
 CAAGCTGCAAGACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGT
 15 GAAGCAGCTGAGCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATC
 CTGAGCAGACTGGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATC
 ACCGGCAGACTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGA
 GCCGCCGAGATCAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAG
 TGCGTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCAC
 20 CTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGA
 CCTACGTGCCCCGCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCA
 CGACGGCAAGGGCCCACTTCCCTAGAGAGGGCGTGTTTCGTGAGCAACGGCAC
 CCACTGGTTCGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACC
 GACAACACCTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGTGAAC
 25 AACACCGTATACGACCCCCCTGCAGCCCGAGCTGGACAGCTTCAAGGAGGAG
 CTGGACAAGTACTTCAAAAACCATACTAGCCCCGACGTGGACCTGGGAGACA
 TCTCTGGCATCAACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCGGCT
 GAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCT
 GGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTC
 30 ATCGCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGAC
 AAGCTGTTGCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAA
 GTTCGACGAGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTA
 CACCTGATGA

SEQ ID NO: 24

35 >SARS-CoV-2 spike protein amino acid sequence
 MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLF
 LPFFSNVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTL
 DSKTQSL LIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSA
 NNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLP
 40 QGFSALEPLVDLPIGINITRFQTLALHRSYLT PGDSSSGWTAGAAAYYVGYLQP
 RTFLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVR

FPNITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASFSTFKCYGVS
 PTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWN
 SNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFP
 LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFN
 5 GLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
 GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAE
 HVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSN
 NSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNR
 ALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLL
 10 FNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQY TSA
 LLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SA
 IGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSR LD
 PPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRV
 DFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPREG
 15 VFVSNGTHWVFTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSF
 KEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQEL
 GKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCCLKGCCSCGSCCKFD
 EDDSEPVLKGVKLHYT

SEQ ID NO: 25

20 >Luciferase CDS
 ATGGAGGACGCCAAGAACATCAAGAAAGGCCCCGCCCCCTTCTACCCCCTG
 GAGGACGGCACCGCCGGCGAGCAGCTGCACAAAGCCATGAAAAGATACGC
 TCTGGTGCCCGGCACCATCGCCTTACCGACGCCCACATCGAGGTGAACAT
 CACCTACGCCGAGTACTTCGAGATGAGCGTGAGACTGGCCGAGGCTATGAA
 25 GAGATACGGCTTGAACACCAACCACAGAATCGTGGTGTGCAGCGAGAACA
 GCCTGCAGTTCTTCATGCCCCGTGCTGGGCGCCCTGTTTCATCGGCGTGGCCGT
 TGCCCCCGCCAACGACATCTACAACGAGAGAGAGCTGCTGAACAGCATGAA
 CATCTCTCAGCCCACCGTGGTTTTCTGTGTCCAAAAGGGCCTGCAGAAGAT
 CCTGAACGTGCAGAAGAAGCTGCCCATCATTGAGAAGATCATCATCATGGAC
 30 AGCAAGACCGACTACCAAGGCTTTCAGAGCATGTACACCTTCGTGACAAGC
 CACCTGCCCCCGGCTTCAACGAGTACGACTTCGTGCCCCGAGAGCTTCGAC
 AGAGACAAGACCATCGCCCTGATCATGAACAGCAGCGGCAGCACCGGCCTG
 CCAAGGGCGTGGCCCTGCCCCACAGAACCGCCTGCGTGAGATTCAGCCAC
 GCTAGAGACCCCATCTTCGGCAATCAGATCATCCCCGACACCGCCATCCTGA
 35 GCGTAGTGCCTTTCCATCATGGCTTCGGCATGTTACCACCCTGGGCTACCT
 GATCTGCGGCTTCAGAGTGGTGCTGATGTACAGATTCGAGGAGGAGCTGTT
 CCTGAGAAGCCTGCAAGACTACAAGATTCAGAGCGCCCTGCTGGTGCCAC
 CCTGTTTCAGCTTCTTCGCCAAGAGCACCTGATCGACAAGTATGACCTGAGC

AACCTGCACGAAATTGCGAGCGGCGGAGCGCCGCTCAGCAAAGAGGTGGG
CGAGGCCGTGGCCAAGAGATTCCACCTGCCCGGCATCAGACAAGGCTACGG
CCTGACCGAGACCACAAGCGCCATCCTGATCACCCCGAGGGCGACGATAA
GCCCCGGCGCCGTGGGCAAGGTGGTACCATTCTTCGAGGCCAAGGTGGTGA
5 CCTGGACACCGGCAAGACCCTGGGCGTGAATCAGAGAGGCGAGCTGTGCG
TGAGAGGCCCCATGATCATGAGCGGCTACGTGAACAACCCCGAGGCCACCA
ACGCCCTCATCGATAAGGATGGGTGGCTGCACAGCGGCGACATCGCCTACT
GGGACGAGGACGAGCACTTCTTCATCGTGGACAGACTCAAGAGCCTGATCA
AGTACAAGGGCTACCAAGTGGCCCCCGCCGAGTTGGAGAGCATCCTGCTGC
10 AGCACCCCAATATCTTCGACGCTGGCGTGGCCGGACTACCCGACGACGACG
CCGGCGAGTTACCCGCCGCCGTGGTGGTGGTGGAGCACGGCAAGACCATGA
CCGAGAAGGAGATCGTGGACTACGTGGCTAGCCAAGTGACCACCGCCAAG
AAGCTGAGAGGCGGCGTGGTGTTCGTTGACGAAGTGCCCAAGGGCCTGAC
CGGCAAGCTGGACGCTAGAAAGATCAGAGAGATCCTGATCAAGGCCAAGA
15 AGGGCGGCAAGAGCAAGCTGTGA

SEQ ID NO: 26

>Luciferase amino acid sequence
MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYA
EYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAVAPANDI
20 YNERELLNSMNISQPTVVVFSKKGLQKILNVQKKLPPIIQKIIIMDSKTDYQGFQS
MYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
FSHARDPIFGNQIIPDTAILSVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFL
RSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKR
FHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVN
25 QRGELCVRGPMIMSGYVNNPEATNALIDKGWLHSGDIAYWDEDEHFFIVDR
KSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKT
MTEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGG
KSKL

SEQ ID NO: 27

30 >LUC1- I' truncation
TAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCA
CTATAAT

SEQ ID NO: 28

>LUC2- I' truncation
35 AGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT
AAT

SEQ ID NO: 29

>LUC3- I' truncation

	GGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT	SEQ ID NO: 30
	>LUC4- I' truncation	
	GCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT	
5		SEQ ID NO: 31
	>LUC5- I' truncation	
	TTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT	
		SEQ ID NO: 32
	>LUC6- I' truncation	
10	AATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT	
		SEQ ID NO: 33
	>LUC7- I' truncation	
	ACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT	
		SEQ ID NO: 34
15	>LUC8- I' truncation	
	GTCATTGCCGAGAAAAGGATAGCACTATAAT	
		SEQ ID NO: 35
	>LUC9- I' truncation	
	TTGCCGAGAAAAGGATAGCACTATAAT	
20		SEQ ID NO: 36
	>LUC10- I' truncation	
	CGAGAAAAGGATAGCACTATAAT	
		SEQ ID NO: 37
	>LUC11- I' truncation	
25	AAAAGGATAGCACTATAAT	
		SEQ ID NO: 38
	>LUC12- I' truncation	
	GGATAGCACTATAAT	
		SEQ ID NO: 39
30	>LUC13- I' truncation	
	AGCACTATAAT	
		SEQ ID NO: 40
	>LUC14- I' truncation	
	CTATAAT	
35		
	>LUC15- I' mutant	
	GAGTTAGGAGTTGCTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGA	
	TAGCACTATAAT	
		SEQ ID NO: 42

>LUC16- I' mutant
GAGTTAGGAGTTGTTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGA
TAGCACTATAAT
SEQ ID NO: 43

5 >5'UTR I general formula
GAGTTAGGAGTTGN₁TGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGA
TAGCACTATAAT
SEQ ID NO: 44

N₁ is selected from the group consisting of A, G, C, and T

10 The following SEQ ID NOs: 45-79 are RNA sequences.
>5'UTR I (63bp)
GAGUUAGGAGUUGAUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAG
GAUAGCACUAUAAU
SEQ ID NO: 45

15 >5'UTR I' (63bp)
GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAG
GAUAGCACUAUAAU
SEQ ID NO: 46

>5'UTR G
20 GCACUUUUCUGAGCAGACGUCCAGAGCAGAGUCAGCCGCCACC
SEQ ID NO: 47

>5'UTR A
ACAUUUGCUUCUGACACAACUGUGUUCACUAGCAACCUCAAACGCCGCCA
CC
25 SEQ ID NO: 48

>5'UTR J
AGAGAGGCAAAGAAACAUUGUGAAAUCUCCAACUCUUAACCUUCAAGCC
GCCACC
SEQ ID NO: 49

30 >3'UTR B
GCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAAGGUUCCUUUGUUCCCUAA
GUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGA
UUCUGCCUAAUAAAAAACAUUUAUUUUCAUUGCAA
SEQ ID NO: 50

35 >3'UTR F
CAGACAAGACAAGCUACUGAAUACGUUCACAUCUGUCUUGAUGCCUAAU
AUUUUUACAUUUCUGUAAACAUAUUUCUGAAAUAUUUUUUGCCUUUCAA
GCGACAGAUGCCUCAUUUUGUGAAAACUUAUGAUGAUUUUGUGUUUAA
GUUCCAAACAUUUGAAUAAAAUAAUUGACAAUA
40 SEQ ID NO: 51

>3'UTR E

CCUGAGGCCUCAGAACACAGGUUUCAGAUUGAUAGGCCUGCAGGUCUCCA
GGCAGCAACCAGCUGAGCGACUAAAGGGCCCAAGGCCAGGGCUCUAGGGA
UGGGGCUCAGCAGAGGCUGGGGUAAGGGGAGCCAGGGAGGAGCUGGGCC
5 UAAUGCAGCACCGGGUCCCCAGGAUGCCGUGUCUCGGCUGGAGGAGGAGA
UGCGGAAGCUCCAGGCCACGGUGCAGGAGCUCCAGAAGCGCUUGGACAGG
CUGGAGGAGACAGUCCAGGCCAAGUAGAGCCCCGCAGGGCCUCCAGCAGG
GUCAGCCAUUCACACCCAUCCACUCACCUCCTCAUUCCCAGCCACAUGGCA
GAGAAAAAAAUCAUAAUAAAAUGGCUUUUUAUUUCUGGUA

10

SEQ ID NO: 52

>3'UTR D

GGGGCCUUCUGACAUGAGUCUGGCCUGGCCCCACCUCCUAGUUCCUCAUA
AUAAAGACAGAUUGCUUCUUCGCUUCUCACUGAGGGGCCUUCUGACAUG
AGUCUGGCCUGGCCCCACCUCCCCAGUUUCUCAUAAUAAAGACAGAUUGC
15 UUCUUCACUUGAAGCUCGCUUUCUUGCUGUCCAAUUCUUAUUAAGGUU
CCUUGUUCUCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGC
CUUGAGCAUCUGGAUUCUGCCUAAUAAAAAACAUUUAUUUUAUUGCAA

SEQ ID NO: 53

20 >B.1.617.2 (Delta) spike protein mRNA

AUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA
CCUGAGGACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAA
GAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCC
ACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGC
25 CAUCCACGUCAGCGGCACCAACGGCACCAAGAGAUUCGACAACCCCGUGC
UGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCGAAAAGAGCAACAUC
AUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACAAAGCCU
GCUGAUCGUUAACAACGCGACCAACGUGGUGAUCAAGGUGUGCGAGUUU
CAGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAA
30 GAGCUGGAUGGAGAGCGGGGUCUACAGCAGCGCCAACAACUGCACCUUCG
AGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAU
UUCAAGAACCUAAGAGAGUUCGUGUUAAGAACAUCGACGGCUACUUCA
AGAUCUACAGCAAGCACACCCCAUCAACCUGGUGAGAGACCUGCCCCAA
GGCUUCAGCGCCCUUGGAGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAU
35 CACAAGAUUUCAGACCCUGCUGGCCCUAGCACCGUAGCUACCUAACCCCGG
GCGAUAGUUCUUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACGUGGGC
UACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAU
CACUGAUGCCGUGGACUGUGCCCUUGGACCCCUAGAGCGAGACCAAGUGCA
CCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUC
40 AGAGUGCAGCCACCGAGAGCAUCGUGAGAUUCCCCAACAUCACCAACCU
GUGCCCCUUCGGCGAGGUGUUAACGCCACAAGAUUCGCUAGCGUGUACG

CGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGCCGACUACAGCGUGCU
 GUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGCGUCAGCCCCA
 CCAAGCUGAACGACCUGUGCUUACCAACGUGUACGCCGACAGCUUCGUG
 AUCAGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCAAACCGGCAAGAU
 5 CGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCG
 CGUGGAACAGCAAUAACCUGGAUUCCAAGGUGGGCGGCAACUACAACUAC
 CGGUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACA
 UCAGCACCGAGAUCUACCAAGCCGGCAGCAAGCCCUGCAACGGCGUGGAG
 GGCUUAACUGCUACUUCCCCCUGCAGAGCUACGGCUUUCAGCCCACCAA
 10 CGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUGAGCUUCGAGCUGC
 UGCACGCCCCCGCAACCGUGUGCGGGCCCCAAAAAGAGCACCAACCUGGUG
 AAGAACAAGUGCGUGAACUUCAACUUAUUGGCCUUAACCGGCACGGGCG
 UGCUGACCGAGAGCAACAAGAAGUCCUGCCAUUUCAGCAAUUCGGCAGA
 GACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAU
 15 CCUGGACAUCACCCCCUGUAGCUUCGGCGGCGUGAGCGUGAUCACCCCCG
 GCACUAACACAAGCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGC
 ACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAG
 AGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCUGCCUGA
 UCGGCGCCGAGCACGUGAACAACAGCUACGAGUGCGACAUCCCAUCGGC
 20 GCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCGCCGGAGAGC
 UAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCG
 CCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAAC
 UUCACCAUCAGCGUGACCACCGAGAUCCUGCCCGUAAGUAUGACCAAGAC
 AAGCGUGGACUGCACCAUGUACAUCUGCGGGCGACAGCACCGAGUGCAGCA
 25 ACCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAACAGAGCCCUG
 ACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCA
 AGUGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGGACUUCGGCGGCUUCA
 ACUUCAGCCAAAUCCUGCCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUC
 AUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAU
 30 CAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCU
 GCGCUCAGAAGUUAUUGGACUCACCGUGCUCCCCCCCCUGCUGACCGAC
 GAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGGCGGGAACCAUAACAAG
 CGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAUCCCCUUCGCCAUGC
 AGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUAC
 35 GAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAU
 CCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAAACG
 UGGUGAAUCAGAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGC
 AGCAACUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCUGUCUAGACU
 GGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGAC
 40 UGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGCCGCCGAG
 AUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCU

GGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUG
 AGCUUCCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCACGUGACCUA
 CGUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACG
 ACGGCAAGGCCCACUUCCCUAGAGAGGGCGUGUUCGUGAGCAACGGCACC
 5 CACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUACCAC
 CGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGA
 ACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAG
 GAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGUGGACCUGGG
 AGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUCAGAAGGAGAUC
 10 GACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCU
 GCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUUGGUACAUC
 UGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAU
 GCUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGC
 UGUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAA
 15 GGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 54

>B.1.1.529 (Omicron) spike protein mRNA

AUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA
 CCUGACCACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAA
 20 GAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCC
 ACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGU
 GAUCAGCGGCACCAACGGCACCAAGAGAUUCGACAACCCCGUGCUGCCCU
 UCAACGACGGCGUGUACUUCGCUAGCAUCGAAAAGAGCAACAUCaucaga
 GGCUGGAUCUUCGGCACCAACCCUAGACUCCAAGACACAAAGCCUGCUGAU
 25 CGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUUCAGUUC
 UGCAACGACCCCUUCCUGGACCACAAGAACAACAAGAGCUGGAUGGAGAG
 CGAGUUCAGAGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGU
 CUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAAC
 CUAAGAGAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACA
 30 GCAAGCACACCCCAUCAUCGUGAGAGAGCCCGAGGACCUGCCCCAAGGC
 UUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCAC
 AAGAUUUCAGACCCUGCUGGCCCUGCACCGUAGCUACCUAACCCCGGCG
 AUAGUUCCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACGUGGGCUAC
 CUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUCAC
 35 UGAUGCCGUGGACUGUGCCCUGGACCCCUAGAGCGAGACCAAGUGCACCC
 UCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAG
 AGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCAACAUACCAACCGUGU
 GCCCCUUCGACGAGGUGUUAACGCCACCAGAUUUGCUAGCGUGUACGCC
 UGGAACCGGAAGAGAAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU
 40 ACAACCUGGCCCCCUUCUUCACCUUCAAGUGCUAUGGCGUCAGCCCCACC
 AAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAU

CAGAGGCGACGAGGUGAGACAGAUCGCCCCCGGGCAGACCGGCAACAUCG
 CCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCC
 UGGAACAGCAACAAGCUGGAUAGCAAGGUGAGCGGCAACUACAAUACC
 UGUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAU
 5 CAGCACCGAGAUCUACCAAGCCGGCAACAAGCCCUGCAACGGCGUGGCGG
 GCUUCAACUGCUACUUCUUUUUUUCAGGAGCUACAGCUUCAGACCCACCUAC
 GGCGUGGGCCAUCAGCCCUACAGAGUGGUCGUGCUGAGCUUCGAGCUGCU
 GCACGCCCCUGCCACAGUGUGCGGGCCCCAAAAAGAGCACCAACCUGGUGA
 AGAACAAGUGCGUGAACUUAACUUAACGGGCCUGAAGGGCACCGGCGUG
 10 CUGACCGAGAGCAACAAGAAGUUCUGCAUUCAGCAAUUCGGCAGAGA
 CAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCC
 UGGACAUCACCCCCUGUAGCUUCGGCGGGCGUGAGCGUGAUCACCCCCGGC
 ACUAACACAAGCAACCAAGUGGGCCGUGCUGUACCAAGGCGUGAACUGCAC
 CGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGAG
 15 UGUUAUAGCACCGGCAGCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAUC
 GGCGCCGAGUACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCGC
 CGGCAUCUGCGCUAGCUAUCAGACACAGACCAAGAGCCACCGGAGAGCUA
 GAAGCGUGGCUAGCCAAAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCC
 GAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUU
 20 CACCAUCAGCGUGACCACCGAGAUCUCCUGCCCGUAAGUAUGACCAAGACAA
 GCGUGGACUGCACCAUGUACAUCUGCGGGCGACAGCACCGAGUGCAGCAAC
 CUGCUCUCCUGCAGUACGGCAGCUUCUGCACACAGCUGAAGAGAGCCCUGAC
 CGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAAG
 UGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGUACUUCGGCGGCUUCAAC
 25 UUCAGCCAAAUCUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCAU
 CGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUCA
 AGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGC
 GCUCAGAAGUUAAGGGACUCACCGUGCUUUUUUUUUUCUGACCCGACGA
 GAUGAUCGCUCAGUACACCUCUGCCUGCUGGGCGGGAACCAUAACAAGCG
 30 GUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCCAUGCAG
 AUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACGA
 GAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUCC
 AAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGUG
 GUGAACCACAACGCCCAAGCCCUGAACACCCUGGUGAAGCAGCUGAGCAG
 35 CAAGUUCGGCGCCAUCAGCAGCGUGCUGAAUGACAUCUUCUCUAGACUGG
 ACCCCCCCGAGGCCGAGGUGCAGAUCGACAGACUGAUCACCGGCAGACUG
 CAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGCCGCCGAGAU
 CAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUGG
 GGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAG
 40 CUUCCUCAGAGCGCCCCCACGGCGUGGUGUUCUGCACGUGACCUACG
 UGCCCCGCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGAC

GGCAAGGCCACUUCCCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCA
 CUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUCACCACCG
 ACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAAC
 AACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAGGA
 5 GCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGUGGACCUGGGAG
 ACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUCAGAAGGAGAUCGA
 CCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUGC
 AAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCCUGGUACAUCUG
 GCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGC
 10 UGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUG
 UGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGG
 GCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 55

>poly A

15 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAA
 AA
 AAAAAAAAAAAAAAA

SEQ ID NO: 56

>I'-Delta24-B mRNA

20 GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAG
 GAUAGCACUAUAUAUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAG
 CUCUCAGUGCGUGAACCUGAGGACAAGAACACAGCUGCCCCCGCCUACA
 CCAACAGCUUCACAAGAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGC
 AGCGUGCUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGU
 25 GACCUGGUUCCACGCCAUCCACGUCAGCGGCACCAACGGCACCAAGAGAU
 UCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACC
 GAAAAGAGCAACAUAUCAGAGGCUGGAUCUUCGGCACCAACCCUAGACUC
 CAAGACACAAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUA
 AGGUGUGCGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACGUGUACUAC
 30 CACAAGAACAACAAGAGCUGGAUGGAGAGCGGGGUCUACAGCAGCGCCA
 ACAACUGCACCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAG
 GGCAAGCAAGGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUCAAGAACA
 UCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCAUCAACCUGGUG
 AGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCC
 35 CAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUA
 GCUACCUAACCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCC
 GCCUACUACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAA
 CGAGAACGGCACUAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCUGA
 GCGAGACCAAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAU
 40 CAGACAAGCAACUUCAGAGUGCAGCCACCGAGAGCAUCGUGAGAUUCCC
 CAACAUCACCAACCUGUGCCCUUCGGCGAGGUGUUAACGCCACAAGAU

UCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGC
 CGACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCU
 AUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUAC
 GCCGACAGCUUCGUGAUCAGAGGCGACGAGGUGAGACAGAUUGCCCCCGG
 5 GCAAACCGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA
 CCGGCUGCGUGAUCGCGUGGAACAGCAAUAACCUGGAUUCCAAGGUGGGC
 GGCAACUACAACUACCGGUACAGACUGUUCAGAAAGAGCAACCUGAAGCC
 CUUCGAGAGAGACAUCAGCACCGAGAUCUACCAAGCCGGCAGCAAGCCCU
 GCAACGGCGUGGAGGGCUUCAAACUGCUACUUCCCCCUGCAGAGCUACGGC
 10 UUUCAGCCCACCAACGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCU
 GAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGGCCCCAAAAAGA
 GCACCAACCUGGUGAAGAACAAGUGCGUGAACUUCAACUUUAAUGGCCU
 UACCGGCACGGGCGUGCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUC
 AGCAAUUCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCU
 15 CAGACCCUGGAGAUCCUGGACAUCACCCCCUGUAGCUUCGGCGGGCUGAG
 CGUGAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCCGUGCUGUACC
 AAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUG
 ACCCCCACCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAG
 AGCCGGCUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGUGCG
 20 ACAUCCCCAUUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAC
 AGCCGCCGGAGAGCUAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACAC
 CAUGAGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCG
 CCAUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUCUGCCCGUA
 AGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGGGCGACAG
 25 CACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGC
 UGAACAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAA
 GAGGUGUUCGCCCAAGUGAAGCAGAUCUACAAGACCCCCCCCCAUCAAGGA
 CUUCGGCGGCUUCAACUUCAGCCAAAUCCUGCCCCGACCCUAGCAAGCCUA
 GCAAGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCC
 30 GACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGC
 UAGAGACCUGAUCUGCGCUCAGAAGUUAUUGGACUCACCGUGCUCCCCC
 CCCUGCUGACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGGC
 GGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAU
 CCCCUCGCCAUGCAGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACAC
 35 AGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAG
 CGCCAUCGGCAAGAUAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCU
 GCAAGCUGCAAAACGUGGUGAAUCAGAACGCCCAAGCCCUGAACACCCUG
 GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGCGUAUUGAAUG
 ACAUCCUGUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGA
 40 CUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCU
 GAUCAGAGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGA

UGAGCGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAA
 GGGCUACCACCUGAUGAGCUUCCCUCAGAGCGCCCCCACGGCGUGGUGU
 UCCUGCACGUGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCU
 CCGGCCAUCUGCCACGACGGCAAGGCCACUUCCCUAGAGAGGGCGUGUU
 5 CGUGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGC
 CUCAGAUAUCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUG
 GUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGCU
 GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGC
 CCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAA
 10 CAUUCAGAAGGAGAUUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAAC
 GAGAGCCUGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCA
 AGUGGGCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUC
 GUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCUGCC
 UGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGACGA
 15 CAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGACUCG
 AGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCUU
 AAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUG
 GAUUCUGCCUAAUAAAAACAUAUUAUUUUCAUUGCAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAA
 20 AA
 AA

SEQ ID NO: 57

>I'-OC24-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAG
 25 GAUAGCACUAUAAUAUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAG
 CUCUCAGUGCGUGAACCUGACCACAAGAACACAGCUGCCCCCGCCUACA
 CCAACAGCUUCACAAGAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGC
 AGCGUGCUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGU
 GACCUGGUUCCACGUGAUCAGCGGCACCAACGGCACCAAGAGAUUCGACA
 30 ACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCAUCGAAAAG
 AGCAACAUAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGAC
 ACAAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUG
 UGCGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACCACAAGAACAACAA
 GAGCUGGAUGGAGAGCGAGUUCAGAGUGUACAGCAGCGCCAACAACUGC
 35 ACCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCA
 AGGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUCAAGAACAUCGACGGC
 UACUUCAAGAUCUACAGCAAGCACACCCCAUCAUCGUGAGAGAGCCCGA
 GGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCCA
 UCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCCGUAGC
 40 UACCUAACCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGC
 CUACUACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACG

AGAACGGCACUAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGC
 GAGACCAAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCA
 GACAAGCAACUUCAGAGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCA
 ACAUCACCAACCUGUGCCCCUUCGACGAGGUGUUAACGCCACCAGAUUU
 5 GCUAGCGUGUACGCCUGGAACCGGAAGAGAAUCAGCAACUGCGUGGCCGA
 CUACAGCGUGCUGUACAACCUGGCCCCCUUCUUCACCUUCAAGUGCUAUG
 GCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCC
 GACAGCUUCGUGAUCAGAGGCGACGAGGUGAGACAGAUCGCCCCCGGGCA
 GACCGGCAACAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCG
 10 GCUGCGUGAUCGCCUGGAACAGCAACAAGCUGGAUAGCAAGGUGAGCGG
 CAACUACA AUUACCUGUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCU
 UCGAGAGAGACAUCAGCACCGAGAUUCUACCAAGCCGGCAACAAGCCUGC
 AACGGCGUGGCGGGCUUCAAACUGCUACUUCUUCCCCUCAGGAGCUACAGCUU
 CAGACCCACCUACGGCGUGGGCCAUCAGCCCUACAGAGUGGUCGUGCUGA
 15 GCUUCGAGCUGCUGCACGCCCCUGCCACAGUGUGCGGCCCCAAAAAGAGC
 ACCAACCUGGUGAAGAACAAGUGCGUGAACUUAACUUAACGGCCUGAA
 GGGCACCGGCGUGCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGC
 AAUUCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAG
 ACCCUGGAGAUCCUGGACAUCACCCCCUGUAGCUUCGGCGGCGUGAGCGU
 20 GAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCCGUGCUGUACCAAG
 GCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACC
 CCCACCUGGAGAGUGUAUAGCACCGGCAGCAACGUGUUUCAGACAAGAGC
 CGGCUGCCUGAUCGGCGCCGAGUACGUGAACAACAGCUACGAGUGCGACA
 UCCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAGAGC
 25 CACCGGAGAGCUAGAAGCGUGGCUAGCCAAAGCAUCAUCGCCUACACCAU
 GAGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCA
 UCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUCCUGCCCGUAAGU
 AUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGGCGACAGCAC
 CGAGUGCAGCAACCUGCUCCUGCAGUACGGCAGCUUCUGCACACAGCUGA
 30 AGAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAG
 GUGUUCGCCAAGUGAAGCAGAUCAACAAGACCCCCCCCCAUCAAGUACUU
 CGGCGGCUUCAACUUCAGCCAAAUCUGCCCGACCCUAGCAAGCCUAGCA
 AGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGAC
 GCCGGCUUCAUCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAG
 35 AGACCUGAUCUGCGCUCAGAAAGUUAAGGGACUCACCGUGCUCCCCCCCC
 UGCUGACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGCGGGA
 ACCAUACAAGCGGUUGGACA UUGGGCGCCGGGGCCGCACUGCAAAUCCC
 CUUCGCCAUGCAGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGA
 ACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCC
 40 AUCGGCAAGA UCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAA
 GCUGCAAGACGUGGUGAACCACAACGCCCAAGCCCUGAACACCCUGGUGA

AGCAGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUGCUGAAUGACAUC
 UUCUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAU
 CACCGGCAGACUGCAGAGCCUUCAGACCUACGUGACACAGCAGCUGAUA
 GAGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGC
 5 GAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCU
 ACCACCUGAUGAGCUUCCCUCAGAGCGCCCCCACGGCGUGGUGUCCUG
 CACGUGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGC
 CAUCUGCCACGACGGCAAGGCCACUUCCCUAGAGAGGGGCGUGUUCGUGA
 GCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAG
 10 AUCAUCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAU
 CGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACA
 GCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGAC
 GUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUC
 AGAAGGAGAUUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAG
 15 CCUGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGC
 CCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUG
 GUGACCAUCAUCGUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAG
 GUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGA
 GCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGACUCGAGGCUC
 20 GCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCUUAAGUCC
 AACUACUAAACUGGGGGGAUUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCU
 GCCUAAUAAAAAACAUAUUAUUAUUGCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

25 SEQ ID NO: 58

>B.1.351 spike protein mRNA
 AUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA
 CUUCACGACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAA
 GAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCC
 30 ACCCAAGACCUGUUCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGC
 CAUCCACGUGAGCGGCACCAACGGCACCAAGAGAUUCGCCAACCCCGUGC
 UGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCGAAAAGAGCAACAUC
 AUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACAAAGCCU
 GCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUU
 35 CAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACCACAAGAACAA
 GAGCUGGAUGGAGAGCGAGUUCGUGUCUACAGCAGCGCCAACAACUGCA
 CCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAA
 GGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUAAGAACAUCGACGGCU
 ACUUCAAGAUCUACAGCAAGCACACCCCCAUCAACCUGGUGAGAGGCCUG
 40 CCCCAGGCUUCAGCGCCCUGGAGCCCUGGUGGACCUGCCCAUCGGCAU
 CAACAUCACAAGAUUUCAGACCCUGCACCGUAGCUACCUAACCCCGGCG

AUAGUUCCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACGUGGGGCUAC
 CUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUCAC
 UGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCAAGUGCACCC
 UCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAG
 5 AGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCAACAUCACCAACCUGU
 GCCCCUUCGGCGAGGUGUUAACGCCACAAGAUUCGCUAGCGUGUACGCG
 UGGAACAGAAAGAGAAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU
 ACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGCGUCAGCCCCACC
 AAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAU
 10 CAGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCAAACCGGCAACAUCG
 CCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCG
 UGGAACAGCAAUAACCUGGAUUAAGGUGGGCGGCAACUACAACUACCU
 GUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUC
 AGCACCGAGAUUCUACCAAGCCGGCAGCACCCCCUGCAACGGCGUGAAGGG
 15 CUUCAACUGCUACUUCCCCCUGCAGAGCUACGGCUUUCAGCCCACCUACG
 GCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUGAGCUUCGAGCUGCU
 GCACGCCCCCGCAACCGUGUGCGGCCCCAAAAAGAGCACCAACCUGGUGA
 AGAACAAGUGCGUGAACUUAACUUUAAUGGCCUUAACGGCACGGGCGU
 GCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAG
 20 ACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUC
 CUGGACAUCACCCCCUGUAGCUUCGGCGGGCGUGAGCGUGAUCACCCCCGG
 CACUAAACACAAGCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCA
 CCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGA
 GUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAU
 25 CGGCGCCGAGCACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCCGGCG
 CCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCCCCGGAGAGCU
 AGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGU
 CGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACU
 UCACCAUCAGCGUGACCACCGAGAUCCUGCCCGUAAGUAUGACCAAGACA
 30 AGCGUGGACUGCACCAUGUACAUCUGCGGGCAGAGCACCGAGUGCAGCAA
 CCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAACAGAGCCUGA
 CCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAA
 GUGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGGACUUCGGCGGCUUCA
 CUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCA
 35 UCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUC
 AAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUG
 CGCUCAGAAGUUCUAAUGGACUCACCGUGCUCCCCCCCCUGCUGACCGACG
 AGAUGAUCGCUCAGUACACCUCUGCCUGCUGGCGGGAACCAUAACAAGC
 GGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCCAUGCA
 40 GAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACG
 AGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUC

CAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGU
 GGUGAAUCAGAACGCCCAAGCCCUGAACACCCUGGUGAAGCAGCUGAGCA
 GCAACUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCCUGUCUAGACUG
 GACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACU
 5 GCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGCCGCCGAGA
 UCAGAGCUAGCGCCAACCUGGGCCGCCACCAAGAUGAGCGAGUGCGUGCUG
 GGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGA
 GCUUCCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCACGUGACCUAC
 GUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGA
 10 CGGCAAGGCCACUUCCCUGAGAGGGCGUGUUCGUGAGCAACGGCACCC
 ACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUAUCACCACC
 GACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGA
 ACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAG
 GAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGUGGACCUGGG
 15 AGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAAGGAGAUC
 GACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCU
 GCAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCCUUGGUACAUC
 UGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAU
 GCUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGC
 20 UGUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAA
 GGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 59

>B.1.1.7 spike protein mRNA

AUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA
 25 CCUGACCACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAA
 GAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCC
 ACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGC
 CAUCAGCGGCACCAACGGCACCAAGAGAUUCGACAACCCCGUGCUGCCCU
 UCAACGACGGCGUGUACUUCGCUAGCACCGAAAAGAGCAACAUAUCAGA
 30 GGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACAAAGCCUGCUGAU
 CGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUUCAGUUC
 UGCAACGACCCCUUCCUGGGCGUGUACCACAAGAACAACAAGAGCUGGAU
 GGAGAGCGAGUUCGUGUCUACAGCAGCGCCAACAACUGCACCUUCGAGU
 ACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUC
 35 AAGAACCUAAGAGAGUUCGUGUUCAAGAACAUCGACGGCUACUUCAAGA
 UCUACAGCAAGCACACCCCCAUCAACCUGGUGAGAGACCUGCCCCAAGGC
 UUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCAC
 AAGAUUUCAGACCCUGCUGGCCUGCACCGUAGCUACCUAACCCCGGGCG
 AUAGUUCCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACGUGGGCUAC
 40 CUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUCAC
 UGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCAAGUGCACCC

UCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAG
 AGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCAACAUACCAACCUGU
 GCCCCUUCGGCGAGGUGUUAACGCCACAAGAUUCGCUAGCGUGUACGCG
 UGGAACAGAAAGAGAAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU
 5 ACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGCGUCAGCCCCACC
 AAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAU
 CAGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCAAACCGGCAAGAUCG
 CCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCG
 UGGAACAGCAAUAACCUGGAUUCCAAGGUGGGCGGCAACUACAACUACCU
 10 GUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUC
 AGCACCGAGAUUCUACCAAGCCGGCAGCACCCCCUGCAACGGCGUGGAGGG
 CUUCAACUGCUACUUCCCCCUGCAGAGCUACGGCUUUCAGCCCACCUACG
 GCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUAGAGCUUCGAGCUGCU
 GCACGCCCCCGCAACCGUGUGCGGGCCCCAAAAAGAGCACCAACCUGGUGA
 15 AGAACAAGUGCGUGAACUUAACUUUAUUGGCCUUAACGGGCACGGGCGU
 GCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAG
 ACAUCGACGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUC
 CUGGACAUCACCCCCUGUAGCUUCGGCGGGCGUGAGCGUGAUCACCCCCGG
 CACUAACACAAGCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCA
 20 CCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGA
 GUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAU
 CGGCGCCGAGCACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCCGGCG
 CCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCACCGGAGAGCU
 AGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGC
 25 CGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCAUCAACU
 UCACCAUCAGCGUGACCACCGAGAUCUGCCCGUAAGUAUGACCAAGACA
 AGCGUGGACUGCACCAUGUACAUCUGCGGGCGACAGCACCGAGUGCAGCAA
 CCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAACAGAGCCCUGA
 CCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAA
 30 GUGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGGACUUCGGCGGCCUUCAA
 CUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCA
 UCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUC
 AAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUG
 CGCUCAGAAGUUCUAAUGGACUACCCGUGCUCCCCCCCCUGCUGACCGACG
 35 AGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGCGGGAACCAUAACAAGC
 GGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCCAUGCA
 GAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACG
 AGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUC
 CAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGU
 40 GGUGAAUCAGAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCA
 GCAACUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCUGGCUAGACUG

GACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACU
 GCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGCCGCCGAGA
 UCAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUG
 GGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGA
 5 GCUUCCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCACGUGACCUAC
 GUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGA
 CGGCAAGGCCACUUCCCUCAGAGAGGGCGUGUUCGUGAGCAACGGCACCC
 ACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUAUCACCACC
 CACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAA
 10 CAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAGG
 AGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGUGGACCUGGGA
 GACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAAGGAGAUCG
 ACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUG
 CAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGGCCUGGUACAUCU
 15 GGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUG
 CUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCU
 GUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAG
 GGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 60

20 >SARS-CoV-2 spike protein mRNA
 AUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA
 CCUGACCACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAA
 GAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCC
 ACCCAAGACCUGUUCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGC
 25 CAUCCACGUGAGCGGCACCAACGGCACCAAGAGAUUCGACAACCCCGUGC
 UGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCGAAAAGAGCAACAUC
 AUCAGAGGCUGGAUCUUCGGCACCAACCCUAGACUCCAAGACACAAAGCCU
 GCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUU
 CAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAA
 30 GAGCUGGAUGGAGAGCGAGUUCGUGUCUACAGCAGCGCCAACAACUGCA
 CCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAA
 GGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUAAGAACAUCGACGGCU
 ACUUCAAGAUCUACAGCAAGCACACCCCAUCAACCUGGUGAGAGACCUG
 CCCCAGGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCCAUCGGCAU
 35 CAACAUCACAAGAUUUCAGACCCUGCUGGGCCUGCACCGUAGCUACCUAA
 CCCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUAC
 GUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGG
 CACUAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCA
 AGUGCACCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGC
 40 AACUUCAGAGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCAACAUCAC
 CAACCUGUGCCCCUUCGGCGAGGUGUUAACGCCACAAGAUUCGCUAGCG

UGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGCCGACUACAG
 CGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGCGUCA
 GCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGC
 UUCGUGAUCAGAGGGCGACGAGGUGAGACAGAUUGCCCCCGGGCA5AACCG
 5 GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGC
 GUGAUCGCGUGGAACAGCAAUAACCUGGAUUCCAAGGUGGGCGGGCAACU
 ACAACUACCUGUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAG
 AGAGACAUCAGCACCGAGAUCUACCAAGCCGGCAGCACCCCCUGCAACGG
 CGUGGAGGGGCUUCAAACUGCUACUUCCCCCUGCAGAGCUACGGCUUUCAGC
 10 CCACCAAUGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUGAGCUUC
 GAGCUGCUGCACGCCCCCGCAACCGUGUGCGGGCCCCAAAAAGAGCACCAA
 CCUGGUGAAGAACAAGUGCGUGAACUUCAACUUAUGGGCCUUAACCGGC
 ACGGGCGUGCUGACCGAGAGCAACAAGAAGUCCUGCCAUUUCAGCAAUU
 CGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCC
 15 UGGAGAUCUGGACAUCACCCCCUGUAGCUUCGGCGGGCGUGAGCGUGAUC
 ACCCCCGGCACUAAACACAAGCAACCAAGUGGCCGUGCUGUACCAAGACGU
 GAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACCCCCA
 CCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGC
 UGCCUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGUGCGACAUCCC
 20 CAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCCUC
 GGAGAGCUAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGC
 CUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCC
 CACCAACUUCACCAUCAGCGUGACCACCGAGAUCCUGCCCGUAAGUAUGA
 CCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGGGCGACAGCACCGAG
 25 UGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAACAG
 AGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACCCCAAGAGGUGU
 UCGCCCAAGUGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGGACUUCGGC
 GGCUUCAACUUCAGCCAAAUCCUGCCCCGACCCUAGCAAGCCUAGCAAGCG
 GAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCG
 30 GCUUCAUCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGAC
 CUGAUCUGCGCUCAGAAGUUCAUUGGACUCACCGUGCUCCCCCCCCUGCU
 GACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGGCGGGAACCA
 UAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAUCCCCUUC
 GCCAUGCAGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGU
 35 GCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCG
 GCAAGAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUG
 CAAGACGUGGUGAAUCAGAACGCCCAAGCCCUGAACACCCUGGUGAAGCA
 GCUGAGCAGCAACUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCUGA
 GCAGACUGGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAUCACC
 40 GGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGC
 CGCCGAGAUCAAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGU

GCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCA
 CCUGAUGAGCUUCCCUCAGAGCGCCCCCACGGCGUGGUGUUCCUGCACG
 UGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUC
 UGCCACGACGGCAAGGGCCACUUCCCUAGAGAGGGGCGUGUUCGUGAGCAA
 5 CGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUA
 UCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGC
 AUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGCUU
 CAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGUGG
 ACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUCAGAA
 10 GGAGAUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUG
 AUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGGCCUG
 GUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGA
 CCAUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGC
 UGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGU
 15 GCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 61

>Luciferase mRNA

AUGGAGGACGCCAAGAACAUCAAGAAAGGCCCCGCCCCUUCUACCCCC
 UGGAGGACGGCACCGCCGGCGAGCAGCUGCACAAAGCCAUGAAAAGAUAA
 20 CGCUCUGGUGCCCGGCACCAUCGCCUUCACCGACGCCACAUCGAGGUGA
 ACAUCACCUACGCCGAGUACUUCGAGAUGAGCGUGAGACUGGCCGAGGC
 UAUGAAGAGAUACGGCUUGAACACCAACCACAGAAUCGUGGUGUGCAGC
 GAGAACAGCCUGCAGUUCUUCAUGCCCGUGCUGGGCGCCUGUUCAUCG
 GCGUGGCCGUUGCCCCCGCCAACGACAUCUACAACGAGAGAGAGCUGCU
 25 GAACAGCAUGAACAUCUCUCAGCCCACCGUGGUUUUCGUGUCCAAAAAG
 GGCCUGCAGAAGAUCUGAACGUGCAGAAGAAGCUGCCCAUCAUUCAGA
 AGAUCAUCAUCAUGGACAGCAAGACCGACUACCAAGGCUUUCAGAGCAU
 GUACACCUUCGUGACAAGCCACCUGCCCGGCUUCAACGAGUACGAC
 UUCGUGCCCGAGAGCUUCGACAGAGACAAGACCAUCGCCUGAUAUGA
 30 ACAGCAGCGGCAGCACCGGCCUGCCCAAGGGCGUGGCCUGCCCCACAGA
 ACCGCCUGCGUGAGAUUCAGCCACGCUAGAGACCCCAUCUUCGGCAAUC
 AGAUCAUCCCCGACACCGCCAUCCUGAGCGUAGUGCCUUUCCAUCAUGG
 CUUCGGCAUGUUCACCACCCUGGGCUACCUGAUCUGCGGCUUCAGAGUG
 GUGCUGAUGUACAGAUUCGAGGAGGAGCUGUUCUGAGAAGCCUGCAAG
 35 ACUACAAGAUUCAGAGCGCCCUGCUGGUGCCCACCCUGUUCAGCUUCUU
 CGCCAAGAGCACCCUGAUCGACAAGUAUGACCUGAGCAACCUGCACGAA
 AUUGCGAGCGGCGGAGCGCCGCUAGCAAAGAGGUGGGCGAGGCCGUGG
 CCAAGAGAUUCCACCUGCCCCGGCAUCAGACAAGGCUACGGCCUGACCGA

GACCACAAGCGCCAUCCUGAUCACCCCCGAGGGCGACGAUAAGCCCCGGC
 GCCGUGGGCAAGGUGGUACCAUUCUUCGAGGCCAAGGUGGUGGACCUGG
 ACACCGGCAAGACCCUGGGCGUGAAUCAGAGAGGCGAGCUGUGCGUGAG
 AGGCCCCAUGAUGAUGAGCGGCUACGUGAACAACCCCGAGGCCACCAAC
 5 GCCCUCAUCGAUAAGGAUUGGGUGGCUGCACAGCGGCGACAUCGCCUACU
 GGGACGAGGACGAGCACUUCUUCUUCGUGGACAGACUCAAGAGCCUGAU
 CAAGUACAAGGGCUACCAAGUGGGCCCCCGCCGAGUUGGAGAGCAUCCUG
 CUGCAGCACCCCAAUAUCUUCGACGCUGGCGUGGCCGGACUACCCGACG
 ACGACGCCGGCGAGUUACCCGCCGCCGUGGUGGUGCUGGAGCACGGCAA
 10 GACCAUGACCGAGAAGGAGAUCGUGGACUACGUGGCUAGCCAAGUGACC
 ACCGCCAAGAAGCUGAGAGGCGGGCGUGGUGUUCGUUGACGAAGUGCCCA
 AGGGCCUGACCGGCAAGCUGGACGCUAGAAAGAUCAGAGAGAUCUUGAU
 CAAGGCCAAGAAGGGCGGCAAGAGCAAGCUGUGA

SEQ ID NO: 62

15 >LUC1- I' truncation
 UAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUA
 GCACUAUAAU

SEQ ID NO: 63

>LUC2- I' truncation
 20 AGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCAC
 UAUAUU

SEQ ID NO: 64

>LUC3- I' truncation
 GGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUA
 25 AU

SEQ ID NO: 65

>LUC4- I' truncation
 GCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU

SEQ ID NO: 66

30 >LUC5- I' truncation
 UUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU

SEQ ID NO: 67

>LUC6- I' truncation
 AAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU

35 SEQ ID NO: 68

>LUC7- I' truncation
 ACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU

SEQ ID NO: 69

>LUC8- I' truncation
GUCAUUGCCGAGAAAAGGAUAGCACUAUAAU
SEQ ID NO: 70

>LUC9- I' truncation
5 UUGCCGAGAAAAGGAUAGCACUAUAAU
SEQ ID NO: 71

>LUC10- I' truncation
CGAGAAAAGGAUAGCACUAUAAU
SEQ ID NO: 72

10 >LUC11- I' truncation
AAAAGGAUAGCACUAUAAU
SEQ ID NO: 73

>LUC12- I' truncation
GGAUAGCACUAUAAU
15 SEQ ID NO: 74

>LUC13- I' truncation
AGCACUAUAAU
SEQ ID NO: 75

>LUC14- I' truncation
20 CUAUAAU

>LUC15- I' mutant
GAGUUAGGAGUUGCUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAU
25 SEQ ID NO: 77

>LUC16- I' mutant
GAGUUAGGAGUUGUUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAU
SEQ ID NO: 78

30 >5'UTR I general formula
GAGUUAGGAGUUGN₂UGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAG
GAUAGCACUAUAAU
SEQ ID NO: 79

N₂ is selected from the group consisting of A, G, C, and U

35 >Wildtype SARS-CoV-2 spike protein amino acid sequence
MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDL
FLPFFSNVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTT
LDSKTQSLNIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSS

ANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDL
 PQGFSALEPLVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAYYVGYL
 QPRTFLLKYNENGITITDAVDCALDPLSETKCTLKSTVEKGIYQTSNFRVQPTES
 IVRFPNITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLNSASFSTFKCY
 5 GVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVI
 AWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFN
 CYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVN
 FNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGV
 SVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAG
 10 CLIGAEHVNNSECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAEN
 SVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSF
 CTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSK
 RSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE
 MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQ
 15 KLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAIS
 SVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK
 MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTVPAQEKNTTAPAI
 CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVN
 NTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNE
 20 VAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCS
 CLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 80

>Wildtype SARS-CoV-2 spike protein CDS (NC_045512.2)

ATGTTTGTGTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTT
 25 ACAACCAGAACTCAATTACCCCTGCATACACTAATTCTTTCACACGTGGTG
 TTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGAC
 TTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATGTCTCT
 GGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATG
 GTGTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTT
 30 GGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTAC
 TAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGG
 TGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGT
 TTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTAT
 GGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTT
 35 AAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACACGCCTATTAATTTA
 GTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTTGGTAGATTTGC

CAATAGGTATTAACATCACTAGGTTTCAAACCTTACTTGCTTTACATAGAAGT
 TATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTA
 TTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGG
 AACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAA
 5 GTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTTCTAAC
 TTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCTAATATTACAACTT
 GTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTT
 GGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAA
 TTCCGCATCATTTTCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAA
 10 TGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGA
 AGTCAGACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTGATTATAATTAT
 AAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCT
 TGATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGT
 CTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGG
 15 TAGCACACCTTGTAATGGTGTGTAAGGTTTTAATTGTTACTTTCCCTTTACAAT
 CATATGGTTTCCAACCCACTAATGGTGTGGTTACCAACCATACAGAGTAGT
 AGTACTTTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAA
 AAGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTT
 AACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCA
 20 ACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAG
 ACACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTGTCAGTGTTAT
 AACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTATCAGGATGTT
 AACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTT
 GCGGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTA
 25 ATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTG
 CAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACG
 TAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACCTGGTGCAGAAA
 ATTCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTA
 GTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTG
 30 TACAATGTACATTTGTGGTGATTCAACTGAATGCAGCAATCTTTTGTGCAAT
 ATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGA
 ACAAGACAAAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACA
 AAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTTCACAAATATTACCA
 GATCCATCAAAACCAAGCAAGAGGTCATTTATTGAAGATCTACTTTTCAACA
 35 AAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGG
 TGATATTGCTGCTAGAGACCTCATTTGTGCACAAAAGTTTAACGGCCTTACT

GTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACT
 GTTAGCGGGTACAATCACTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTA
 CAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAATGGTATTGGAGTTAC
 ACAGAATGTTCTCTATGAGAACCAAAAATTGATTGCCAACCAATTTAATAGT
 5 GCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGTGCACCTTGGAA
 AACTTCAAGATGTGGTCAACCAAAATGCACAAGCTTTAAACACGCTTGTTA
 AACAACCTTAGCTCCAATTTTGGTGAATTTCAAGTGTTTTAAATGATATCCTT
 TCACGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACA
 GGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTG
 10 CAGAAATCAGAGCTTCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGT
 ACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATG
 TCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGT
 CCCTGCACAAGAAAAGAACTTCACAACCTGCTCCTGCCATTTGTCATGATGG
 AAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTG
 15 GTTTGTAAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAAC
 ACATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAG
 TTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAA
 TATTTTAAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATT
 AATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTG
 20 CCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGA
 GCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGA
 TTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGT
 TGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAAGACG
 ACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAA

25 SEQ ID NO: 81

>Wildtype SARS-CoV-2 spike protein mRNA

AUGUUUGUUUUUCUUGUUUUAUUGCCACUAGUCUCUAGUCAGUGUGUUA
 AUCUUACAACCAGAACUCAAUUACCCCCUGCAUACACUAAUUCUUUCAC
 ACGUGGUGUUUAUUACCCUGACAAAGUUUUCAGAUCCUCAGUUUUACAU
 30 UCAACUCAGGACUUGUUCUUACCUUUCUUUUCCAAUGUUACUUGGUUCC
 AUGCUAUACAUGUCUCUGGGACCAAUGGUACUAAGAGGUUUGAUAAACC
 UGUCCUACCAUUUAAUGAUGGUGUUUAUUUUGCUUCCACUGAGAAGUCU
 AACAUAAUAAGAGGCUGGAUUUUUGGUACUACUUUAGAUUCGAAGACCC
 AGUCCCUACUUUAUUGUUAUAUAACGCUACUAAUGUUGUUAUUAAGUCUG
 35 UGAAUUUCAAUUUUGUAAUGAUCCAUUUUUGGGUGUUUAUUACCACAA
 AAACAACAAAAGUUGGAUGGAAAGUGAGUUCAGAGUUUAUUCUAGUGC

GAAUAAUUGCACUUUUGAAUAUGUCUCUCAGCCUUUUCUUAUGGACCUU
 GAAGGAAAACAGGGUAAUUUCAAAAAUCUUAGGGAAUUUGUGUUUAAG
 AAUAUUGAUGGUUAUUUUAAAAUAUAUUCUAAGCACACGCCUAUUAUU
 UUAGUGCGUGAUCUCCUCAGGGUUUUUCGGCUUUAGAACCAUUGGUAG
 5 AUUUGCCAAUAGGUAAUUAACAUCACUAGGUUUCAAACUUUACUUGCUIIU
 ACAUAGAAGUUAUUUGACUCCUGGUGAUUCUUCUUCAGGUUGGACAGCU
 GGUGCUGCAGCUUAUUAUGUGGGUUAUCUUCAACCUAGGACUUUUCUAU
 UAAAAUAUAAUGAAAAUGGAACCAUUAACAGAUGCUGUAGACUGUGCACU
 UGACCCUCUCUCAGAAACAAAGUGUACGUUGAAAUCCUUCACUGUAGAA
 10 AAAGGAAUCUAUCAACUUCUAACUUUAGAGUCCAACCAACAGAAUCUA
 UUGUUAGAUAUCCUAAUAUUACAAACUUGUGCCCUUUUGGUGAAGUUUU
 UAACGCCACCAGAUUUGCAUCUGUUUAUGCUUGGAACAGGAAGAGAAUC
 AGCAACUGUGUUGCUGAUUAUUCUGUCCUAUAUAAUUCCGCAUCAUUUU
 CCACUUUUUAAAGUGUUAUGGAGUGUCUCCUACUAAAUAAAUGAUCUCUG
 15 CUUUACUAAUGUCUAUGCAGAUUCAUUGUAAUUAAGAGGUGAUGAAGU
 CAGACAAAUCGCUCCAGGGCAAACUGGAAAGAUUGCUGAUUAUAAUUAU
 AAUUAACCAGAUGAUUUUACAGGCUGCGUUAUAGCUUGGAUUCUAACA
 AUCUUGAUUCUAAGGUUGGUGGUAAUUAUAAUUAACUGUAUAGAUUGU
 UUAGGAAGUCUAAUCUCAAAACCUUUUGAGAGAGAUAAUUAACUGAAAU
 20 CUAUCAGGCCGGUAGCACACCUUGUAAUGGUGUUGAAGGUUUUAAUUGU
 UACUUUCCUUUACAAUCAUAUGGUUUCCAACCCACUAAUGGUGUUGGUU
 ACCAACCAUACAGAGUAGUAGUACUUCUUUUGAACUUCUACAUGCACC
 AGCAACUGUUUGUGGACC UAAAAAGUCUACUAAUUGGUUAAAAACAAA
 UGUGUCAAUUUAACUUCUAAUGGUUUAAACAGGCACAGGUGUUCUACUG
 25 AGUCUAACAAAAAGUUUCUGCCUUCCAACAAUUGGCAGAGACAUUGC
 UGACACUACUGAUGCUGUCCGUGAUCCACAGACACUUGAGAUUCUUGAC
 AUUACACCAUGUUCUUUUUGGUGGUGUCAGUGUUUAUAAACACCAGGAACAA
 AUACUUCUAACCAGGUUGCUGUUCUUUAUCAGGAUGUUAACUGCACAGA
 AGUCCCUGUUGCUAUUCAUGCAGAUCAACUACUCCUACUUGGCGUGUU
 30 UAUUCUACAGGUUCUAAUGUUUUUCAAAACACGUGCAGGCUGUUUAAUAG
 GGGCUGAACAUUGUCAACAACUCAUAUGAGUGUGACAUAACCAUUGGUGC
 AGGUUAUUGCGCUAGUUAUCAGACUCAGACUAAUUCUCCUCGGCGGGCA
 CGUAGUGUAGCUAGUCAAUCCAUCAUUGCCUACACUAUGUCACUUGGUG
 CAGAAAAUUCAGUUGCUIUACUCUAAUAACUCUAUUGCCAUAACCCACAAA
 35 UUUUACUAUUAGUGUUACCACAGAAAUUCUACCAGUGUCUAUGACCAAG
 ACAUCAGUAGAUAUGUACAAUGUACAUAUUGUGGUGAUUUAACUGAAUGCA

GCAAUCUUUUGUUGCAAUAUGGCAGUUUUUGUACACAAUUAACCGUGC
 UUUAACUGGAAUAGCUGUUGAACAAAGACAAAAACACCCAAGAAGUUUUU
 GCACAAGUCAAAACAAAUUUACAAAACACCACCAAUUAAAGAUUUUGGUG
 GUUUUAAUUUUUCACAAAUUUUACCAGAUCCAUCAAAACCAAGCAAGAG
 5 GUCAUUUAUUGAAGAUCUACUUUUCAACAAAGUGACACUUGCAGAUGCU
 GGCUUCAUCAAAACAAUAUGGUGAUUGCCUUGGUGAUUUGCUGCUAGAG
 ACCUCAUUUGUGGCACAAAAGUUUAACGGCCUACUGUUUUGCCACCUUU
 GCUCACAGAUGAAAUGAUUGCUCAAUACACUUCUGCACUGUUAGCGGGU
 ACAUUCACUUCUGGUUGGACCUUUGGUGCAGGUGCUGCAUUACAAAUAC
 10 CAUUUGCUAUGCAAUUGGCUUAUAGGUUUAAUGGUUUGGAGUUACAC
 AGAAUGUUCUCUAUGAGAACCACAAAAUUGAUUGCCAACCAAUUAAUAG
 UGCUAUUGGCAAAAUUCAAGACUCACUUUCUCCACAGCAAGUGCACUU
 GGAAAACUUCAAGAUGUGGUCAACCAAAAUGCACAAGCUUUAACACGC
 UUGUUAAACAACUUAGCUCCAUUUUUGGUGCAAUUUCAAGUGUUUUAAA
 15 UGAUAUCCUUUCACGUCUUGACAAAGUUGAGGCUGAAGUGCAAUUGAU
 AGGUUGAUCACAGGCAGACUUCAAAGUUUGCAGACAUUUGUGACUCAAC
 AAUUAUUUAGAGCUGCAGAAAUCAGAGCUUCUGCUAAUCUUGCUGCUAC
 UAAAAUGUCAGAGUGUGUACUUGGACAAUCAAAAAGAGUUGAUUUUUUG
 UGGAAAGGGCUAUCUUAUGUCCUCCCCUCAGUCAGCACCUCAUGGU
 20 GUAGUCUUCUUGCAUGUGACUUAUGUCCCUGCACAAGAAAAGAACUUC
 CAACUGCUCCUGCCAUUUGUCAUGAUGGAAAAGCACACUUUCCUCGUGA
 AGGUGUCUUUGUUUCAAUUGGCACACACUGGUUUUGUAACACAAAGGAAU
 UUUUAUGAACCACAAAUCAUUACUACAGACAACACAUUUGUGUCUGGUA
 ACUGUGAUGUUGUAAUAGGAAUUGUCAACAACACAGUUUAUGAUCCUUU
 25 GCAACCUGAAUUAGACUCAUUAAGGAGGAGUUAGAUAUUUUUAA
 GAAUCAUACAUCACCAGAUGUUGAUUUUAGGUGACAUCUCUGGCAUUAU
 GCUUCAGUUGUAAACAUUCAAAAAGAAAUUGACCGCCUCAUUGAGGUUG
 CCAAGAAUUUAAAUGAAUCUCUCAUCGAUCUCCAAGAACUUGGAAAGUA
 UGAGCAGUAUAUAAAAUGGCCAUGGUACAUUUGGCUAGGUUUUAUAGC
 30 UGGCUUGAUUGCCAUAUAGUAAUGGUGACAAUUAUGCUUUGCUGUAUGACC
 AGUUGCUGUAGUUGUCUCAAGGGCUGUUGUUCUUGUGGAUCCUGCUGCA
 AAUUUGAUGAAGACGACUCUGAGCCAGUGCUCAAAGGAGUCAAAUUACA
 UUACACAUAA

SEQ ID NO: 82

35 >Influenza A virus H1N1 (A/Wisconsin/588/2019) HA amino acid sequence
 MKAILVVMPLYTFTTANADTLICIGYHANNSTDVDTVLEKNVTVTHSVNLLED

KHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTARSWSYIVETSNSDNG
TCYPGDFINYEELREQLSSVSSFERFEIFPKTSSWPNHDSNDNGVTAACPHAGAK
SFYKNLIWLKKGKSYPKINQTYINDKGKEVLVLWGIHHPPTIADQQSLYQNA
DAYVFGVTSRYSKKFKPEIATRPKVRDQEGRMNYYWTLVEPGDKITFEATGNL
5 VAPRYAFTMERDAGSGIIISDTPVHDCNTTCQTPEGAINSTLPPFQNVHPITIGKCP
KYVKSTKLRLATGLRNVPSIQSRGLFGAIAAGFIEGGWTGMVDGWYGYHHQNE
QGSGYAADLKSTQNAIDKITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNK
KVDDGFLDIWTYNAELLVLENERTLDYHDSNVKNLYEKVRNQLKNNAKEIG
NGCFEFYHKCDNTCMESVKNGTYPKYSEEAKLNREKIDGVKLDSTRIYQIL
10 AIYSTVASSLVVSLGAISFWMCSNGSLQCRICI

SEQ ID NO: 83

>Influenza A virus H3N2 (A/Cambodia/e0826360/2020) HA amino acid sequence
MKTIIALSYILCLVFAQKIPGNDNSTATLCLGHHAVPNGTIVKTITNDRIEVTNAT
ELVQNSSIGEICDSPHQILDGGNCTLIDALLGDPQCDGFQNKEDLFVERSRAN
15 SNCYPYDVPDYASLRSLVASSGTLEFKNESFNWTGVKQNGTSSACIRGSSSSFFS
RLNWLTHLNYTYPALNVTMPNNEQFDKLYIWGVHHPSTDQISLFAQPSGRI
TVSTKRSQQAVIPNIGSRPRIRDIPSRSISYWTIVKPGDILLINSTGNLIAPRGYFKI
RSGKSSIMRSDAPIGKCKSECITPNGSIPNDKPFQNVNRITYGACPRYVKQSTLK
LATGMRNVPEKQTRGIFGAIAGFIENGWEGMVDGWYGFRHQNSEGRGQAAD
20 LKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRVQDLEKYVEDTKIDL
WSYNAELLVALENQHTIDLTDSMNKLFKTKKQLRENAEDMGNGCFKIYHK
CDNACIGSIRNETYDHNVYRDEALNNRFQIKGVELKSGYKDWILWISFAMSCF
LLCIIALLGFIMWACQKGNIRCNICI

SEQ ID NO: 84

25 >Influenza B virus Victoria lineage (B/Washington/02/2019) HA amino acid sequence
MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATQGEVNVTVGIPLTTTPTKSH
FANLKGTETRGKLCPKCLNCTDLDVALGRPCKTGKIPSARVSILHEVRPVTSGC
FPIMHDRTKIRQLPNLLRGYEHVRLSTHNVINAEDAPGRPYEIGTSGSCPNTNG
NGFFATMAWAVPKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLY
30 GDSKPQKFTSSANGVTTHYVSQIGGFNPQTEDGGLPQSGRIVVDYMVQKSGKT
GTITYQRGILLPQKVWCASGRSKVIKGSPLIGEADCLHEKYGGLNLSKPYT
GEHAKAIGNCPIWVKTPKLKLANGTKYRPPAKLLKERGFFGAIAGFLEGGWEG
MIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGA
MDELHNEILELDEKVDDLADTISSQIELAVLLSNEGIINSEDEHLLALERKLLK
35 MLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNTAASLN
DDGLDNHTILLYYSTAASSLAVTLMIAIFVYVMVSRDNVSCSICL

SEQ ID NO: 85

>Influenza B virus Yamagata lineage (B/Phuket/3073/2013) HA amino acid sequence
MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATQGEVNVTVIPLTTTPTKSY
FANLKGTRTRGKLCPDCLNCTDLDVALGRPMC VGTTPSAKASILHEVRPVTSG
5 CFPIMHDRTKIRQLPNLLRGYEKIRLSTQNVIDAEKAPGGPYRLGTSGSCP NATS
KIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMK
SLYGDSNPQKFTSSANGVTTHYVSQIGDFPDQTEDGGLPQSGRIVVDYMMQKP
GKTGTIVYQRGVLLPQKVWCASGRSKVIKGS LPLIGEADCLHEEYGG LNKSKP
YYTGKHAKAIGNCPIWVK TPLKLANGTKYRPPAKLLKERGFFGAIAGFLEGG
10 WEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQR
LSGAMDELHNEILELDEKVDDL RADTISSQIELAVLLSNEGIINSEDEHLLALER
KLKKMLGPSAVDIGNGCFETKHKCNQTCLDR IAAGTFNAGEFSLPTFDSL NITA
ASLNDDGLDNHTILLYSTAASSLAVTLM LAIFIVYMVSRDNVSCSICL

SEQ ID NO: 86

15 >Influenza A virus H1N1 (A/Wisconsin/588/2019) HA CDS
ATGAAGGCCATCCTGGTGGTGATGCTGTACACCTTCACCACCGCCAACGCC
GACACCCTGTGCATCGGCTACCACGCCAACACAGCACCGACACCGTGGAC
ACCGTGCTGGAGAAGAACGTGACCGTGACCCACAGCGTG AACCTGCTGGA
GGACAAGCAC AACGGCAAGCTGTGCAAGCTGCGGGGCGTG GCCCCCTGC
20 ACCTGGGCAAGTGCAACATCGCCGGCTGGATCCTGGGCAACCCCGAGTGCG
AGTCCCTGAGCACCGCCCGCTCCTGGAGCTACATCGTG GAGACCTCCAACA
GCGACAACGGCACCTGCTACCCCGGGGACTTCATCAACTACGAGGAGCTGC
GGGAGCAGCTGTCCAGCGTGTCAGCTTCGAGCGCTTCGAGATCTTCCCCA
AGACCTCCAGCTGGCCCAACCACGACTCCGACAACGGCGTGACCGCCGCC
25 TGCCCCCACGCCGGGGCCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTG
AAGAAGGGCAAGTCCTACCCCAAGATCAACCAGACCTACATCAACGACAA
GGGGAAGGAGGTCCTCGTGCTGTGGGGCATCCACCACCCCCCACCATCGC
CGACCAGCAGAGCCTCTACCAGAACGCCGACGCCTACGTCTTCGTGGGGAC
CTCCCGGTACAGCAAGAAGTTCAAGCCCGAGATCGCCACCCGCCCCAAGGT
30 CCGGGACCAGGAGGGCCGCATGAACTACTACTGGACCCTGGTGGAGCCGG
GGGACAAGATCACCTTCGAGGCCACCGGCAACCTCGTCGCCCCGCGGTACG
CCTTCACGATGGAGCGCGACGCCGGGTCCGGCATCATCATCAGCGACACCC
CCGTGCACGACTGCAACACGACCTGCCAGACGCCCGAGGGGGCGATCAAC
ACCTCCCTGCCGTTCCAGAACGTCCACCCGATCACGATCGGCAAGTGCCCC
35 AAGTACGTGAAGTCCACCAAGCTCCGGCTGGCCACGGGGCTCCGCAACGTC
CCCTCCATCCAGAGCCGGGGCCTGTTCCGGGGCGATCGCCGGCTTCATCGAG

GGGGGCTGGACCGGGATGGTGGACGGCTGGTACGGGTACCACCACCAGAA
CGAGCAGGGCTCCGGGTACGCGGCCGACCTCAAGTCGACGCAGAACGCGA
TCGACAAGATCACCAACAAGGTCAACAGCGTGATCGAGAAGATGAACACG
CAGTTCACCGCCGTGCGCAAGGAGTTCAACCACCTGGAGAAGCGCATCGA
5 GAACCTCAACAAGAAGGTGGACGACGGGTTCCTGGACATCTGGACGTACA
ACGCGGAGCTCCTGGTCCTCCTGGAGAACGAGCGGACCCTCGACTACCACG
ACTCCAACGTGAAGAACCTGTACGAGAAGGTCCGCAACCAGCTCAAGAAC
AACGCCAAGGAGATCGGCAACGGGTGCTTCGAGTTCTACCACAAGTGCGAC
AACACGTGCATGGAGTCGGTGAAGAACGGCACCTACGACTACCCGAAGTAC
10 AGCGAGGAGGCGAAGCTGAACCGGGAGAAGATCGACGGGGTCAAGCTCGA
CTCCACGCGCATCTACCAGATCCTGGCCATCTACTCGACCGTGGCGAGCTCC
CTCGTCTGGTGGTCTCGCTCGGCGCCATCAGCTTCTGGATGTGCTCCAACG
GGTCGCTGCAGTGCCGGATCTGCATCTGA

SEQ ID NO: 87

15 >Influenza A virus H3N2 (A/Cambodia/e0826360/2020) HA CDS
ATGAAGACCATCATCGCCCTGAGCTACATCCTGTGCCTGGTGTTCGCCCAGA
AGATCCCCGGCAACGACAACAGCACCGCCACCCTGTGCCTGGGCCACCAC
GCCGTGCCCAACGGCACCATCGTGAAGACCATCACCAACGACCGGATCGAG
GTGACCAACGCCACCGAGCTGGTGCAGAACTCCAGCATCGGCGAGATCTGC
20 GACTCCCCCACCAGATCCTGGACGGCGGCAACTGCACCCTGATCGACGCC
CTGCTGGGCGACCCCCAGTGCGACGGGTTCAGAAACAAGGAGTGGGACCT
GTTCGTGGAGCGGAGCCGCGCCAACTCCAAGTGTACCCCTACGACGTGCC
CGACTACGCCAGCCTGCGGTCCCTGGTGGCCAGCTCCGGCACCCCTGGAGTT
CAAGAACGAGAGCTTCAACTGGACCGGGGTCAAGCAGAACGGCACCTCCA
25 GCGCCTGCATCCGCGGGTCCAGCTCCAGCTTCTTCTCCCGGCTGAACTGGCT
CACCCACCTGAACTACACCTACCCCGCCCTCAACGTGACCATGCCCAACAA
CGAGCAGTTCGACAAGCTGTACATCTGGGGCGTCCACCACCCAGCACCGA
CAAGGACCAGATCTCCCTCTTCGCCCAGCCCAGCGGGCGCATCACCGTGTC
CACGAAGCGGAGCCAGCAGGCCGTCATCCCCAACATCGGCTCCCGCCCCGCG
30 GATCCGCGACATCCCCAGCCGGATCTCCATCTACTGGACGATCGTGAAGCCG
GGGGACATCCTGCTCATCAACAGCACCGGCAACCTGATCGCCCCCGCGGG
TACTTCAAGATCCGGTCCGGCAAGAGCTCGATCATGCGCTCCGACGCCCCGA
TCGGGAAGTGCAAGAGCGAGTGCATACCCCCAACGGCTCGATCCCGAACG
ACAAGCCCTTCCAGAACGTCAACCGGATCACGTACGGGGCCTGCCCGCGCT
35 ACGTGAAGCAGTCCACGCTCAAGCTGGCCACCGGCATGCGGAACGTCCCCG
AGAAGCAGACCCGCGGGATCTTCGGCGCCATCGCGGGGTTCATCGAGAACG

GCTGGGAGGGGATGGTGGACGGCTGGTACGGGTTCGGGCACCAGAACAGC
 GAGGGCCGCGGGCAGGCCGCGGACCTCAAGTCGACGCAGGCCGCGATCGA
 CCAGATCAACGGCAAGCTGAACCGGCTCATCGGGAAGACGAACGAGAAGT
 TCCACCAGATCGAGAAGGAGTTCTCCGAGGTCGAGGGCCGCGTGCAGGAC
 5 CTGGAGAAGTACGTCGAGGACACCAAGATCGACCTCTGGAGCTACAACGCC
 GAGCTGCTCGTGGCGCTGGAGAACCAGCACACCATCGACCTCACGGACTCG
 GAGATGAACAAGCTGTTTCGAGAAGACGAAGAAGCAGCTCCGGGAGAACGC
 CGAGGACATGGGGAACGGCTGCTTCAAGATCTACCACAAGTGCGACAACGC
 GTGCATCGGGTCCATCCGCAACGAGACCTACGACCACAACGTCTACCGGGA
 10 CGAGGCCCTGAACAACCGCTTCCAGATCAAGGGCGTGGAGCTCAAGAGCG
 GGTACAAGGACTGGATCCTGTGGATCTCGTTCGCGATGTCCTGCTTCCTCCT
 GTGCATCGCCCTCCTGGGCTTCATCATGTGGGCGTGCCAGAAGGGGAACAT
 CCGGTGCAACATCTGCATCTGA

SEQ ID NO: 88

15 >Influenza B virus Victoria lineage (B/Washington/02/2019) HA CDS
 ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGG
 ATCTGCACCGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCC
 ACCCAGGGCGAGGTGAACGTGACCGGCGTGATCCCCCTGACCACCACCCCC
 ACCAAGTCCCCTTCGCCAACCTGAAGGGCACCGAGACCCGGGGCAAGCT
 20 GTGCCCCAAGTGCTGAACTGCACCGACCTGGACGTGGCCCTGGGCCGCCC
 CAAGTGCACCGGCAAGATCCCCAGCGCCCGGGTGTCCATCCTGCACGAGGT
 GCGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCACCAAGAT
 CCGGCAGCTGCCAACCTGCTGCGGGGGTACGAGCACGTCCGCCTGTCCAC
 CCACAACGTGATCAACGCCGAGGACGCCCCCGGCCGGCCCTACGAGATCGG
 25 GACCAGCGGCTCCTGCCCCAACATCACCACGGGAACGGCTTCTTCGCCAC
 CATGGCCTGGGCCGTCCCCAAGAACAAGACCGCCACCAACCCCCTGACCAT
 CGAGGTGCCGTACATCTGCACCGAGGGGGAGGACCAGATCACCGTCTGGGG
 CTTCCACAGCGACAACGAGACGCAGATGGCCAAGCTGTACGGGGACTCCA
 AGCCCCAGAAGTTCACGAGCTCCGCCAACGGCGTGACCACCCACTACGTCA
 30 GCCAGATCGGGGGCTTCCCGAACCAGACGGAGGACGGGGGCCTGCCCCAG
 TCCGGGCGCATCGTGGTCGACTACATGGTGCAGAAGTCCGGCAAGACGGGG
 ACCATCACGTACCAGCGGGGCATCCTGCTGCCGCAGAAGGTCTGGTGCGCC
 TCCGGGCGCAGCAAGGTGATCAAGGGCTCCCTGCCCCTCATCGGGGAGGCC
 GACTGCCTGCACGAGAAGTACGGCGGGCTCAACAAGAGCAAGCCGTACTA
 35 CACCGGCGAGCACGCCAAGGCCATCGGGAACCTGCCCCATCTGGGTCAAGAC
 CCCGCTGAAGCTCGCCAACGGCACGAAGTACCGGCCCCCGGCCAAGCTGCT

CAAGGAGCGCGGGTTCTTCGGCGCCATCGCCGGGTTCCTGGAGGGCGGGTG
 GGAGGGCATGATCGCCGGGTGGCACGGCTACACGTCCCACGGGGCCACG
 GCGTGGCGGTGCGCCGCGGACCTCAAGAGCACCCAGGAGGCCATCAACAAG
 ATCACGAAGAACCTGAACTCCCTCAGCGAGCTGGAGGTGAAGAACCTCCA
 5 GCGCCTGTCCGGGGCGATGGACGAGCTCCACAACGAGATCCTGGAGCTCGA
 CGAGAAGGTCGACGACCTGCGGGCCGACACCATCAGCTCGCAGATCGAGCT
 CGCGGTGCTGCTCTCCAACGAGGGGCATCATCAACAGCGAGGACGAGCACCT
 GCTCGCCCTGGAGCGGAAGCTCAAGAAGATGCTGGGGCCCTCGGCGGTG
 AGATCGGCAACGGGTGCTTCGAGACCAAGCACAAAGTGCAACCAGACGTGC
 10 CTCGACCGCATCGCCGCGGGCACGTTTCGACGCCGGGGAGTTCTCCCTGCCG
 ACCTTCGACAGCCTCAACATCACGGCGGCCTCGCTGAACGACGACGGCCTC
 GACAACCACACCATCCTGCTCTACTACTCCACCGCGGCCAGCTCGCTGGCG
 GTGACGCTCATGATCGCCATCTTCGTCGTGTACATGGTCTCCCGCGACAACG
 TGAGCTGCTCGATCTGCCTGTGA

15 SEQ ID NO: 89

>Influenza B virus/Yamagata lineage (B/Phuket/3073/2013) HA CDS
 ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGG
 ATCTGCACCGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCC
 ACCCAGGGCGAGGTGAACGTGACCGGCGTGATCCCCCTGACCACCACCCCC
 20 ACCAAGTCCTACTTCGCCAACCTGAAGGGCACCCGGACCCGCGGCAAGCTG
 TGCCCCGACTGCCTGAACTGCACCGACCTGGACGTGGCCCTGGGCGGGCCC
 ATGTGCGTGGGCACCACCCCCAGCGCCAAGGCCTCCATCCTGCACGAGGTG
 CGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCACCAAGATC
 CGGCAGCTGCCAACCTGCTGCGGGGGTACGAGAAGATCCGCCTGTCCACC
 25 CAGAACGTCATCGACGCCGAGAAGGCCCCCCGGCGGGCCCTACCGGCTGGG
 CACCAGCGGGTCCTGCCCCAACGCCACCAGCAAGATCGGCTTCTTCGCCAC
 CATGGCCTGGGCGGTGCCCAAGGACAACCTACAAGAAGGCCACCAACCCCCCT
 GACCGTCGAGGTGCCCTACATCTGCACCGAGGGGGAGGACCAGATCACGGT
 CTGGGGCTTCCACTCCGACAACAAGACCCAGATGAAGTCCCTGTACGGGGA
 30 CTCCAACCCGCAGAAGTTCACGAGCTCCGCCAACGGCGTGACCACGCACTA
 CGTCAGCCAGATCGGGGACTTCCCGGACCAGACCGAGGACGGCGGGCTGC
 CCCAGTCCGGCCGCATCGTGGTCGACTACATGATGCAGAAGCCCGGGAAGA
 CGGGCACCATCGTGTACCAGCGGGGGGTCTGCTGCCGCAGAAGGTGTGGT
 GCGCCAGCGGCCGCTCCAAGGTCATCAAGGGGAGCCTGCCCCCTCATCGGCG
 35 AGGCCGACTGCCTGCACGAGGAGTACGGGGGCCTCAACAAGTCCAAGCCG
 TACTACACGGGGAAGCACGCCAAGGCCATCGGCAACTGCCCGATCTGGGTG

AAGACCCCCCTGAAGCTCGCCAACGGGACGAAGTACCGGCCGCCCCGCAA
 GCTGCTCAAGGAGCGCGGCTTCTTCGGGGCCATCGCCGGCTTCCTGGAGGG
 GGGCTGGGAGGGGATGATCGCCGGCTGGCACGGGTACACCAGCCACGGCG
 CCCACGGGGTTCGCGGTGGCCGCGGACCTCAAGTCCACGCAGGAGGCCATC
 5 AACAAGATCACCAAGAACCTGAACAGCCTCTCCGAGCTGGAGGTCAAGAA
 CCTCCAGCGCCTGAGCGGCGCGATGGACGAGCTCCACAACGAGATCCTGGA
 GCTCGACGAGAAGGTGGACGACCTGCGGGCCGACACGATCTCGTCCCAGAT
 CGAGCTCGCGGTCTGCTCAGCAACGAGGGGATCATCAACTCGGAGGACGA
 GCACCTGCTCGCCCTGGAGCGGAAGCTCAAGAAGATGCTGGGCCCCGTCCGC
 10 GGTGGACATCGGGAACGGCTGCTTCGAGACCAAGCACAAGTGCAACCAGA
 CGTGCCTCGACCGCATCGCCGCGGGGACCTTCAACGCCGGCGAGTTCAGCC
 TGCCACGTTCGACTCGCTCAACATCACCGCGGCCTCCCTGAACGACGACG
 GGCTCGACAACCACACGATCCTGCTCTACTACAGCACCGCGGCCTCGTCCC
 TGGCGGTACGCTCATGCTGGCCATCTTCATCGTGTACATGGTCAGCCGCGA
 15 CAACGTGTCGTGCTCCATCTGCCTCTGA

SEQ ID NO: 90

>5'UTR-Moderna

GGGAAATAAGAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGACCCCCGGCGC
 CGCCACC

20 SEQ ID NO: 91

>3'UTR-Moderna

GCTGGAGCCTCGGTGGCCTAGCTTCTTGCCCCCTGGGCCTCCCCCAGCCCC
 TCCTCCCCTTCCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTG
 GGCGGC

25 SEQ ID NO: 92

>Influenza A virus H1N1 (A/Wisconsin/588/2019) HA Moderna

ATGAAGGCCATCCTGGTCGTGATGCTGTACACCTTCACCACCGCCAACGCCG
 ACACCCTGTGCATCGGCTACCACGCCAACAACAGCACCGACACCGTGGACA
 CCGTGCTGGAGAAGAACGTGACCGTGACCCACAGCGTGAACCTGCTGGAG
 30 GACAAGCACAACGGCAAGCTGTGCAAGCTGAGGGGAGTGGCACCCCTGCA
 CCTGGGCAAGTGCAACATCGCCGGCTGGATCCTGGGCAACCCCGAGTGCGA
 GAGCCTGAGCACAGCCCGGAGCTGGAGCTACATCGTGGAGACCAGCAACA
 GCGACAACGGCACCTGTTACCCCGGCGACTTCATCAACTACGAGGAGCTGC
 GGGAGCAGCTGAGCAGCGTGAGCAGCTTCGAGCGGTTCGAGATCTTCCCCA
 35 AGACCAGCAGCTGGCCCAACCACGACAGCGACAACGGCGTGACAGCAGCC
 TGTCCACACGCCGGAGCCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTG

AAGAAGGGCAAGAGCTACCCCAAGATCAACCAGACCTACATCAACGACAA
 GGGCAAGGAGGTGCTGGTGTGTGGGGCATCCACCACCCACCTACCATCGC
 CGACCAGCAGAGCCTGTACCAGAACGCCGACGCCTACGTGTTCTGTGGGCAC
 CAGCCGGTACAGCAAGAAGTTCAAGCCAGAGATCGCCACCCGGCCCAAGG
 5 TGAGAGACCAGGAGGGCCGGATGAACTACTACTGGACCCTGGTGGAGCCC
 GGAGACAAGATTACCTTCGAGGCCACCGGCAACCTGGTGGCCCCTCGGTAC
 GCCTTCACCATGGAACGGGACGCTGGCAGCGGCATCATCATCAGCGACACT
 CCCGTGCACGACTGCAACACCACCTGCCAGACTCCCGAGGGGCGCTATCAAC
 ACCAGCCTGCCCTTCCAGAACGTGCACCCCATCACCATCGGCAAGTGCCCC
 10 AAGTACGTAAAGAGCACCAAATTGCGGCTGGCCACCGGACTCAGGAACGT
 GCCCAGCATCCAAAGCCGGGGCCTGTTTGGCGCAATCGCCGGCTTCATCGA
 GGGCGGCTGGACTGGCATGGTGGACGGCTGGTACGGCTACCACCACCAGA
 ACGAACAGGGGAGCGGCTACGCAGCTGACCTGAAGAGCACCCAGAACGCC
 ATCGACAAGATCACCAACAAGGTGAACAGCGTGATCGAGAAGATGAACAC
 15 CCAGTTCACCGCCGTGGGCAAGGAGTTCAACCACCTGGAGAAGCGGATCG
 AGAACCTGAACAAGAAGGTGGACGACGGCTTCCTGGACATCTGGACCTAC
 AACGCCGAGCTGCTGGTTCTGCTGGAGAACGAGCGGACCCTGGACTATCAC
 GACAGCAACGTGAAGAACCTGTACGAGAAGGTGCGGAACCAGCTGAAGAA
 CAACGCCAAGGAGATCGGCAACGGCTGCTTCGAGTTCTACCACAAGTGCGA
 20 CAACACCTGCATGGAGAGCGTGAAGAACGGCACCTACGACTACCCCAAGTA
 CAGCGAGGAGGCCAAGCTGAACCGGGAGAAGATCGACGGCGTGAAGCTGG
 ACAGCACCCGGATCTACCAGATCCTGGCCATCTACAGCACCGTGGCCAGCA
 GCCTGGTGTGCTGGTGGTGTGAGCCTGGGCGCCATCAGCTTCTGGATGTGCAGCA
 ACGGCAGCCTGCAGTGCCGGATCTGCATC

25 SEQ ID NO: 93

>Influenza A virus H3N2 (A/Cambodia/e0826360/2020) HA Moderna
 ATGAAGACCATCATCGCCCTGAGCTACATCCTGTGCCTGGTGTTCGCCCAGA
 AGATCCCCGGCAACGATAACAGCACCGCCACCCTGTGTCTGGGACACCACG
 CCGTGCCCAACGGCACCATCGTGAAGACTATCACCAACGACCGGATCGAGG
 30 TGACCAACGCCACCGAGCTGGTGCAGAACAGCAGCATCGGCGAGATCTGC
 GACAGCCCTCACCAGATCCTGGACGGCGGCAACTGCACCCTGATCGACGCA
 CTGCTGGGCGACCCTCAGTGCGACGGCTTTCAGAACAAAGGAGTGGGACCT
 GTTCGTGGAGAGATCGCGGGCCAACAGCAACTGCTACCCCTACGACGTCCC
 CGACTACGCAAGCCTGAGAAGCCTCGTGGCCTCAAGCGGCACCCTGGAGTT
 35 CAAGAACGAGAGCTTCAACTGGACCGGCGTGAAGCAGAACGGCACCTCAA
 GCGCCTGCATCCGGGGCTCCAGCAGCAGCTTCTTCTCACGGCTGAACTGGC

TGACCCACCTGAACTACACCTACCCCGCCCTGAACGTGACCATGCCCCAACA
ACGAGCAGTTTCGACAAGCTGTACATCTGGGGAGTGCACCATCCCAGCACCG
ACAAGGACCAGATTAGCCTGTTTCGCCCAGCCCAGCGGCCGGATCACCGTGA
GCACCAAGCGGAGCCAGCAGGCCGTGATCCCCAACATCGGCTCTCGGCCCA
5 GAATCCGGGACATCCCCAGCCGGATCAGCATCTACTGGACCATTGTGAAGCC
CGGCGACATCCTGCTGATCAACTCCACCGGCAACCTGATCGCCCCCTCGGGG
CTATTTCAAGATCCGGAGCGGCAAGAGCAGCATCATGCGGAGCGACGCCCC
TATCGGCAAGTGCAAGAGCGAGTGCATCACACCCAACGGAAGCATCCCCAA
CGACAAGCCCTTCCAGAACGTGAACCGGATAACCTACGGCGCCTGCCCTAG
10 ATACGTGAAGCAGAGCACCTGAAGCTGGCCACCGGCATGCGGAACGTGCC
CGAGAAGCAGACTCGGGGCATCTTCGGCGCCATCGCCGGCTTCATCGAGAA
CGGCTGGGAGGGCATGGTGGACGGCTGGTACGGCTTCCGGCACCAGAACTC
TGAGGGCAGAGGACAGGCCGCAGACCTGAAGAGCACCCAGGCCGCCATCG
ACCAGATCAACGGCAAGCTGAACCGGCTGATCGGCAAGACCAACGAGAAG
15 TTCCACCAGATCGAGAAGGAGTTCAGCGAGGTGGAGGGCAGGGTACAGGA
CCTGGAGAAGTACGTGGAGGACACCAAGATCGACCTGTGGAGCTACAACG
CCGAGCTGCTGGTAGCCCTGGAGAACCAGCACACCATCGACCTGACCGACA
GCGAGATGAACAAGCTGTTCGAGAAGACCAAGAAGCAGCTGCGGGAGAAC
GCCGAGGACATGGGCAACGGCTGCTTCAAGATCTACCACAAGTGCGACAAC
20 GCCTGCATCGGCAGCATCCGGAACGAGACCTACGACCACAACGTGTACCGG
GACGAGGCCCTGAACAACCGGTTCCAGATCAAGGGCGTGGAGCTGAAGAG
CGGCTACAAGGACTGGATCCTGTGGATCAGCTTCGCCATGTCCTGCTTCCTG
CTGTGCATCGCCCTGCTGGGTTTCATCATGTGGGCCTGCCAGAAGGGCAACA
TCCGGTGCAACATCTGCATC

25 SEQ ID NO: 94

>Influenza B virus Victoria lineage (B/Washington/02/2019) HA CDS Moderna
ATGAAGGCCATCATCGTGCTGTTAATGGTGGTGACCAGCAACGCCGACCGG
ATCTGCACCGGCATCACCTCTAGCAACAGCCCTCACGTGGTGAAGACCGCC
ACACAGGGCGAGGTGAACGTGACCGGCGTGATTCCCCTGACCACCACCCCT
30 ACCAAGAGCCACTTCGCCAACCTGAAGGGAACCGAGACCCGGGGCAAGCT
GTGTCCCAAGTGCCTGAACTGCACCGACCTGGACGTGGCCCTGGGCAGACC
CAAGTGCACCGGCAAGATCCCCAGCGCCCGGGTGTCTATCCTGCACGAAGT
GCGGCCCCGTGACTAGCGGCTGCTTCCCCATCATGCACGACCGGACCAAGAT
CCGGCAGCTGCCCCAACCTGCTGCGGGGCTACGAGCACGTGCGGCTGAGCA
35 CCCACAACGTGATCAACGCCGAAGACGCACCCGGGAGACCATACGAGATCG
GCACCAGCGGCTCTTGCCCCAACATCACCAACGGCAACGGCTTCTTCGCTA

CCATGGCCTGGGCCGTGCCAAAGAACAAGACTGCCACCAACCCTCTGACCA
 TCGAGGTGCCCTACATCTGCACCGAGGGCGAGGACCAGATCACCGTGTGGG
 GCTTCCACAGCGACAGCGAGACCCAGATGGCCAAGCTGTACGGCGACAGC
 AAGCCCCAGAAGTTCACCAGCAGCGCCAACGGCGTGACCACCCACTACGT
 5 GAGCCAGATCGGCGGCTTCCCCAACCAGACCGAGGACGGCGGCTTACCCCA
 GAGCGGCCGGATCGTGGTGGACTACATGGTGCAGAAGAGCGGCAAGACCG
 GCACCATCACCTACCAGCGGGGCATCCTGCTGCCACAGAAGGTGTGGTGCG
 CCTCAGGGCGGTCAAAGGTGATCAAGGGCAGCCTGCCACTGATTGGCGAGG
 CCGACTGCCTGCACGAGAAGTACGGCGGCCTGAACAAGAGCAAGCCCTACT
 10 ACACCGGCGAGCACGCCAAGGCAATCGGCAACTGCCCCATCTGGGTGAAG
 ACACCCCTGAAGCTGGCCAACGGCACCAAGTACCGGCCACCCGCCAACT
 GCTGAAGGAGCGGGGCTTCTTCGGCGCCATTGCCGGCTTCCTCGAAGGCGG
 TTGGGAGGGCATGATCGCCGGCTGGCACGGCTACACTAGCCACGGCGCACA
 CGGAGTAGCAGTGGCCGCCGACCTGAAGAGCACCCAGGAGGCCATCAACA
 15 AGATCACCAAGAACCTGAACAGCCTGAGCGAGCTGGAGGTGAAGAATCTG
 CAGCGGCTGTCTGGCGCTATGGACGAGCTGCACAACGAGATCCTGGAGCTG
 GACGAGAAGGTGGACGACTTACGGGGCCGACACCATCAGCAGCCAGATCGA
 GCTGGCCGTGCTGCTCTCCAACGAGGGGCATCATCAACAGCGAGGACGAGCA
 CCTGCTCGCCCTGGAGCGGAAGCTGAAGAAGATGCTGGGGCCCTTCTGCCGT
 20 GGAGATCGGTAACGGCTGCTTCGAGACCAAGCACAAGTGCAACCAGACCT
 GCCTGGATCGGATCGCAGCCGGCACCTTTGACGCCGGGGAGTTCAGCCTGC
 CCACCTTCGACAGCCTGAACATCACCGCCGCCAGCCTGAACGACGACGGCC
 TGGACAACCACACCATCCTGCTGTACTACTCTACAGCCGCTAGCAGCCTGGC
 CGTGACCCTGATGATCGCCATCTTCGTGGTGTACATGGTGAGCCGGGACAAC
 25 GTGAGCTGCAGCATCTGCCTG

SEQ ID NO: 95

>Influenza B virus Yamagata lineage (B/Phuket/3073/2013) HA CDS Moderna
 ATGAAGGCCATCATCGTGCTACTGATGGTGGTGACCAGCAACGCCGACCGG
 ATCTGCACCGGCATCACCAGCAGCAACAGCCCGCACGTGGTGAAGACCGCC
 30 ACCCAAGGCGAGGTGAACGTGACCGGCGTGATCCCACTGACCACCACTCCC
 ACCAAGAGCTACTTCGCCAACCTGAAGGGCACACGGACTCGGGGCAAGCT
 GTGCCCCGACTGCCTGAACTGCACCGACCTGGACGTGGCCCTGGGCAGACC
 CATGTGCGTGGGCACCAACCCTTCTGCCAAGGCCAGCATCCTGCACGAGGT
 GAGACCCGTGACCAGCGGGTGCTTCCCCATCATGCACGACCGGACCAAGAT
 35 CCGGCAGCTGCCCAACCTGCTGCGGGGCTACGAGAAGATCCGGCTGAGCAC
 CCAGAACGTGATCGACGCCGAGAAGGCCCTGGAGGTCCCTACCGGCTGG

GCACCAGCGGAAGCTGCCCCAACGCCACCAGCAAGATCGGCTTCTTCGCCA
 CCATGGCCTGGGCTGTGCCCAAGGACAACACTACAAGAACGCCACCAATCCCC
 TGACCGTGGAGGTGCCCTACATCTGCACCGAGGGCGAGGACCAGATCACCG
 TGTGGGGCTTCCACAGCGACAACAAGACCCAGATGAAGAGCCTGTACGGC
 5 GACAGCAATCCCCAGAAGTTCACAAGCAGCGCCAACGGCGTGACCACCCA
 CTACGTGAGCCAGATCGGCGACTTCCCCGACCAGACCGAGGACGGAGGGC
 TGCCTCAGAGTGGCCGGATCGTGGTGGACTACATGATGCAGAAGCCCGGCA
 AGACCGGCACCATCGTGTACCAGCGGGGCGTGCTGTTGCCTCAGAAAGTTT
 GGTGTGCCAGCGGCAGGAGCAAGGTGATCAAGGGCAGCCTGCCCCTGATC
 10 GCGGAGGCAGACTGCCTCCACGAGGAGTACGGCGGCCTGAACAAGAGCAA
 GCCCTACTACACCGGCAAGCACGCCAAGGCCATCGGCAACTGCCCCATCTG
 GGTGAAGACCCCTCTGAAGCTGGCCAACGGCACCAAGTACCGGCCACCAG
 CCAAGCTGCTGAAGGAGCGGGGCTTCTTTGGCGCCATTGCCGGCTTCCTCG
 AGGGAGGCTGGGAGGGCATGATCGCCGGCTGGCACGGCTACACAAGCCAC
 15 GGCGCACACGGAGTGGCTGTGGCTGCCGACCTGAAGAGCACCCAGGAGGC
 CATCAACAAGATCACCAAGAACCTGAACAGCCTGAGCGAGCTGGAGGTGA
 AGAACCTGCAGCGGCTGTCAGGCGCCATGGACGAGCTGCACAACGAGATC
 CTGGAGCTGGACGAGAAGGTGGACGACCTGCGTGCCGACACCATCAGCAG
 CCAGATCGAGCTGGCCGTGCTGCTGAGCAACGAGGGCATCATCAACAGCGA
 20 GGACGAGCACCTGCTGGCCCTGGAGCGGAAACTGAAGAAGATGCTGGGAC
 CCTCTGCCGTGGACATCGGCAACGGCTGCTTCGAGACCAAGCACAAGTGCA
 ACCAGACCTGCCTGGATCGGATCGCCGCCGGAACCTTCAACGCCGGCGAGT
 TCAGCCTGCCACCTTCGACAGCCTGAACATCACCGCCGCCAGCCTGAACG
 ACGACGGCCTGGACAACCACACCATCCTGCTGTACTACAGCACTGCCGCCT
 25 CAAGCCTGGCCGTGACCCTGATGCTGGCCATCTTCATCGTGTACATGGTGAG
 CCGGGACAACGTGAGCTGCAGCATCTGCCTG

SEQ ID NO: 96

> SARS-CoV-2 BA.4 spike protein CDS

ATGTTTCGTGTTCTTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACC
 30 TGATCACAAGAACACAGTCGTACACCAACAGCTTCACAAGAGGCGTGTACT
 ACCCCGACAAGGTGTTTCTAGAAGCAGCGTGCTGCACTCCACCCAAGACCTGT
 TTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCAGCGGCACCAA
 CGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTCAACGACGGCGTGTA
 CTTTCGCTAGCACCGAAAAGAGCAACATCATCAGAGGCTGGATCTTCGGCAC
 35 CACCCTAGACTCCAAGACACAAAGCCTGCTGATCGTTAACAACGCGACCAA
 CGTGGTGATCAAGGTGTGCGAGTTTTCAGTTCTGCAACGACCCCTTCCTGGA

CGTGTACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCAGAG
 TCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCT
 GATGGACCTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTCG
 TGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCCA
 5 TCAACCTGGGGAGAGACCTGCCCCAAGGCTTCAGCGCCCTGGAGCCCCTGG
 TGGACCTGCCCATCGGCATCAACATCACAAGATTTAGACCCTGCTGGCCCT
 GCACCGTAGCTACCTAACCCCCGGCGATAGTTCCCTCTGGCTGGACGGCCGG
 CGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAG
 TACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCC
 10 CTGAGCGAGACCAAGTGCACCCTCAAGAGCTTCACCGTGGAGAAGGGCAT
 CTATCAGACAAGCAACTTCAGAGTGCAGCCCACCGAGAGCATCGTGAGATT
 CCCCACATCACCAACCTGTGCCCTTCGACGAGGTGTTCAACGCCACAAG
 ATTCGCTAGCGTGTACGCGTGGAACAGAAAGAGAATCAGCAACTGCGTGGC
 CGACTACAGCGTGCTGTACAACTTCGCTCCCTTCTTCGCCTTCAAGTGCTAT
 15 GGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCC
 GACAGCTTCGTGATCAGAGGCAACGAGGTGAGCCAGATTGCCCCCGGGCA
 AACC GGCAACATCGCCGACTACA ACTACAAGCTGCCCGACGACTTCACCGG
 CTGCGTGATCGCGTGGAACAGCAATAAGCTGGATTCCAAGGTGGGCGGCAA
 CTACA ACTACCGGTACAGACTGTT CAGAAAGAGCAACCTGAAGCCCTTCGA
 20 GAGAGACATCAGCACCAGATCTACCAAGCCGGCAACAAGCCCTGCAACG
 GCGTGGCGGGCGTCAACTGCTACTTCCCCCTGCAGAGCTACGGCTTTAGGC
 CCACCTACGGCGTGGGCCATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCG
 AGCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCAAAAAGAGCACCAAC
 CTGGTGAAGAACAAGTGCGTGA ACTTCAACTTTAATGGCCTTAAGGGCACG
 25 GGCGTGCTGACCGAGAGCAACAAGAAGTTCCTGCCATTTAGCAATTCGGC
 AGAGACATCGCCGACACCACCGACGCGGTCAGAGACCCTCAGACCCTGGA
 GATCCTGGACATCACCCCTGTAGCTTCGGCGGCGTGAGCGTGATCACCCCC
 GGC ACTAACACAAGCAACCAAGTGGCCGTGCTGTACCAAGGCGTGA ACTG
 CACCGAGGTGCCCGTGGCCATCCACGCCGATCAGCTGACCCCCACCTGGAG
 30 AGTATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGAT
 CGGCGCCGAGTACGTGAACAACAGCTACGAGTGCGACATCCCCATCGGCGC
 CGGCATCTGCGCTAGCTATCAGACACAGACCAAGAGCCACCGGAGAGCTAG
 AAGCGTGGCTTCTCAGAGCATCATCGCCTACACCATGAGCCTGGGCGCCGA
 GAACAGCGTGGCCTACAGCAACAACAGCATCGCCATCCCCACCAACTTCAC
 35 CATCAGCGTGACCACCGAGATCCTGCCCCGTAAGTATGACCAAGACAAGCGT
 GGACTGCACCATGTACATCTGCGGCGACAGCACCAGGTGCAGCAACCTGCT

GCTGCAGTACGGCAGCTTCTGCACACAGCTGAAGAGAGCCCTGACCGGCAT
 CGCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTCCGCCAAGTGAAGC
 AGATCTACAAGACCCCCCCCATCAAGTACTTCGGCGGCTTCAACTTCAGCCA
 AATCCTGCCCCGACCCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCT
 5 GCTGTTCAACAAGGTGACCCTGGCCGACGCCGGCTTCATCAAGCAGTACGG
 CGACTGCCTAGGCGACATCGCAGCTAGAGACCTGATCTGCGCTCAGAAGTT
 CAATGGACTCACCGTGCTCCCCCCCCCTGCTGACCGACGAGATGATCGCTCA
 GTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATTTGG
 CGCCGGGGGCCGCACTGCAAATCCCCTTCGCCATGCAGATGGCCTACAGATTC
 10 AACGGCATCGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATC
 GCCAATCAGTTCAACAGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGC
 ACCGCTAGCGCCCTGGGCAAGCTGCAAGACGTGGTGAATCACAACGCCCA
 AGCCCTGAACACCCTGGTGAAGCAGCTGAGCAGCAAGTTCGGCGCCATCA
 GCAGCGTATTGAATGACATCCTGTCTAGACTGGACCCCCCGAGGCCGAGG
 15 TGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCAGACCTACG
 TGACACAGCAGCTGATCAGAGCCGCCGAGATCAGAGCTAGCGCCAACCTGG
 CCGCCACCAAGATGAGCGAGTGCGTGCTGGGGCAGAGCAAGAGAGTGGAC
 TTCTGCGGCAAGGGCTACCACCTGATGAGCTTCCCTCAGAGCGCCCCCAC
 GGCGTGGTGTTCCTGCACGTGACCTACGTGCCCCGCCAAGAGAAGAACTTC
 20 ACCACCGCTCCCGCCATCTGCCACGACGGCAAGGCCCACTTCCCTAGAGAG
 GGCGTGTTCGTGAGCAACGGCACCCACTGGTTCGTGACACAGAGAACTTC
 TACGAGCCTCAGATCATCACCACCGACAACACCTTCGTGAGCGGCAACTGC
 GACGTGGTGATCGGCATAGTGAACAACACCGTATACGACCCCCTGCAGCCC
 GAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACT
 25 AGCCCCGACGTGGACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTG
 AACATTCAGAAGGAGATCGACCGGCTGAATGAGGTGGCCAAGAACCTGAA
 CGAGAGCCTGATCGACCTGCAAGAGCTGGGCAAGTACGAGCAGTACATCAA
 GTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCTGATCGCCATCGTG
 ATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGCAGCTGCCTGAAAG
 30 GTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGACAGCGAGC
 CCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 97

>Influenza A virus H1N1 HA0 mutant amino acid sequence

MKAILVVMlyTFTTANADTLcIGYHANNSTDTVDTVLEKNVTVTHSVNLLED
 35 KHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTARSWSYIVETSNSDNG
 TCYPGDFINYEELREQLSSVSSFERFEIFPKTSSWPNHDS DNGVTAACPHAGAK

SFYKNLIWLVLKKGKSYPKINQTYINDKGKEVLVLWGIHHPPTIADQQSLYQNA
DAYVFGVTSRYSKKFKPEIATRPKVRDQEGRMNYYWTLVEPGDKITFEATGNL
VAPRYAFTMERDAGSGIIISDTPVHDCNTTCQTPEGAINSTLPPFQNVHPITIGKCP
KYVKSTKLRLATGLRNVPSIQSGLFGAIAAGFIEGGWTGMVDGWYGYHHQNE
5 QGSGYAADLKSTQNAIDKITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNK
KVDDGFLDIWTYNAELLVLENERTLDYHDSNVKNLYEKVRNQLKNNAKEIG
NGCFEFYHKCDNTCMESVKNGTYPKYSEEAKLNREKIDGVKLDSTRIYQIL
AIYSTVASSLVLVVSLGAISFWMCSNGSLQCRICI

SEQ ID NO: 98

10 >Influenza A virus H3N2 HA0 mutant amino acid sequence
MKTIIALSYILCLVFAQKIPGNDNSTATLCLGHHAVPNGTIVKTITNDRIEVTNAT
ELVQNSSIGEICDSPHQILDGGNCTLIDALLGDPQCDGFQNKEDLDFVERSRAN
SNCYPYDVPDYASLRSLVASSGTLEFKNESFNWTGVKQNGTSSACIRGSSSSFFS
RLNWLTHLNYTYPALNVTMPNNEQFDKLYIWGVHHPSTDKDQISLFAQPSGRI
15 TVSTKRSQQAVIPNIGSRPRIRDIPSRSIYWTVKPGDILLINSTGNLIAPRGYFKI
RSGKSSIMRSDAPIGKCKSECITPNGSIPNDKPFQNVNRITYGACPRYVKQSTLK
LATGMRNVPEKQTQGIFGAIAAGFIENGWEGMVDGWYGFRHQNSEGRGQAAD
LKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRVQDLEKYVEDTKIDL
WSYNAELLVALENQHTIDLTSEMKNLFEKTKKQLRENAEDMGNGCFKIYHK
20 CDNACIGSIRNETYDHNVYRDEALNNRFQIKGVELKSGYKDWILWISFAMSCF
LLCIIALLGFIMWACQKGNIRCNICI

SEQ ID NO: 99

>Influenza B virus/Victoria HA0 mutant amino acid sequence
MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATQGEVNVTVGIPLTTTPTKSH
25 FANLKGTTETRGKLCPCLNCTDLDVALGRPCKTGKIPSARVSILHEVRPVTSGC
FPIMHDKIRQLPNLLRGYEHVRLSTHNVINAEDAPGRPYEIGTSGSCPNTNG
NGFFATMAWAVPKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLY
GDSKPQKFTSSANGVTTHYVSQIGGFNPQTEDGGLPQSGRIVVDYMVQKSGKT
GTITYQRGILLPQKVWCASGRSKVIKGSPLIGEADCLHEKYGGLNKSCKPYT
30 GEHAKAIGNCPIWVKTPKLKLANGTKYRPPAKLLKEQGFFGAIAAGFLEGGWEG
MIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGA
MDELHNEILELDEKVDDLADTSSQIELAVLLSNEGIINSEDEHLLALERKLLK
MLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNTAASLN
DDGLDNHTILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSICL

35 SEQ ID NO: 100

>Influenza B virus/Yamagata HA0 mutant amino acid sequence

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATATQGEVNVTVGIPLTTTPTKSY
 FANLKGTRTRGKLCPCDCLNCTDLDVALGRPMC VGTTTPSAKASILHEVRPVTSG
 CFPIMHDRTKIRQLPNLLRGYEKIRLSTQNVIDAEKAPGGPYRLGTSGSCP NATS
 KIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMK
 5 SLYGDSNPQKFTSSANGVTTHYVSQIGDFPDQTEDGGLPQSGRIVVDYMMQKP
 GKTGTIVYQRGVLLPQKVWCASGRSKVIKGSPLIGEADCLHEEYGGLNKSKP
 YYTGKHAKAIGNCPIWVKTPCLKLANGTKYRPPAKLLKEQGFFGAIA GFLEGG
 WEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQR
 LSGAMDELHNEILELDEKVDDL RADTISSQIELAVLLSNEGIINSEDEHLLALER
 10 KLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSL NITA
 ASLNDDGLDNHTILLYYSTAASSLAVTLM LAIFIVYMVSRDNVSCSICL

SEQ ID NO: 101

>Influenza A virus H1N1 HA0 mutant CDS

ATGAAGGCCATCCTGGTGGTGATGCTGTACACCTTCACCACCGCCAACGCC
 15 GACACCCTGTGCATCGGCTACCACGCCAACACAGCACCGACACCGTGGAC
 ACCGTGCTGGAGAAGAACGTGACCGTGACCCACAGCGTGAACCTGCTGGA
 GGACAAGCACAACGGCAAGCTGTGCAAGCTGCGGGGCGTGGCCCCCCTGC
 ACCTGGGCAAGTGCAACATCGCCGGCTGGATCCTGGGCAACCCCGAGTGCG
 AGTCCCTGAGCACCGCCCGCTCCTGGAGCTACATCGTGGAGACCTCCAACA
 20 GCGACAACGGCACCTGCTACCCCGGGGACTTCATCAACTACGAGGAGCTGC
 GGGAGCAGCTGTCCAGCGTGTCCAGCTTCGAGCGCTTCGAGATCTTCCCCA
 AGACCTCCAGCTGGCCCAACCACGACTCCGACAACGGCGTGACCGCCGCC
 TGCCCCCACGCCGGGGCCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTG
 AAGAAGGGCAAGTCCTACCCCAAGATCAACCAGACCTACATCAACGACAA
 25 GGGGAAGGAGGTCCTCGTGCTGTGGGGCATCCACCACCCCCCACCATCGC
 CGACCAGCAGAGCCTCTACCAGAACGCCGACGCCTACGTCTTCGTGGGGAC
 CTCCCGGTACAGCAAGAAGTTCAAGCCCGAGATCGCCACCCGCCCCAAGGT
 CCGGGACCAGGAGGGCCGCATGA ACTACTACTGGACCCTGGTGGAGCCGG
 GGGACAAGATCACCTTCGAGGCCACCGGCAACCTCGTCGCCCCGCGGTACG
 30 CCTTCACGATGGAGCGCGACGCCGGGTCCGGCATCATCATCAGCGACACCC
 CCGTGACGACTGCAACACGACCTGCCAGACGCCCGAGGGGGCGATCAAC
 ACCTCCCTGCCGTTCCAGAACGTCCACCCGATCACGATCGGCAAGTGCCCC
 AAGTACGTGAAGTCCACCAAGCTCCGGCTGGCCACGGGGCTCCGCAACGTC
 CCCTCCATCCAGAGCCAGGGCCTGTTCCGGGGCGATCGCCGGCTTCATCGAG
 35 GGGGGCTGGACCGGGATGGTGGACGGCTGGTACGGGTACCACCACCAGAA
 CGAGCAGGGCTCCGGGTACGCGGCCGACCTCAAGTCGACGCAGAACGCGA

TCGACAAGATCACCAACAAGGTCAACAGCGTGATCGAGAAGATGAACACG
CAGTTCACCGCCGTCGGCAAGGAGTTCAACCACCTGGAGAAGCGCATCGA
GAACCTCAACAAGAAGGTGGACGACGGGTTCTGGACATCTGGACGTACA
ACGCGGAGCTCCTGGTCCTCCTGGAGAACGAGCGGACCCTCGACTACCACG
5 ACTCCAACGTGAAGAACCTGTACGAGAAGGTCCGCAACCAGCTCAAGAAC
AACGCCAAGGAGATCGGCAACGGGTGCTTCGAGTTCTACCACAAGTGCGAC
AACACGTGCATGGAGTCGGTGAAGAACGGCACCTACGACTACCCGAAGTAC
AGCGAGGAGGCGAAGCTGAACCGGGAGAAGATCGACGGGGTCAAGCTCGA
CTCCACGCGCATCTACCAGATCCTGGCCATCTACTCGACCGTGGCGAGCTCC
10 CTCGTCTGGTGGTCTCGCTCGGCGCCATCAGCTTCTGGATGTGCTCCAACG
GGTCGCTGCAGTGCCGGATCTGCATCTGA

SEQ ID NO: 102

>Influenza A virus H3N2 HA0 mutant CDS

ATGAAGACCATCATCGCCCTGAGCTACATCCTGTGCCTGGTGTTCGCCCAGA
15 AGATCCCCGGCAACGACAACAGCACCGCCACCCTGTGCCTGGGCCACCAC
GCCGTGCCCAACGGCACCATCGTGAAGACCATCACCAACGACCGGATCGAG
GTGACCAACGCCACCGAGCTGGTGCAGAACTCCAGCATCGGCGAGATCTGC
GACTCCCCCACCAGATCCTGGACGGCGGCAACTGCACCCTGATCGACGCC
CTGCTGGGCGACCCCCAGTGCGACGGGTTCCAGAACAAGGAGTGGGACCT
20 GTTCGTGGAGCGGAGCCGCGCCAACCTCAACTGCTACCCCTACGACGTGCC
CGACTACGCCAGCCTGCGGTCCCTGGTGGCCAGCTCCGGCACCCCTGGAGTT
CAAGAACGAGAGCTTCAACTGGACCGGGGTCAAGCAGAACGGCACCTCCA
GCGCCTGCATCCGCGGGTCCAGCTCCAGCTTCTTCTCCCGGCTGAACTGGCT
CACCCACCTGAACTACACCTACCCCGCCCTCAACGTGACCATGCCCAACAA
25 CGAGCAGTTCGACAAGCTGTACATCTGGGGCGTCCACCACCCAGCACCGA
CAAGGACCAGATCTCCCTCTTCGCCCAGCCCAGCGGGCGCATCACCGTGTC
CACGAAGCGGAGCCAGCAGGCCGTCATCCCCAACATCGGCTCCCGCCCGCG
GATCCGCGACATCCCCAGCCGGATCTCCATCTACTGGACGATCGTGAAGCCG
GGGGACATCCTGCTCATCAACAGCACCGGCAACCTGATCGCCCCCGCGGG
30 TACTTCAAGATCCGGTCCGGCAAGAGCTCGATCATGCGCTCCGACGCCCCGA
TCGGGAAGTGCAAGAGCGAGTGCATACCCCCAACGGCTCGATCCCGAACG
ACAAGCCCTTCCAGAACGTCAACCGGATCACGTACGGGGCCTGCCCCGCGCT
ACGTGAAGCAGTCCACGCTCAAGCTGGCCACCGGCATGCGGAACGTCCCCG
AGAAGCAGACCCAGGGGATCTTCGGCGCCATCGCGGGGTTCATCGAGAACG
35 GCTGGGAGGGGATGGTGGACGGCTGGTACGGGTTCGGGCACCAGAACAGC
GAGGGCCGCGGGCAGGCCGCGGACCTCAAGTCGACGCAGGCCGCGATCGA

CCAGATCAACGGCAAGCTGAACCGGCTCATCGGGAAGACGAACGAGAAGT
 TCCACCAGATCGAGAAGGAGTTCTCCGAGGTCGAGGGCCGCGTGCAGGAC
 CTGGAGAAGTACGTCGAGGACACCAAGATCGACCTCTGGAGCTACAACGCC
 GAGCTGCTCGTGGCGCTGGAGAACCAGCACACCATCGACCTCACGGACTCG
 5 GAGATGAACAAGCTGTTCGAGAAGACGAAGAAGCAGCTCCGGGAGAACGC
 CGAGGACATGGGGAACGGCTGCTTCAAGATCTACCACAAGTGCGACAACGC
 GTGCATCGGGTCCATCCGCAACGAGACCTACGACCACAACGTCTACCGGGA
 CGAGGCCCTGAACAACCGCTTCCAGATCAAGGGCGTGGAGCTCAAGAGCG
 GGTACAAGGACTGGATCCTGTGGATCTCGTTCGCGATGTCCTGCTTCCTCCT
 10 GTGCATCGCCCTCCTGGGCTTCATCATGTGGGCGTGCCAGAAGGGGAACAT
 CCGGTGCAACATCTGCATCTGA

SEQ ID NO: 103

>Influenza B virus/Victoria HA0 mutant CDS

ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGG
 15 ATCTGCACCGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCC
 ACCCAGGGCGAGGTGAACGTGACCGGCGTGATCCCCCTGACCACCACCCCC
 ACCAAGTCCCCTTCGCCAACCTGAAGGGCACCGAGACCCGGGGCAAGCT
 GTGCCCCAAGTGCTGAACCTGCACCGACCTGGACGTGGCCCTGGGCGCCG
 CAAGTGCACCGGCAAGATCCCCAGCGCCCGGGTGTCCATCCTGCACGAGGT
 20 GCGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCACCAAGAT
 CCGGCAGCTGCCAACCTGCTGCGGGGGTACGAGCACGTCCGCCTGTCCAC
 CCACAACGTGATCAACGCCGAGGACGCCCCCGGCCGGCCCTACGAGATCGG
 GACCAGCGGCTCCTGCCCCAACATCACCACGGGAACGGCTTCTTCGCCAC
 CATGGCCTGGGCGTCCCCAAGAACAAGACCGCCACCAACCCCCTGACCAT
 25 CGAGGTGCCGTACATCTGCACCGAGGGGGAGGACCAGATCACCCTGTGGGG
 CTTCCACAGCGACAACGAGACGCAGATGGCCAAGCTGTACGGGGACTCCA
 AGCCCCAGAAGTTCACGAGCTCCGCCAACGGCGTGACCACCCACTACGTCA
 GCCAGATCGGGGGCTTCCCGAACCAGACGGAGGACGGGGGCCTGCCCCAG
 TCCGGGCGCATCGTGGTCGACTACATGGTGCAGAAGTCCGGCAAGACGGGG
 30 ACCATCACGTACCAGCGGGGCATCCTGCTGCCGCAGAAGGTCTGGTGCGCC
 TCCGGGCGCAGCAAGGTGATCAAGGGCTCCCTGCCCCTCATCGGGGAGGCC
 GACTGCCTGCACGAGAAGTACGGCGGGCTCAACAAGAGCAAGCCGTACTA
 CACCGGCGAGCACGCCAAGGCCATCGGGAACCTGCCCCATCTGGGTCAAGAC
 CCCGCTGAAGCTCGCCAACGGCACGAAGTACCGGCCCCCGGCCAAGCTGCT
 35 CAAGGAGCAGGGGTTCTTCGGCGCCATCGCCGGGTTCTTGGAGGGCGGGTG
 GGAGGGCATGATCGCCGGGTGGCACGGCTACACGTCCCACGGGGGCCACG

GCGTGGCGGTGCGCCGCGGACCTCAAGAGCACCCAGGAGGCCATCAACAAG
ATCACGAAGAACCTGAACTCCCTCAGCGAGCTGGAGGTGAAGAACCTCCA
GCGCCTGTCCGGGGCGATGGACGAGCTCCACAACGAGATCCTGGAGCTCGA
CGAGAAGGTGACGACCTGCGGGCCGACACCATCAGCTCGCAGATCGAGCT
5 CGCGGTGCTGCTCTCCAACGAGGGGCATCATCAACAGCGAGGACGAGCACCT
GCTCGCCCTGGAGCGGAAGCTCAAGAAGATGCTGGGGCCCTCGGCGGTGCG
AGATCGGCAACGGGTGCTTCGAGACCAAGCACAAAGTGCAACCAGACGTGC
CTCGACCGCATCGCCGCGGGCACGTTCGACGCCGGGGAGTTCTCCCTGCCG
ACCTTCGACAGCCTCAACATCACGGCGGCCTCGCTGAACGACGACGGCCTC
10 GACAACCACACCATCCTGCTCTACTACTCCACCGCGGCCAGCTCGCTGGCG
GTGACGCTCATGATCGCCATCTTCGTCGTGTACATGGTCTCCCGCGACAACG
TGAGCTGCTCGATCTGCCTGTGA

SEQ ID NO: 104

>Influenza B virus/Yamagata HA0 mutant CDS

15 ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGG
ATCTGCACCGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCC
ACCCAGGGCGAGGTGAACGTGACCGGCGTGATCCCCCTGACCACCACCCCC
ACCAAGTCCTACTTCGCCAACCTGAAGGGCACCCGGACCCGCGGCAAGCTG
TGCCCCGACTGCCTGAACTGCACCGACCTGGACGTGGCCCTGGGCGGGCCC
20 ATGTGCGTGGGCACCACCCCCAGCGCCAAGGCCTCCATCCTGCACGAGGTG
CGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCACCAAGATC
CGGCAGCTGCCCCAACCTGCTGCGGGGGTACGAGAAGATCCGCCTGTCCACC
CAGAACGTCATCGACGCCGAGAAGGCCCCCCGGCGGGCCCTACCGGCTGGG
CACCAGCGGGTCCTGCCCCAACGCCACCAGCAAGATCGGCTTCTTCGCCAC
25 CATGGCCTGGGCGTGCCCCAAGGACAACCTACAAGAACGCCACCAACCCCCT
GACCGTCGAGGTGCCCTACATCTGCACCGAGGGGGAGGACCAGATCACGGT
CTGGGGCTTCCACTCCGACAACAAGACCCAGATGAAGTCCCTGTACGGGGA
CTCCAACCCGCAGAAGTTCACGAGCTCCGCCAACGGCGTGACCACGCACTA
CGTCAGCCAGATCGGGGACTTCCCGGACCAGACCGAGGACGGCGGGCTGC
30 CCCAGTCCGGCCGCATCGTGGTCTGACTACATGATGCAGAAGCCCGGGAAGA
CGGGCACCATCGTGTACCAGCGGGGGGTCTGCTGCCGCAGAAGGTGTGGT
GCGCCAGCGGCCGCTCCAAGGTCATCAAGGGGAGCCTGCCCCCTCATCGGCG
AGGCCGACTGCCTGCACGAGGAGTACGGGGGCCTCAACAAGTCCAAGCCG
TACTACACGGGGAAGCACGCCAAGGCCATCGGCAACTGCCCCGATCTGGGTG
35 AAGACCCCCCTGAAGCTCGCCAACGGGACGAAGTACCGGCCGCCCCGCCAA
GCTGCTCAAGGAGCAGGGCTTCTTCGGGGCCATCGCCGGCTTCTTGGAGGG

GGGCTGGGAGGGGATGATCGCCGGCTGGCACGGGTACACCAGCCACGGCG
 CCCACGGGGTTCGCGGTGGCCGCGGACCTCAAGTCCACGCAGGAGGCCATC
 AACAAGATCACCAAGAACCTGAACAGCCTCTCCGAGCTGGAGGTCAAGAA
 CCTCCAGCGCCTGAGCGGCGCGATGGACGAGCTCCACAACGAGATCCTGGA
 5 GCTCGACGAGAAGGTGGACGACCTGCGGGCCGACACGATCTCGTCCCAGAT
 CGAGCTCGCGGTCTTGTCTCAGCAACGAGGGGATCATCAACTCGGAGGACGA
 GCACCTGCTCGCCCTGGAGCGGAAGCTCAAGAAGATGCTGGGCCCCGTCCGC
 GGTGGACATCGGGAACGGCTGCTTCGAGACCAAGCACAAGTGCAACCAGA
 CGTGCTCGACCGCATCGCCGCGGGGACCTTCAACGCCGGCGAGTTCAGCC
 10 TGCCCACGTTCGACTCGCTCAACATCACCGCGGCCTCCCTGAACGACGACG
 GGCTCGACAACCACACGATCCTGCTCTACTACAGCACCGCGGCCTCGTCCC
 TGGCGGTACGCTCATGCTGGCCATCTTCATCGTGTACATGGTCAGCCGCGA
 CAACGTGTCGTGCTCCATCTGCCTCTGA

SEQ ID NO: 105

15 >I'-Delta24-B-120A mRNA
 GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAG
 GAUAGCACUAUAUAUGUUCGUGUUCUGGGUGCUGCUGCCCCUGGUGAG
 CUCUCAGUGCGUGAACCUGAGGACAAGAACACAGCUGCCCCCGCCUACA
 CCAACAGCUUCACAAGAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGC
 20 AGCGUGCUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGU
 GACCUGGUUCCACGCCAUCCACGUCAGCGGCACCAACGGCACCAAGAGAU
 UCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACC
 GAAAAGAGCAACAUAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUC
 CAAGACACAAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUA
 25 AGGUGUGCGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACGUGUACUAC
 CACAAGAACAACAAGAGCUGGAUGGAGAGCGGGGUCUACAGCAGCGCCA
 ACAACUGCACCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAG
 GGCAAGCAAGGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUAAGAACA
 UCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCAUCAACCUGGUG
 30 AGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCC
 CAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUA
 GCUACCUAACCCCGGCGAUAGUUCUUCUGGCUGGACGGCCGGCGCUGCC
 GCCUACUACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAA
 CGAGAACGGCACUAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCUGA
 35 GCGAGACCAAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAU
 CAGACAAGCAACUUCAGAGUGCAGCCACCGAGAGCAUCGUGAGAUUCCC
 CAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUAACGCCACAAGAU
 UCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGC
 CGACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCU

AUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUAC
 GCCGACAGCUUCGUGAUCAGAGGCGACGAGGUGAGACAGAUUGCCCCCGG
 GCAAACCGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUA
 CCGGCUUGCGUGAUCGCGUGGAACAGCAAUAACCUGGAUUCCAAGGUGGGC
 5 GGCAACUACAACUACCGGUACAGACUGUUCAGAAAGAGCAACCUGAAGCC
 CUUCGAGAGAGACAUCAGCACCGAGAUCUACCAAGCCGGCAGCAAGCCCU
 GCAACGGCGUGGAGGGCUUCAAACUGCUACUUCCCCCUGCAGAGCUACGGC
 UUUCAGCCCACCAACGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCU
 GAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCCAAAAAGA
 10 GCACCAACCUGGUGAAGAACAAGUGCGUGAACUUAACUUUAAUGGCCU
 UACCGGCACGGGCGUGCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUC
 AGCAAUUCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCU
 CAGACCCUGGAGAUCUGGACAUCACCCCCUGUAGCUUCGGCGGGCGUGAG
 CGUGAUCACCCCCGGCACUAACACAAGCAACCAAGUGGGCCGUGCUGUACC
 15 AAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUG
 ACCCCCACCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAG
 AGCCGGCUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGUGCG
 ACAUCCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAC
 AGCCGCCGGAGAGCUAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACAC
 20 CAUGAGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCG
 CCAUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUCUGCCCGUA
 AGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGGGCGACAG
 CACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGC
 UGAACAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAA
 25 GAGGUGUUCGCCCAAGUGAAGCAGAUCUACAAGACCCCCCCCCAUAAGGA
 CUUCGGCGGCUUCAACUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUA
 GCAAGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCC
 GACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGC
 UAGAGACCUGAUCUGCGCUCAGAAGUUAUAUGGACUCACCGUGCUCCCCC
 30 CCCUGCUGACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGGC
 GGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAU
 CCCCUCGCCAUGCAGAUGGCCUACAGAUUAACGGCAUCGGCGUGACAC
 AGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAG
 CGCCAUCGGCAAGAUAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGG
 35 GCAAGCUGCAAAACGUGGUGAAUCAGAACGCCCAAGCCCUGAACACCCUG
 GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGCGUAUUGAAUG
 ACAUCCUGUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGA
 CUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCU
 GAUCAGAGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGA
 40 UGAGCGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAA
 GGGCUACCACCUGAUGAGCUUCCUCAGAGCGCCCCCACGGCGUGGUGU

UCCUGCACGUGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCU
CCCGCCAUCUGCCACGACGGCAAGGCCACUUCCCUAGAGAGGGCGUGUU
CGUGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGC
CUCAGAUCAUACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUG
5 GUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCCGAGCU
GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGC
CCCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAA
CAUUCAGAAGGAGAUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAAC
GAGAGCCUGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCA
10 AGUGGGCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUC
GUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCUGCC
UGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGACGA
CAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGACUCG
AGGCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAAGGUUCCUUUGUCCCCU
15 AAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUG
GAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAAAAAAAA
AA
AA
AAAAAAAAAAAA

20

SEQ ID NO: 106

>I'-OC24-B-120A mRNA
GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAG
GAUAGCACUAUAUAUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAG
CUCUCAGUGCGUGAACCUGACCACAAGAACACAGCUGCCCCCGCCUACA
25 CCAACAGCUUCACAAGAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGC
AGCGUGCUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGU
GACCUGGUUCCACGUGAUCAGCGGCACCAACGGCACCAAGAGAUUCGACA
ACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCAUCGAAAAG
AGCAACAUCAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGAC
30 ACAAAGCCUGCUGAUCGUUACAACGCGACCAACGUGGUGAUAAGGUG
UGCGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACCACAAGAACAACA
GAGCUGGAUGGAGAGCGAGUUCAGAGUGUACAGCAGCGCCAACAACUGC
ACCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCA
AGGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUAAGAACAUCGACGGC
35 UACUUCAAGAUCUACAGCAAGCACACCCCAUCAUCGUGAGAGAGCCCGA
GGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCCA
UCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUAGC
UACCUAACCCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGC
CUACUACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACG
40 AGAACGGCACUAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCGAGC
GAGACCAAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCA

GACAAGCAACUUCAGAGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCA
ACAUCACCAACCUGUGCCCCUUCGACGAGGUGUUAACGCCACCAGAUUU
GCUAGCGUGUACGCCUGGAACCGGAAGAGAAUCAGCAACUGCGUGGCCGA
CUACAGCGUGCUGUACAACCUGGCCCCCUUCUUCACCUUCAAGUGCUAUG
5 GCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCC
GACAGCUUCGUGAUCAGAGGCGACGAGGUGAGACAGAUCGCCCCCGGGCA
GACCGGCAACAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCG
GCUGCGUGAUCGCCUGGAACAGCAACAAGCUGGAUAGCAAGGUGAGCGG
CAACUACAAUUAACUGUACAGACUGUUCAGAAAGAGCAACCUGAAGCCCU
10 UCGAGAGAGACAUCAGCACCGAGAUCUACCAAGCCGGCAACAAGCCUCG
AACGGCGUGGCGGGCUUCAAACUGCUACUUCCCCCUACAGGAGCUACAGCUU
CAGACCCACCUACGGCGUGGGCCAUCAGCCCUACAGAGUGGUCGUGCUGA
GCUUCGAGCUGCUGCACGCCCCUGCCACAGUGUGCGGCCCCAAAAAGAGC
ACCAACCUGGUGAAGAACAAGUGCGUGAACUUAACUUAACGGCCUGAA
15 GGGCACCGGCGUGCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGC
AAUUCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAG
ACCCUGGAGAUCUGGACAUCACCCCCUGUAGCUUCGGCGGCGUGAGCGU
GAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCCGUGCUGUACCAAG
GCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACC
20 CCCACCUGGAGAGUGUAUAGCACCGGCAGCAACGUGUUUCAGACAAGAGC
CGGCUGCCUGAUCGGCGCCGAGUACGUGAACAACAGCUACGAGUGCGACA
UCCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAGAGC
CACCGGAGAGCUAGAAGCGUGGCUAGCCAAAGCAUCAUCGCCUACACCAU
GAGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCA
25 UCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUCUGCCCGUAAGU
AUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGGCGACAGCAC
CGAGUGCAGCAACCUGCUCCUGCAGUACGGCAGCUUCUGCACACAGCUGA
AGAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACCCCAAGAG
GUGUUCGCCAAGUGAAGCAGAUCAACAAGACCCCCCCCCAUCAAGUACUU
30 CGGCGGCUUCAACUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCA
AGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGAC
GCCGGCUUCAUCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAG
AGACCUGAUCUGCGCUCAGAAGUUAAGGGACUCACCGUGCUCCCCCCCC
UGCUGACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGCGGGA
35 ACCAUACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCC
CUUCGCCAUGCAGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGA
ACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCC
AUCGGCAAGAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAA
GCUGCAAGACGUGGUGAACCACAACGCCCAAGCCCUGAACACCCUGGUGA
40 AGCAGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUGCUGAAUGACAUC
UUCUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAU

CACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCA
 GAGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGC
 GAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCU
 ACCACCUGAUGAGCUUCCCUCAGAGCGCCCCCACGGCGUGGUGUCCUG
 5 CACGUGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGC
 CAUCUGCCACGACGGCAAGGCCACUUCCCUAGAGAGGGGCGUGUUCGUGA
 GCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAG
 AUCAUACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAU
 CGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACA
 10 GCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGAC
 GUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUC
 AGAAGGAGAU CGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAG
 CCUGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGC
 CCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUG
 15 GUGACCAUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAG
 GUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGA
 GCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGACUCGAGGCUC
 GCUUUCUUGCUGUCCAAUUUCUAUUAAGGUUCCUUGUUCCCUAAGUCC
 AACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCU
 20 GCCUAAUAAAAACAUUUAUUUUAUUGCAAAAAAAAAAAAAAAAAAAAAA
 AA
 AA
 AAAA

SEQ ID NO: 107

25 >I'-BA4-B-120A mRNA
 GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAAUAUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGA
 GCUCUCAGUGCGUGAACCUGAUCACAAGAACACAGUCGUACACCAACAG
 CUUCACAAGAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUG
 30 CUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGUGACCU
 GGUUCCACGCCAUCAGCGGCACCAACGGCACCAAGAGAUUCGACAACCC
 CGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCAGAAAGAGC
 AACAUCAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACAC
 AAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUG
 35 CGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAG
 AACAACAAGAGCUGGAUGGAGAGCGAGUUCAGAGUCUACAGCAGCGCCA
 ACAACUGCACCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGA
 GGGCAAGCAAGGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUAAGAAC
 AUCGACGGCUACUUCAAGAUCUACAGCAAGCACACCCCCAUCAACCUGG

GGAGAGACCUGCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCU
GCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCCUGCAC
CGUAGCUACCUAACCCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCG
CUGCCGCCUACUACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAA
5 GUACAACGAGAACGGCACUAUCACUGAUGCCGUGGACUGUGCCCUGGAC
CCCCUGAGCGAGACCAAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGG
GCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACCGAGAGCAUCGU
GAGAUUCCCCAACAUACCAACCUGUGCCCCUUCGACGAGGUGUUAAC
GCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCA
10 ACUGCGUGGCCGACUACAGCGUGCUGUACAACUUCGCUCCCUUCUUCGC
CUUCAAGUGCUAUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUC
ACCAACGUGUACGCCGACAGCUUCGUGAUCAGAGGCAACGAGGUGAGCC
AGAUUGCCCCCGGGCAAACCGGCAACAUCGCCGACUACAACUACAAGCU
GCCCCGACGACUUCACCGGCUGCGUGAUCGCGUGGAACAGCAAUAAGCUG
15 GAUUCCAAGGUGGGCGGCAACUACAACUACCGGUACAGACUGUUCAGAA
AGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUCUACCA
AGCCGGCAACAAGCCCUGCAACGGCGUGGGCGGGCGUCAACUGCUACUUC
CCCCUGCAGAGCUACGGCUUUAGGCCCACCUACGGCGUGGGCCAUCAGC
CCUACAGAGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCCGCAAC
20 CGUGUGCGGCCCCAAAAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG
AAUUAACUUAUAAUGGCCUUAAGGGCACGGGCGUGCUGACCGAGAGCA
ACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGCCGACAC
CACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCUGGACAUCACC
CCCUGUAGCUUCGGCGGGCGUGAGCGUGAUCACCCCCGGCACUAACACAA
25 GCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCC
CGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGAGUAUACAGC
ACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAUCGGCGCCG
AGUACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAU
CUGCGCUAGCUAUCAGACACAGACCAAGAGCCACCGGAGAGCUAGAAGC
30 GUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGA
ACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACC
AUCAGCGUGACCACCGAGAUCUGCCCGUAAGUAUGACCAAGACAAGCG
UGGACUGCACCAUGUACAUCUGCGGGCAGCAGCACCGAGUGCAGCAACCU
GCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAAGAGAGCCCUGACC
35 GGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAAG
UGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGUACUUCGGCGGCUUCAA

CUUCAGCCAAAUCCUGCCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUC
AUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCA
UCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAU
CUGCGCUCAGAAGUUCAAUGGACUCACCGUGCUCCCCCCCCCUGCUGACCG
5 ACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGGCGGGAACCAUAAC
AAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCC
AUGCAGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGC
UGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGG
CAAGAUCCAAGACAGCCUGAGCAGCACCCGCUAGCGCCCUGGGCAAGCUG
10 CAAGACGUGGUGAAUCACAACGCCCAAGCCCUGAACACCCUGGUGAAGC
AGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCU
GUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUCGACAGACUGAUC
ACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUC
GAGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAG
15 CGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGC
UACCACCUGAUGAGCUUCCUCAGAGCGCCCCCACGGCGUGGUGUUC
UGCACGUGACCUACGUGCCCCGCCCAAGAGAAGAACUUCACCACCGCUCCC
GCCAUCUGCCACGACGGCAAGGCCCACUUCCCUAGAGAGGGCGUGUUCG
UGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGCC
20 UCAGAUCAUACACCACGACAACACCUUCGUGAGCGGCAACUGCGACGUG
GUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGC
UGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUA AAAACCAUACUAG
CCCCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUG
AACAUUCAGAAGGAGAU CGACCGGCUGAAUGAGGUGGCCAAGAACCUGA
25 ACGAGAGCCUGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAU
CAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCC
AUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCU
GCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGA
CGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA
30 CUCGAGGCUCGCUUUCUUGCUGUCCAAUUAUUAUUAAGGUUCCUUGU
UCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGC
AUCUGGAUUCUGCCUAAUAAAAACA UUAUUAUUAUUGCAAAAAAAAAA
AA
AA
35 AAAAAAAAAAAAAAAAAA

SEQ ID NO: 108

>I'-BA4-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAUAUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGA
GCUCUCAGUGCGUGAACCUGAUCACAAGAACACAGUCGUACACCAACAG
5 CUUCACAAGAGGCGUGUACUACCCCGACAAGGUGUUCAGAAGCAGCGUG
CUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUCAGCAACGUGACCU
GGUUCCACGCCAUCAGCGGCACCAACGGCACCAAGAGAUUCGACAACCC
CGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCAGAAAAGAGC
AACAUCAUCAGAGGCUGGAUCUUCGGCACCAACCCUAGACUCCAAGACAC
10 AAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUG
CGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAG
AACAACAAGAGCUGGAUGGAGAGCGAGUUCAGAGUCUACAGCAGCGCCA
ACAACUGCACCUUCGAGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGA
GGGCAAGCAAGGCAAUUUCAAGAACCUAAGAGAGUUCGUGUUCAAGAAC
15 AUCGACGGCUACUUCAAGAUCUACAGCAAGCACACCCCCAUCAACCUGG
GGAGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCU
GCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCAC
CGUAGCUACCUAACCCCCGGCGAUAGUUCUUCUGGCUGGACGGCCGGCG
CUGCCGCCUACUACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAA
20 GUACAACGAGAACGGCACUAUCACUGAUGCCGUGGACUGUGCCCUGGAC
CCCCUGAGCGAGACCAAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGG
GCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACCGAGAGCAUCGU
GAGAUUCCCCAACAUCACCAACCUGUGCCCCUUCGACGAGGUGUUAAC
GCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCA
25 ACUGCGUGGCCGACUACAGCGUGCUGUACAACUUCGCUCCCUUCUUCGC
CUUCAAGUGCUAUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUC
ACCAACGUGUACGCCGACAGCUUCGUGAUCAGAGGCAACGAGGUGAGCC
AGAUUGCCCCCGGGCAAACCGGCAACAUCGCCGACUACAACUACAAGCU
GCCCCGACGACUUCACCGGCUGCGUGAUCGCGUGGAACAGCAAUAAGCUG
30 GAUUCCAAGGUGGGCGGCAACUACAACUACCGGUACAGACUGUUCAGAA
AGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUCUACCA
AGCCGGCAACAAGCCCUGCAACGGCGUGGCGGGCGUCAACUGCUACUUC
CCCCUGCAGAGCUACGGCUUUAGGCCCAACUACGGCGUGGGCCAUCAGC
CCUACAGAGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCCGCAAC
35 CGUGUGCGGCCCCAAAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG
AACUUCAACUUUAAUGGCCUUAAGGGCACGGGCGUGCUGACCGAGAGCA

ACAAGAAGUUCCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGCCGACAC
 CACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCUGGACAUCACC
 CCCUGUAGCUUCGGCGGCGUGAGCGUGAUCACCCCCGGCACUAACACAA
 GCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCC
 5 CGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGAGUAUACAGC
 ACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAUCGGCGCCG
 AGUACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAU
 CUGCGCUAGCUAUCAGACACAGACCAAGAGCCACCGGAGAGCUAGAAGC
 GUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGA
 10 ACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACC
 AUCAGCGUGACCACCGAGAUCUGCCCGUAAGUAUGACCAAGACAAGCG
 UGGACUGCACCAUGUACAUCUGCGGGCAGCAGCACCGAGUGCAGCAACCU
 GCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAAGAGAGCCCUGACC
 GGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAAG
 15 UGAAGCAGAUUCUACAAGACCCCCCCCCAUCAAGUACUUCGGCGGCUUCAA
 CUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUC
 AUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCA
 UCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAU
 CUGCGCUCAGAAGUUCAAUGGACUCACCGUGCUCUUUUUUUUUUGCUGACCG
 20 ACGAGAUGAUCGCUCAGUACACCUCUGCCUGCUGGCGGGAACCAUAAC
 AAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCC
 AUGCAGAUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGC
 UGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGG
 CAAGAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUG
 25 CAAGACGUGGUGAAUCACAACGCCCAAGCCCUGAACACCCUGGUGAAGC
 AGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCU
 GUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUCGACAGACUGAUC
 ACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUC
 GAGCCGCCGAGAUCAAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAG
 30 CGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGC
 UACCACCUGAUGAGCUUCCCUAGAGCGCCCCCACGGCGUGGUGUUC
 UGCACGUGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCC
 GCCAUCUGCCACGACGGCAAGGCCACUUCUUUAGAGAGGGCGUGUUCG
 UGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUACGAGCC
 35 UCAGAUCAUCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUG
 GUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGC

UGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAG
 CCCCAGCUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUG
 AACAUUCAGAAGGAGAUUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGA
 ACGAGAGCCUGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAU
 5 CAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCC
 AUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCU
 GCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGA
 CGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA
 CUCGAGGCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAGGUUCCUUGU
 10 UCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGC
 AUCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAA
 AA
 AAAAAAAA

15

SEQ ID NO: 109

>I'-H1N1-B mRNA

GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCU
 UCACCACCGCCAACGCCGACACCCUGUGCAUCGGCUACCACGCCAACAAAC
 20 AGCACCGACACCGUGGACACCGUGCUGGAGAAGAACGUGACCGUGACCC
 ACAGCGUGAACCUGCUGGAGGACAAGCACAAACGGCAAGCUGUGCAAGCU
 GCGGGGCGUGGGCCCCCUGCACCUGGGCAAGUGCAACAUCGCCGGCUGG
 AUCCUGGGCAACCCCGAGUGCGAGUCCCUGAGCACCGCCCGCUCCUGGA
 GCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGG
 25 GGACUUCAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCC
 AGCUUCGAGCGCUUCGAGAUCUUCCCCAAGACCUCAGCUGGCCCCAACC
 ACGACUCCGACAACGGCGUGACCGCCGCCUGCCCCACGCCGGGGCCAAG
 AGCUUCUACAAGAACCUGAUCUGGCUGGUGAAGAAGGGCAAGUCCUACC
 CCAAGAUCAACCAGACCUACAUCAACGACAAGGGGAAGGAGGUCCUCGU
 30 GCUGUGGGGCAUCCACCACCCCCACCAUCGCCGACCAGCAGAGCCUCU
 ACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCGCCGUACAGCAA
 GAAGUUCAAGCCCGAGAUCGCCACCCGCCCCAAGGUCCGGGACCAGGAG
 GGCCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUA
 CCUUCGAGGCCACCGGCAACCUCGUCGCCCCGCGGUACGCCUUCACGAUG
 35 GAGCGCGACGCCGGGUCCGGCAUCAUCAUCAGCGACACCCCGUGCACG
 ACUGCAACACGACCUGCCAGACGCCCGAGGGGGCGAUAACACCUCUU

GCCGUUCCAGAACGUCCACCCGAUCACGAUCGGCAAGUGCCCCAAGUAC
 GUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGCAACGUCCCCUC
 CAUCCAGAGCCGGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGG
 GGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACG
 5 AGCAGGGGUCCGGGUACGCGGGCCGACCUCAAGUCGACGCAGAACGCGAU
 CGACAAGAUCACCAACAAGGUCAACAGCGUGAUCGAGAAGAUGAACACG
 CAGUUCACCGCCGUCGGCAAGGAGUUAACCAACCUGGAGAAGCGCAUCG
 AGAACCUCAACAAGAAGGUGGACGACGGGUUCCUGGACAUCUGGACGUA
 CAACGCGGAGCUCCUGGUUCCUUGGAGAACGAGCGGACCCUCGACUAC
 10 CACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCGCAACCAGCUCA
 AGAACAAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACCACAA
 GUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUCGACUAC
 CCGAAGUACAGCGAGGAGGCGAAGCUGAACCGGGAGAAGAUCGACGGGG
 UCAAGCUCGACUCCACGCGCAUCUACCAGAUCCUGGCCAUCUACUCGACC
 15 GUGGCGAGCUCCUCGUCCUGGUGGUCUCGCUCGGCGCCAUCAGCUUCU
 GGAUGUGCUCCAACGGGUCGCUGCAGUGCCGGAUCUGCAUCUGAUGACU
 CGAGGCUCGCUUUCUUGCUGUCCA AUUUCUAUUAAGGUUCCUUGUUC
 CCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAU
 CUGGAUUCUGCCUAAUAAAAACA UUUUAUUUUCAUUGCAAAAAAAAAA
 20 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAA

SEQ ID NO: 110

>I'-H3N2-B mRNA

25 GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAAUAUGAAGACCAUCAUCGCCUGAGCUACAUCUGU
 GCCUGGUGUUCGCCAGAAGA UCCCCGGCAACGACAACAGCACCGCCACC
 CUGUGCCUGGGCCACCACGCCGUGCCCAACGGCACCAUCGUGAAGACCA
 UCACCAACGACCGGAUCGAGGUGACCAACGCCACCGAGCUGGUGCAGAA
 30 CUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCAGAUCCUGGACGGCG
 GCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCAUGCGACGG
 GUUCCAGAACAAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAAC
 UCCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCU
 GGUGGCCAGCUCCGGCACCCUGGAGUUAAGAACGAGAGCUUCAACUGG
 35 ACCGGGGUCAAGCAGAACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCA
 GCUCCAGCUUCUUCUCCCGGCUGAACUGGCUCACCCACCUGAACUACACC

UACCCCGCCCCUAACGUGACCAUGCCCAACAACGAGCAGUUCGACAAGC
 UGUACAUCUGGGGCGUCCACCACCCAGCACCGACAAGGACCAGAUCUC
 CCUCUUCGCCCAGCCCAGCGGGCGCAUCACCGUGUCCACGAAGCGGAGCC
 AGCAGGCCGUAUCCCCAACAUCGGCUCGCGCCGCGGAUCCGCGACAUC
 5 CCCAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUC
 UGCUCAUAACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUCAA
 GAUCCGGUCCGGCAAGAGCUCGAUCAUGCGCUCCGACGCCCCGAUCGGG
 AAGUGCAAGAGCGAGUGCAUCACCCCCAACGGCUCGAUCCCGAACGACA
 AGCCCUUCCAGAACGUAACCGGAUCACGUACGGGGGCCUGCCCGCGCUA
 10 CGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCGGAACGUCCCC
 GAGAAGCAGACCCGCGGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCGAGA
 ACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAGAA
 CAGCGAGGGCCGCGGGCAGGCCGCGGACCUCAAGUCGACGCAGGCCGCG
 AUCGACCAGAUCAACGGCAAGCUGAACCGGCUCAUCCGGAAGACGAACG
 15 AGAAGUUCCACCAGAUUCGAGAAGGAGUUCUCCGAGGUCGAGGGCCGCGU
 GCAGGACCUGGAGAAGUACGUCGAGGACACCAAGAUCGACCUCUGGAGC
 UACAACGCCGAGCUGCUCGUGGCGCUGGAGAACCAGCACACCAUCGACC
 UCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGAAGAAGCAGCU
 CCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUCUACCAC
 20 AAGUGCGACAACGCGUGCAUCGGGUCCAUCGCAACGAGACCUACGACC
 ACAACGUCUACCGGGACGAGGCCCUGAACAACCGCUUCCAGAUCAAGGG
 CGUGGAGCUCAAGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUC
 GCGAUGUCCUGCUUCCUCCUGUGCAUCGCCCUCUGGGCUUCAUCAUGU
 GGGCGUGCCAGAAGGGGAACAUCGCGUGCAACAUCUGCAUCUGAUGACU
 25 CGAGGCUCGCUUUCUUGCUGUCCAUAUUAUUAAGGUUCCUUUGUUC
 CCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAU
 CUGGAUUCUGCCUAAUAAAAACAUAUUAUUUCAUUGCAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAA
 AAA
 30 AAAAAA

SEQ ID NO: 111

>I'-B-Victoria-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
 35 CCAGCAACGCCGACCGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCC
 CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGA

UCCCCUGACCACCACCCCCACCAAGUCCCACUUCGCCAACCUGAAGGGC
 ACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCUGAACUGCACCGACC
 UGGACGUGGCCCUGGGCCGCCCCAAGUGCACCGGCAAGAUGCCCAGCGCC
 CGGGUGUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC
 5 CCAUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCCAACCUGCUGCGG
 GGGUACGAGCACGUCCGCCUGUCCACCCACAACGUGAUAACGCCGAGG
 ACGCCCCCGGCCGGCCCUACGAGAUCGGGACCAGCGGCUCCUGCCCCAAC
 AUCACCAACGGGAACGGCUUCUUCGCCACCAUGGCCUGGGCCGUCCCCA
 AGAACAAGACCGCCACCAACCCCCUGACCAUCGAGGUGCCGUACAUCUG
 10 CACCGAGGGGGAGGACCAGAUCACCGUCUGGGGCUUCCACAGCGACAAC
 GAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAAGUUCA
 CGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGG
 CUUCCCGAACCAGACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUC
 GUGGUCGACUACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGU
 15 ACCAGCGGGGCAUCCUGCUGCCGCAGAAAGGUCUGGUGCGCCUCCGGGCG
 CAGCAAGGUGAUAAGGGCUCCUGCCCCUCAUCGGGGAGGCCGACUGC
 CUGCACGAGAAGUACGGCGGGCUACAACAAGAGCAAGCCGUACUACACCG
 GCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUGGGGUCAAGACCCC
 GCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCUGCUC
 20 AAGGAGCGCGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGU
 GGGAGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCCACGGGGCCCA
 CGGCGUGGCGGUCGCCGCGGACCUCAAGAGCACCCAGGAGGCCAUCAAC
 AAGAUACGAAGAACCUGAACUCCCUCAGCGAGCUGGAGGUGAAGAACC
 UCCAGCGCCUGUCCGGGGCGAUGGACGAGCUCCACAACGAGAUCUGGA
 25 GCUCGACGAGAAGGUCGACGACCUGCGGGCCGACACCAUCAGCUCGCAG
 AUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAUAACAGCGAGG
 ACGAGCACCGCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCC
 CUCGGCGGUCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGC
 AACCAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCCGGGG
 30 AGUUCUCCCUGCCGACCUUCGACAGCCUCAACAUCACGGCGGGCCUCGCUG
 AACGACGACGGCCUCGACAACCACACCAUCCUGCUCUACUACUCCACCGC
 GGCCAGCUCGCUGGCGGUGACGCUCAUGAUCGCCAUCUUCGUCGUGUAC
 AUGGUCUCCCGCGACAACGUGAGCUGCUCGAUCUGCCUGUGAUGACUCG
 AGGCUCGCUUUCUUGCUGUCCA AUUUCUAUUAAGGUUCCUUGUUGUCCC
 35 UAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAUCU
 GGAUUCUGCCUAAUAAAAAACA UUUUAUUUUCAUUGCAAAAAAAAAA

AAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAA
AA
AAAAAA

SEQ ID NO: 112

5 >I'-B-Yamagata-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
CCAGCAACGCCGACCGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCC
CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGA
10 UCCCCUGACCACCACCCCCACCAAGUCCUACUUCGCCAACCUGAAGGGC
ACCCGGACCCGCGGCAAGCUGUGCCCCGACUGCCUGAACUGCACCGACCU
GGACGUGGGCCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCCCAGCGCCA
AGGCCUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCC
AUCAUGCACGACCGCACCAAGAUCGCGCAGCUGCCCAACCUGCUGCGGG
15 GGUACGAGAAGAUCGCCCUGUCCACCCAGAACGUCAUCGACGCCGAGAA
GGCCCCCGGCGGGCCCUACCGGCUGGGCACCAGCGGGUCCUGCCCCAACG
CCACCAGCAAGAUCGGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCAA
GGACAACUACAAGAACGCCACCAACCCCCUGACCGUCGAGGUGCCCUAC
AUCUGCACCGAGGGGGAGGACCAGAUCACGGUCUGGGGGCUUCCACUCCG
20 ACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGCAGAA
GUUCACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAUC
GGGGACUUCCCGGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCC
GCAUCGUGGUCGACUACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAU
CGUGUACCAGCGGGGGGUCCUGCUGCCGCAGAAGGUGUGGUGCGCCAGC
25 GGCCGCUCCAAGGUCAUCAAGGGGAGCCUGCCCCUCAUCGGCGAGGCCG
ACUGCCUGCACGAGGAGUACGGGGGGCCUCAACAAGUCCAAGCCGUACUA
CACGGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAUCUGGGUGAAG
ACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCU
GCUCAAGGAGCGCGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGG
30 GGCUGGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCG
CCCACGGGGUCGCGGUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAU
CAACAAGAUCACCAAGAACCUGAACAGCCUCUCCGAGCUGGAGGUCAAG
AACCUCAGCGCCUGAGCGGCGCGAUGGACGAGCUCCACAACGAGAUCC
UGGAGCUCGACGAGAAGGUGGACGACCUGCGGGCCGACACGAUCUCGUC
35 CCAGAUCGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAUCAUCAACUCG
GAGGACGAGCACCUGCUCGCCCUUGGAGCGGAAGCUCAAGAAGAUGCUGG

GCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAA
 GUGCAACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUUCAACGCC
 GGCGAGUUCAGCCUGCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUC
 CCUGAACGACGACGGGCUCGACAACCACACGAUCCUGCUCUACUACAGC
 5 ACCGCGGCCUCGUCCCUGGCGGUCACGCUCAUGCUGGCCAUCUUCAUCG
 UGUACAUGGUCAGCCGCGACAACGUGUCGUGCUCCAUCUGCCUCUGAUG
 ACUCGAGGCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAAGGUUCCUUUG
 UUCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAG
 CAUCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAA
 10 AAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAA
 AAA
 AAAAAAAAAA

SEQ ID NO: 113

>I'-H1N1-HA0-mut-B mRNA

15 GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCU
 UCACCACCGCCAACGCCGACACCCUGUGCAUCGGCUACCACGCCAACAAC
 AGCACCGACACCGUGGACACCGUGCUGGAGAAGAACGUGACCGUGACCC
 ACAGCGUGAACCUGCUGGAGGACAAGCACAAACGGCAAGCUGUGCAAGCU
 20 GCGGGGCGUGGCCCCCUGCACCUGGGCAAGUGCAACAUCGCCGGCUGG
 AUCCUGGGCAACCCCGAGUGCGAGUCCCUGAGCACCGCCCGCUCCUGGA
 GCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGG
 GGACUUCAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCC
 AGCUUCGAGCGCUUCGAGAUCUUCCCCAAGACCUCAGCUGGCCCAACC
 25 ACGACUCCGACAACGGCGUGACCGCCGCCUGCCCCACGCCGGGGCCAAG
 AGCUUCUACAAGAACCUGAUCUGGCUGGUGAAGAAGGGCAAGUCCUACC
 CCAAGAUCAACCAGACCUACAUCAACGACAAGGGGAAGGAGGUCCUCGU
 GCUGUGGGGCAUCCACCACCCCCACCAUCGCCGACCAGCAGAGCCUCU
 ACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCCCGGUACAGCAA
 30 GAAGUUCAAGCCCAGAUCCGCCACCCGCCCAAGGUCCGGGACCAGGAG
 GGCCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUA
 CCUUCGAGGCCACCGGCAACCUCGUCGCCCCGCGGUACGCCUUCACGAUG
 GAGCGCGACGCCGGGUCCGGCAUCAUCAUCAGCGACACCCCGUGCACG
 ACUGCAACACGACCUGCCAGACGCCCCAGGGGGGCGAUAACACCUCCCU
 35 GCCGUUCCAGAACGUCCACCCGAUCACGAUCGGCAAGUGCCCCAAGUAC
 GUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGCAACGUCCCCUC

CAUCCAGAGCCAGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGG
 GGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACG
 AGCAGGGCUCCGGGUACGCGGCCGACCUCAAGUCGACGCAGAACGCGAU
 CGACAAGAUCACCAACAAGGUCAACAGCGUGAUCGAGAAGAUGAACACG
 5 CAGUUCACCGCCGUCGGCAAGGAGUUCAACCACCUGGAGAAGCGCAUCG
 AGAACCUCAACAAGAAGGUGGACGACGGGUUCCUGGACAUCUGGACGUA
 CAACGCGGAGCUCCUGGUCCUCCUGGAGAACGAGCGGACCCUCGACUAC
 CACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCGCAACCAGCUCA
 AGAACAAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACCACAA
 10 GUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUCGACUAC
 CCGAAGUACAGCGAGGAGGCGAAGCUGAACCGGGAGAAGAUCGACGGGG
 UCAAGCUCGACUCCACGCGCAUCUACCAGAUCUCCUGGCCAUCUACUCGACC
 GUGGCGAGCUCCUCGUCCUGGUGGUCUCGCUCCGGCGCCAUCAGCUUCU
 GGAUGUGCUCCAACGGGUCGCUAGUGCCGGAUCUGCAUCUGAUGACU
 15 CGAGGCUCGCUUUCUUGCUGUCCAUAUUAUUAAGGUUCCUUGUUC
 CCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAU
 CUGGAUUCUGCCUAAUAAAAAACAUAUUAUUAUUGCAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAA
 AA
 20 AAAAAAA

SEQ ID NO: 114

>I'-H3N2-HA0-mut-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAAUAUGAAGACCAUCAUCGCCUGAGCUACAUCUGU
 25 GCCUGGUGUUCGCCCAGAAGAUCGCCGCAACGACAACAGCACCGCCACC
 CUGUGCCUGGGCCACCACGCCGUGCCCAACGGCACCAUCGUGAAGACCA
 UCACCAACGACCGGAUCGAGGUGACCAACGCCACCGAGCUGGUGCAGAA
 CUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCAGAUCUGGACGGCG
 GCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCAGUGCGACGG
 30 GUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAAC
 UCCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCCU
 GGUGGCCAGCUCCGGCACCCUGGAGUUCAAGAACGAGAGCUUCAACUGG
 ACCGGGGUCAAGCAGAACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCA
 GCUCCAGCUUCUUCUCCCGGCUGAACUGGCUCACCCACCUGAACUACACC
 35 UACCCCGCCCUC AACGUGACCAUGCCCAACAACGAGCAGUUCGACAAGC
 UGUACAUCUGGGGCGUCCACCACCCAGCACCGACAAGGACCAGAUCUC

CCUCUUCGCCCAGCCCAGCGGGCGCAUCACCGUGUCCACGAAGCGGAGCC
 AGCAGGCCGUAUCCCCAACAUCGGCUCGCGCCGCGGAUCCGCGACAUC
 CCCAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUC
 UGCUCAUAACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUA
 5 GAUCCGGUCCGGCAAGAGCUCGAUCAUGCGCUCCGACGCCCCGAUCGGG
 AAGUGCAAGAGCGAGUGCAUCACCCCCAACGGCUCGAUCCCGAACGACA
 AGCCCUUCCAGAACGUAACCGGAUCACGUACGGGGCCUGCCCGCGCUA
 CGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCAGAACGUCCCC
 GAGAAGCAGACCCAGGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCGAGA
 10 ACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACAGAA
 CAGCGAGGGCCGCGGGCAGGCCGCGGACCUCAGUGACGCAGGCCGCG
 AUCGACCAGAUCAACGGCAAGCUGAACCGGCUCAUCGGGAAGACGAACG
 AGAAGUUCCACCAGAUUCGAGAAGGAGUUCUCCGAGGUCGAGGGCCGCGU
 GCAGGACCUGGAGAAGUACGUCGAGGACACCAAGAUCGACCUCUGGAGC
 15 UACAACGCCGAGCUCGUCGUGGCGCUGGAGAACCAGCACACCAUCGACC
 UCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGAAGAAGCAGCU
 CCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUCUACCAC
 AAGUGCGACAACGCGUGCAUCGGGUCCAUCGCAACGAGACCUACGACC
 ACAACGUCUACCGGGACGAGGCCUGAACAACCGCUUCCAGAUCAAGGG
 20 CGUGGAGCUCAAGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUC
 GCGAUGUCCUGCUUCCUCCUGUGCAUCGCCCUCUGGGCUUCAUCAUGU
 GGGCGUGCCAGAAGGGGAACAUCGGUGCAACAUCUGCAUCUGAUGACU
 CGAGGCUCGCUUUCUUGCUGUCCAUAUUAAGGUUCCUUGUUC
 CCUAAGUCCAACUACUAAACUGGGGGAUUAUUAUGAAGGGCCUUGAGCAU
 25 CUGGAUUCUGCCUAAUAAAAACAUAUUAUUAUUGCAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAA
 AAA
 AAAAAAA

SEQ ID NO: 115

30 >I'-B-Victoria-HA0-mut-B mRNA
 GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
 CCAGCAACGCCGACCGGAUCUGCACC GGCAUCACCAGCUCCAACAGCCCC
 CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGA
 35 UCCCCCUGACCACCACCCCCACCAAGUCCCAUUCGCCAACCUGAAGGGC
 ACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCUGAACUGCACCGACC

UGGACGUGGCCCUGGGCCGCCCCAAGUGCACCGGCAAGAUGCCCCAGCGCC
CGGGUGUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC
CCAUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGG
GGGUACGAGCACGUCCGCCUGUCCACCCACAACGUGAUAACGCCGAGG
5 ACGCCCCCGGCCGGCCCUACGAGAUCGGGACCAGCGGCUCCUGCCCCAAC
AUCACCAACGGGAACGGCUUCUUCGCCACCAUGGCCUGGGCCGUCCCCA
AGAACAAGACCGCCACCAACCCCCUGACCAUCGAGGUGCCGUACAUCUG
CACCGAGGGGGAGGACCAGAUCACCGUCUGGGGCUUCCACAGCGACAAC
GAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAAGUUA
10 CGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGG
CUUCCCGAACCAGACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUC
GUGGUCGACUACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGU
ACCAGCGGGGCAUCCUGCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGGCG
CAGCAAGGUGAUAAGGGCUCCCUGCCCCUCAUCGGGGAGGCCGACUGC
15 CUGCACGAGAAGUACGGCGGGCUCAACAAGAGCAAGCCGUACUACACCG
GCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUGGGUCAAGACCCC
GCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCUGCUC
AAGGAGCAGGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGU
GGGAGGGCAUGAUCGCCGGGUGGCACGGCUACACGUUCCACGGGGCCCA
20 CGGCGUGGCGGUCGCCGCGGACCUCAAGAGCACCCAGGAGGCCAUCAAC
AAGAUACGAAGAACCUGAACUCCCUCAGCGAGCUGGAGGUGAAGAACC
UCCAGCGCCUGUCCGGGGCGAUGGACGAGCUCCACAACGAGAUCUGGA
GCUCGACGAGAAGGUCGACGACCUGCGGGCCGACACCAUCAGCUCGCAG
AUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAUAACAGCGAGG
25 ACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCC
CUCGGCGGUCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGC
AACCAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCCGGGG
AGUUCUCCCUGCCGACCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUG
AACGACGACGGCCUCGACAACCACACCAUCCUGCUCUACUACUCCACCGC
30 GGCCAGCUCGCUGGCGGUGACGCUCAUGAUCGCCAUCUUCGUCGUGUAC
AUGGUCUCCCGCGACAACGUGAGCUGCUCGAUCUGCCUGUGAUGACUCG
AGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCUCC
UAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAUCU
GGAUUCUGCCUAAUAAAAACAUUUAUUUUAUUGCAAAAAAAAAAAAAA
35 AAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAA
AA

AAAAAA

SEQ ID NO: 116

>I'-B-Yamagata-HA0-mut-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
5 GGAUAGCACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
CCAGCAACGCCGACCGGAUCUGCACC GGCAUCACCAGCUCCAACAGCCCC
CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGA
UCCCCCUGACCACCACCCCCACCAAGUCCUACUUCGCCAACCUGAAGGGC
ACCCGGACCCGCGGCAAGCUGUGCCCCGACUGCCUGAACUGCACCGACCU
10 GGACGUGGCCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCCCAGCGCCA
AGGCCUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCC
AUCAUGCACGACCGCACCAAGAUCGGCAGCUGCCCAACCUGCUGCGGG
GGUACGAGAAGAUCCGCCUGUCCACCCAGAACGUCAUCGACGCCGAGAA
GGCCCCCGGCGGGCCCUACCGGCUGGGCACCAGCGGGUCCUGCCCCAACG
15 CCACCAGCAAGAUCGGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCCAA
GGACAACUACAAGAACGCCACCAACCCCCUGACCGUCGAGGUGCCCUAC
AUCUGCACCGAGGGGGAGGACCAGAUACGGUCUGGGGCUUCCACUCCG
ACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGCAGAA
GUUCACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAU
20 GGGGACUUCCCGGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCC
GCAUCGUGGUCGACUACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAU
CGUGUACCAGCGGGGGGUCCUGCUGCCGCAGAAGGUGUGGUGCGCCAGC
GGCCGCUCCAAGGUCAUCAAGGGGAGCCUGCCCCUCAUCGGCGAGGCCG
ACUGCCUGCACGAGGAGUACGGGGGCCUCAACAAGUCCAAGCCGUACUA
25 CACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCC GAUCUGGGUGAAG
ACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCU
GCUCAAGGAGCAGGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGG
GGCUGGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCG
CCCACGGGGUCGCGGUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAU
30 CAACAAGAUACCAAGAACCUGAACAGCCUCUCCGAGCUGGAGGUCAAG
AACCUC CAGCGCCUGAGCGGCGCGAUGGACGAGCUCCACAACGAGAUCC
UGGAGCUCGACGAGAAGGUGGACGACCUGCGGGCCGACACGAUCUCGUC
CCAGAU CGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAUCAUCAACUCG
GAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGG
35 GCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAA
GUGCAACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUUCAACGCC

GGCGAGUUCAGCCUGCCCACGUUCGACUCGCUCAACAUCACCGCGGGCCUC
 CCUGAACGACGACGGGCUCGACAACCACACGAUCCUGCUCUACUACAGC
 ACCGCGGGCCUCGUCCCUGGCGGUCACGCUCAUGCUGGCCAUCUUCAUCG
 UGUACAUGGUCAGCCGCGACAACGUGUCGUGCUCCAUCUGCCUCUGAUG
 5 ACUCGAGGCUCGCUUUCUUGCUGUCCA AUUUCUAUUAAAGGUUCCUUUG
 UUCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAG
 CAUCUGGAUUCUGCCUAAUAAAAACA UUUUUUUAUUGCAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA
 10 AAAAAAAAAA

SEQ ID NO: 117

>I'-H1N1-B-120A mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCU
 15 UCACCACCGCCAACGCCGACACCCUGUGCAUCGGCUACCACGCCAACAAC
 AGCACCGACACCGUGGACACCGUGCUGGAGAAGAACGUGACCGUGACCC
 ACAGCGUGAACCUGCUGGAGGACAAGCACAACGGCAAGCUGUGCAAGCU
 GCGGGGCGUGGGCCCCCUGCACCUGGGCAAGUGCAACAUCGCCGGCUGG
 AUCCUGGGCAACCCCGAGUGCGAGUCCCUGAGCACCGCCCGCUCCUGGA
 20 GCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGG
 GGACUUCAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCC
 AGCUUCGAGCGCUUCGAGAUCUUCCTCAAGACCUCAGCUGGCCCCAACC
 ACGACUCCGACAACGGCGUGACCGCCGCCUGCCCCACGCCGGGGCCAAG
 AGCUUCUACAAGAACCUGAUCUGGCUGGUGAAGAAGGGCAAGUCCUACC
 25 CCAAGAUCAACCAGACCUACAUCAACGACAAGGGGAAGGAGGUCCUCGU
 GCUGUGGGGCAUCCACCACCCCCCACCAUCGCCGACCAGCAGAGCCUCU
 ACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCCTGGUACAGCAA
 GAAGUUCAAGCCCGAGAUCCGACCCGCCCAAGGUCCGGGACCAGGAG
 GGCCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUA
 30 CCUUCGAGGCCACCGGCAACCUCGUCGCCCCGCGGUACGCCUUCACGAUG
 GAGCGCGACGCCGGGUCCGGCAUCAUCAUCAGCGACACCCCGUGCACG
 ACUGCAACACGACCUGCCAGACGCCCGAGGGGGCGAUCAACACCUCCCU
 GCCGUUCCAGAACGUCCACCCGAUCACGAUCGGCAAGUGCCCCAAGUAC
 GUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGCAACGUCCCCUC
 35 CAUCCAGAGCCGGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGG
 GGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACG

AGCAGGGCUCCGGGUACGCGGCCGACCUCAAGUCGACGCAGAACGCGAU
 CGACAAGAUCACCAACAAGGUCAACAGCGUGAUCGAGAAGAUGAACACG
 CAGUUCACCGCCGUCGGCAAGGAGUUCAACCACCUGGAGAAGCGCAUCG
 AGAACCUCAACAAGAAGGUGGACGACGGGUUCCUGGACAUCUGGACGUA
 5 CAACGCGGAGCUCCUGGUCCUCCUGGAGAACGAGCGGACCCUCGACUAC
 CACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCGCAACCAGCUCA
 AGAACAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACCACAA
 GUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUACGACUAC
 CCGAAGUACAGCGAGGAGGCGAAGCUGAACCAGGAGAGAAGAUUCGACGGGG
 10 UCAAGCUCGACUCCACGCGCAUCUACCAGAUCCUGGCCAUCUACUCGACC
 GUGGCGAGCUCCUCGUCCUGGUGGUCUCGCUCGGCGCCAUCAGCUUCU
 GGAUGUGCUCCAACGGGUCGCUGCAGUGCCGGAUCUGCAUCUGAUGACU
 CGAGGCUCGCUUUCUUGCUGUCCA AUUUCUAUUAAGGUUCCUUGUUC
 CCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAU
 15 CUGGAUUCUGCCUAAUAAAAAACA UUUUAUUUUCAUUGCAAAAAAAAAA
 AA
 AA
 AAAAAAAAAAAAAAAAAA

SEQ ID NO: 118

20 >I'-H3N2-B-120A mRNA
 GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
 GGAUAGCACUAUAUAUGAAGACCAUCAUCGCCUGAGCUACAUCUGU
 GCCUGGUGUUCGCCCAGAAGAUCCCCGGCAACGACAACAGCACCGCCACC
 CUGUGCCUGGGCCACCACGCCGUGCCCAACGGCACCAUCGUGAAGACCA
 25 UCACCAACGACCGGAUCGAGGUGACCAACGCCACCGAGCUGGUGCAGAA
 CUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCAGAUCCUGGACGGCG
 GCAACUGCACCCUGAUCGACGCCUGGUGGCGACCCCAUGUGCGACGG
 GUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAAC
 UCCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCCU
 30 GGUGGCCAGCUCCGGCACCCUGGAGUUCAAGAACGAGAGCUUCAACUGG
 ACCGGGGUCAAGCAGAACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCA
 GCUCCAGCUUCUUCUCCCGGCUGAACUGGCUCACCCACCUGAACUACACC
 UACCCCGCCCUCAACGUGACCAUGCCCAACAACGAGCAGUUCGACAAGC
 UGUACAUCUGGGGCGUCCACCACCCAGCACCGACAAGGACCAGAUUC
 35 CCUCUUCGCCCAGCCAGCGGGCGCAUACCGUGUCCACGAAGCGGAGCC
 AGCAGGCCGUAUCCCCAACAUCGGCUCCCGCCCGCGGAUCCGCGACAUC

```

CCCAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUCC
UGCUCAUAACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUCAA
GAUCCGGUCCGGCAAGAGCUCGAUCAUGCGCUCCGACGCCCCGAUCGGG
AAGUGCAAGAGCGAGUGCAUACCCCCAACGGCUCGAUCCCGAACGACA
5 AGCCCUUCCAGAACGUCAACCGGAUCACGUACGGGGGCCUGCCCGCGCUA
CGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCGGAACGUCCCC
GAGAAGCAGACCCGCGGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCGAGA
ACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAGAA
CAGCGAGGGCCGCGGGCAGGCCGCGGACCUCAAGUCGACGCAGGCCGCG
10 AUCGACCAGAUCAACGGCAAGCUGAACCGGCUCAUCGGGAAGACGAACG
AGAAGUUCACCAGAUUCGAGAAGGAGUUCUCCGAGGUCGAGGGCCGCGU
GCAGGACCUGGAGAAGUACGUCGAGGACACCAAGAUCGACCUCUGGAGC
UACAACGCCGAGCUGCUCGUGGCGCUGGAGAACCAGCACACCAUCGACC
UCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGAAGAAGCAGCU
15 CCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUCUACCAC
AAGUGCGACAACGCGUGCAUCGGGUCCAUCCGCAACGAGACCUACGACC
ACAACGUCUACCGGGACGAGGCCCUGAACAACCGCUUCCAGAUCAAGGG
CGUGGAGCUCAAGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUC
GCGAUGUCCUGCUUCCUCCUGUGCAUCGCCCUCUGGGCUUCAUCAUGU
20 GGGCGUGCCAGAAGGGGAACAUCGGUGCAACAUCUGCAUCUGAUGACU
CGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUC
CCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAU
CUGGAUUCUGCCUAAUAAAAACAUAUUAUUUUCAUUGCAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
25 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAA

```

SEQ ID NO: 119

>I'-B-Victoria-B-120A mRNA

```

GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
30 GGAUAGCACUAUAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
CCAGCAACGCCGACCGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCC
CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGA
UCCCCCUGACCACCACCCCCACCAAGUCCCACUUCGCCAACCUGAAGGGC
ACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCUGAACUGCACCGACC
35 UGGACGUGGCCCUUGGGCCGCCCAAGUGCACCGGCAAGAUCCCCAGCGCC
CGGGUGUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC

```

CCAUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGG
GGGUACGAGCACGUCCGCCUGUCCACCCACAACGUGAUAACGCCGAGG
ACGCCCCCGGCCGGCCCUACGAGAUCGGGACCAGCGGCUCCUGCCCCAAC
AUCACCAACGGGAACGGCUUCUUCGCCACCAUGGCCUGGGCCGUCCTCA
5 AGAACAAGACCGCCACCAACCCCCUGACCAUCGAGGUGCCGUACAUCUG
CACCGAGGGGGAGGACCAGAUCACCGUCUGGGGCUUCCACAGCGACAAC
GAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAAGUUCA
CGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGG
CUUCCCGAACCAGACGGAGGACGGGGGCCUGCCCCAGUCCGGGGCGCAUC
10 GUGGUCGACUACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGU
ACCAGCGGGGCAUCCUGCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGGCG
CAGCAAGGUGAUAAGGGCUCCUGCCCCUCAUCGGGGAGGCCGACUGC
CUGCACGAGAAGUACGGCGGGCUCAACAAGAGCAAGCCGUACUACACCG
GCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUGGGUCAAGACCCC
15 GCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCUGCUC
AAGGAGCGCGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGU
GGGAGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCACGGGGCCCA
CGGCGUGGCGGUCGCCGCGGACCUCAAGAGCACCCAGGAGGCCAUCAAC
AAGAUCACGAAGAACCUGAACUCCCUCAGCGAGCUGGAGGUGAAGAACC
20 UCCAGCGCCUGUCCGGGGCGAUGGACGAGCUCCACAACGAGAUCUGGA
GCUCGACGAGAAGGUCGACGACCUGCGGGCCGACACCAUCAGCUCGCAG
AUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAUAACAGCGAGG
ACGAGCACCGCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCC
CUCGGCGGUCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGC
25 AACCAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCCGGGG
AGUUCUCCCUGCCGACCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUG
AACGACGACGGCCUCGACAACCACACCAUCCUGCUCUACUACUCCACCGC
GGCCAGCUCGCUGGCGGUGACGCUCAUGAUCGCCAUCUUCGUCGUGUAC
AUGGUCUCCCGCGACAACGUGAGCUGCUCGAUCUGCCUGUGAUGACUCG
30 AGGCUCGCUUUCUUGCUGUCCA AUUUCUAUUAAAGGUUCCUUGUUCUCC
UAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAUCU
GGAUUCUGCCUAAUAAAAACA UUUUAUUUUCAUUGCAAAAAAAAAAAAA
AA
AA
35 AAAAAAAAAAAAAA

SEQ ID NO: 120

>I'-B-Yamagata-B-120A mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
CCAGCAACGCCGACCGGAUCUGCACC GGCAUCACCAGCUCCAACAGCCCC
5 CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGGCGUGA
UCCCCCUGACCACCACCCCCACCAAGUCCUACUUCGCCAACCUGAAGGGC
ACCCGGACCCGCGGCAAGCUGUGCCCCGACUGCCUGAACUGCACCAGCCU
GGACGUGGCCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCCCAGCGCCA
AGGCCUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCC
10 AUCAUGCACGACCGCACCAAGA UCCGGCAGCUGCCCAACCUGCUGCGGG
GGUACGAGAAGA UCCGCCUGUCCACCCAGAACGUCAUCGACGCCGAGAA
GGCCCCCGGCGGGCCCUACCGGCUGGGCACCAGCGGGUCCUGCCCCAACG
CCACCAGCAAGAUCGGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCCAA
GGACAACUACAAGAACGCCACCAACCCCCUGACCGUCGAGGUGCCCUAC
15 AUCUGCACCAGGGGGGAGGACCAGAUCACGGUCUGGGGCUUCCACUCCG
ACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGCAGAA
GUUCACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAU
GGGGACUUCCCGGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCC
GCAUCGUGGUCGACUACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAU
20 CGUGUACCAGCGGGGGGUCCUGCUGCCGCAGAAGGUGUGGUGCGCCAGC
GGCCGCUCCAAGGUCAUCAAGGGGAGCCUGCCCCUCAUCGGCGAGGCCG
ACUGCCUGCACGAGGAGUACGGGGGCCUCAACAAGUCCAAGCCGUACUA
CACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAUCUGGGUGAAG
ACCCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCU
25 GCUCAAGGAGCGCGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGG
GGCUGGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCG
CCCACGGGGUCGCGGUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAU
CAACAAGAUCAACCAAGAACCUGAACAGCCUCUCCGAGCUGGAGGUCAAG
AACCUC CAGCGCCUGAGCGGCGCGAUGGACGAGCUCCACAACGAGAUCC
30 UGGAGCUCGACGAGAAGGUGGACGACCUGCGGGCCGACACGAUCUCGUC
CCAGAU CGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAUCAUCAACUCG
GAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGG
GCCCCUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAA
GUGCAACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUUCAACGCC
35 GGCGAGUUCAGCCUGCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUC
CCUGAACGACGACGGGCUCGACAACCACACGAUCCUGCUCUACUACAGC

ACCGCGGCCUCGUCCCUGGCGGUCACGCUCAUGCUGGCCAUCUUCAUCG
UGUACAUGGUCAGCCGCGACAACGUGUCGUGCUGCAUCUGCCUCUGAUG
ACUCGAGGCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAAGGUUCCUUUG
UUCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAG
5 CAUCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAA
AA
AA
AAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 121

10 >I'-H1N1-HA0-mut-B-120A mRNA
GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCU
UCACCACCGCCAACGCCGACACCCUGUGCAUCGGCUACCACGCCAACAAC
AGCACCGACACCGUGGACACCGUGCUGGAGAAGAACGUGACCGUGACCC
15 ACAGCGUGAACCUGCUGGAGGACAAGCACAACGGCAAGCUGUGCAAGCU
GCGGGGCGUGGGCCCCCUGCACCUGGGCAAGUGCAACAUCGCCGGCUGG
AUCCUGGGCAACCCCGAGUGCGAGUCCCUGAGCACCGCCCGCUCCUGGA
GCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGG
GGACUUCAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCC
20 AGCUUCGAGCGCUUCGAGAUCUUCCCCAAGACCUCAGCUGGCCCAACC
ACGACUCCGACAACGGCGUGACCGCCGCCUGCCCCACGCCGGGGCCAAG
AGCUUCUACAAGAACCUGAUCUGGCUGGUGAAGAAGGGCAAGUCCUACC
CCAAGAUAACCAGACCUACAUAACGACAAGGGGAAGGAGGUCCUCGU
GCUGUGGGGCAUCCACCACCCCCCACC AUCCGCCGACCAGCAGAGCCUCU
25 ACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCGCCGUACAGCAA
GAAGUUCAAGCCCGAGAUCGCCACCCGCCCAAGGUCCGGGACCAGGAG
GGCCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUA
CCUUCGAGGCCACCGGCAACCUCGUCGCCCCGCGGUACGCCUUCACGAUG
GAGCGCGACGCCGGGUCCGGCAUCAUCAUCAGCGACACCCCGUGCACG
30 ACUGCAACACGACCUGCCAGACGCCCGAGGGGGCGAUCAACACCUCCCU
GCCGUUCCAGAACGUCCACCCGAUCACGAUCGGCAAGUGCCCCAAGUAC
GUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGCAACGUCCCCUC
CAUCCAGAGCCAGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGG
GGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACG
35 AGCAGGGCUCCGGGUACGCGGGCCGACCUCAAGUCGACGCAGAACGCGAU
CGACAAGAUCACCAACAAGGUCAACAGCGUGAUCGAGAAGAUGAACACG

CAGUUCACCGCCGUCGGCAAGGAGUUCAACCACCUGGAGAAGCGCAUCG
 AGAACCUCAACAAGAAGGUGGACGACGGGUUCCUGGACAUCUGGACGUA
 CAACGCGGAGCUCCUGGUCCUCCUGGAGAACGAGCGGACCCUCGACUAC
 CACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCGCAACCAGCUCA
 5 AGAACAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACCACAA
 GUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUCGACUAC
 CCGAAGUACAGCGAGGAGGCGAAGCUGAACCAGGAGAGAAGAUCCGACGGG
 UCAAGCUCGACUCCACGCGCAUCUACCAGAUCCUGGCCAUCUACUCGACC
 GUGGCGAGCUCCUCGUCUCCUGGUGGUCUCGUCGCGGCCAUCAGCUUCU
 10 GGAUGUGCUCCAACGGGUCGUCGAGUGCCGGAUCUGCAUCUGAUGACU
 CGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUC
 CCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAU
 CUGGAUUCUGCCUAAUAAAAACAUAUUAUUUUCAUUGCAAAAAAAAAA
 AA
 15 AA
 AAAAAAAAAAAAAAAAAA

SEQ ID NO: 122

>I'-H3N2-HA0-mut-B-120A mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA
 20 GGAUAGCACUAUAUAUGAAGACCAUCAUCGCCUGAGCUACAUCUGU
 GCCUGGUGUUCGCCCAGAAGAUCGCCGCAACGACAACAGCACCGCCACC
 CUGUGCCUGGGCCACCACGCCGUGCCCAACGGCACCAUCGUGAAGACCA
 UCACCAACGACCGGAUCGAGGUGACCAACGCCACCGAGCUGGUGCAGAA
 CUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCAGAUCCUGGACGGCG
 25 GCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCAGUGCGACGG
 GUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAAC
 UCCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCCU
 GGUGGCCAGCUCCGGCACCCUGGAGUUAAGAACGAGAGCUUCAAACUGG
 ACCGGGGUCAAGCAGAACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCA
 30 GCUCCAGCUUCUUCUCCCGGCUGAACUGGCUCACCCACCUGAACUACACC
 UACCCCGCCCUC AACGUGACCAUGCCCAACAACGAGCAGUUCGACAAGC
 UGUACAUCUGGGGCGUCCACCACCCAGCACCGACAAGGACCAGAUCUC
 CCUCUUCGCCCAGCCAGCGGGCGCAUACCGUGUCCACGAAGCGGAGCC
 AGCAGGCCGUAUCCCCAACAUCCGGCUCCCGCCCGCGGAUCCGCGACAUC
 35 CCCAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUC
 UGCUCAUCAACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUCAA

GAUCCGGUCCGGCAAGAGCUCGAUCAUGCGCUCCGACGCCCCGAUCGGG
AAGUGCAAGAGCGAGUGCAUCACCCCAACGGCUCGAUCCCGAACGACA
AGCCCUUCCAGAACGUCAACCGGAUCACGUACGGGGCCUGCCCGCGCUA
CGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCAGGAACGUCCCC
5 GAGAAGCAGACCCAGGGGAUCUUCGGGCGCCAUCGCGGGGUUCAUCGAGA
ACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAGAA
CAGCGAGGGCCGCGGGCAGGCCGCGGACCUC AAGUCGACGCAGGCCGCG
AUCGACCAGAUCAACGGCAAGCUGAACCGGCUC AUCCGGAAGACGAACG
AGAAGUUCCACCAGAU CGAGAAGGAGUUCUCCGAGGUCGAGGGCCGCGU
10 GCAGGACCUGGAGAAGUACGUCGAGGACACCAAGAUCGACCUCUGGAGC
UACAACGCCGAGCUGCUCGUGGCGCUGGAGAACCAGCACACCAUCGACC
UCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGAAGAAGCAGCU
CCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUCUACCAC
AAGUGCGACAACGCGUGCAUCGGGUCCAUCCGCAACGAGACCUACGACC
15 ACAACGUCUACCGGGACGAGGCCCUGAACAACCGCUUCCAGAUCAAGGG
CGUGGAGCUCAAGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUC
GCGAUGUCCUGCUUCCUCCUGUGCAUCGCCCUCUGGGCUUCAUCAUGU
GGGCGUGCCAGAAGGGGAACAUCGGUGCAACAUCUGCAUCUGAUGACU
CGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUC
20 CCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAU
CUGGAUUCUGCCUAAUAAAAACA UUUUUUUAUUGCAAAAAAAAAA
AAA
AAA
AAAAAAAAAAAAAAAAA

25

SEQ ID NO: 123

>I'-B-Victoria-HA0-mut-B-120A mRNA

GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAA
GGAUAGCACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
CCAGCAACGCCGACCGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCC
30 CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGA
UCCCCCUGACCACCACCCCCACCAAGUCCCACUUCGCCAACCUGAAGGGC
ACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCUGAACUGCACCGACC
UGGACGUGGCCCUUGGGCCGCCCAAGUGCACCGGCAAGAUCCCCAGCGCC
CGGGUGUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC
35 CCAUCAUGCACGACCGCACCAAGAUCGGCAGCUGCCCAACCUGCUGCGG
GGGUACGAGCACGUCCGCCUGUCCACCCACAACGUGAUCAACGCCGAGG

ACGCCCCCGGCCGGCCCUACGAGAUCGGGACCAGCGGCUCCUGCCCCAAC
 AUCACCAACGGGAACGGCUUCUUCGCCACCAUGGCCUGGGCCGUCCCCA
 AGAACAAGACCGCCACCAACCCCCUGACCAUCGAGGUGCCGUACAUCUG
 CACCGAGGGGGGAGGACCAGAUCACCGUCUGGGGCUUCCACAGCGACAAC
 5 GAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAAGUUCA
 CGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGG
 CUUCCCGAACCAGACGGGAGGACGGGGGGCCUGCCCCAGUCCGGGCGCAUC
 GUGGUCGACUACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGU
 ACCAGCGGGGCAUCCUGCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGGCG
 10 CAGCAAGGUGAUCAAGGGCUCCCUGCCCCUCAUCGGGGGAGGCCGACUGC
 CUGCACGAGAAGUACGGCGGGCUCAACAAGAGCAAGCCGUACUACACCG
 GCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUGGGGUCAAGACCCC
 GCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCUGCUC
 AAGGAGCAGGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGU
 15 GGGAGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCCACGGGGCCCA
 CGGCGUGGCGGUCGCCGCGGACCUCAAGAGCACCCAGGAGGCCAUCAAC
 AAGAUACGAAGAACCUGAACUCCCUCAGCGAGCUGGAGGUGAAGAACC
 UCCAGCGCCUGUCCGGGGCGAUGGACGAGCUCCACAACGAGAUCUGGA
 GCUCGACGAGAAGGUCGACGACCUGCGGGCCGACACCAUCAGCUCGCAG
 20 AUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAUAACAGCGAGG
 ACGAGCACCUUGCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCC
 CUCGGCGGUCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGC
 AACCAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCCGGGG
 AGUUCUCCCUGCCGACCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUG
 25 AACGACGACGGCCUCGACAACCACACCAUCCUGCUCUACUACUCCACCGC
 GGCCAGCUCGCUGGCGGUGACGCUCAUGAUCGCCAUCUUCGUCGUGUAC
 AUGGUCUCCCGCGACAACGUGAGCUGCUCGAUCUGCCUGUGAUGACUCG
 AGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAAGGUUCCUUGUUCUCC
 UAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCAUCU
 30 GGAUUCUGCCUAAUAAAAACAUUUAUUUUAUUGCAAAAAAAAAAAAAA
 AA
 AA
 AAAAAAAAAAAAAAA

SEQ ID NO: 124

35 >I'-B-Yamagata-HA0-mut-B-120A mRNA
 GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAA

GGAUAGCACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGA
 CCAGCAACGCCGACCGGAUCUGCACC GGCAUCACCAGCUCCAACAGCCCC
 CACGUGGUGAAGACCGCCACCCAGGGCGAGGUGAACGUGACCGGGCGUGA
 UCCCCCUGACCACCACCCCCACCAAGUCCUACUUCGCCAACCUGAAGGGC
 5 ACCCGGACCCGCGGCAAGCUGUGCCCCGACUGCCUGAACUGCACCGACCU
 GGACGUGGCCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCCCAGCGCCA
 AGGCCUCCAUCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCC
 AUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGGG
 GGUACGAGAAGAUCCGCCUGUCCACCCAGAACGUCAUCGACGCCGAGAA
 10 GGCCCCCGGCGGGCCCUACCGGCUGGGCACCAGCGGGUCCUGCCCCAACG
 CCACCAGCAAGAUCGGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCAA
 GGACAACUACAAGAACGCCACCAACCCCCUGACCGUCGAGGUGCCCUAC
 AUCUGCACCGAGGGGGAGGACCAGAUACGGUCUGGGGGCUUCCACUCCG
 ACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGCAGAA
 15 GUUCACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAU
 GGGGACUUCCCGGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCC
 GCAUCGUGGUCGACUACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAU
 CGUGUACCAGCGGGGGGUGCCUGCUGCCGCAGAAGGUGUGGUGCGCCAGC
 GGCCGCUCCAAGGUCAUCAAGGGGAGCCUGCCCCUCAUCGGCGAGGCCG
 20 ACUGCCUGCACGAGGAGUACGGGGGGCCUCAACAAGUCCAAGCCGUACUA
 CACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAUCUGGGUGAAG
 ACCCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGGCCGCCGCAAGCU
 GCUCAAGGAGCAGGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGG
 GGCUGGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCG
 25 CCCACGGGGUGCGGGUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAU
 CAACAAGAUACCAAGAACCUGAACAGCCUCUCCGAGCUGGAGGUCAAG
 AACCUCAGCGCCUGAGCGGCGCGAUGGACGAGCUCCACAACGAGAUCC
 UGGAGCUCGACGAGAAGGUGGACGACCUGCGGGCCGACACGAUCUCGUC
 CCAGAU CGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAUCAUCAACUCG
 30 GAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGG
 GCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAA
 GUGCAACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUUCAACGCC
 GGCGAGUUCAGCCUGCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUC
 CCUGAACGACGACGGGCUCGACAACCACACGAUCCUGCUCUACUACAGC
 35 ACCGCGGCCUCGUCCUGGGCGGUCACGCUCAUGCUGGCCAUUCUUAUCG
 UGUACAUGGUCAGCCGCGACAACGUGUCGUGCUCCAUCUGCCUCUGAUG

ACUCGAGGCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAAGGUUCCUUUG
 UUCCCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAG
 CAUCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAA
 AA
 5 AA
 AAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 125

>Influenza A virus H1N1 (A/Wisconsin/588/2019) HA mRNA
 AUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUCACCACCGCCAACG
 10 CCGACACCCUGUGCAUCGGCUACCACGCCAACACAGCACCGACACCGUG
 GACACCGUGCUGGAGAAGAACGUGACCGUGACCCACAGCGUGAACCUGC
 UGGAGGACAAGCACAACGGCAAGCUGUGCAAGCUGCGGGGCGUGGCCCC
 CCUGCACCUUGGGCAAGUGCAACAUCGCCGGCUGGAUCCUGGGCAACCCC
 GAGUGCGAGUCCUGAGCACCGCCCGCUCCUGGAGCUACAUCGUGGAGA
 15 CCUCCAACAGCGACAACGGCACCUGCUACCCCGGGGACUUCAUCAACUAC
 GAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGAGCGCUUCG
 AGAUCUUCCCCAAGACCUCAGCUGGCCCAACCACGACUCCGACAACGGC
 GUGACCGCCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCU
 GAUCUGGCUGGUGAAGAAGGGCAAGUCCUACCCCAAGAUCAACCAGACC
 20 UACAUCAACGACAAGGGGAAGGAGGUCCUCGUGCUGUGGGGCAUCCACC
 ACCCCCCACCAUCGCCGACCAGCAGAGCCUCUACCAGAACGCCGACGCC
 UACGUCUUCGUGGGGACCUCGCGUACAGCAAGAAGUUAAGCCCGAGA
 UCGCCACCCGCCCCAAGGUCCGGGACCAGGAGGGCCGCAUGAACUACUA
 CUGGACCCUGGUGGAGCCGGGGGACAAGAUCACCUUCGAGGCCACCGGC
 25 AACCUCGUCGCCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGU
 CCGGCAUCAUCAUCAGCGACACCCCGUGCACGACUGCAACACGACCUGC
 CAGACGCCCCGAGGGGGCGAUCAACACCUCCCUGCCGUUCCAGAACGUCC
 ACCCGAUCACGAUCGGCAAGUGCCCCAAGUACGUGAAGUCCACCAAGCU
 CCGGCUGGCCACGGGGCUCCGCAACGUCCCCUCCAUCCAGAGCCGGGGCC
 30 UGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGGGGCGUGGACCGGGAUGGU
 GGACGGCUGGUACGGGUACCACCACCAGAACGAGCAGGGCUCCGGGUAC
 GCGGCCGACCUCAAGUCGACGCAGAACGCGAUCGACAAGAUCACCAACA
 AGGUCAACAGCGUGAUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGG
 CAAGGAGUUAACCACCUGGAGAAGCGCAUCGAGAACCUCAACAAGAAG
 35 GUGGACGACGGGUUCCUGGACAUCUGGACGUACAACGCGGAGCUCCUGG
 UCCUCCUGGAGAACGAGCGGACCCUCGACUACCACGACUCCAACGUGAA

GAACCUGUACGAGAAGGUCCGCAACCAGCUCAAGAACAACGCCAAGGAG
AUCGGCAACGGGUGCUUCGAGUUCUACCACAAGUGCGACAACACGUGCA
UGGAGUCGGUGAAGAACGGCACCUACGACUACCCGAAGUACAGCGAGGA
GGCGAAGCUGAACCGGGAGAAGAUCGACGGGGUCAAGCUCGACUCCACG
5 CGCAUCUACCAGAUCCUGGCCAUCUACUCGACCGUGGGCGAGCUCCCUCG
UCCUGGUGGUCUCGCUCGGCGCCAUCAGCUUCUGGAUGUGCUCCAACGG
GUCGCUGCAGUGCCGGAUCUGCAUCUGA

SEQ ID NO: 126

>Influenza A virus H3N2 (A/Cambodia/e0826360/2020) HA mRNA
10 AUGAAGACCAUCAUCGCCCUGAGCUACAUCUGUGCCUGGUGUUCGCCC
AGAAGAUCCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCAC
CACGCCGUGCCCAACGGCACCAUCGUGAAGACCAUCACCAACGACCGGA
UCGAGGUGACCAACGCCACCGAGCUGGUGCAGAACUCCAGCAUCGGCGA
GAUCUGCGACUCCCCCACCAGAUCCUGGACGGCGGCAACUGCACCCUGA
15 UCGACGCCUGCUGGGCGACCCCCAGUGCGACGGGUUCCAGAACAAGGA
GUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACUCCAACUGCUACCCC
UACGACGUGCCCGACUACGCCAGCCUGCGGUCCCUGGUGGCCAGCUCCG
GCACCCUGGAGUUCAAGAACGAGAGCUUCAAACUGGACCGGGGUCAAGCA
GAACGGCACCUCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCU
20 CCCGGCUGAACUGGCUCACCCACCUGAACUACACCUACCCCGCCCUCAAC
GUGACCAUGCCCAACAACGAGCAGUUCGACAAGCUGUACAUCUGGGGCG
UCCACCACCCAGCACCGACAAGGACCAGAUUCUCCUCUUCGCCCAGCCC
AGCGGGCGCAUACCCGUGUCCACGAAGCGGAGCCAGCAGGCCGUAUCC
CCAACAUCGGCUCCCGCCCGCGGAUCCGCGACAUCCCCAGCCGGAUCUCC
25 AUCUACUGGACGAUCGUGAAGCCGGGGGACAUCCUGCUCAUACAACAGCA
CCGGCAACCUGAUCGCCCCCGCGGGUACUUAAGAUCCGGUCCGGCAA
GAGCUCGAUCAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAG
UGCAUACCCCCAACGGCUCGAUCCCGAACGACAAGCCCUUCCAGAACGU
CAACCGGAUCACGUACGGGGCCUGCCCGCGCUACGUGAAGCAGUCCACG
30 CUCAAGCUGGCCACCGGCAUGCGGAACGUCCCCGAGAAGCAGACCCGCG
GGAUCUUCGGCGCCAUCGCGGGGUUCAUCGAGAACGGCUUGGGAGGGGAU
GGUGGACGGCUGGUACGGGUUCCGGCACCAAGACGAGGGGCCGCGGG
CAGGCCGCGGACCUCAAGUCGACGCAGGCCGCGAUCCAGCAGAUCAACG
GCAAGCUGAACCGGCUCAUCGGGAAGACGAACGAGAAGUCCACCAGAU
35 CGAGAAGGAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAG
UACGUCGAGGACACCAAGAUCGACCUCUGGAGCUACAACGCCGAGCUGC

UCGUGGCGCUGGAGAACCAGCACACCAUCGACCUCACGGACUCGGAGAU
 GAACAAGCUGUUCGAGAAGACGAAGAAGCAGCUCCGGGAGAACGCCGAG
 GACAUGGGGAACGGCUGCUUCAAGAUCUACCACAAGUGCGACAACGCGU
 GCAUCGGGUCCAUCCGCAACGAGACCUACGACCACAACGUCUACCGGGA
 5 CGAGGCCCUGAACAACCGCUUCCAGAUCAAGGGCGUGGAGCUCAAGAGC
 GGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCC
 UCCUGUGCAUCGCCCUCUGGGGCUUCAUCAUGUGGGCGUGCCAGAAGGG
 GAACAUCCGGUGCAACAUCUGCAUCUGA

SEQ ID NO: 127

10 >Influenza B virus Victoria lineage (B/Washington/02/2019) HA mRNA
 AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACC
 GGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGAC
 CGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACC
 ACCCCACCAAGUCCACUUCGCCAACCUGAAGGGCACCGAGACCCGGGG
 15 CAAGCUGUGCCCCAAGUGCCUGAACUGCACCAGCUGGACGUGGGCCUG
 GGCCGCCCCAAGUGCACCGGCAAGAUGCCAGCGCCCGGGUGUCCAUCCU
 GCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGCACGACC
 GCACCAAGAUCCGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGCACGU
 CCGCCUGUCCACCCACAACGUGAUCAACGCCGAGGACGCCCCGGCCGGC
 20 CCUACGAGAUCCGGGACCAGCGGCUCUGCCCCAACAUACCAACGGGAA
 CGGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCCAAGAACAAGACCGCCA
 CCAACCCCCUGACCAUCGAGGUGCCGUACAUCUGCACCGAGGGGGAGGA
 CCAGAUACCGUCUGGGGCUUCCACAGCGACAACGAGACGCAGAUGGCC
 AAGCUGUACGGGGACUCCAAGCCCCAGAAGUUCACGAGCUCCGCCAACG
 25 GCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCUUCCCGAACAGAC
 GGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGACUACAUG
 GUGCAGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCC
 UGCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUCAA
 GGGCUCCCUGCCCCUCAUCGGGGAGGCCGACUGCCUGCACGAGAAGUAC
 30 GGCGGGCUCAACAAGAGCAAGCCGUACUACACCGGCGAGCACGCCAAGG
 CCAUCGGGAACUGCCCCAUCUGGGUCAAGACCCCGCUGAAGCUCGCCAA
 CGGCACGAAGUACCGGCCCCCGGCCAAGCUGCUCAAGGAGCGCGGGUUC
 UUCGGCGCCAUCGCCGGGUUCCUGGAGGGGCGGGUGGGAGGGCAUGAUCG
 CCGGGUGGCACGGCUACACGUCCCACGGGGCCCACGGCGUGGCGGUUCG
 35 CGCGGACCUCAAGAGCACCCAGGAGGCCAUCAACAAGAUCACGAAGAAC
 CUGAACUCCUCAGCGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCG

GGGCGAUGGACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGU
 CGACGACCUGCGGGCCGACACCAUCAGCUCGCAGAUCGAGCUCGCGGUG
 CUGCUCUCCAACGAGGGCAUCAUCAAACAGCGAGGACGAGCACCUGCUCG
 CCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCCCUCGGCGGUCGAGAU
 5 CGGCAACGGGUGCUUCGAGACCAAGCACAAGUGCAACCAGACGUGCCUC
 GACCGCAUCGCCGCGGGCACGUUCGACGCCGGGGAGUUCUCCCUGCCGA
 CCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUGAACGACGACGGCCUC
 GACAACCACACCAUCCUGCUCUACUACUCCACCGCGGCCAGCUCGCUGGC
 GGUGACGCUCAUGAUCGCCAUCUUCGUCGUGUACAUGGUCUCCCGCGAC
 10 AACGUGAGCUGCUCGAUCUGCCUGUGA

SEQ ID NO: 128

>Influenza B virus /Yamagata lineage (B/Phuket/3073/2013) HA mRNA

AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACC
 GGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGAC
 15 CGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACC
 ACCCCACCAAGUCCUACUUCGCCAACCUGAAGGGCACCCGGACCCGCGG
 CAAGCUGUGCCCCGACUGCCUGAACUGCACCGACCUGGACGUGGCCCCUG
 GGCCGGCCCAUGUGCGUGGGCACCACCCCAGCGCCAAGGCCUCCAUCCU
 GCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGCACGACC
 20 GCACCAAGAUCCGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGAAGAU
 CCGCCUGUCCACCCAGAACGUCAUCGACGCCGAGAAGGCCCCCGGCGGGC
 CCUACCGGCUGGGCACCAAGCGGGUCCUGCCCCAACGCCACCAGCAAGAUC
 GGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCAAGGACAACUACAAGA
 ACGCCACCAACCCCCUGACCGUCGAGGUGCCCUACAUCUGCACCGAGGGG
 25 GAGGACCAGAUACGGUCUGGGGCUUCCACUCCGACAACAAGACCCAGA
 UGAAGUCCCUGUACGGGGACUCCAACCCGCAGAAGUUCACGAGCUCCGC
 CAACGGCGUGACCACGCACUACGUCAGCCAGAUCCGGGGACUUCCCGGAC
 CAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUCGACU
 ACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAUCGUGUACCAGCGGGG
 30 GGUCCUGCUGCCGCAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUC
 AUCAAGGGGAGCCUGCCCCUCAUCGGCGAGGCCGACUGCCUGCACGAGG
 AGUACGGGGGCCUCAACAAGUCCAAGCCGUACUACACGGGGAAGCACGC
 CAAGGCCAUCGGCAACUGCCC GAUCUGGGUGAAGACCCCCUGAAGCUC
 GCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCUGCUCAAGGAGCGCG
 35 GCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCUGGGAGGGGAU
 GAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUUCGCG

GUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAUCAACAAGAUACCA
AGAACCUGAACAGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCU
GAGCGGCGCGAUGGACGAGCUCCACAACGAGAUCUGGAGCUCGACGAG
AAGGUGGACGACCUGCGGGGCCGACACGAUCUCGUCCCAGAUCGAGCUCG
5 CGGUCCUGCUCAGCAACGAGGGGAUCAUAACUCGGAGGACGAGCACCU
GCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGCCCGUCCGCGGUG
GACAUCGGGAACGGCUGCUUCGAGACCAAGCACAAGUGCAACCAGACGU
GCCUCGACCGCAUCGCCGCGGGGACCUUCAACGCCGGCGAGUUCAGCCU
GCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUCCCUGAACGACGACG
10 GGCUCGACAACCACACGAUCCUGCUCUACUACAGCACCGCGGCCUCGUCC
CUGGCGGUCACGCUCAUGCUGGCCAUCUUCAUCGUGUACAUGGUCAGCC
GCGACAACGUGUCGUGCUCCAUCUGCCUCUGA

SEQ ID NO: 129

15 >SARS-CoV-2 BA.4 spike protein mRNA
AUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGA
ACCUGAUCACAAGAACACAGUCGUACACCAACAGCUUCACAAGAGGCGU
GUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAA
GACCUGUUUCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCA
20 GCGGCACCAACGGCACCAAGAGAUUCGACAACCCCGUGCUGCCCUUCAA
CGACGGCGUGUACUUCGCUAGCACCGAAAAGAGCAACAUCAGAGGC
UGGAUCUUCGGCACCAACCUAGACUCCAAGACACAAAGCCUGCUGAUCG
UUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUUCAGUUCUG
CAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAAGAGCUGG
25 AUGGAGAGCGAGUUCAGAGUCUACAGCAGCGCCAACAACUGCACCUUCG
AGUACGUGUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAA
UUUCAAGAACCUAAGAGAGUUCGUGUUAAGAACAUCGACGGCUACUUC
AAGAUCUACAGCAAGCACACCCCCAUCAACCUGGGGAGAGACCUGCCCC
AAGGCUUCAGCGCCCUUGGAGCCCCUGGUGGACCUGCCCAUCGGCAUCAA
30 CAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUAGCUACCUAACC
CCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACG
UGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGG
CACUAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACC
AAGUGCACCCUCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAA
35 GCAACUUCAGAGUGCAGCCCACCGAGAGCAUCGUGAGAUUCCCCAACAU
CACCAACCUGUGCCCCUUCGACGAGGUGUUAACGCCACAAGAUUCGCU

AGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGCCGACU
 ACAGCGUGCUGUACAACUUCGCUCCCUUCUUCGCCUUCAAGUGCUAUGG
 CGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCC
 GACAGCUUCGUGAUCAGAGGCAACGAGGUGAGCCAGAUUGCCCCCGGGC
 5 AAACCGGCAACAUCGCCGACUACAACUACAAGCUGCCCCGACGACUUCAC
 CGGCUGCGUGAUCGCGUGGAACAGCAAUAAGCUGGAUUCCAAGGUGGGC
 GGCAACUACAACUACCGGUACAGACUGUUCAGAAAGAGCAACCUGAAGC
 CCUUCGAGAGAGACAUCAGCACCGAGAUCUACCAAGCCGGCAACAAGCC
 CUGCAACGGCGUGGGCGGGCGUCAACUGCUACUUCCCCCUGCAGAGCUAC
 10 GGCUUUAGGCCACCACGCGUGGGCCAUCAGCCCUACAGAGUGGUGG
 UGCUGAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCAA
 AAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUAACUUUAU
 GGCCUUAAGGGCACGGGCGUGCUGACCGAGAGCAACAAGAAGUUCCUGC
 CAUUUCAGCAAUUCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAG
 15 AGACCCUCAGACCCUGGAGAUCCUGGACAUCACCCCCUGUAGCUUCGGC
 GGCGUGAGCGUGAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCCG
 UGCUGUACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGC
 CGAUCAGCUGACCCCCACCUGGAGAGUAUACAGCACCGGAAGCAACGUG
 UUUCAGACAAGAGCCGGCUGCCUGAUCGGCGCCGAGUACGUGAACAACA
 20 GCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUA
 GACACAGACCAAGAGCCACCGGAGAGCUAGAAGCGUGGCUUCUCAGAGC
 AUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACAGCGUGGCCUACA
 GCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCACC
 GAGAUCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGU
 25 ACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGCUGCUGCAGUACGG
 CAGCUUCUGCACACAGCUGAAGAGAGCCCUGACCGGCAUCGCCGUGGAG
 CAAGACAAGAACACCCAAGAGGUGUUCGCCCAAGUGAAGCAGAUCUACA
 AGACCCCCCCCCAUCAAGUACUUCGGCGGCUUCAACUUCAGCCAAAUCCU
 GCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGGACCUGCUG
 30 UUCAACAAGGUGACCCUGGCCGACGCCGGCUUCAUCAAGCAGUACGGCG
 ACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUU
 CAAUGGACUCACCGUGCUCCCCCCCCUGCUGACCGACGAGAUGAUCGCUC
 AGUACACCUCUGCCCUGCUGGGCGGGAACCAUAACAAGCGGUUGGACAUU
 UGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCCAUGCAGAUGGCCUAC
 35 AGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACGAGAAUCAGA
 AGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUCCAAGACAG

CCUGAGCAGCACCGCUAGCGCCCUUGGGCAAGCUGCAAGACGUGGUGAAU
 CACAACGCCCAAGCCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAAGU
 UCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCCUGUCUAGACUGGACCC
 CCCCAGGGCCGAGGUGCAGAU CGACAGACUGAUCACCGGCAGACUGCAG
 5 AGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGCCGCCGAGAUA
 GAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUGGG
 GCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGC
 UUCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCACGUGACCUACG
 UGCCCCGCCAAGAGAAGAACUUCACCACCGCUCCCCGCCAUCUGCCACGAC
 10 GGCAAGGCCCCACUCCCCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCC
 ACUGGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUCACCAC
 CGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUG
 AACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGCUUCAAGG
 AGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGUGGACCU
 15 GGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAAGGAG
 AUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCG
 ACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUUGGUA
 CAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACC
 AUCAUGCUGUGCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCU
 20 GCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGU
 GCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 130

>Influenza A virus H1N1 HA0 mutant mRNA

AUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUCACCACCGCCAACG
 25 CCGACACCCUGUGCAUCGGCUACCACGCCAACACAGCACCGACACCGUG
 GACACCGUGCUGGAGAAGAACGUGACCGUGACCCACAGCGUGAACCUGC
 UGGAGGACAAGCACAACGGCAAGCUGUGCAAGCUGCGGGGCGUGGCCCC
 CCUGCACCUUGGGCAAGUGCAACAUCGCCGGCUGGAUCCUGGGCAACCCC
 GAGUGCGAGUCCCUGAGCACCGCCCCGCUCCUGGAGCUACAUCGUGGAGA
 30 CCUCCAACAGCGACAACGGCACCUGCUACCCCGGGGACUUCAUCAACUAC
 GAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGAGCGCUUCG
 AGAUCUUCCCCAAGACCUCCAGCUGGCCCAACCACGACUCCGACAACGGC
 GUGACCGCCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCU
 GAUCUGGCUGGUGAAGAAGGGCAAGUCCUACCCCAAGAUAACCAGACC
 35 UACAUCAACGACAAGGGGAAGGAGGUCCUCGUGCUGUGGGGCAUCCACC
 ACCCCCCACCAUCGCCGACCAGCAGAGCCUCUACCAGAACGCCGACGCC

UACGUCUUCGUGGGGACCUCCTGGUACAGCAAGAAGUUCAAGCCCGAGA
 UCGCCACCCGCCCCAAGGUCCGGGACCAGGAGGGCCGCAUGAACUACUA
 CUGGACCCUGGUGGAGCCGGGGGACAAGAUCACCUUCGAGGCCACCGGC
 AACCUCGUCGCCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGU
 5 CCGGCAUCAUCAUCAGCGACACCCCCGUGCACGACUGCAACACGACCUGC
 CAGACGCCCCGAGGGGGCGAUCAACACCUCCCUGCCGUUCCAGAACGUCC
 ACCCGAUCACGAUCGGCAAGUGCCCCAAGUACGUGAAGUCCACCAAGCU
 CCGGCUGGCCACGGGGCUCCGCAACGUCCCCUCCAUCCAGAGCCAGGGCC
 UGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGGGGCGUGGACCGGGAUGGU
 10 GGACGGCUGGUACGGGUACCACCACCAGAACGAGCAGGGCUCGCGGUAC
 GCGGCCGACCUCAAGUCGACGCAGAACGCGAUCGACAAGAUCACCAACA
 AGGUCAACAGCGUGAUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGG
 CAAGGAGUUCAACCACCUGGAGAAGCGCAUCGAGAACCUCAACAAGAAG
 GUGGACGACGGGUUCCUGGACAUCUGGACGUACAACGCGGAGCUCCUGG
 15 UCCUCCUGGAGAACGAGCGGACCCUCGACUACCACGACUCCAACGUGAA
 GAACCUGUACGAGAAGGUCCGCAACCAGCUCAAGAACAACGCCAAGGAG
 AUCGGCAACGGGUGCUUCGAGUUCUACCACAAGUGCGACAACACGUGCA
 UGGAGUCGGUGAAGAACGGCACCUACGACUACCCGAAGUACAGCGAGGA
 GGCGAAGCUGAACCAGGGAGAAGAUCGACGGGGUCAAGCUCGACUCCACG
 20 CGCAUCUACCAGAUCCUGGCCAUCUACUCGACCGUGGCGAGCUCCCUCG
 UCCUGGUGGUCUCGCUCGGCGCCAUCAGCUUCUGGAUGUGCUCCAACGG
 GUCGCUGCAGUGCCGGAUCUGCAUCUGA

SEQ ID NO: 131

>Influenza A virus H3N2 HA0 mutant mRNA

25 AUGAAGACCAUCAUCGCCUGAGCUACAUCUGUGCCUGGUGUUCGCCC
 AGAAGAUCCTCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCAC
 CACGCCGUGCCCAACGGCACCAUCGUGAAGACCAUCACCAACGACCGGA
 UCGAGGUGACCAACGCCACCGAGCUGGUGCAGAACUCCAGCAUCGGCGA
 GAUCUGCGACUCCCCCACCAGAUCCUGGACGGCGGCAACUGCACCCUGA
 30 UCGACGCCCUGCUGGGCGACCCCCAGUGCGACGGGUUCCAGAACAAGGA
 GUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACUCCAACUGCUACCCC
 UACGACGUGCCCGACUACGCCAGCCUGCGGUCCCUGGUGGCCAGCUCCG
 GCACCCUGGAGUUAAGAACGAGAGCUUCAAACUGGACCGGGGUCAAGCA
 GAACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCU
 35 CCCGGCUGAACUGGCUCACCCACCUGAACUACACCUACCCCGCCCUAAC
 GUGACCAUGCCCAACAACGAGCAGUUCGACAAGCUGUACAUCUGGGGCG

UCCACCACCCCAGCACCGACAAGGACCAGAUUCCCCUCUUCGCCCAGCCC
 AGCGGGCGCAUCACCGUGUCCACGAAGCGGAGCCAGCAGGCCGUCAUCC
 CCAACAUCGGCUCCCGCCCGCGGAUCCGCGACAUCCCCAGCCGGAUCUCC
 AUCUACUGGACGAUCGUGAAGCCGGGGGACAUCCUGCUCAUCAACAGCA
 5 CCGGCAACCUGAUCGCCCCCGCGGGUACUUCAAGAUCCGGUCCGGCAA
 GAGCUCGAUCAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAG
 UGCAUCACCCCCAACGGCUCGAUCCCGAACGACAAGCCCUUCCAGAACGU
 CAACCGGAUCACGUACGGGGCCUGCCCCGCGCUACGUGAAGCAGUCCACG
 CUCAAGCUGGCCACCGGCAUGCGGAACGUCCCCGAGAAGCAGACCCAGG
 10 GGAUCUUCGGCGCCAUCGCGGGGUUCAUCGAGAACGGCUUGGGAGGGGAU
 GGUGGACGGCUUGGUACGGGUUCCGGCACCAGAACAGCGAGGGCCGCGGG
 CAGGCCGCGGACCUCAAGUCGACGACGGCCGCGAUUCGACCAGAUCAACG
 GCAAGCUGAACCGGCUCAUCGGGAAGACGAACGAGAAGUUCCACCAGAU
 CGAGAAGGAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAG
 15 UACGUCGAGGACACCAAGAUCGACCUCUGGAGCUACAACGCCGAGCUGC
 UCGUGGCGCUGGAGAACCAGCACACCAUCGACCUCACGGACUCGGAGAU
 GAACAAGCUGUUCGAGAAGACGAAGAAGCAGCUCCGGGAGAACGCCGAG
 GACAUGGGGAACGGCUGCUUCAAGAUCUACCACAAGUGCGACAACGCGU
 GCAUCGGGUCCAUCCGCAACGAGACCUACGACCACAACGUCUACCGGGA
 20 CGAGGCCCUGAACAACCGCUUCCAGAUCAAGGGCGUGGAGCUCAAGAGC
 GGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCC
 UCCUGUGCAUCGCCCUCUGGGCUUCAUCAUGUGGGCGUGCCAGAAGGG
 GAACAUCCGGUGCAACAUCUGCAUCUGA

SEQ ID NO: 132

25 >Influenza B virus/Victoria HA0 mutant mRNA
 AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACC
 GGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGAC
 CGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACC
 ACCCCACCAAGUCCACUUCGCCAACCUGAAGGGCACCGAGACCCGGGG
 30 CAAGCUGUGCCCCAAGUGCCUGAACUGCACCGACCUGGACGUGGCCCUG
 GGCCGCCCCAAGUGCACCGGCAAGAUCCCAGCGCCCGGGUGUCCAUCCU
 GCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGCACGACC
 GCACCAAGAUCCGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGCACGU
 CCGCCUGUCCACCCACAACGUGAUCAACGCCGAGGACGCCCCGGCCGGC
 35 CCUACGAGAUCGGGACCAGCGGCUCCUGCCCCAACAUACCAACGGGAA
 CGGCUUCUUCGCCACCAUGGCCUGGGCCGUCCCCAAGAACAAGACCGCCA

CCAACCCCCUGACCAUCGAGGUGCCGUACAUCUGCACCGAGGGGGAGGA
 CCAGAUACACCGUCUGGGGCUUCCACAGCGACAACGAGACGCAGAUGGCC
 AAGCUGUACGGGGACUCCAAGCCCCAGAAGUUCACGAGCUCCGCCAACG
 GCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCUUCCCGAACAGAC
 5 GGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGACUACAUG
 GUGCAGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCC
 UGCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUCAA
 GGGCUCCCUGCCCCUCAUCGGGGAGGCCGACUGCCUGCACGAGAAGUAC
 GGCGGGCUCAACAAGAGCAAGCCGUACUACACCGGCGAGCACGCCAAGG
 10 CCAUCGGGAACUGCCCCAUCUGGGUCAAGACCCCGCUGAAGCUCGCCAA
 CGGCACGAAGUACCGGGCCCCCGGCCAAGCUGCUCAAGGAGCAGGGGUUC
 UUCGGCGCCAUCGCCGGGUUCCUGGAGGGGCGGGUGGGAGGGCAUGAUCG
 CCGGGUGGCACGGCUACACGUCCCCACGGGGCCCACGGCGUGGCGGUCGC
 CGCGGACCUCAAGAGCACCCAGGAGGCCAUAACAAGAUCACGAAGAAC
 15 CUGAACUCCCUCAGCGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCG
 GGGCGAUGGACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGU
 CGACGACCUGCGGGCCGACACCAUCAGCUCGCAGAU CGAGCUCGCGGUG
 CUGCUCUCCAACGAGGGCAUCAUAACAGCGAGGACGAGCACCUGCUCG
 CCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCCCUCGGCGGUCGAGAU
 20 CGGCAACGGGUGCUUCGAGACCAAGCACAAAGUGCAACCAGACGUGCCUC
 GACCGCAUCGCCGCGGGCACGUUCGACGCCGGGGAGUUCUCCCUGCCGA
 CCUUCGACAGCCUCAACAUCACGGCGGGCCUCGCUGAACGACGACGGCCUC
 GACAACCACACCAUCCUGCUCUACUACUCCACCGCGGCCAGCUCGCUGGC
 GGUGACGCUCAUGAUCGCCAUCUUCGUCGUGUACAUGGUCUCCCGCGAC
 25 AACGUGAGCUGCUCGAUCUGCCUGUGA

SEQ ID NO: 133

>Influenza B virus/Yamagata HA0 mutant mRNA

AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACC
 GGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGAC
 30 CGCCACCCAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCCUGACCACC
 ACCCCACCAAGUCCUACUUCGCCAACCUGAAGGGCACCCGGACCCGCGG
 CAAGCUGUGCCCCGACUGCCUGAACUGCACCGACCUGGACGUGGGCCUG
 GGCCGGCCCAUGUGCGUGGGCACCAACCCAGCGCCAAGGCCUCCAUCCU
 GCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGCACGACC
 35 GCACCAAGAUCCGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGAAGAU
 CCGCCUGUCCACCCAGAACGUCAUCGACGCCGAGAAGGCCCCCGGCGGGC

CCUACCGGCUGGGCACCAGCGGGUCCUGCCCCAACGCCACCAGCAAGAUC
GGCUUCUUCGCCACCAUGGCCUGGGCCGUGCCCAAGGACAACUACAAGA
ACGCCACCAACCCCCUGACCGUCGAGGUGCCCUACAUCUGCACCGAGGGG
GAGGACCAGAUCACGGUCUGGGGCUUCCACUCCGACAACAAGACCCAGA
5 UGAAGUCCCUGUACGGGGACUCCAACCCGCAGAAGUUCACGAGCUCCGC
CAACGGCGUGACCACGCACUACGUCAGCCAGAUCGGGGACUUCCCGGAC
CAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUCGACU
ACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAUCGUGUACCAGCGGGG
GGUCCUGCUGCCGCAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUC
10 AUCAAGGGGAGCCUGCCCCUCAUCGGCGAGGCCGACUGCCUGCACGAGG
AGUACGGGGGCCUCAACAAGUCCAAGCCGUACUACACGGGGGAAGCACGC
CAAGGCCAUCGGCAACUGCCCAGUCCGGGUGAAGACCCCCUGAAGCUC
GCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCUGCUCUAAGGAGCAGG
GCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCUGGGAGGGGAU
15 GAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUUCGCG
GUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAUCAACAAGAUACCA
AGAACCUGAACAGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCU
GAGCGGCGCGAUGGACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAG
AAGGUGGACGACCUGCGGGCCGACACGAUCUCGUCCCAGAUCGAGCUCG
20 CGGUCCUGCUCAGCAACGAGGGGAUCAUCAACUCGGAGGACGAGCACC
GCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGCCCGUCCGCGGUG
GACAUCGGGAACGGCUGCUUCGAGACCAAGCACAAGUGCAACCAGACGU
GCCUCGACCGCAUCGCCGCGGGGACCUUCAACGCCGCGGAGUUCAGCCU
GCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUCCUGAACGACGACG
25 GGCUCGACAACCACACGAUCCUGCUCUACUACAGCACCGCGGCCUCCUG
CUGGCGGUCACGCUCAUGCUGGCCAUCUUCAUCGUGUACAUGGUCAGCC
GCGACAACGUGUCGUGCUCCAUCUGCCUCUGA

SEQ ID NO: 134

>Poly A

30 AA
AAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 135

>SARS-CoV-2 BA.4 spike protein amino acid sequence

35 MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPF
FSNVTWFHAISGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKT

QSLIVNNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESEFRVYSSANNCT
 FEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLGRDLPQGFS
 ALEPLVDLPIGINITRFQTLALHRSYLTGDXSSGWTAGAAAYYVGYLQPRTF
 LLKYNENGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSNFRVQPTESIVRFP
 5 NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADYSVLYNFAPFFAFKCYGVSP
 TKLNDLCFTNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVIAWN
 SNKLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIQAGNKPCNGVAGVNCYF
 PLQSYGFRPTYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNF
 NGLKGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVI
 10 TPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLI
 GAEYVNNSYECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTMSLGAENSV
 YSNNIAIPTNFTISVTTEILPVXMTKTSVDCTMYICGDSTECSNLLLQYGSFCT
 QLKRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRS
 FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMI
 15 AQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLI
 ANQFNSAIGKIQDSLSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVL
 NDILSRDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSEC
 VLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVITYVPAQEKNFTTAPAICHDG
 KAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVY
 20 DPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKN
 LNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKG
 CCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 136

CLAIMS

1. A nucleic acid construct, comprising:
 - (a) an open reading frame (ORF), and
 - (b) a 5' untranslated region element (5'UTR),
- 5 wherein the 5'UTR is derived from an ARHGAP gene;
preferably, the ARHGAP is ARHGAP15;
more preferably, the 5'UTR comprises the nucleotide sequence set forth in SEQ ID NO: 44 or a truncation thereof.
- 10 2. The nucleic acid construct according to claim 1, wherein the 5'UTR comprises at least one point mutation capable of being used to inhibit an ATG within the UTR from initiating translation;
preferably, the point mutation is selected from the group consisting of a mutation at any one or more sites of A, T, or G of the ATG sequence in the 5'UTR;
- 15 more preferably, the mutation is to mutate A into G, C, or T.
3. The nucleic acid construct according to claim 1 or 2, wherein the 5'UTR truncation still maintains a function of regulating expression of an ORF-encoded protein; or, the 5'UTR truncation has an enhanced function of regulating expression of an
- 20 ORF-encoded protein as compared to a natural 5'UTR;
preferably, a truncation mode for the 5'UTR truncation comprises: in a sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in a sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end.
- 25 4. The nucleic acid construct according to any one of claims 1-3, wherein the 5'UTR truncation is formed by, in a sequence direction from the 5' end to the 3' end, deleting a sequence of contiguous nucleotides from the 5' end; and/or, in a sequence direction from the 3' end to the 5' end, deleting a sequence of contiguous nucleotides from the 3' end;
- 30 preferably, the 5'UTR truncation comprises a sequence of at least 5 contiguous nucleotides from the sequence set forth in SEQ ID NO: 44, more preferably comprises a sequence of at least 7 contiguous nucleotides from the sequence set forth in SEQ ID NO: 2;
- preferably, the nucleotide at the 5' end of the 5'UTR truncation is the nucleotide at any
- 35 one of positions 1-57, according to natural counting, of the sequence set forth in SEQ ID NO: 44, more preferably the nucleotide at any one of positions 1-13 or 17-29, according to natural counting, of the sequence set forth in SEQ ID NO: 2;
- preferably, the 5'UTR truncation comprises CTATAAT or the nucleotide sequence set

forth in SEQ ID NO: 193.

5. The nucleic acid construct according to any one of claims 1-3, wherein the 5'UTR comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 2, 28-43, 76,
5 and 137-196 or comprises CTATAAT; or a nucleotide sequence having at least 80% identity to any one of the aforementioned sequences.

6. The nucleic acid construct according to any one of claims 1 to 5, further comprising:
(c) a 3' untranslated region element (3'UTR),
10 wherein preferably, the 3'UTR comprises a 3'UTR derived from any one of the gene HBB, ARHGAP, CORO1A, or HPX;
more preferably, the 3'UTR comprises a 3'UTR derived from ARHGAP15.

7. The nucleic acid construct according to claim 6, wherein:
15 the 3'UTR comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 7-10, or, the 3'UTR comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 7-10.

8. The nucleic acid construct according to any one of claims 1 to 7, further comprising:
20 (d) a polyadenylic acid (poly-A) tail.

9. The nucleic acid construct according to claim 8, wherein:
the poly-A tail is selected from the group consisting of HGH polyA, SV40polyA, BGH polyA, rbGlob polyA, and SV40late polyA;
25 preferably, the poly-A tail comprises the nucleotide sequence set forth in SEQ ID NO: 16 or SEQ ID NO: 135 or a nucleotide sequence having at least 80% identity thereto.

10. The nucleic acid construct according to any one of claims 1 to 9, wherein the ORF encodes a viral antigen; preferably, the ORF encodes a coronavirus antigen or an
30 influenza virus antigen;
the coronavirus is preferably SARS-COV-2, and the coronavirus antigen is preferably a spike protein;
more preferably, the spike protein is selected from the group consisting of a spike protein of any one of the viral strain SARS-COV-2, SARS-COV-2 Alpha,
35 SARS-COV-2 Beta, SARS-COV-2 Gamma, SARS-COV-2 Kappa, SARS-COV-2 Delta, or SARS-COV-2 Omicron;
the influenza virus is preferably influenza A virus or influenza B virus, and the antigen

is a hemagglutinin protein (HA) and/or a ceramidase (NA);
more preferably, the influenza virus antigen is selected from the group consisting of a
hemagglutinin protein and/or a neuraminidase of any one of the viral strains influenza
A virus H1N1, influenza A virus H3N2, influenza B virus Victoria, and influenza B
5 virus Yamagata.

11. The nucleic acid construct according to any one of claims 1 to 10, wherein:
a polypeptide sequence encoded by the ORF comprises the amino acid sequence set
forth in any one of SEQ ID NOs: 12, 14, 21, 23, 25, 80, 83-86, 98-101, and 136;
10 preferably, the ORF comprises the nucleotide sequence set forth in any one of SEQ ID
NOs: 13, 15, 20, 22, 24, 81, 87-90, 97, and 102-105, or a nucleotide sequence having
at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to the nucleotide
sequence set forth in any one of SEQ ID NOs: 13, 15, 20, 22, 24, 81, 87-90, 97, and
102-105.

15 12. The nucleic acid construct according to any one of the preceding claims,
comprising the nucleotide sequences set forth in SEQ ID NOs: 18-19 or nucleotide
sequences having at least 80% identity thereto.

20 13. An RNA molecule, comprising:
(a) an open reading frame (ORF),
(b) a 5' untranslated region element (5'UTR), and
(c) a 3' untranslated region element (3'UTR),
wherein the 5'UTR is selected from the group consisting of a 5'UTR derived from any
25 one of the gene ARHGAP, HSPB1, HBB, or CCL13, and the 3'UTR is selected from
the group consisting of a 3'UTR derived from any one of the gene HBB, ARHGAP,
CORO1A, or HPX;
the ARHGAP is preferably ARHGAP15, and the 5'UTR preferably comprises the
nucleotide sequence set forth in SEQ ID NO: 79 or a truncation thereof.

30 14. The RNA molecule according to claim 13, wherein the 5'UTR and 3'UTR are
selected from the group consisting of any one of the following combinations:
1) the 5'UTR is derived from a 5'UTR of ARHGAP, and the 3'UTR is derived from a
3'UTR of any one of HBB, ARHGAP, CORO1A, or HPX;
35 2) the 5'UTR is derived from a 5'UTR of HSPB1, and the 3'UTR is derived from a
3'UTR of any one of HBB, ARHGAP, CORO1A, or HPX;

- 3) the 5'UTR is derived from a 5'UTR of HBB, and the 3'UTR is derived from a 3'UTR of any one of HBB, ARHGAP, CORO1A, or HPX; and
- 4) the 5'UTR is derived from a 5'UTR of CCL13, and the 3'UTR is derived from a 3'UTR of any one of HBB, ARHGAP, CORO1A, or HPX;
- 5 preferably,
- the 5'UTR of ARHGAP comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 45-46, 63-75, and 77-78 and CUAUAAU, or comprises CUAUAAU, or a nucleotide sequence having at least 80% identity to any one of the aforementioned sequences;
- 10 the 5'UTR derived from HSPB1 comprises the nucleotide sequence set forth in SEQ ID NO: 47 or a nucleotide sequence having at least 80% identity thereto;
- the 5'UTR derived from HBB comprises the nucleotide sequence set forth in SEQ ID NO: 48 or a nucleotide sequence having at least 80% identity thereto;
- the 5'UTR derived from CCL13 comprises the nucleotide sequence set forth in SEQ ID
- 15 NO: 49 or a nucleotide sequence having at least 80% identity thereto;
- the 3'UTR derived from HBB comprises the nucleotide sequence set forth in SEQ ID NO: 50 or a nucleotide sequence having at least 80% identity thereto;
- the 3'UTR derived from ARHGAP comprises the nucleotide sequence set forth in SEQ ID NO: 51 or a nucleotide sequence having at least 80% identity thereto;
- 20 the 3'UTR derived from CORO1A comprises the nucleotide sequence set forth in SEQ ID NO: 52 or a nucleotide sequence having at least 80% identity thereto;
- the 3'UTR derived from HPX comprises the nucleotide sequence set forth in SEQ ID NO: 53 or a nucleotide sequence having at least 80% identity thereto.
- 25 15. The RNA molecule according to any one of claims 13 to 14, wherein the ORF encodes a viral antigen; preferably, the ORF encodes a coronavirus antigen or an influenza virus antigen;
- the coronavirus is preferably SARS-COV-2, and the coronavirus antigen is preferably a spike protein;
- 30 more preferably, the spike protein is selected from the group consisting of a spike protein of any one of the viral strain SARS-COV-2, SARS-COV-2 Alpha, SARS-COV-2 Beta, SARS-COV-2 Gamma, SARS-COV-2 Kappa, SARS-COV-2 Delta, or SARS-COV-2 Omicron; or,
- the influenza virus is selected from the group consisting of influenza A virus and
- 35 influenza B virus, and the influenza virus antigen is a hemagglutinin protein (HA) and/or a neuraminidase (NA);

more preferably, the influenza virus antigen is selected from the group consisting of a hemagglutinin protein and/or a neuraminidase of any one of the viral strains influenza A virus H1N1, influenza A virus H3N2, influenza B virus Victoria, and influenza B virus Yamagata.

5

16. The RNA molecule according to claim 15, wherein:

a polypeptide sequence encoded by the ORF comprises the amino acid sequence set forth in any one of SEQ ID NOs: 12, 14, 21, 23, 25, 80, 83-86, 98-101, and 136;

preferably, the nucleotide sequence of the ORF comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 54-55, 59-61, 82, and 126-134, or a nucleotide sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to the nucleotide sequence set forth in any one of SEQ ID NOs: 54-55, 59-61, 82, and 126-134.

15 17. The RNA molecule according to any one of claims 13 to 16, further comprising:
(d) a polyadenylic acid (poly-A) tail.

18. The RNA molecule according to claim 17, wherein:

the poly-A tail is selected from the group consisting of HGH polyA, SV40polyA, BGH polyA, rbGlob polyA, and SV40late polyA;

preferably, the poly-A tail comprises the nucleotide sequence set forth in SEQ ID NO: 56 or SEQ ID NO: 135 or a nucleotide sequence having at least 80% identity thereto.

19. The RNA molecule according to any one of claims 13 to 18, comprising (a) an open reading frame (ORF), (b) a 5' untranslated region element (5'UTR), (c) a 3' untranslated region element (3'UTR), and (d) a polyadenylic acid (poly-A) tail, wherein preferably, the RNA molecule comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 57-58 and 106-125.

20. The RNA molecule according to any one of claims 13 to 19, further comprising:
(e) a 5' cap structure (5'Cap).

21. The RNA molecule according to claim 20, wherein:

the 5'Cap is selected from the group consisting of Cap0, Cap1, Cap2, Cap3, Cap4, ARCA, modified ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and

2-azido-guanosine;

preferably, the 5'Cap is selected from the group consisting of ARCA, 3'OMe-m7G(5')ppp(5')G, m7G(5')ppp(5')(2'OMeA)pU, m7Gppp(A2'O-MOE)pG, m7G(5')ppp(5')(2'OMeA)pG, m7G(5')ppp(5')(2'OMeG)pG,

5 m7(3'OMeG)(5')ppp(5')(2'OMeG)pG, and m7(3'OMeG)(5')ppp(5')(2'OMeA)pG; preferably, the 5'Cap is m7G(5')ppp(5')(2'OMeA)pG.

22. The RNA molecule according to any one of claims 13 to 21, wherein the RNA molecule is an mRNA.

10

23. The RNA molecule according to any one of claims 13 to 22, further comprising one or more modifications,

wherein preferably, the modifications include a backbone modification, a sugar modification, a base modification, and/or a lipid modification;

15 more preferably, the base modification is a pseudouridine modification.

24. An RNA molecule, comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 54-55, 59-61, 82, and 126-134 or a nucleotide sequence having at least 80% identity to the nucleotide sequence set forth in any one of SEQ ID NOs: 54-55, 59-61, 82, and 126-134,

20

wherein preferably, the RNA molecule further comprises one or more modifications; more preferably, the modifications are pseudouridine modifications.

25. Use of the nucleic acid construct according to any one of claims 1-12 or the RNA molecule according to any one of claims 13-24 in any one of the following: (1) preparing a vaccine; (2) encoding a viral antigen in a subject or *in vitro*; (3) preparing a medicament encoding a viral antigen in a subject or *in vitro*; and (4) preparing a medicament.

25

26. A vaccine, comprising the RNA molecule according to any one of claims 13 to 24, wherein preferably, the RNA molecule encodes one or more antigens of one or more viral strains.

30

27. A vector, comprising the nucleic acid construct according to any one of claims 1 to 12, or the RNA molecule according to any one of claims 13 to 24.

35

28. A host cell, comprising the vector according to claim 27.
29. A preparation method for the RNA molecule according to any one of claims 13-24, comprising: reverse-transcribing the nucleic acid construct according to any one of
5 claims 1 to 12 or the vector according to claim 27 to obtain an RNA molecule, wherein preferably, the method further comprises adding a 5'Cap to the 5' end of the RNA molecule.
30. A delivery system, comprising the nucleic acid construct according to any one of
10 claims 1 to 12 or the RNA molecule according to any one of claims 13 to 24, wherein a delivery vehicle of the delivery system is a cationic lipid delivery particle; preferably, the delivery vehicle is a lipid nanoparticle.
31. A pharmaceutical composition, comprising:
15 a pharmaceutically acceptable carrier, diluent, or excipient, and any one selected from the group consisting of the following:
the nucleic acid construct according to any one of claims 1 to 12, the RNA molecule according to any one of claims 13 to 24, the vaccine according to claim 26, and/or the delivery system according to claim 30.
20
32. A product or kit, comprising the nucleic acid construct according to any one of claims 1 to 12, the RNA molecule according to any one of claims 13 to 24, the vaccine according to claim 26, the delivery system according to claim 30, and/or the pharmaceutical composition according to claim 31.
25
33. Use for preparing a medicament for treating and/or preventing a disease, comprising administering to a subject in need thereof an effective amount of the nucleic acid construct according to any one of claims 1 to 12, the RNA molecule according to any one of claims 13 to 24, the vaccine according to claim 26, the
30 delivery system according to claim 30, the pharmaceutical composition according to claim 31, and/or the product or kit according to claim 32, wherein the disease is a viral infectious disease or a respiratory disease associated with viral infection; preferably, the virus is a coronavirus or an influenza virus; more preferably, the virus is SARS-CoV-2, influenza A virus, or influenza B virus;
35 preferably, the respiratory disease associated with viral infection includes simple infection, fever, cough, sore throat, rhinitis, headache, pneumonia, acute respiratory

infection, severe acute respiratory infection (SARI), hypoxemic respiratory failure, acute respiratory distress syndrome, sepsis, septic shock, and severe acute respiratory syndrome (SARS).

5 34. A method for treating and/or preventing a disease, comprising administering to a subject in need thereof an effective amount of the nucleic acid construct according to any one of claims 1 to 12, the RNA molecule according to any one of claims 13 to 24, the vaccine according to claim 26, the delivery system according to claim 30, the pharmaceutical composition according to claim 31, and/or the product or kit according to claim 32, wherein the disease is a viral infectious disease or a respiratory disease associated with viral infection;

preferably, the virus is a coronavirus or an influenza virus; more preferably, the virus is SARS-CoV-2, influenza A virus, or influenza B virus;

15 preferably, the respiratory disease associated with viral infection includes simple infection, fever, cough, sore throat, rhinitis, headache, pneumonia, acute respiratory infection, severe acute respiratory infection (SARI), hypoxemic respiratory failure, acute respiratory distress syndrome, sepsis, septic shock, and severe acute respiratory syndrome (SARS).

20 35. A method for inducing a neutralizing antibody reaction and/or a T cell immune response in a subject, comprising administering to a subject in need thereof an effective amount of the nucleic acid construct according to any one of claims 1 to 12, the RNA molecule according to any one of claims 13 to 24, the vaccine according to claim 26, the delivery system according to claim 30, the pharmaceutical composition according to claim 31, and/or the product or kit according to claim 32,

25 wherein preferably, the neutralizing antibody reaction is a neutralizing antibody reaction against a viral antigen, and the T cell immune response includes a CD4+ and/or CD8+ T cell immune reaction;

30 preferably, the virus is a coronavirus or an influenza virus; more preferably, the virus is SARS-CoV-2, influenza A virus, or influenza B virus.