

Circular RNA and preparation method thereof

Technical field

The present invention relates to the field of biomedicine, in particular, to an engineered circular RNA molecule, a pharmaceutical composition comprising the same, and use thereof. The present invention also relates to the vector as well as the method used for preparing the same.

Background

Circular RNA is a common type of RNA in eukaryotes. Naturally occurring circular RNAs are primarily produced by a molecular mechanism within cells called "back-splicing". Eukaryotic circular RNAs have been found to have a variety of molecular and cellular regulatory functions. For example, circular RNAs can regulate the expression of target genes by binding to microRNAs. Circular RNAs can also regulate gene expression by directly binding to target proteins.

Due to its circular nature, circular RNA has a longer half-life than linear mRNA, so it is speculated that circular RNA synthesized *in vitro* may have higher stability. Methods of forming circular RNAs *in vitro* include the chemical method, the enzymatic catalysis method, and the ribozyme catalysis method. Chemical methods are expensive and the size of the circular RNA molecules that can be produced is limited. The enzymatic method mainly utilizes T4 RNA ligase to catalyze the circularization of linear RNA, and the size of RNA payload that can achieve circularization is also limited. Ribozyme catalysis (e.g., based on Group I introns) is a promising method for the preparation of circular RNAs.

The natural Group I Intron system can undergo cleavage and ligation reactions to form circular intronic RNAs. A specific cleavage site conserved sequence located in the 5' exon E1 is cleaved by the nucleophilic attack of the free 3' hydroxyl group of guanosine triphosphate, resulting in a naked 3' hydroxyl group, while the guanylate binds on the cleaved 5' exon E1. After that, the exposed 3' hydroxyl group at the 5' end of the intron attacks the conserved sequence between the 3' end of the intron and the exon E2, the exon E2 is excised, and the intron undergoes a circularizing reaction to obtain a circular intronic RNA. An improved ribozyme-catalyzed approach derived from Anabaena tRNA introns has been reported for the formation of circular RNAs *in vitro*, termed the "Group I permuted intron-exon self-splicing system, PIE system". This method can excise the intron to form circular RNA containing exons. Therefore, this method has the potential to form expressible circular RNAs. The basic design principle of the PIE system is to connect the exon E1 sequence and E2 sequence end to end by molecular cloning to form a continuous circular plasmid. The intron is cleaved by restriction endonuclease to obtain a linear plasmid. Then, *in vitro* transcription is performed through the T7 promoter upstream of the

permuted 3' intron to obtain a linear RNA containing a 3' intron-E2-E1-5' intron structure. Similar to the natural Group I intron system, the specific cleavage site conserved sequence of exon E1 is cleaved by the nucleophilic attack of the free 3' hydroxyl of guanylate, and the exon E1 produces a naked 3' hydroxyl, while the guanylate binds to the cleaved 5' intron. Subsequently, the exposed 3' hydroxyl of exon E1 attacks the conserved sequence between the 3' intron and exon E2, the 3' intron is excised, and exon E2 and E1 undergo a circularizing reaction to obtain circular E1-E2 RNA.

However, Group I Introns available for *in vitro* circularization in the prior art are limited and only a few of them have been proved to be useful in the PIE system to form circular RNAs. Therefore, there is a need in the art to develop new Group I Introns which can be used or *in vitro* RNA circularization with high efficiency.

Summary of the invention

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment; and
- b) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA through the self-splicing of the circular RNA precursor,

wherein the self-splicing intron is selected from the Group I introns and Group II introns, for example, the self-splicing intron is selected from Group I introns,

wherein the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

In aspects and embodiments of the present invention, the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At).

In aspects of the present invention, the circular RNA precursor further comprises a homology arm sequence I upstream of the 3' self-splicing intron fragment and a homology arm sequence II downstream of the 5' self-splicing intron fragment.

In one aspect, it is an object of the present invention to provide a circular RNA precursor

comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a first residual circularizing element;
- c) a nucleotide sequence of interest;
- d) a second residual circularizing element; and
- e) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element through the self-splicing of the circular RNA precursor.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a circularizing sequence fragment II comprising a ribozyme recognition site II at its 5' end and a circularizing sequence fragment I comprising a ribozyme recognition site I at its 3' end,

wherein the circularizing sequence fragment I is derived from a 5' terminal portion of a circularizing sequence, and the circularizing sequence fragment II is derived from a 3' terminal portion of a circularizing sequence, and

wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the circularizing sequence fragment I is connected with the circularizing sequence fragment II to form the circularizing sequence.

In embodiments, the circularizing sequence is a protein-coding sequence, a non-coding sequence, a translation initiation element (TIE), or a non-TIE functional element.

In embodiments, the nucleotide sequence of interest further comprises one or more sequences selected from the following group between the circularizing sequence fragment II and the circularizing sequence fragment I: at least one protein coding sequences, translation initiation element (TIE), a non-TIE functional element, at least one non-coding sequence, or any

combination thereof.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisting of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment ;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES), and a protein-coding sequence fragment I;

wherein the protein-coding sequence fragment I is derived from a 5' terminal portion of at least one protein-coding sequences, and the protein-coding sequence fragment II is derived from a 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment ;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES),

and a protein-coding sequence fragment I;

wherein the protein-coding sequence fragment I is the 5' terminal portion of the protein-coding sequences, and the protein-coding sequence fragment II is the 3' terminal portion of the protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the protein-coding sequence.

In embodiments, the at least one protein-coding sequence is a sequences encoding a diagnostic or therapeutic protein. In embodiments, the protein encoded by the protein-coding sequence is a mammalian protein, for example a human protein. In embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more proteins. In embodiments, the at least one protein-coding sequence a sequence encoding two or more diagnostic or therapeutic protein. Such at least two protein-coding sequences are operably linked with each other by spacers, 2A sequence or further translation initiation elements. In embodiments, a further translation initiation element is operably linked between any two adjacent protein-coding sequences. In embodiments, the translation initiation element is selected from an IRES sequence, a 5' UTR sequence, a Kozak sequence, a sequence containing m⁶A modification, a sequence complementary to ribosome 18S rRNA, or any combination thereof. In embodiments, the translation initiation element is an IRES sequence.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisting of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a translation

initiation element fragment II, at least one protein-coding sequence, and a translation initiation element fragment I;

wherein the translation initiation element fragment I is derived from a 5' terminal portion of a translation initiation element, and the translation initiation element fragment II is derived from a 3' terminal portion of the translation initiation element, and

the translation initiation element fragment I comprises a ribozyme recognition site I located at its 3' end, the translation initiation element fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the translation initiation element fragment I is connected with the translation initiation element fragment II to form the translation initiation element which can recruit ribosome to initiate translation.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a translation initiation element fragment II, at least one protein-coding sequence, and a translation initiation element fragment I;

wherein the translation initiation element fragment I is the 5' terminal portion of a translation initiation element, and the translation initiation element fragment II is the 3' terminal portion of the translation initiation element, and

the translation initiation element fragment I comprises a ribozyme recognition site I located at its 3' end, the translation initiation element fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the translation initiation element fragment I is connected with the translation initiation element fragment II to form the translation initiation element which can recruit ribosome to initiate translation.

In embodiments, the at least one protein-coding sequence is a sequence encoding a diagnostic or therapeutic protein. In embodiments, the protein encoded by the protein-coding sequence is a mammalian protein, for example a human protein. In embodiments, the at least one protein-coding

sequence comprises at least two protein-coding sequences. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more proteins. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more diagnostic or therapeutic proteins. Such at least two protein-coding sequences may be operably linked with each other by spacers, 2A sequence or further translation initiation elements. In embodiments, the translation initiation element is selected from an IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof. In embodiments, the translation initiation element is an IRES sequence.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisting of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest;
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a non-TIE functional element fragment II, a translation initiation element, and at least one protein-coding sequence, and a non-TIE functional element fragment I;

wherein the non-TIE functional element fragment I is derived from a 5' terminal portion of a non-TIE functional element, and the non-TIE translation initiation element fragment II is derived from a 3' terminal portion of the non-TIE functional element, and the non-TIE functional element fragment I comprises a ribozyme recognition site I located at its 3' end, the non-TIE functional element fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-TIE functional element fragment I is connected with the non-TIE functional element fragment II to form the non-TIE functional element.

In embodiments, the non-TIE functional element is a 3' UTR or a replicon. In embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences. Such at least two protein-coding sequences may connect with each other by spacers, 2A sequence or further translation initiation elements. In embodiments, the translation initiation element is selected from

an IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m⁶A modification, a sequence complementary to ribosome 18S rRNA, or any combination thereof. In embodiments, the translation initiation element is an IRES sequence.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisting of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest;
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a non-coding sequence fragment II, and a non-coding sequence fragment I;

wherein the non-coding sequence fragment I is derived from a 5' terminal portion of a non-coding sequence, and the non-coding sequence fragment II is derived from a 3' terminal portion of the non-coding sequence, and the non-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the non-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-coding sequence fragment I is connected with the non-coding sequence fragment II to form the complete non-coding sequence.

In embodiments, the non-coding sequence is an endogenous circular RNA. In embodiments, the nucleotide sequence of interest further comprises one or more non-coding sequences between the non-coding sequence fragment I and the non-coding sequence fragment II.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisting of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In aspects of the present invention, the protein-coding sequence is a sequence encoding an aquaporin protein, such as AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein. In

embodiments, the AQP-1 protein is a mammalian AQP-1 protein, in particular, a human AQP-1 protein.

In another aspect, it is an object of the present invention to provide a nucleic acid vector for generating a circular RNA molecule, said vector comprises a coding sequence of the circular RNA precursor as described above.

In another aspect, it is an object of the present invention to provide a circular RNA polynucleotide, which is prepared from the circular RNA precursor as described above or from the nucleic acid vector as described above. In embodiments, the circular RNA polynucleotide of the present invention has a duration of in vivo expression of at least 24 hours (1 day), for example, 36 hours, 48 hours, 72 hours, 96 hours, 5 days, 10 days, 15 days, 16 days, 17 days, 18 days, 19 days, 20 days or more. In another aspect, it is an object of the present invention to provide use of the circular RNA precursor as described above and/or the circular RNA as described above as an expression vector. In embodiments, the expression vector is used to express one or more diagnostic or therapeutic proteins.

In another aspect, it is an object of the present invention to provide a pharmaceutical composition comprising the circular RNA precursor as described above and/or the nucleic acid vector as described above and/or the circular RNA as described above, and a pharmaceutically acceptable carrier.

In another aspect, it is an object of the present invention to provide a method for preparing a circular RNA, comprising:

- 1) providing the circular RNA precursor as described above or obtaining a circular RNA precursor by transcribing from the nucleic acid vector as described above;
- 2) incubating the circular RNA precursor in the presence of a divalent metal cation at a temperature at which RNA circularization occurs; and
- 3) harvesting the circular RNA obtained in step 2).

In another aspect, it is an object of the present invention to provide a method for preparing a circular RNA, comprising

a) providing a nucleic acid vector comprising a self-splicing intron-based RNA circularizing elements as a transcription template; and

b) incubating the nucleic acid vector in an in vitro transcription system comprising a divalent metal cation and an RNA polymerase for a first time period during which the linear RNA produced by in vitro transcription is self-circularized under the action of the RNA circularizing elements to produce a circular RNA.

In another aspect, it is an object of the present invention to provide a method for purifying a circular RNA, comprising:

a) contacting a mixture comprising the circular RNA and the circular RNA precursor with a circular RNA-specific probe under a condition that allows the circular RNA-specific probe to specifically bind to and form a complex with the circular RNA;

b) separating the complex from one or more components in the mixture that are not bound to the circular RNA-specific probe; and

c) releasing the circular RNA from the complex.

In another aspect, it is an object of the present invention to provide a method for purifying a circular RNA, comprising:

i) contacting a mixture comprising the circular RNA and the circular RNA precursor to a linear circular RNA precursor-specific probe under a condition that allows the linear circular RNA precursor-specific probe to specifically bind to and form a complex with the linear circular RNA precursor; and

ii) removing the complex formed by the linear circular RNA precursor-specific probe with the linear circular RNA precursor from the mixture,

iii) collecting the circular RNA-containing mixture obtained in step ii), and

optionally, steps i)-iii) are performed multiple times, e.g., 2 times, 3 times, 4 times or more.

In another aspect, it is an object of the present invention to provide a method for purifying a circular RNA, comprising:

i) adding a linear RNA-specific tag to the linear RNA in a mixture comprising circular RNA and linear RNA;

ii) contacting the mixture comprising circular RNA and linear RNA with a linear RNA probe that specifically binds to the tag under a condition that allows the probe to specifically bind to and form a complex with the linear RNA; and

iii) removing the complex formed by the linear RNA probe with the linear RNA from the mixture,

iv) collecting the circular RNA-containing mixture obtained in step iii),

optionally, steps ii)-iv) are performed multiple times, e.g., 2 times, 3 times, 4 times or more.

In another aspect, it is an object of the present invention to provide a product expressed by the circular RNA as described above, the circular RNA prepared by the method as described above or the circular RNA purified by the method as described above.

In yet another aspect, it is an object of the present invention to provide a pharmaceutical composition comprising the circular RNA precursor as described above and/or the nucleic acid vector as described above and/or the circular RNA as described above and/or the circular RNA prepared by the method as described above and/or the circular RNA purified by the method as described above, and a pharmaceutically acceptable carrier.

In yet another aspect, it is an object of the present invention to provide a method for preventing or treating a disease or condition in a subject, comprising administering to a subject the circular RNA as described above, the circular RNA prepared by the method as described above or the circular RNA purified by the method as described above, or the pharmaceutical composition as described above.

In another aspect, it is an object of the present invention to provide use of the circular RNA as described above, the circular RNA prepared by the method as described above or the circular RNA purified by the method as described above, or the pharmaceutical composition as described above in the manufacture of a medicament for preventing or treating a disease or condition in a subject.

In some aspects, it is an object of the present invention to provide a method of screening a target open reading frame (ORF) sequence encoding a target polypeptide for a nucleotide motif, comprising steps of:

1) extracting amino acid units in the direction of N terminal to C terminal in the target polypeptide containing m amino acids, and each group of amino acid units includes an amino acid with the number of n ; with a step of 1 amino acid, slide along the N terminal to the C terminal, n is an integer and $n \geq 2$, m is an integer and ≥ 1 ;

2) listing all nucleotide sequences corresponding to each group of amino acid units according to their degenerate codons; and

3) identifying nucleotide sequences obtained from step 2) that match the nucleotide motif;

wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU, and

4) optionally, selecting the target ORF as the circularizing sequence of a circular RNA precursor.

In some aspects, it is an object of the present invention to provide a system for screening a target open reading frame (ORF) for a nucleotide motif, wherein the target open reading frame (ORF) is a nucleotide sequence encoding a target polypeptide, and the system comprises:

a) an amino acid extraction module, which is used for extracting amino acid units in the direction of N terminal to C terminal in the target polypeptide containing m amino acids, wherein each group of amino acid units comprise an amino acid with a number of n , with a step of 1 amino acid, sliding along the N terminal to the C terminal, and n is an integer and $n \geq 2$, m is an integer and ≥ 1 ;

b) a conversion module, which is used for determining the set of codon sequences of each amino acid unit, each set of codon sequences includes all degenerate codon sequences corresponding to each set of amino acid units; and

c) a target codon sequence filtering module, which is used for determining whether the target

codon sequence contain a nucleotide sequence that exactly matches the nucleotide motif;

wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some aspects, it is an object of the present invention to provide a method for screening a target internal ribosome entry site (IRES) for a nucleotide motif, comprising steps of:

1) identifying regions of non-conserved sequences of the target IRES by comparing the target IRES with the other IRES sequences from the same viral subgenus, through multiple sequence alignment tools, for example: Clustal Omega, or MUSCLE;

2) based on the results of multiple sequence alignment, determining a low-structure region of the target IRES through secondary structure prediction software, such as RNAfold, Mfold, RNAstructure or RNAalifold; and

3) editing 0, 1, or 2 bases in the low-structure region in the target IRES to match the nucleotide motif without affecting the translation initiation activity of said IRES;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some aspects, it is an object of the present invention to provide a system for screening a target internal ribosome entry site (IRES) for a nucleotide motif, wherein the system comprises:

a) a multi-sequence alignment module, which is used for comparing the target IRES with the other IRES sequences from the same viral subgenus to identify regions of non-conservative sequences;

b) a low-structure prediction module, which is used for determining a low-structure region within the target IRES according to the result of the multiple sequence alignments of the target IRES; and

c) a base editing module, which is used for editing the bases at any one or more positions in the region to obtain an edited region matching the nucleotide motif,

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some aspects, it is an object of the present invention to provide a screening method for screening a target non-TIE functional element for the nucleotide motif in ribozyme recognition site I, comprising a step of:

a) editing 0, 1, or 2 bases on the non-TIE functional element so that the non-TIE functional element sequence (such as 3' UTR or a replicon) comprises an edited region that matches the nucleotide motif whilst maintaining the non-TIE function;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some aspects, it is an object of the present invention to provide a system for screening a target region sequence in a non-TIE functional element sequence (such as 3' UTR or a replicon) for a nucleotide motif, wherein the screening system comprises:

a) a base editing module, which is used for editing the bases at any one or more positions in the region to obtain an edited region matching the nucleotide motif;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

It is to be understood that each of the method of prevention or treatment embodiments herein can also be formulated as corresponding use type embodiments.

Brief description of the figures

Figure 1. Different PIE designs of *Azoarcus* Group I Intron (Azop).

Figure 2. The circularization efficiency of Azop with different PIEs.

Figure 3. PIE design of *Scytalidium dimidiatum* Group I Intron (S.d).

Figure 4. The circularization efficiency of S.d with different PIEs.

Figure 5. Different PIE designs of *Staphylococcus phage Twort* Group I Intron (Twort).

Figure 6. The circularization efficiency of Twort with different PIEs.

Figure 7. Different PIE designs of *Scytonema hofmanii* Group I Intron (Sh).

Figure 8. The circularization efficiency of *Scytonema hofmanii* (Sh).

Figure 9. Circularization efficiency of IL2 with different Group I Introns of the present invention.

Figure 10. The RNA circularization efficiency of *Agrobacterium-tumefaciens* Group I Intron (At).

Figure 11. Homology arm (HA) design: structure prediction of precursors with different HAs.

Figure 12. The circularization efficiency of AnaX-IL2 with different Homology arms.

Figure 13. Analysis of the correlation between circularization efficiency and average base MFE of Homology arm.

Figure 14. Design of CLEAN-PIE on ORF or IRES from PIE.

Figure 15. Design of CLEAN-PIE on functional element from PIE.

Figure 16. Sequence and structure of CLEAN-PIE designs for IL2.

Figure 17. The circularization efficiency of AnaX Clean PIE for IL2.

Figure 18. Different CLEAN-PIE Designs for circular RNA polynucleotides expressing the AQP1 protein (CLEAN-PIE circAQP1).

Figure 19. The circularization efficiency, translation and immunogenicity of Clean PIE for circAQP1 (CLEAN-PIE circAQP1).

Figure 20. The effect of miRNA target sites in circRNA on expression level.

Figure 21. The effect of UTR on circRNA translation.

Figure 22. CLEAN-PIE circAQP1 expression in salivary glands in rats.

Figure 23. Representative images of IHC staining for AQP1 expression in salivary glands.

Detailed description of the invention

In the present invention, unless indicated otherwise, the scientific and technological terminologies used herein refer to meanings commonly understood by a person skilled in the art. Also, the terminologies and experimental procedures used herein relating to protein and nucleotide chemistry, molecular biology, cell and tissue cultivation, microbiology, immunology, all belong to terminologies and conventional methods generally used in the art. For example, the standard DNA recombination and molecular cloning technology used herein are well known to a person skilled in the art, and are described in details in the following references: Sambrook, J., Fritsch, Efland Maniatis, T., Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989. In the meantime, in order to better understand the present invention, definitions and explanations for the relevant terminologies are provided below.

As used herein, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. It is further noted that the claims may be drafted to exclude any optional element. For example, a circular RNA precursor refers to one or more circular RNA precursors. As such, the terms "a", "an", "one or more" and "at least one" can be used interchangeably. For example, the term "at least one" refers to one, two or more. This statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only" and the like in connection with the recitation of claim elements, or use of a "negative" limitation. Similarly the terms "comprising", "including" and "having" can be used interchangeably.

As used herein, the term "and/or" encompasses all combinations of items connected by the term, and each combination should be regarded as individually listed herein. For example, "A and/or B" covers "A", "A and B", and "B". For example, "A, B, and/or C" covers "A", "B", "C", "A and B", "A and C", "B and C", and "A and B and C".

As used herein, "about," "approximately," "substantially," and "significantly" will be understood by persons of ordinary skill in the art and will vary to some extent on the context in which they are used. If there are uses of these terms which are not clear to persons of ordinary skill in the art given the context in which they are used, "about" and "approximately" will mean plus or minus <10% of the particular term and "substantially" and "significantly" will mean plus or minus >10% of the particular term.

In the present application, "optional" or "optionally" means that the subsequently described

event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances in which it does not.

"Polynucleotide", "nucleic acid sequence", "nucleotide sequence", or "nucleic acid fragment" are used interchangeably to refer to a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide. Although the nucleotide sequences herein may be represented as DNA sequences (comprising T(s)), when referring to RNA, one skilled in the art can readily determine the corresponding RNA sequence (i.e., replacing T with U).

For example, the nucleotide sequence of interest in the present invention can be a non-coding sequence, such as an antisense RNA, aptamer, guide RNA, or non-coding RNA existing in any organism.

Sequence "identity" has recognized meaning in the art, and the percentage of sequence identity between two nucleic acids or polypeptide molecules or regions can be calculated using the disclosed techniques. Sequence identity can be measured along the entire length of a polynucleotide or polypeptide or along a region of the molecule. (See, for example, Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). There are many methods for determining sequence identity. An example of algorithms suitable for determining percent sequence identity is the algorithm used in the Basic Local Alignment Search Tool (hereinafter "BLAST"), see e.g., Altschul et al., J. Mol. Biol. 215:403-410, 1990 and Altschul et al, Nucleic Acids Res., 15:3389-3402, 1997. Software for performing BLAST analysis is publicly available through the National Center for Biotechnology Information (hereafter "NCBI"). Default parameters used to determine sequence identity using software available from NCBI (such as BLASTN for nucleic acid sequences) are described in McGinnis et al. Nucleic Acids Res., 32: W20-W25, 2004.

A polynucleotide sequence (e.g. a RNA polynucleotide sequence or its fragments) "derived from" a designated polynucleotide sequence means the former polynucleotide sequence is originated from the latter. In some embodiments, the polynucleotide sequence which is derived

from a particular polynucleotide sequence has a polynucleotide sequence that is identical, essentially identical or homologous to that particular sequence or a fragment thereof. Polynucleotide sequences derived from a particular polynucleotide sequence may be variants of that particular sequence or a fragment thereof. For example, it will be understood by one of ordinary skill in the art that the circular RNA molecules suitable for use herein may be altered such that they vary in sequence from the sequences from which they were derived, while retaining the desirable activity thereof. "Circular RNA precursor" herein refers to a linear RNA molecule capable of forming a covalently linked closed circular RNA molecule, e.g., by self-splicing. The circular RNA precursor may be produced by transcription from a nucleic acid vector comprising a coding sequence of the circular RNA precursor. Alternatively, the circular RNA precursor may also be obtained by chemical synthesis.

In some embodiments, the circular RNA precursor is capable of forming a covalently linked closed circular RNA molecule by self-splicing under the action of the self-splicing intron fragments and the residual circularizing elements.

As used herein, the term "self-splicing intron" refers to an intron having self-splicing ribozyme activity and capable of excising itself and joining two flanking exons. In some embodiments, the splicing is autocatalytic splicing.

"Self-splicing introns" include, but are not limited to, Group I introns and Group II introns. Group I introns contain 14 subgroups, while most of the Group I introns belong to the IC3 subgroup. For example, the Group I intron may be a Group I intron of the cyanobacterium *Anabaena* belonging to the IC3 subgroup or a Group I intron from a T4 phage Group I intron belonging to the IA2 subgroup or a Group I intron from *Azoarcus* sp. BH72 belonging to the IC3 subgroup. Additional examples of self-splicing introns useful in the present invention include, but are not limited to, self-splicing introns derived from the following organisms: Enterobacteriophage T4, Bacteriophage Twort, Bacteriophage SPO1, Bacteriophage S3b, *Bacillus anthracis*, *Clostridium botulinum*, *Tetrahymena thermophila*, *Dunaliella parva*, *Pneumocystis carinii*, *Physarum polycephalum*, *Anabaena* sp. PCC7120, *Scytonema hofmanni*, *Agrobacterium tumefaciens*, *Synechocystis* PCC 6803, *Synechococcus elongatus* PCC 6301, *Neurospora crassa*, *Candida albicans*, *Scytalidium cerradiumydiaes*, *Pediadiaces Chlamydomonas nivalis*, *Chlorella vulgaris*, *Amoebidium parasiticum*, *Neurospora crassa*, *Emericella nidulans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Neochloris aquatica*, *Dunaliella parva*, *Symkania negevensis*, *Emericella nidulans*. See e.g., Vicens, Q., et al., (2008). Toward predicting self-splicing and protein-facilitated splicing of group I introns. *RNA* 14: 2013-2029; Tanner, A. M., et al., (1996). Activity and thermostability of the small self-splicing group I intron in the pre-tRNA^{Ile} of the purple bacterium *Azoarcus*. *RNA* 2:74-83.

In some embodiments, the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet Group I Intron (Sh), or *Agrobacterium-tumefaciens* Group I Intron (At). For example, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

In some embodiments, the 3' self-splicing intron fragment and the 5' self-splicing intron fragment are derived from a same self-splicing intron. In some embodiments, the 3' self-splicing intron fragment is derived from or contains a 3' terminal portion of the self-splicing intron (a native self-splicing intron), and accordingly the 5' self-splicing intron fragment is derived from or contains a 5' terminal portion of the self-splicing intron (a native self-splicing intron). In some embodiments, the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of the self-splicing intron (a native self-splicing intron).

In some embodiments, the self-splicing intron is a Group I intron. In some embodiments, the self-splicing intron is a Group I intron of the IA2 or IC3 subgroup, preferably IC3 subgroup.

In some embodiments, the 3' self-splicing intron fragment is a 3' Group I intron fragment. In some embodiments, the 5' self-splicing intron fragment is a 5' Group I intron fragment.

As used herein, a 3' self-splicing intron fragment (e.g., a 3' Group I intron fragment) is a sequence that is at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identical to the 3' terminal portion of a native self-splicing intron (e.g., a Group I intron). A 5' self-splicing intron fragment (e.g., a 5' Group I intron fragment) is a sequence that is at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identical to the 5' terminal portion of a native self-splicing intron (e.g., a Group I intron).

It is generally believed that in order to achieve circularization, the native Group I intron needs to be split at an internal site to form a so called "Group I permuted intron-exon self-splicing system, PIE system". An internal split site as used herein refers to a location where the native Group I intron is split into two portions, and the internal split site is selected such that the generated two

separate portions of native Group I intron (the 3' terminal portion and the 5' terminal portion) together can maintain ribozyme activity necessary for the self-splicing, even after permutation. It is believed that the two separate portions of native Group I intron maintaining the overall conformation of the native Group I intron can maintain ribozyme activity necessary for the self-splicing. As used herein, the internal split site is located between an intron internal split site I and an intron internal split site II, the intron internal split site I refers to the first nucleotide downstream of the internal split site and the intron internal split site II refers to the first nucleotide upstream of the internal split site. As used herein, "PIE" or "PIE method" or "PIE system" refers to the above described "Group I permuted intron-exon self-splicing system, PIE system".

In some embodiments, the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

Only a few Group I introns of which the exact internal split site have been determined by a person skilled in the art, e.g., by referenced to Puttaraju M., et al., (1992) Group I permuted intron-exon (PIE) sequences self-splice to produce circular exons; and/or Puttaraju M., et al., (1996) Circular ribozymes generated in *Escherichia coli* using group I self-splicing permuted intron-exon sequences. For example, for a Group I intron, especially an *Anabaena* Group I intron, it can usually be split at a specific site in its P6 region to form the PIE system.

The 3' terminal portion of a native Group I intron may have a length of about 5% to about 95%, such as about 5%, about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95% of the full length of the native Group I intron. Accordingly, the 5' terminal portion of a native Group I intron may have a length of about 5% to about 90%, such as about 5%, about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95% of the full length of the native Group I intron.

In some embodiments, the combination of the 3' Group I intron fragment and the 5' Group I intron fragment substantially has the self-splicing activity of the corresponding native Group I intron.

"Circularization efficiency" as used herein may refer to the ratio of outcome circular RNA to input precursor in a given time period. Alternatively, "Circularization efficiency" as used herein may refer to the ratio of desired circular RNA to linear circRNA precursor or the ratio of circRNA to the sum of circRNA and linear circRNA precursor in the final product in a given time period. The circularizing efficiency can be determined by methods well known in the art, such as those described in Example 2, 4, 6 and 8 of the present application.

In some embodiments, the circularizing efficiency of the circular RNA is increased by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 200%, at least about 300%, at least about 400%, at least about 500% or more.

"Reduced immunogenicity" as used herein may refer to that the circular RNA, upon contacted with cells, elicits a reduced immune response, i.e., an immune response at a level lower than a control circular RNA or control linear RNA. For example, reduced immune response refers to reduced expression of cytokines. The cytokines include, but are not limited to, IFN β , TNF α , IL6 and/or RIG-I. In some embodiments, the immunogenicity of the circular RNA is reduced by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or more.

As used herein, "spacer" refers to any contiguous nucleotide sequence that at least does not negatively interfere with the function of the elements connected by it. Generally, if it is desired to avoid the interaction of two near or adjacent elements, a spacer can be inserted between the two elements. The spacer sequences described herein can serve two functions: (1) to facilitate circularization and (2) to facilitate functionality by allowing correct folding of the residual circularizing element and the nucleotide sequence of interest (e.g., IRES). In some embodiments, the spacer is no more than 150, no more than 100, no more than 50, no more than 30, no more than 10, no more than 5, or no more than 3 nucleotides in length. In some embodiments, the spacer is 5 nucleotides in length. In some embodiments, the spacer is 4 nucleotides in length. In some embodiments, the spacer is 3 nucleotides in length. In some embodiments, the first spacer may be absent. In some embodiments, the second spacer may be absent. In some embodiments, the first spacer and the second spacer may be absent.

In one aspect, the present invention provides a circular RNA precursor comprising from 5' to 3' direction :

- a) a 3' self-splicing intron fragment; and
- b) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA through the self-splicing of the circular RNA precursor,

wherein the self-splicing intron is selected from the Group I introns and Group II introns, for example, the self-splicing intron is selected from Group I introns.

In embodiments of the present invention, the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I

Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At).

In some embodiments of the present invention, the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

In some embodiments of the present invention, when the self-splicing intron is cyanobacterium *Anabaena* Group I Intron such as AnaX, the intron internal split site II and the intron internal split site I are the nucleotide A at position 115 and the nucleotide A at position 116 from 5' end of the AnaX intron (SEQ ID NO:1) respectively, and when connected form the nucleotide sequence of AA.

In some embodiments of the present invention, when the self-splicing intron is *Azoarcus* Group I Intron (Azo) (SEQ ID NO:2),

1) the intron internal split site II and the intron internal split site I are the nucleotide A at position 24 and the nucleotide A at position 25 from 5' end of the Azo intron respectively, and when connected form the nucleotide sequence of AA (P2);

2) the intron internal split site II and the intron internal split site I are the nucleotide A at position 85 and the nucleotide A at position 86 from 5' end of the Azo intron respectively, and when connected form the nucleotide sequence of AA (P4-5);

3) the intron internal split site II and the intron internal split site I are the nucleotide C at position 71 and the nucleotide C at position 72 from 5' end of the Azo intron respectively, and when connected form the nucleotide sequence of CC (P5);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 108 and the nucleotide U at position 109 from 5' end of the Azo intron respectively, and when connected form the nucleotide sequence of CU (P6);

5) the intron internal split site II and the intron internal split site I are the nucleotide A at position 154 and the nucleotide A at position 155 from 5' end of the Azo intron respectively, and when connected form the nucleotide sequence of AA (P8); or

6) the intron internal split site II and the intron internal split site I are the nucleotide C at position 198 and the nucleotide A at position 199 from 5' end of the Azo intron respectively, and when connected form the nucleotide sequence of CA (P9).

In some embodiments of the present invention, when the self-splicing intron is *Scytalidium dimidiatum* Group I Intron (Sd) (SEQ ID NO:4),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position

88 and a nucleotide G at position 89 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide G at position 174 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of UG (P5);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 204 and a nucleotide G at position 205 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P6);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 252 and a nucleotide G at position 253 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P8); or

5) the intron internal split site II and the intron internal split site I are a nucleotide C at position 370 and a nucleotide G at position 371 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P9).

In some embodiments of the present invention, when the self-splicing intron is *Staphylococcus* phage Twort Group I Intron (Twort) (SEQ ID NO:3),

1) the intron internal split site II and the intron internal split site I are a nucleotide U at position 70 and a nucleotide A at position 71 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P5);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 103 and a nucleotide A at position 104 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P6);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 132 and a nucleotide U at position 133 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AU (P7);

4) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide C at position 174 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UC (P8);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 206 and a nucleotide A at position 207 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P9);

6) the intron internal split site II and the intron internal split site I are a nucleotide G at position 226 and a nucleotide U at position 227 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of GU (P9.1);

7) the intron internal split site II and the intron internal split site I are a nucleotide U at position

23 and a nucleotide A at position 24 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P2); or

8) the intron internal split site II and the intron internal split site I are a nucleotide A at position 82 and a nucleotide A at position 83 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P4-5).

In some embodiments of the present invention, when the self-splicing intron is *Scytonema-hofmani* tRNA fMet group I intron (Sh) (SEQ ID NO:5),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 129 and a nucleotide A at position 130 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P6).

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 60 and a nucleotide A at position 61 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P2).

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 109 and a nucleotide A at position 110 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P4-5).

4) the intron internal split site II and the intron internal split site I are a nucleotide A at position 213 and a nucleotide A at position 214 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of UA (P9).

In some embodiments of the present invention, when the self-splicing intron is *Agrobacterium-tumefaciens* group I intron (At) (SEQ ID NO:6), the intron internal split site II and the intron internal split site I are a nucleotide A at position 107 and a nucleotide A at position 108 from 5' end of the At intron respectively, and when connected form a nucleotide sequence of AA (P6).

In some embodiments of the present invention, the 3' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 3' terminal portion of the self-splicing intron, and the 5' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 5' terminal portion of the self-splicing intron.

In some embodiments of the present invention, the self-splicing intron is the cyanobacterium

Anabaena Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% or 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

In some embodiments of the present invention, the self-splicing intron is AnaX (SEQ ID NO:1), wherein the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8.

In some embodiments of the present invention, the self-splicing intron is the Azoarcus Group I Intron (Azo), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 9, 11, 13, 15, 17 or 19 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 9, 11, 13, 15, 17 or 19, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 10, 12, 14, 16, 18 or 20 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 10, 12, 14, 16, 18 or 20.

In some embodiments of the present invention, the self-splicing intron is the Azoarcus Group I Intron (Azo) (SEQ ID NO:2), wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment respectively are: SEQ ID NO: 9 and SEQ ID NO: 10, SEQ ID NO: 11 and SEQ ID NO: 12, SEQ ID NO: 13 and SEQ ID NO: 14, SEQ ID NO: 15 and SEQ ID NO: 16, SEQ ID NO: 17 and SEQ ID NO: 18.

In some embodiments of the present invention, the self-splicing intron is the Scytalidium dimidiatum Group I Intron (Sd) (SEQ ID NO:4), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 37, 39, 41, 43 or 45 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least

85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 37, 39, 41, 43 or 45, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 38, 40, 42, 44 or 46 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 38, 40, 42, 44 or 46.

In some embodiments of the present invention, the self-splicing intron is the *Scytalidium dimidiatum* Group I Intron (Sd) (SEQ ID NO:4), wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment respectively are: SEQ ID NO: 37 and SEQ ID NO: 38, SEQ ID NO: 39 and SEQ ID NO: 40, SEQ ID NO: 41 and SEQ ID NO: 42, SEQ ID NO: 43 and SEQ ID NO: 44, SEQ ID NO: 45 and SEQ ID NO: 46.

In some embodiments of the present invention, the self-splicing intron is the *Staphylococcus phage Twort* Group I Intron (Twort) (SEQ ID NO:3), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36.

In some embodiments of the present invention, the self-splicing intron is the *Staphylococcus phage Twort* Group I Intron (Twort) (SEQ ID NO:3), wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment respectively are: SEQ ID NO: 21 and SEQ ID NO: 22, SEQ ID NO: 23 and SEQ ID NO: 24, SEQ ID NO: 25 and SEQ ID NO: 26, SEQ ID NO: 27 and SEQ ID NO: 28, SEQ ID NO: 29 and SEQ ID NO: 30, SEQ ID NO: 31 and SEQ ID NO: 32, SEQ ID NO: 33 and SEQ ID NO: 34, SEQ ID NO: 35 and SEQ ID NO: 36.

In some embodiments of the present invention, the self-splicing intron is the *Scytonema hofmani* Trna fMet group I intron (Sh) (SEQ ID NO:5), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 47, 49, 51, 53, or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%,

at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 47, 49, 51, 53, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 48, 50, 52, 54, or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 48, 50, 52, 54.

In some embodiments of the present invention, the self-splicing intron is the *Scytonema hofmani* Trna fMet group I intron (Sh) (SEQ ID NO:5), wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment respectively are: SEQ ID NO: 47 and SEQ ID NO: 48, SEQ ID NO: 49 and SEQ ID NO: 50, SEQ ID NO: 51 and SEQ ID NO: 52, SEQ ID NO: 53 and SEQ ID NO: 54.

In some embodiments of the present invention, the self-splicing intron is the *Agrobacterium tumefaciens* group I intron (At) (SEQ ID NO:6), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 55 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 55, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 56 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 56.

In some embodiments of the present invention, the self-splicing intron is the *Agrobacterium tumefaciens* group I intron (At) (SEQ ID NO:6), wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment respectively are: SEQ ID NO: 55 and SEQ ID NO: 56.

In one aspect, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a first residual circularizing element;
- c) optionally, a nucleotide sequence of interest;
- d) a second residual circularizing element; and
- e) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in

combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element through the self-splicing of the circular RNA precursor,

wherein the total length of the first residual circularizing element and the second residual circularizing element is about 2 to about 200 nucleotides.

As used herein, a "residual circularizing element" refers to an exon-derived sequence that is additionally introduced to facilitate the circularization process. The residual circularizing element is a sequence derived from the native exon of the self-splicing intron (the exon flanking the self-splicing intron) and is involved in or required for circularization by the self-splicing intron and participates in circularization together with the self-splicing intron but is retained in the final circular RNA. In particular, the residual circularizing element is capable of being recognized and/or spliced by the self-splicing intron (or a combination of the first self-splicing intron fragment and the second self-splicing intron fragment), and thus is required for circularization.

As used herein, "exon region" in the residual circularizing element is a sequence derived from the native exon of the self-splicing intron (the exon flanking the self-splicing intron) and capable of being recognized and/or spliced by the self-splicing intron (or a combination of the first self-splicing intron fragment and the second self-splicing intron fragment), and thus is required for circularization. An "exon region" can also be referred as a "splicing site sequence" herein.

In some embodiments, the first residual circularizing element comprises or consists of from 5' to 3' direction a 3' exon region and optionally a first spacer. In the circular RNA precursor, the 5' end of the 3' exon region is directly connected to the 3' end of the 3' self-splicing intron fragment.

In some embodiments, the 3' exon region is derived from the native 3' exon of the self-splicing intron (the exon flanking (downstream of) the 3' end of the self-splicing intron) or a contiguous fragment thereof starting from the 5' terminal nucleotide.

In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is no greater than about 500 nucleotides, e.g., no greater than about 200, about 100, about 90, about 80, about 70, about 60, about 50, about 40, about 30, about 20, about 15, about 10, about 5 nucleotides, about 4, nucleotides, about 3 nucleotides, about 2 nucleotides, about 1 nucleotide or 0 nucleotide.

In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is from about 2 to about 200 nucleotides, or any integer number therebetween. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is about 2 to about 30 nucleotides, or any integer number therebetween. In some embodiments, the total length of the first residual

circularizing element and the second residual circularizing element is from about 2 to about 20 nucleotides, or any integer number therebetween. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is from about 2 to about 15 nucleotides, or any integer number therebetween. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is about 2 to about 10 nucleotides, or any integer number therebetween.

In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is at least 5 nucleotides in length. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is at least 10 nucleotides in length. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is at least 15 nucleotides in length.

In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is from about 5 to about 100 nucleotides, or any integer number therebetween. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is from about 10 to about 100 nucleotides, or any integer number therebetween. In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is from about 15 to about 100 nucleotides, or any integer number therebetween.

In some embodiments, the total length of the first residual circularizing element and the second residual circularizing element is 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200 nucleotides.

In some embodiments, the first residual circularizing element and the second residual circularizing element are configured such that the circular RNA comprising them has reduced immunogenicity.

In some embodiments, the first residual circularizing element and the second residual circularizing element are configured such that the circular RNA or circular RNA precursor comprising them has comparable or increased circularization efficiency.

In some embodiments, the first residual circularizing element comprises or consists of from 5' to 3' direction a 3' exon region and optionally a first spacer. In the circular RNA precursor, the 5' end of the 3' exon region is directly connected to the 3' end of the 3' self-splicing intron fragment.

As used herein, "exon region" in the residual circularizing element is a sequence derived from

the native exon of the self-splicing intron (the exon flanking the self-splicing intron) and capable of being recognized and/or spliced by the self-splicing intron (or a combination of the first self-splicing intron fragment and the second self-splicing intron fragment), and thus is required for circularization. An “exon region” can also be referred as a “splicing site sequence” herein.

In some embodiments, the 3' exon region is derived from the native 3' exon of the self-splicing intron (the exon flanking (downstream of) the 3' end of the self-splicing intron) or a contiguous fragment thereof starting from the 5' terminal nucleotide.

In some embodiments, the 3' exon region is the entire native 3' exon of the self-splicing intron, or has at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% or at least 99% sequence identity with the entire native 3' exon of the self-splicing intron, or has 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more nucleotide substitutions, deletions or additions compared to the entire native 3' exon of the self-splicing intron.

In some embodiments, the 3' exon region is a contiguous fragment starting from the 5' terminal nucleotide of the native 3' exon. In some embodiments, the 3' exon region has at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% or at least 99% sequence identity with a contiguous fragment starting from the 5' terminal nucleotide of the native 3' exon. In some embodiments, the 3' exon region has 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more nucleotide substitutions, deletions or additions compared to a continuous fragment starting from the 5' terminal nucleotide of the native 3' exon.

In some embodiments, the contiguous fragment starting from the 5' terminal nucleotide of the native 3' exon comprises or consists of at least 1%, at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% nucleotides of the native 3' exon. In some embodiments, the contiguous fragment starting from the 5' terminal nucleotide of the native 3' exon is at least 1 nucleotide in length, such as at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15 nucleotides, at least 20, at least 25, at least 50 or more nucleotides in length. In some embodiments, the contiguous fragment starting from the 5' terminal nucleotide of the native 3' exon is 1 nucleotide in length or up to 2, up to 3, up to 4, up to 5, up to 6, up to 7, up to 8, up to 9, up to 10, up to 15, up to 20, up to 25, up to 50 nucleotides in length or up to the total length of the native 3' exon.

In some embodiments, where the self-splicing intron is a Group I intron, the 3' exon region at least comprises a sequence (such as a sequence of about 1 to about 20 nucleotides) at it 5'

terminus which can pair with the P1 region of the corresponding Group I intron to form a P10 duplex region.

It is believed that for self-splicing of Group I introns, the consecutive one or more nucleotides (such as at least about 1 to about 7 nucleotides) from the 5' end of the native 3' exon can pair with the P1 region to form a P10 duplex region, and thus plays an important role in self-splicing. Definitions of the P1 and P10 regions of Group I introns are known in the art and can be determined, for example, with reference to the following documents: Burke, J. M., et al., (1987) Structural conventions for group I introns; Stahley, R. M., et al (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis; and/or Woodson, A. S., (2005) Structure and assembly of group I introns.

In some embodiments, the second residual circularizing element comprises or consists of from 3' to 5' direction a 5' exon region and optionally a second spacer. In the circular RNA precursor, the 3' end of the 5' exon region is directly connected to the 5' end of the second self-splicing intron fragment.

In some embodiments, the 5' exon region is derived from the native 5' exon of the self-splicing intron (the exon flanking (downstream of) the 5' end of the self-splicing intron) or a contiguous fragment thereof starting from the 3' terminal nucleotide.

In some embodiments, the 5' exon region is the entire native 5' exon of the self-splicing intron, or has at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% or at least 99% sequence identity with the entire native 5' exon of the Group I intron, or has 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more nucleotide substitutions, deletions or additions compared to the entire native 5' exon of the self-splicing intron.

In some embodiments, the 5' exon region is a contiguous fragment starting from the 3' terminal nucleotide of the native 5' exon. In some embodiments, the 5' exon region has at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% or at least 99% sequence identity with a contiguous fragment starting from the 3' terminal nucleotide of the native 5' exon. In some embodiments, the 5' exon region has 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more nucleotide substitutions, deletions or additions compared to a continuous fragment starting from the 3' terminal nucleotide of the native 5' exon.

In some embodiments, the contiguous fragment starting from the 3' terminal nucleotide of the native 5' exon comprises or consists of at least 1%, at least 5%, at least 10%, at least 20%, at least

30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% nucleotides of the native 5' exon. In some embodiments, the contiguous fragment starting from the 3' terminal nucleotide of the native 5' exon is at least 1 nucleotide in length, such as at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15 nucleotides, at least 20, at least 25, at least 50 or more nucleotides in length. In some embodiments, the contiguous fragment starting from the 3' terminal nucleotide of the native 5' exon is 1 nucleotide in length or up to 2, up to 3, up to 4, up to 5, up to 6, up to 7, up to 8, up to 9, up to 10, up to 15, up to 20, up to 25, up to 50 nucleotides in length or up to the total length of the native 5' exon.

In some embodiments, where the self-splicing intron is a Group I intron, the 5' exon region comprises a sequence (for example, a sequence of about 3 to about 8 consecutive nucleotides) at its 3' terminus which can pair with the internal guide sequence (IGS) of the corresponding Group I intron to form a P1 double-stranded region.

It is believed that for self-splicing of Group I introns, about 3 to about 8 consecutive nucleotides from the 3' end of the native 5' exon can pair with the internal guide sequence (IGS) of the intron to form the P1 double-stranded region, thus playing an important role in self-splicing. Definitions of IGS and/or P1 region of Group I introns are known in the art and can be determined, for example, with reference to the following documents: Burke, J. M., et al., (1987) Structural conventions for group I introns; Stahley, R. M., et al (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis; and/or Woodson, A. S., (2005) Structure and assembly of group I introns.

In some embodiments, the first residual circularizing element and the second residual circularizing element comprise spacers of different sequences, or one of them comprises a spacer and the other does not.

In some embodiments, the first residual circularizing element and the second residual circularizing element are configured to be capable of forming a stem-loop structure. In some embodiments, the loop of the stem-loop structure comprises the splicing junction.

In some embodiments, the presence of the stem-loop structure can be predicted and/or determined by the nucleotide sequences of the 3' self-splicing intron fragment (e.g., 3' Group I intron fragment), the first residual circularizing element, the second residual circularizing element, and the 5' self-splicing intron fragment (e.g., 5' Group I intron fragment) involved in circularization. In some embodiments, the presence of a stem-loop structure can be predicted and/or determined from the nucleotide sequence by RNA structure prediction tools such as RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>) or RNAstructure (<https://rna.urmc.rochester.edu/RNAstructureWeb/index.html>).

In some embodiments, the first residual circularizing element comprises the sequence

structure of the following formula: 5'-first loop sequence-first pairing sequence-first non-pairing sequence-3'; and the second residual circularizing element comprises the sequence structure of the following formula: 5'-second non-pairing sequence-second pairing sequence-second loop sequence-3',

wherein the first non-pairing sequence or the second non-pairing sequence may be independently present or absent,

the first pairing sequence and the second pairing sequence can complementarily pair to each other to form the stem of the stem-loop structure, wherein the first loop sequence and the second loop sequence can form the loop of the stem-loop structure, e.g., through self-splicing for circularization.

Typically, the sequences forming the loop of the stem-loop structure are derived from the 3' exon region and/or the 5' exon region.

In some embodiments, where the self-splicing intron is a Group I intron, the first loop sequence comprises or consists of one or more nucleotides (for example, about 1 to about 20 nucleotides) which can pair with the P1 region of the corresponding Group I intron (or the structure formed by the 3' self-splicing intron fragment and the 5' self-splicing intron fragment) to form a P10 duplex region during the circularization.

In some embodiments, the first loop sequence may comprise or consist of a nucleotide sequence of (N)_n, wherein N represents any nucleotides (A, G, U, or C), n represents an integer from 1-20, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some specific embodiments, n is 2, 4, or 5.

In some embodiments, the first loop sequence comprises or consists of the about 1 to about 7 consecutive nucleotides starting from the 5' terminal nucleotide of the native 3' exon of the Group I intron.

In some embodiments, the first loop sequence for example comprises or consists of AAAA, AA, UUUU, UAAA, CAAA, or GAAA.

In some embodiments, where the self-splicing intron is a Group I intron, the second loop sequence comprises or consists of one or more nucleotides (about 3 to about 8 nucleotides) which can pair with the internal guide sequence (IGS) of the corresponding Group I intron (or the structure formed by the 3' self-splicing intron fragment and the 5' self-splicing intron fragment) to form a P1 duplex region during the circularization.

In some embodiments, the second loop sequence comprises or consists of the about 3 to about 8 consecutive nucleotides starting from the 3' terminal nucleotide of the native 5' exon of the Group I intron.

In some embodiments, the second loop sequence for example comprises or consists of CUU

or CUC.

In some embodiments, a loop with a sequence of CUUAAAA, CUUUUUU, CUUAA, CUUGAAA, CUUUAAA, CUUCAAA or CUCAAAA can be formed after circularization.

In some embodiments, the first loop sequence comprises or consists of AAAA and the second loop sequence comprises or consists of CUU. In some specific embodiments, a loop with a sequence of CUUAAAA is formed after circularization.

The pairing sequences forming the stem of the stem-loop structure may be derived from the exon regions, however, it may also be derived from the spacer sequences. Alternatively, the pairing sequence may be derived from an exon region and a spacer sequence, i.e., the pairing sequence comprises at least a portion of an exon region and at least a portion of the spacer.

Without being bound by any theory, the RNA circularization efficiency based on intron self-splicing (e.g., Group I intron self-splicing) is related to the number of base pairs or the type or composition of base pairs in the stem portion of the stem-loop structure formed by the residual circularizing element. The stability of the stem-loop structure (e.g., as can be predicted from calculated free energies) may affect the circularization efficiency.

In some embodiments, the stem portion of the stem-loop structure comprises at least 2 base pairs, such as at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15 or more base pairs, preferably consecutive matched base pairs. In some embodiments, the stem portion of the stem-loop structure comprises 2-15 or more consecutive matched base pairs. In some embodiments, the stem portion of the stem-loop structure comprises 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 base pairs, preferably consecutive matched base pairs. In some embodiments, the stem portion of the stem-loop structure comprises 5 base pairs, preferably consecutive matched base pairs. In some embodiments, the stem portion of the stem-loop structure comprises 6 base pairs, preferably consecutive matched base pairs. In some embodiments, the stem portion of the stem-loop structure comprises 7 base pairs, preferably consecutive matched base pairs.

In some embodiments, the stem portion in the stem-loop structure comprises up to 2 base mismatches, or up to 1 base mismatch, preferably, the stem portion comprises no base mismatches.

In some embodiments, the first pairing sequence comprises only G and the second pairing sequence comprises only Cs. In some embodiments, the first pairing sequence comprises only C and the second pairing sequence comprises only G. In some embodiments, the first pairing sequence includes only A and the second pairing sequence includes only U.

In some embodiments, the circular RNA precursor further comprises a 5' homology arm sequence and a 3' homology arm sequence capable of complementary pairing to form a homology arm double-stranded region. In some embodiments, the 5' homology arm sequence is upstream of

the 5' end of the 3' self-splicing intron fragment and the 3' homology arm sequence is downstream of the 3' end of the 5' self-splicing intron fragment.

In some embodiments, the nucleotide sequence of interest provided herein comprises at least one protein-coding sequence and a translation initiation element such as an internal ribosome entry site (IRES) operably linked thereto. Here, "operably linked" refers to that the translation initiation element such as IRES can mediate translation of the encoded protein. In some embodiments, in the circular RNA precursor, the translation initiation element, such as an IRES, is located upstream of the 5' end of the at least one protein-coding sequence, or the translation initiation element, such as an IRES, is located downstream of the 3' end of the at least one protein-coding sequence.

Protein-coding sequences can encode proteins of eukaryotic, prokaryotic or viral origin. In certain embodiments, the protein can be any protein for therapeutic or diagnostic use. For example, the protein coding region can encode human proteins, antigens, antibodies, gene editing enzymes such as CRISPR nucleases, and the like. For example, the encoded protein can be a chimeric antigen receptor, an immunomodulatory protein, and/or a transcription factor, and the like. Some specific examples include, but are not limited to, *EGF*, *FGF1*, *RBD*, *G6PC*, *PAH*, *HGF*, and the like.

In some embodiments, the at least one protein-coding sequence can encode IL-2 or an aquaporin, such as AQP-1 protein or a protein having the same activity as the naturally occurring AQP-1 protein and derived from the naturally occurring AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the naturally occurring AQP-1 protein. In embodiments, the protein-coding sequence of the present invention encodes the AQP-1 protein. In embodiments, the AQP-1 protein is a mammalian AQP-1 protein, in particular, a human AQP-1 protein.

In some embodiments, the at least one protein-coding sequence of the invention comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

In some embodiments, the translation initiation element of the invention comprises: IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof.

The IRES sequence may be selected from, but is not limited to, the following IRES sequences: Taura syndrome virus, blood-sucking bug virus, Tyler's encephalomyelitis virus, simian virus 40, red fire ant virus 1, cereal constrictor virus, reticulovirus Endothelial hyperplasia virus, Forman

poliovirus 1, soybean inchworm virus, Kashmir bee virus, human rhinovirus 2, glass leafhopper virus-1, human immunodeficiency virus type 1, glass leafhopper virus-1, lice P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinovirus, Tea inchworm-like virus, Encephalomyocarditis virus (EMCV), *Drosophila* C virus, Cruciferous tobacco Virus, cricket paralysis virus, bovine viral diarrhea virus 1, black queen cell virus, aphid lethal paralysis virus, avian encephalomyelitis virus, acute bee paralysis virus, hibiscus yellow ring spot virus, swine fever virus, human FGF2, human SFTPA1, Human AML1/RUNX1, *Drosophila* Antennae, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 α , Human n.myc, mouse Gtx, human p27kipl, human PDGF2/c-sis, human p53, human Pim-1, mouse Rbm3, *Drosophila* reaper, canine Scamper, *Drosophila* Ubx, human UNR, mouse UtrA, human VEGF-A, human XIAP, *Drosophila* hairless, *Saccharomyces cerevisiae* TFIID, *Saccharomyces cerevisiae* YAP1, human c-src, human FGF-1, simian picornavirus, turnip crepe disease virus, eIF4G aptamer, Coxsackie Virus B3 (CVB3) or Coxsackie virus A (CVA1/2). Wild-type IRES sequences can also be modified and used in the present invention. Preferably, the IRES is CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2.

For example, the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% sequence identity to one of SEQ ID NOs: 120-526.

In some embodiments, the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

In some embodiments, the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

In some embodiments, the homology arm sequence I or the homology arm sequence II is about 5-500, about 5-450, about 5-400, about 5-350, about 5-300, about 5-250, about 5-200, about 5-150, about 5-100, about 5-80, about 5-60, about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-15 or about 5-10 nucleotides in length, preferably, about 20, 25, 40, 50, 100, or 300 nucleotides in length.

In some embodiments, the two homology arm sequences may be polyA and polyT, respectively, or polyG and polyC, respectively.

In some embodiments, one of the homology arm sequence I and the homology arm sequence II of the invention comprises a nucleotide sequence of any one of SEQ ID NO: 71-94 or a

nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 71-94, while the other homology arm sequence comprises the corresponding complementary sequence.

In some embodiments, the circular RNA precursor of the invention comprises a nucleotide sequence selected from any one of SEQ ID NO: 527, 532 and 537-542, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 527, 532 and 537-542.

In some embodiments, the nucleotide sequence of interest is a non-coding sequence. For example, the non-coding sequence can be POLR2A, antisense RNA, aptamer, guide RNA, or non-protein-coding RNA existing in any organism, and the like. The non-coding sequence may or may not contain a specific secondary structure.

In one aspect, provided herein is a nucleic acid vector for generating a circular RNA molecule, said vector comprises a coding sequence of the circular RNA precursor of the present invention.

As used herein, "vector" refers to a DNA derived from a virus, plasmid or cell of a higher organism into which a foreign DNA fragment can be or has been inserted for cloning and/or expression purposes. In certain embodiments, the vector can be stably maintained in the organism. The vector may contain, for example, an origin of replication, a selectable marker or a reporter gene, such as antibiotic resistance or GFP, and/or a multiple cloning site (MCS). The term includes linear DNA fragments (e.g., PCR products, linear plasmid fragments), plasmid vectors, viral vectors, cosmid, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), and the like.

In some embodiments, the nucleic acid vector further comprises an RNA polymerase promoter sequence operably linked to the coding sequence of the circular RNA precursor. The operably linked promoter allows in vivo and/or in vitro transcription of the circular RNA precursor. The promoter is, for example, a T7 RNA polymerase promoter, a T6 viral RNA polymerase promoter, a SP6 viral RNA polymerase promoter, a T3 viral RNA polymerase promoter or a T4 viral RNA polymerase promoter.

There is provided herein a novel Clean PIE system for preparation of circular RNA, which can improve the sequence accuracy of the circular RNA, reduce the change of the secondary structure of circular RNA, thereby decreasing the immunogenicity of the circular RNA, increasing the stability of the circular RNA in cells and reducing the safety risk in clinical applications.

In one aspect, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a circularizing sequence fragment II comprising a ribozyme recognition site II at its 5' end and a circularizing sequence fragment I comprising a ribozyme recognition site I at its 3' end,

wherein the circularizing sequence fragment I is derived from a 5' terminal portion of a circularizing sequence, and the circularizing sequence fragment II is derived from a 3' terminal portion of a circularizing sequence, and

wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the circularizing sequence fragment I is connected with the circularizing sequence fragment II to form the circularizing sequence.

In one aspect, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a circularizing sequence fragment II comprising a ribozyme recognition site II at its 5' end and a circularizing sequence fragment I comprising a ribozyme recognition site I at its 3' end,

wherein the circularizing sequence fragment I is the 5' terminal portion of a circularizing sequence, and the circularizing sequence fragment II is the 3' terminal portion of a circularizing sequence, and

wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the circularizing sequence fragment I is connected with the

circularizing sequence fragment II to form the circularizing sequence.

In embodiments, the nucleotide sequence of interest comprises one or more sequences selected from the following group between the circularizing sequence fragment II and the circularizing sequence fragment I: at least one protein coding sequences, translation initiation element (TIE), a non-TIE functional element or any combination thereof.

As used herein, in the Clean-PIE system, "circularizing sequence" refers to a nucleotide sequence which is part of the nucleotide sequence of interest. The circularization sequence is involved in or required for circularization by the self-splicing intron and participates in circularization together with the self-splicing intron. In embodiments, the circularizing sequence is a protein-coding sequence, a non-coding sequence, a translation initiation element (TIE), or a non-TIE functional element. As used herein, in the Clean-PIE system, "circularizing sequence fragment I" and "circularizing sequence fragment II" each is part of the circularizing sequence. In the Clean-PIE system of the present invention, circularizing sequence fragments I and II are connected to form the circularizing sequence and such a circularizing sequence formed will retain its biological functions or activities.

As used herein, in the Clean-PIE system, "ribozyme recognition site" refers to a segment of nucleotide sequence which belongs to a part of the circularizing sequence. The ribozyme recognition site is involved in or required for circularization by the self-splicing intron and participates in circularization together with the self-splicing intron but is retained as a part of the circularizing sequence in the final circular RNA. In embodiments, the ribozyme recognition site I is connected with ribozyme recognition II to form the ribozyme recognition site. In embodiments, the ribozyme recognition site comprises 2 to 50 nucleotides, for example 2 to 40, 2 to 30, 2 to 25, 2 to 20, 2 to 15, 2 to 10 nucleotides, in particular 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more nucleotides. In embodiments, the ribozyme recognition site I is comprised of 1 to 10 nucleotides, in particular 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides. In embodiments, the ribozyme recognition site II is comprised of 1 to 10 nucleotides, in particular 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides.

As used herein, "a nucleotide motif" refers to a nucleotide sequence pattern that is related to biological function of the ribozyme recognition site. In the present invention, the nucleotide motif is located at the 3' end of ribozyme recognition site I.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisting of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

Only for the purpose of illustrating the present invention, it is provided herein the following examples of the "Clean-PIE system" of the invention. These particular examples include "Clean-

PIE-Cargo system”, “Clean-PIE-TIE system”, “Clean-PIE-NonTIE system” and “Clean-PIE-Non-coding Cargo system”. A person skilled in the art understands that any variations, deviations, modifications etc. that are made to the “Clean-PIE system” illustrated as follows do not deviate from the present invention and are thus comprised in the scope of the present invention.

For Clean-PIE-Cargo system, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES), and a protein-coding sequence fragment I;

wherein the protein-coding sequence fragment I is derived from a 5' terminal portion of at least one protein-coding sequence, and the protein-coding sequence fragment II is derived from a 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence.

The present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES), and a protein-coding sequence fragment I;

wherein the protein-coding sequence fragment I is the 5' terminal portion of at least one protein-coding sequence, and the protein-coding sequence fragment II is the 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence.

In embodiments, the at least one protein-coding sequence is a sequence encoding a diagnostic or therapeutic protein. In embodiments, the protein encoded by the protein-coding sequence is a mammalian protein, for example a human protein. In embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more proteins. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more diagnostic or therapeutic protein. In some embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences, and another translation initiation element may be linked between any two adjacent protein-coding sequences, and such another translation initiation element comprises: IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof;

In some embodiments, the at least one protein-coding sequence comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

It is understood that, in the Clean-PIE-Cargo system of the invention, ribozyme recognition site I and ribozyme recognition site II are respectively a segment of nucleotide sequence which belongs to a part of a cargo sequence, and configured such that the circular RNA comprising them has reduced immunogenicity. In some embodiments, the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA precursor comprising them has comparable or increased circularization efficiency.

In some embodiments, the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to

about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the ribozyme recognition site I comprises the sequence structure of the following formula: 5'-third non-pairing sequence-third pairing sequence-third loop sequence-3'; and the ribozyme recognition site II comprises the sequence structure of the following formula: 5'-fourth loop sequence-fourth pairing sequence-fourth non-pairing sequence-3'.

In some examples, the third non-pairing sequence or the fourth non-pairing sequence may be independently present or absent.

In some embodiments, the ribozyme recognition site I comprises or consists of a 'nucleotide motif as defined above. In embodiments, the ribozyme recognition site I comprises or consists of

from 5' to 3' direction a nucleotide motif, wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide motif having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70; preferably, the ribozyme recognition site I comprises a nucleotide sequence of GAGGUCUU, AGAGACUU, UCCAGCUU or UCUGGCAU.

In some embodiments, the ribozyme recognition site II comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69; preferably, the ribozyme recognition site II comprises a nucleotide sequence of AAACCUC, AAUCUCU, AACUGGA or AACCAGC.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69.

In embodiments, the ribozyme recognition site I is selected from a nucleotide sequence of UGUAACUU, GAGGUCUU, AGAGACUU, UCCAGCUU, or UCUGGCAU.

In embodiments, the ribozyme recognition site II is selected from a nucleotide sequence of AGAAGUAAC, AAAAUUACA, AAACCUC, AAUCUCU, AACUGGA, or AACCAGC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AGAAGUAAC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AAAAUUACA.

In embodiments, the ribozyme recognition site I is GAGGUCUU, and the ribozyme

recognition site II is AAACCUC.

In embodiments, the ribozyme recognition site I is AGAGACUU, and the ribozyme recognition site II is AAUCUCU.

In embodiments, the ribozyme recognition site I is UCCAGCUU, and the ribozyme recognition site II is AACUGGA.

In embodiments, the ribozyme recognition site I is UCUGGCAU, and the ribozyme recognition site II is AACCAGC.

For Clean-PIE-TIE system, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a translation initiation element fragment II, at least one protein-coding sequence, and a translation initiation element fragment I;

wherein the translation initiation element fragment I is derived from a 5' terminal portion of a translation initiation element, and the translation initiation element fragment II is derived from a 3' terminal portion of the translation initiation element, and

the translation initiation element fragment I comprises a ribozyme recognition site I located at its 3' end, the translation initiation element fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the translation initiation element fragment I is connected with the translation initiation element fragment II to form the translation initiation element which can recruit ribosome to initiate translation.

The present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the

circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a translation initiation element fragment II, at least one protein-coding sequence, and a translation initiation element fragment I;

wherein the translation initiation element fragment I is the 5' terminal portion of a translation initiation element, and the translation initiation element fragment II is the 3' terminal portion of the translation initiation element, and

the translation initiation element fragment I comprises a ribozyme recognition site I located at its 3' end, the translation initiation element fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the translation initiation element fragment I is connected with the translation initiation element fragment II to form the translation initiation element which can recruit ribosome to initiate translation.

In embodiments, the at least one protein-coding sequence is a sequence encoding a diagnostic or therapeutic protein. In embodiments, the protein encoded by the protein-coding sequence is a mammalian protein, for example a human protein. In embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more proteins. In embodiments, the at least one protein-coding sequence is a sequence encoding two or more diagnostic or therapeutic protein. In some embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences, and another translation initiation element may be linked between any two adjacent protein-coding sequences, and such another translation initiation element comprises: IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof.

In some embodiments, the at least one protein-coding sequence of the invention comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

It will be understood that, in the Clean-PIE-TIE system of the invention, ribozyme recognition site I and ribozyme recognition site II are respectively a nucleotide sequence that belongs to a part of an IRES sequence, and configured such that the circular RNA comprising them has reduced immunogenicity. In some embodiments, the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA precursor comprising them

has comparable or increased circularization efficiency.

In some embodiments, the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the ribozyme recognition site I comprises the sequence structure of the following formula: 5'-third non-pairing sequence-third pairing sequence-third loop sequence-3'; and the ribozyme recognition site II comprises the sequence structure of the following formula:

5'-fourth loop sequence-fourth pairing sequence-fourth non-pairing sequence-3'.

In some embodiments, the third non-pairing sequence or the fourth non-pairing sequence may be independently present or absent.

In some embodiments, the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

In some embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consists of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In some embodiments, the ribozyme recognition site II can be selected from various sequence.

In some embodiments, the ribozyme recognition site II comprises or consists of a 3' exon region as defined above, and the ribozyme recognition site I comprises or consists of a 5' exon region as defined above.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70; preferably, the ribozyme recognition site I comprises a nucleotide sequence of GAGGUCUU, AGAGACUU, UCCAGCUU or UCUGGCAU.

In some embodiments, the ribozyme recognition site II comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69; preferably, the ribozyme recognition site II comprises a nucleotide sequence of AAACCUC, AAUCUCU, AACUGGA or AACCAGC.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69.

In embodiments, the ribozyme recognition site I is selected from a nucleotide sequence of

UGUAACUU, UGUAACUU, GAGGUCUU, AGAGACUU, UCCAGCUU, or UCUGGCAU.

In embodiments, the ribozyme recognition site II is selected from a nucleotide sequence of AGAAGUAAC, AAAAUUACA, AAACCUC, AAUCUCU, AACUGGA, or AACCAGC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AGAAGUAAC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AAAAUUACA.

In embodiments, the ribozyme recognition site I is GAGGUCUU, and the ribozyme recognition site II is AAACCUC.

In embodiments, the ribozyme recognition site I is AGAGACUU, and the ribozyme recognition site II is AAUCUCU.

In embodiments, the ribozyme recognition site I is UCCAGCUU, and the ribozyme recognition site II is AACUGGA.

In embodiments, the ribozyme recognition site I is UCUGGCAU, and the ribozyme recognition site II is AACCAGC.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest;
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a non-TIE functional element fragment II, a translation initiation element, and at least one protein-coding sequence, and a non-TIE functional element fragment I;

wherein the non-TIE functional element fragment I is derived from a 5' terminal portion of a non-TIE functional element, and the non-TIE translation initiation element fragment II is derived from a 3' terminal portion of the non-TIE functional element, and

the non-TIE functional element fragment I comprises a ribozyme recognition site I located at its 3' end, the non-TIE functional fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-TIE functional element fragment I is connected with the non-TIE functional element fragment II to form the functional element.

As used herein, the term “non-TIE functional element” refers to a functional element that is not involved in translation initiation, such as a nucleotide sequence for regulating RNA translation, RNA circularization or RNA transcription.

In embodiments, the non-TIE functional element is a 3' UTR or a replicon. In embodiments, the at least one protein-coding sequence comprises at least two protein-coding sequences. Such at least two protein-coding sequences may connect with each other by spacers or further translation initiation elements. In embodiments, the translation initiation element is selected from an IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A modification, a sequence complementary to ribosome 18S rRNA, or any combination thereof. In embodiments, the translation initiation element is an IRES sequence.

In some embodiments, the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about

100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

In some embodiments, the ribozyme recognition site II can be selected from various sequences.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide motif as defined above. In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif, wherein the nucleotide motif consists of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70; preferably, the ribozyme recognition site I comprises a nucleotide sequence of GAGGUCUU, AGAGACUU, UCCAGCUU or UCUGGCAU.

In some embodiments, the ribozyme recognition site II comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69; preferably, the ribozyme recognition site II comprises a nucleotide sequence of AAACCUC, AAUCUCU, AACUGGA or AACCAGC.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69.

In embodiments, the ribozyme recognition site I is selected from a nucleotide sequence of UGUAACUU, UGUAACUU, GAGGUCUU, AGAGACUU, UCCAGCUU, or UCUGGCAU.

In embodiments, the ribozyme recognition site II is selected from a nucleotide sequence of AGAAGUAAC, AAAAUUACA, AAACCUC, AAUCUCU, AACUGGA, or AACCAGC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AGAAGUAAC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AAAAUUACA.

In embodiments, the ribozyme recognition site I is GAGGUCUU, and the ribozyme recognition site II is AAACCUC.

In embodiments, the ribozyme recognition site I is AGAGACUU, and the ribozyme recognition site II is AAUCUCU.

In embodiments, the ribozyme recognition site I is UCCAGCUU, and the ribozyme recognition site II is AACUGGA.

In embodiments, the ribozyme recognition site I is UCUGGCAU, and the ribozyme recognition site II is AACCAGC.

In one aspect, it is an object of the present invention to provide a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest;
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a non-coding sequence fragment II, and a non-coding sequence fragment I;

wherein the non-coding sequence fragment I is derived from a 5' terminal portion of a non-coding sequence, and the non-coding sequence fragment II is derived from a 3' terminal portion of the non-coding sequence, and the non-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the non-coding sequence fragment II comprises a ribozyme

recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-coding sequence fragment I is connected with the non-coding sequence fragment II to form the complete non-coding sequence.

In embodiments, the non-coding sequence is an endogenous circRNA.

In some embodiments, the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88,

89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

In some embodiments, the ribozyme recognition site II can be selected from various sequences.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide motif as defined above. In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif, wherein the nucleotide motif consists of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70; preferably, the ribozyme recognition site I comprises a nucleotide sequence of GAGGUCUU, AGAGACUU, UCCAGCUU or UCUGGCAU.

In some embodiments, the ribozyme recognition site II comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69; preferably, the ribozyme recognition site II comprises a nucleotide sequence of AAACCUC, AAUCUCU, AACUGGA or AACCAGC.

In some embodiments, the ribozyme recognition site I comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69.

In embodiments, the ribozyme recognition site I is selected from a nucleotide sequence of

UGUAACUU, UGUAACUU, GAGGUCUU, AGAGACUU, UCCAGCUU, or UCUGGCAU.

In embodiments, the ribozyme recognition site II is selected from a nucleotide sequence of AGAAGUAAC, AAAAUUACA, AAACCUC, AAUCUCU, AACUGGA, or AACCAGC.

In embodiments, the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consists of any of SEQ ID NO: 95-116, in particular said nucleotide motif consists of any one of CUU, CUC or CAU.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AGAAGUAAC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AAAAUUACA.

In embodiments, the ribozyme recognition site I is GAGGUCUU, and the ribozyme recognition site II is AAACCUC.

In embodiments, the ribozyme recognition site I is AGAGACUU, and the ribozyme recognition site II is AAUCUCU.

In embodiments, the ribozyme recognition site I is UCCAGCUU, and the ribozyme recognition site II is AACUGGA.

In embodiments, the ribozyme recognition site I is UCUGGCAU, and the ribozyme recognition site II is AACCAGC.

In some embodiments, the nucleotide sequence of interest is at least 10, 20, 40, 60, 80, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 10000, 20000 nucleotides in length.

In some embodiments, the circular RNA is at least 10, 20, 40, 60, 80, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 10000, 20000 nucleotides in length. In some embodiments, the circular RNA is at least about 10 nucleotides in length. In some embodiments, the circular RNA is about 50 nt or less. In some embodiments, the circular RNA is about 80 nt or less. In some embodiments, the circular RNA is about 100 nt or less. In some embodiments, the circular RNA is about 150 nt or less. In some embodiments, the circular RNA is about 200 nt or less. In some embodiments, the circular RNA is about 250 nt or less. In some embodiments, the circular RNA is about 300 nt or less. In some embodiments, the circular RNA is about 350 nt or less. In some embodiments, the circular RNA is about 400 nt or less. In some embodiments, the circular RNA is about 450 nt or less. In some embodiments, the circular RNA is about 500 nt or less. In some embodiments, the circular RNA is about 600 nt or less. In some embodiments, the circular RNA is about 700 nt or less. In some embodiments, the circular RNA is about 800 nt or less. In some embodiments, the circular RNA is about 900 nt or less. In some embodiments, the circular RNA is at least about 1000 nt.

In further aspect, the present invention provides a pharmaceutical composition comprising the circular RNA precursor of the invention and/or the nucleic acid vector of the invention and/or the circular RNA of the invention, and a pharmaceutically acceptable carrier.

In some embodiments, a pharmaceutically acceptable carrier may include, but are not limited to, buffers, excipients, stabilizers or preservatives. Examples of pharmaceutically acceptable carriers are physiologically compatible solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, such as salts, buffers, carbohydrates, antioxidants, aqueous or non-aqueous carriers, preservatives, wetting agents, surfactants or emulsifying agents or combinations thereof. The amount of a pharmaceutically acceptable carrier in a pharmaceutical composition can be determined experimentally based on the activity of the carrier and the desired properties of the formulation, such as stability and/or minimal oxidation.

In another aspect, the present invention provides a method for preparing a circular RNA, the method comprises:

- 1) providing a circular RNA precursor of the present invention or obtaining a circular RNA precursor by transcribing from the nucleic acid vector of the present invention;
- 2) incubating the circular RNA precursor in the presence of a divalent metal cation at a temperature at which RNA circularization occurs; and
- 3) harvesting the circular RNA obtained in step 2).

In some embodiments, the divalent metal cation is Mg^{2+} and/or Mn^{2+} .

In some embodiments, the concentration of the divalent metal cation is at least about 5 mM, e.g., about 5 mM to about 550 mM, e.g., at least about 5 mM, about 10 mM, about 15 mM, at least about 20 mM, at least about 30 mM, at least about 40 mM, at least about 50 mM, at least about 60 mM, at least about 70 mM, at least about 80 mM, at least about 90 mM, at least about 100 mM, at least about 125 mM, at least about 150 mM, at least about 175 mM, at least about 200 mM, at least about 250 mM, at least about 300 mM, at least about 350 mM, at least about 400 mM, at least about 450 mM, at least about 500 mM, at least about 550 mM or higher.

In another aspect, the present invention provides a method for preparing a circular RNA, the method comprises

- a) providing a nucleic acid vector comprising a self-splicing intron-based RNA circularizing elements as a transcription template; and
- b) incubating the nucleic acid vector in an in vitro transcription system comprising a divalent metal cation and an RNA polymerase for a first time period during which the linear RNA produced by in vitro transcription is self-circularized under the action of the RNA circularizing elements to produce a circular RNA.

In some embodiments, the nucleic acid vector is the nucleic acid vector of the present invention.

In some embodiments, the divalent metal cation in the in vitro transcription system is Mg^{2+} .

In some embodiments, the in vitro transcription system further comprises a monovalent metal cation and/or a monovalent anion.

In some embodiments, the monovalent metal cation is Na^+ or K^+ .

In some embodiments, the monovalent metal anion is Cl^- or CH_3COO^- (OAc).

In some embodiments, the concentration of the divalent metal cation in the system during the first time period is from about 5 mM to about 50 mM, e.g., about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM. In some embodiments, the concentration of the divalent metal cation in the system during the first time period is 30 mM.

In some embodiments, the concentration of the monovalent metal cation in the system during the first time period is from about 5 mM to about 50 mM, e.g., about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM, about 60 mM, about 70 mM, about 80 mM, about 90 mM, about 100 mM. In some embodiments, the monovalent metal cation is Na^+ , the concentration in the system during the first time period is 15 mM. In some embodiments, the monovalent metal cation is K^+ , the concentration in the system during the first time period is 90 mM.

In some embodiments, the concentration of the monovalent anion in the system during the first time period is from about 5 mM to about 50 mM, e.g., about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM, about 100 mM, about 150 mM. In some embodiments, the monovalent anion is Cl^- , the concentration in the system during the first time period is 90 mM. In some embodiments, the monovalent anion is CH_3COO^- (OAc), the concentration in the system during the first time period is 125 mM.

In some embodiments, the in vitro transcription and self-circularization occur in the same reaction system.

In some embodiments, the method does not include a step of isolating and/or purifying the linear RNA produced by the in vitro transcription.

Those skilled in the art would know that the in vitro transcription system also comprises various components required for transcription, such as buffers, rATP, rCTP, rUTP, rGTP and the like.

In some embodiments, the buffer of the in vitro transcription system is Tris-HCl buffer, or HEPES buffer, or MES buffer, or citrate buffer, or phosphate buffer. In some embodiments, the

buffer of the in vitro transcription system is HEPES buffer.

In some embodiments, the in vitro transcription system has a pH of about 5- about 8, such as a pH of about 5, about 5.5, about 6, about 6.5, about 7, about 7.5, or about 8. In some embodiments, the in vitro transcription system has a pH of about 7.5.

The RNA polymerase depends on the promoter used in the nucleic acid vector to drive transcription. The RNA polymerase may include, but is not limited to, a T7 RNA polymerase, a T6 viral RNA polymerase, a SP6 viral RNA polymerase, a T3 viral RNA polymerase, or a T4 viral RNA polymerase. In some embodiments, the RNA polymerase is T7 RNA polymerase.

In some embodiments, the first time period is at least 0.5 hours, such as about 0.5 hours to about 24 hours, such as about 0.5 hours, about 1 hour, about 1.5 hours, about 2 hours, about 2.5 hours, about 3 hours, about 3.5 hours, about 4 hours, about 5 hours, about 10 hours, or about 24 hours. In some embodiments, the first time period is 3 hours.

In some embodiments, the incubation for the first time period is carried out at about 16°C to about 60°C, such as at about 16°C, about 17°C, about 18°C, about 19°C, about 20°C °C, about 21°C, about 22°C, about 23°C, about 24°C, about 25°C, about 26°C, about 27°C, about 28°C, about 29°C, about 30°C °C, about 31°C, about 32°C, about 33°C, about 34°C, about 35°C, about 36°C, about 37°C, about 38°C, about 39°C, about 40°C °C, about 41°C, about 42°C, about 43°C, about 44°C, about 45°C, about 46°C, about 47°C, about 48°C, about 49°C, about 50°C °C, about 51°C, about 52°C, about 53°C, about 54°C, about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, or about 60°C. In some embodiments, the incubation for the first time period is carried out at about 37°C.

In some embodiments, after incubation for the first time period in step b), the method further comprises step c):

adding an additional amount of metal cation to the in vitro transcription system and incubating for a second time period; or

changing the buffer of the system, adding a metal cation and incubating for a second time period.

In some embodiments, the metal cation added for the incubation of the second time period is a divalent metal cation, such as Mg^{2+} or Mn^{2+} .

In some embodiments, during the incubation of the second time period, the metal cation is added to a final concentration of at least about 5 mM, such as about 5 mM to about 550 mM, such as at least about 5 mM, at least about 10 mM, at least about 15 mM, at least about 20 mM, at least about 30 mM, at least about 40 mM, at least about 50 mM, at least about 60 mM, at least about 70 mM, at least about 80 mM, at least about 90 mM, at least about 100 mM, at least about 125 mM, at least about 150 mM, at least about 175 mM, at least about 200 mM, at least about 250 mM,

at least about 300 mM, at least about 350 mM, at least about 400 mM, at least about 450 mM, at least about 500 mM, at least about 550 mM or higher.

In some embodiments, the buffer of the system is changed to Tris-HCl buffer, or HEPES buffer, or MES buffer, or citrate buffer, or phosphate buffer during the second time period.

In some embodiments, the pH of the system in the second time period is 5-8, such as pH 5, pH 5.5, pH 6, pH 6.5, pH 7, pH 7.5, or pH 8.

In some embodiments, the second time period is at least 5 minutes, such as about 5 minutes to about 2 hour, such as about 5 minutes, about 10 minutes, about 15 minutes, about 20 minutes, about 25 minutes, about 30 minutes, about 45 minutes, about 50 minutes, about 55 minutes, about 60 minutes, about 120 minutes or more.

In some embodiments, wherein the reaction for the second time period is carried out at about 25°C to about 75°C, such as at about 25°C, about 26°C, about 27°C, about 28°C, about 29°C, about 30°C, about 31°C, about 32°C, about 33°C, about 34°C, about 35°C, about 36°C, about 37°C, about 38°C, about 39°C, about 40°C, about 41°C, about 42°C, about 43°C, about 44°C, about 45°C, about 50°C, about 55°C, about 60°C, about 65°C, about 70°C, or about 75°C.

In some embodiments, the method further comprises a step of recovering or purifying the circular RNA as produced.

In one aspect, the present invention provides a circular RNA produced by the method of the invention.

In one aspect, the present invention provides an in vitro transcription method comprising:

- a) providing a nucleic acid vector as a template for in vitro transcription;
- b) incubating the nucleic acid vector in an in vitro transcription system comprising a divalent metal cation and an RNA polymerase for a first period of time; and
- c) i) adding an additional amount of metal cation to the system, or ii) changing the buffer of the system and adding a metal cation to the system, and incubating the system for a second period of time.

In some embodiments, the in vitro transcription system during the first time period further comprises a monovalent metal cation and/or a monovalent anion.

In some embodiments, the divalent metal cation in the in vitro transcription system during the first time period is Mg^{2+} .

In some embodiments, the monovalent metal cation during the first time period is Na^+ or K^+ .

In some embodiments, the monovalent metal anion during the first time period is Cl^- or CH_3COO^- (OAc).

The RNA polymerase depends on the promoter used in the nucleic acid vector to drive transcription. The RNA polymerase may include, but is not limited to, a T7 RNA polymerase, a

T6 viral RNA polymerase, a SP6 viral RNA polymerase, a T3 viral RNA polymerase, or a T4 viral RNA polymerase. In some embodiments, the RNA polymerase is T7 RNA polymerase.

In some embodiments, the concentration of the divalent metal cation in the system for the first time period is from about 5 mM to about 50 mM, e.g., about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM.

In some embodiments, the concentration of the monovalent metal cation in the system for the first time period is from about 5 mM to about 100 mM, e.g., about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM, about 60 mM, about 70 mM, about 80 mM, about 90 mM, about 100 mM.

In some embodiments, the concentration of the monovalent anion in the system for the first time period is from about 5 mM to about 150 mM, e.g., about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM, about 100 mM, about 150 mM.

In some embodiments, the buffer of the in vitro transcription system for the first time period is Tris-HCl buffer, or HEPES buffer, or MES buffer, or citrate buffer, or phosphate buffer.

In some embodiments, the pH of the in vitro transcription system for the first time period is 5-8, such as pH 5, pH 5.5, pH 6, pH 6.5, pH 7, pH 7.5, or pH 8.

In some embodiments, the first time period is at least 0.5 hours, such as about 0.5 hours to about 24 hours, such as about 0.5 hours, about 1 hour, about 1.5 hours, about 2 hours, about 2.5 hours, about 3 hours, about 3.5 hours, about 4 hours, about 5 hours, about 10 hours, or about 24 hours.

In some embodiments, the incubation for the first time period is carried out at about 16°C to about 60°C, such as at about 16°C, about 17°C, about 18°C, about 19°C, about 20°C, about 21°C, about 22°C, about 23°C, about 24°C, about 25°C, about 26°C, about 27°C, about 28°C, about 29°C, about 30°C, about 31°C, about 32°C, about 33°C, about 34°C, about 35°C, about 36°C, about 37°C, about 38°C, about 39°C, about 40°C, about 41°C, about 42°C, about 43°C, about 44°C, about 45°C, about 46°C, about 47°C, about 48°C, about 49°C, about 50°C, about 51°C, about 52°C, about 53°C, about 54°C, about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, or about 60°C.

In some embodiments, the added metal cation in the system for the second time period is a divalent metal cation, such as Mg^{2+} or Mn^{2+} .

In some embodiments, during the second time period, the metal cation is added to a final concentration of at least about 5 mM, such as about 5 mM to about 550 mM, such as at least about 5 mM, at least about 10 mM, at least about 15 mM, at least about 20 mM, at least about 30 mM,

at least about 40 mM, at least about 50 mM, at least about 60 mM, at least about 70 mM, at least about 80 mM, at least about 90 mM, at least about 100 mM, at least about 125 mM, at least about 150 mM, at least about 175 mM, at least about 200 mM, at least about 250 mM, at least about 300 mM, at least about 350 mM, at least about 400 mM, at least about 450 mM, at least about 500 mM, at least about 550 mM or higher.

In some embodiments, the second time period is at least 5 minutes, such as about 5 minutes to about 2 hour, such as about 5 minutes, about 10 minutes, about 15 minutes, about 20 minutes, about 25 minutes, about 30 minutes, about 45 minutes, about 50 minutes, about 55 minutes, about 60 minutes, about 120 minutes or more.

In some embodiments, the incubation for the second time period is carried out at about 25°C to about 75°C, such as at about 25°C, about 26°C, about 27°C, about 28°C, about 29°C, about 30°C, about 31°C, about 32°C, about 33°C, about 34°C, about 35°C, about 36°C, about 37°C, about 38°C, about 39°C, about 40°C, about 41°C, about 42°C, about 43°C, about 44°C, about 45°C, about 50°C, about 55°C, about 60°C, about 65°C, about 70°C, or about 75°C.

In some embodiments, the method further comprises a step of recovering and/or purifying the RNA obtained in step c).

In one aspect, the present invention provides an RNA produced by the method of the invention.

In another aspect, the present invention provides a method for purifying a circular RNA, the method comprises:

- a) contacting a mixture comprising circular RNA and the circular RNA precursor with a circular RNA-specific probe under a condition that allows the circular RNA-specific probe to specifically bind to and form a complex with the circular RNA;
- b) separating the complex from one or more components in the mixture that are not bound to the circular RNA-specific probe; and
- c) releasing the circular RNA from the complex.

In some embodiments, the circular RNA is prepared by circularizing a linear circular RNA precursor. In some embodiments, the circular RNA is prepared by ligating both ends of a linear circular RNA precursor with an RNA ligase, such as a T4 RNA ligase. In some embodiments, the circular RNA is prepared by the self-splicing ribozyme activity of Group I intron-based circularizing elements contained in the linear circular RNA precursor, e.g., the circular RNA is the circular RNA described in Section I herein or any clause below and/or prepared by the method described in Section II herein or any clause below.

In some embodiments, the circular RNA-specific probe is a single-stranded DNA probe.

In some embodiments, the circular RNA-specific probe is a single-stranded RNA probe.

In some embodiments, the circular RNA-specific probe specifically hybridizes to a region

flanking the circularization junction of the circular RNA.

In some embodiments, the circular RNA-specific probe is at least 10 nucleotides, at least 12 nucleotides, at least 14 nucleotides, at least 16 nucleotides, at least 18 nucleotides in length acid, at least 20 nucleotides, at least 25 nucleotides, at least 30 nucleotides, at least 35 nucleotides in length or longer, for example, the circular RNA-specific probe is 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 25, 27, 28, 29 or 30 nucleotides in length.

In some embodiments, the circular RNA-specific probe is immobilized on a support such as a solid support, e.g., the circular RNA-specific probe is immobilized on the support after binding to the circular RNA or the circular RNA-specific probe is pre-immobilized on the support.

In some embodiments, the condition in step a) include denaturing the RNA at between about 60°C and about 95°C (e.g., about 60°C, about 62°C, about 64°C, about 66°C, about 68°C, about 70°C, about 75°C, about 80°C, about 85°C, about 90°C, about 95°C) for about 2 minutes to about 10 minutes (e.g., about 2, 3, 4, 5, 6, 7, 8, 9, or 10 minutes), then gradually reducing the temperature to below about 40°C (e.g., below about 35°C, below about 30°C, below about 25°C, below about 20°C or less) to allow the circular RNA annealing to the circular RNA-specific probe.

In some embodiments, the condition in step a) include a high salt concentration range of 0.25M-2M, e.g., 0.25M, 0.5M, 0.75M, 1M, 1.25M, 1.5M, 1.75M or 2M. In some embodiments, the salt is NaCl or a guanidine salt (e.g., Guanidine hydrochloride).

In some embodiments, in step b), the one or more components are removed by washing the complex with a washing buffer.

In some embodiments, step c) is performed by increasing the temperature to about 60°C to about 95°C (e.g., about 60°C, about 62°C, about 64°C, about 66°C, about 68°C) C, about 70°C, about 75°C, about 80°C, about 85°C, about 90°C, about 95°C) to release the circular RNA.

In some embodiments, in step c) the circular RNA is released by elution with an elution buffer. In some embodiments, the elution buffer is a low salt buffer, e.g., a buffer with salt concentration lower than 0.5M. In some embodiment, the elution buffer is Tris-EDTA buffer (TE buffer) or water.

In some embodiments, the method further includes the following steps:

i) contacting the mixture comprising circular RNA and the circular RNA precursor to a linear circular RNA precursor-specific probe in a condition that allows the linear circular RNA precursor-specific probe to specifically bind to and form a complex with the linear circular RNA precursor;

ii) removing the complex formed by the linear circular RNA precursor-specific probe with the linear circular RNA precursor from the mixture, and

iii) collecting the circular RNA-containing mixture obtained in step ii).

In some embodiments, steps i)-iii) are performed before step a), for example, steps i)-iii) may be performed multiple times before step a), e.g., 2, 3, 4 or more times. In some embodiments, steps

i)-iii) are performed concurrently with steps a)-c).

In one aspect, the present invention provides a method for purifying circular RNA, the method comprises:

i) contacting a mixture comprising circular RNA and the circular RNA precursor to a linear circular RNA precursor-specific probe under a condition that allows the linear circular RNA precursor-specific probe to specifically bind to and form a complex with the linear circular RNA precursor; and

ii) removing the complex formed by the linear circular RNA precursor-specific probe with the linear circular RNA precursor from the mixture,

iii) collecting the circular RNA-containing mixture obtained in step ii), and

optionally, steps i)-iii) are performed multiple times, e.g., 2 times, 3 times, 4 times or more.

In some embodiments, the linear circular RNA precursor-specific probe specifically binds to the linear circular RNA precursor and does not substantially bind to the circular RNA.

In some embodiments, the linear circular RNA precursor-specific probe is immobilized on a support, such as a solid support, for example, the linear circular RNA precursor-specific probe is then immobilized on the support after binding to the linear circular RNA precursor, or the linear circular RNA precursor-specific probe is pre-immobilized on the support.

In some embodiments, the circular RNA-specific probe specifically hybridizes to at least a portion of the first residual circularizing element and a portion of the second residual circularizing element.

In some embodiments, the linear precursor RNA-specific probe hybridizes to a portion of the linear circular RNA precursor outside the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element.

In some embodiments, the linear precursor RNA-specific probe hybridizes to the 3' self-splicing intron fragment or a portion thereof or a 5' flanking sequence thereof, or the 5' self-splicing intron fragment or a portion thereof or a 3' flanking sequence thereof.

In some specific embodiments, the linear circular RNA precursor contains a sequence selected from SEQ ID NOs:600-610 or a complement sequence thereof, preferably, SEQ ID NO:606 or a complement sequence thereof outside the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element, to which the linear precursor RNA-specific probe specifically hybridizes.

In some embodiments of the method for purifying circular RNA of the invention, the molar ratio of the probe to the RNA molecules in the mixture is from about 1:1 to about 100,000:1.

In some embodiments, the linear circular RNA precursor-specific probe hybridizes to a 3' homology arm sequence or portion thereof on the linear circular RNA precursor, or a 5' homology

arm sequence or portion thereof on the linear circular RNA precursor. In some embodiments, the homology arm sequence comprises polyA, polyU, polyC or polyG. In some embodiments, the homology arm sequence is about 10-about 200 nt in length. Accordingly, in some embodiments, the linear precursor RNA-specific probe comprises polyU, poly A, polyG or polyC. In some embodiments, the linear precursor RNA-specific probe is about 10-about 200 nt in length.

In one aspect, the present invention provides a method for purifying circular RNA, the method comprises:

i) adding a linear RNA-specific tag to a mixture comprising circular RNA and linear RNA, wherein the linear RNA-specific tag can be linked to the linear RNA;

ii) contacting the mixture comprising circular RNA and linear RNA with a linear RNA probe that specifically binds to the tag under a condition that allows the probe to specifically bind to and form a complex with the linear RNA; and

iii) removing the complex formed by the linear RNA probe with the linear RNA from the mixture,

iv) collecting the circular RNA-containing mixture obtained in step iii),

optionally, steps ii)-iv) are performed multiple times, e.g., 2 times, 3 times, 4 times or more.

In some embodiments, the tag comprises a polyA, polyG, polyU, or polyC sequence. Accordingly, the probe includes a polyT/polyU, polyC, polyA, or polyG sequence. Alternatively, the tag can be a random sequence, and the probe comprises a sequence complementarily paired with said random sequence.

In some embodiments, the tag may be about 10-200nt in length. In some embodiments, the probe may be about 10-200nt in length.

In some embodiments, the tag is added to the linear RNA by adding a PolyA/T/C/G polymerase or a ligase to the mixture. In some embodiments, adding rATP, rGTP, rUTP, rCTP or rNTP, or a random tag sequence of about 10-200nt to the mixture is also included.

In some embodiments, the linear RNA probe specifically binds to the added tag without substantially binding to the circular RNA. In some embodiments, the linear RNA probe is a single-stranded DNA probe, or a single-stranded RNA probe.

In some embodiments, the linear RNA probe is immobilized on a support such as a solid support, for example, the linear RNA probe is immobilized on the support after binding to the linear RNA, or, the linear RNA probe is pre-immobilized on the support.

In some embodiments, the present invention provides a method of screening a target open reading frame (ORF) sequence encoding a target polypeptide for a nucleotide motif, the method comprises steps of:

1) extracting amino acid units in the direction of N terminal to C terminal in the target

polypeptide containing m amino acids, and each group of amino acid units includes an amino acid with the number of n ; with a step of 1 amino acid, slide along the N terminal to the C terminal, n is an integer and $n \geq 2$, m is an integer and ≥ 1 ;

2) listing all nucleotide sequences corresponding to each group of amino acid units according to their degenerate codon; and

3) identifying nucleotide sequences obtained from step 2) that matches the nucleotide motif; wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU, and

4) optionally, selecting the target ORF as the circularizing sequence of a circular RNA precursor.

In some embodiments, the present invention provides a system for screening a target open reading frame (ORF) for a nucleotide motif, wherein the target open reading frame (ORF) is a nucleotide sequence encoding a target polypeptide, and the system comprises:

a) an amino acid extraction module, which is used for extracting amino acid units in the direction of N terminal to C terminal in the target polypeptide containing m amino acids, wherein each group of amino acid units comprise an amino acid with a number of n , with a step of 1 amino acid, sliding along the N terminal to the C terminal, and n is an integer and $n \geq 2$, m is an integer and ≥ 1 ;

b) a conversion module, which is used for determining the set of codon sequences of each amino acid unit, each set of codon sequences includes all degenerate codon sequence corresponding to each set of amino acid units; and

c) a target codon sequence filtering module, which is used for determining whether the target codon sequence contain a nucleotide sequence that exactly matches the nucleotide motif;

wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some embodiments, the present invention provides a method of screening a target non-coding sequence for a nucleotide motif, comprising a step of:

1) editing 0, 1, or 2 bases on the non-coding sequence so that the non-coding sequence comprises an edited region that matches the nucleotide motif whilst maintaining the non-coding sequence's function;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-

116, preferably CUU, CUC or CAU.; and

2) optionally, selecting the target non-coding sequence as the circularizing sequence of a circular RNA precursor, wherein the target non-coding sequence is an endogenous circRNA.

In some embodiments, the present invention provides a system for screening a target non-coding sequence for a nucleotide motif, wherein the target non-coding sequence is an endogenous circRNA, and the system comprises:

a) a base editing module, which is used for editing the bases at any one or more positions in the target non-coding sequence to obtain an edited region matching the nucleotide motif,

wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some embodiments, the present invention provides a method for screening a target internal ribosome entry site (IRES) for a nucleotide motif, comprising steps of:

1) identifying regions of non-conserved sequences of the target IRES by comparing the target IRES with the other IRES sequences from the same viral subgenus, through multiple sequence alignment tools, for example: Clustal Omega, or MUSCLE;

2) based on the results of multiple sequence alignment, determining a low-structure region of the target IRES through secondary structure prediction software, such as RNAfold, Mfold, RNAstructure or RNAalifold; and

3) editing 0, 1, or 2 bases in the low-structure region in the target IRES to match the nucleotide motif without affecting the translation initiation activity of said IRES;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

As used herein, the term “low-structure region” refers to a region that does not have a stem-loop structure, a region having high free energy, or a single strand region.

In some embodiments, the present invention provides a system for screening a target internal ribosome entry site (IRES) for a nucleotide motif, wherein the system comprises:

a) a multi-sequence alignment module, which is used for comparing the target IRES with the other IRES sequences from the same viral subgenus to identify regions of non-conservative sequences;

b) a low-structure prediction module, which is used for determining a low-structure region within the target IRES according to the result of the multiple sequence alignments of the target

IRES; and

c) a base editing module, which is used for editing the bases at any one or more positions in the region to obtain an edited region matching the nucleotide motif,

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some embodiments, the present invention provides a screening method for screening a target non-TIE functional element for the nucleotide motif in ribozyme recognition site I, comprising a step of:

a) editing 0, 1, or 2 bases on the non-TIE functional element so that the non-TIE functional element sequence (such as 3' UTR or a replicon) comprises an edited region that matches the nucleotide motif whilst maintaining the non-TIE function;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In some embodiments, the present invention provides a system for screening a target region sequence in a non-TIE functional element sequence (such as 3'UTR or a replicon) for a nucleotide motif, wherein the screening system comprises:

a) a base editing module, which is used for editing the bases at any one or more positions in the region to obtain an edited region matching the nucleotide motif;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

In another aspect, there is provided a method for preventing or treating a disease or condition in a subject, comprising administering to a subject the circular RNA of the invention, or the pharmaceutical composition of the invention.

In some embodiments, the disease or condition comprises but not limited to xerostomia, xerophthalmia, Sjogren's syndrome or the like.

Sjogren's syndrome is a systemic autoimmune disease in which immune cells attack and destroy the exocrine glands that produce saliva and tears. Sjögren's syndrome can also affect multiple organs, including kidneys and lungs. Sjögren's syndrome can occur in all age groups of both women and men. Sjögren's syndrome can occur independently, referred to as primary Sjögren's syndrome, or may develop years after the onset of an associated rheumatic disorder, referred to as secondary Sjögren's syndrome.

Xerostomia (dry mouth) and xerophthalmia (conjunctivitis sicca, dry eyes) are hallmarks of Sjögren's syndrome. Immunologically-activated or apoptotic glandular epithelial cells that expose autoantigens in predisposed individuals might drive autoimmune-mediated tissue injury. Immune

activation is typically presented as focal, mononuclear (T, B and macrophage) cell infiltrates proximal to the ductal epithelial cells (epithelitis) and forms sialadenitis. Though the pathogenetic mechanism for this autoimmune exocrinopathy has not been fully elucidated, it has been shown that CD4⁺ T-lymphocytes constitute 60-70 percent of the mononuclear cells infiltrating the glands. Abnormal activation of proinflammatory Th1 and Th17 cells have been reported to be central to induction of SS in either human or animal models.

The present invention provides a novel therapy to prevent or treat Sjögren's syndrome-related xerostomia in a subject. The inventors have discovered that administration of a circular RNA comprising a nucleotide sequence that encodes an aquaporin-1 (AQP-1) protein or variants thereof to a subject can prevent or treat Sjögren's syndrome-related xerostomia in a subject.

In embodiments, a protein useful in the present invention is an AQP-1 protein or variants thereof, wherein such AQP1 protein or variants thereof refers to a naturally occurring AQP-1 protein, or a protein having the same activity as the naturally occurring AQP-1 protein and derived from the naturally occurring AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the naturally occurring AQP-1 protein.

In some embodiments, an AQP-1 protein or variants thereof comprises a sequence having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to the entire amino acid sequence of a naturally occurring AQP-1 protein. In embodiments, an AQP-1 protein is a portion of the amino acid sequence of an AQP-1 protein, wherein such portion of an AQP-1 protein retains the ability to form a channel in a cell membrane that allows the passage of water. It is also known in the art that several isoforms of AQP-1 protein exist. Thus, in embodiments, an AQP-1 protein is an isoform of an AQP-1 protein, wherein such isoform retains the ability to form a channel that allows the passage of water. In embodiments, an AQP-1 protein is a portion of an isoform or other naturally-occurring variant of an AQP-1 protein, wherein such portion retains the ability to form a channel in a membrane that allows the passage of water. Methods to produce functional portions and variants of AQP-1 proteins, such as conservative variants, of AQP-1 protein are known to those skilled in the art.

In some embodiments, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest, comprising from 5' to 3' direction: protein-coding

sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES) and a protein-coding sequence fragment I; and

c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the protein-coding sequence fragment II, the translation initiation element, and the protein-coding sequence fragment I through the self-splicing of the circular RNA precursor;

wherein the protein-coding sequence fragment I is derived from a 5' terminal portion of at least one protein-coding sequence, and the protein-coding sequence fragment II is derived from a 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence;

wherein the at least one protein-coding sequence encodes an aquaporin, such as AQP-1 protein or variants thereof.

In some embodiments, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

a) a 3' self-splicing intron fragment;

b) a nucleotide sequence of interest, comprising from 5' to 3' direction: protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES) and a protein-coding sequence fragment I; and

c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the protein-coding sequence fragment II, the translation initiation element, and the protein-coding sequence fragment I through the self-splicing of the circular RNA precursor;

wherein the protein-coding sequence fragment I is the 5' terminal portion of at least one protein-coding sequence, and the protein-coding sequence fragment II is the 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located

at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence;

wherein the at least one protein-coding sequence encodes an aquaporin, such as AQP-1 protein or variants thereof.

In embodiments, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a translation initiation element (TIE) fragment II;
- c) a nucleotide sequence that encodes an aquaporin-1 (AQP-1) protein or variants thereof;
- d) a translation initiation element fragment I and
- e) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the TIE fragment I is derived from a 5' terminal portion of a TIE sequence, and the TIE fragment II is derived from a 3' terminal portion of a TIE sequence, and

the TIE fragment I comprises a ribozyme recognition site I located at its 3' end, the TIE fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the TIE I is connected with the TIE II to form the TIE sequence which is able to recruit ribosome to initiate translation.

In embodiments, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a translation initiation element (TIE) fragment II;
- c) a nucleotide sequence that encodes an aquaporin-1 (AQP-1) protein or variants thereof;
- d) a translation initiation element fragment I and
- e) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a

circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the TIE fragment I is the 5' terminal portion of a TIE sequence, and the TIE fragment II is the 3' terminal portion of a TIE sequence, and

the TIE fragment I comprises a ribozyme recognition site I located at its 3' end, the TIE fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the TIE I is connected with the TIE II to form the TIE sequence which is able to recruit ribosome to initiate translation.

In some embodiments, the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At).

In some embodiments, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

In some embodiments, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1 and 3, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1 and 3.

In some embodiments, the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

In some embodiments, the self-splicing intron is the Staphylococcus phage Twort Group I Intron (Twort), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 27 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 27, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 28 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 28.

In some embodiments, the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the total length of the ribozyme recognition site II is about 0 to about

100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 80 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

In some embodiments, the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 58, 60 or 64.

In some embodiments, the ribozyme recognition site I comprises a nucleotide sequence of UCCAGCUU, UCUGGCAU or UGUAACUU.

In some embodiments, the ribozyme recognition site II comprises a nucleotide sequence of AACUGGA, AACCAGC or AAAAUUACA.

In embodiments, the ribozyme recognition site I is UCCAGCUU, and the ribozyme recognition site II is AACUGGA.

In embodiments, the ribozyme recognition site I is UCUGGCAU, and the ribozyme recognition site II is AACCAGC.

In embodiments, the ribozyme recognition site I is UGUAACUU, and the ribozyme recognition site II is AAAAUUACA.

In some embodiments, the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

In some embodiments, the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

In some embodiments, the homology arm sequence I or the homology arm sequence II is about 5-200 nucleotides in length, about 5-150 nucleotides in length, about 5-100 nucleotides in length, about 5-80 nucleotides in length, about 5-50 nucleotides in length, about 5-40 nucleotides in length, about 5-30 nucleotides in length, about 5-20 nucleotides in length, about 5-10 nucleotides in length, preferably, about 5, 10, 15, 20, 25, 30, 35 or 40 nucleotides in length.

In some embodiments, one of the homology arm sequence I and the homology arm sequence II comprises a nucleotide sequence of any one of SEQ ID NO: 83-84 or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 83-84 , while the other homology arm sequence comprises the corresponding complementary sequence.

In some embodiments, the translation initiation element is an internal ribosome entry site (IRES) which is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2, preferably CVB3.

In embodiments, the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% sequence identity to one of SEQ ID NOs: 120-526.

In some embodiments, the IRES comprises a nucleotide sequence of SEQ ID NO: 121, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to SEQ ID NO: 121.

In some embodiments, the at least one protein-coding sequence comprises an AQP-1 protein coding sequence, or a nucleotide sequence encoding a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

In some embodiments, an AQP-1 protein encoding sequence comprises a nucleotide sequence of SEQ ID NO: 118, or a nucleotide sequence having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to SEQ ID NO: 118. In particular, the AQP-1 protein coding sequence is SEQ ID NO: 118, or a nucleotide sequence having 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to SEQ ID NO: 118.

In some embodiments, the circular RNA precursor of the invention comprises a nucleotide sequence selected from any one of SEQ ID NO: 533-536, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%,

at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 533-536.

In embodiments, the present invention provides a circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a non-translation initiation element (non-TIE) fragment II;
- c) a nucleotide sequence that encodes an aquaporin-1 (AQP-1) protein, or a variant thereof;
- d) a non-translation initiation element (non-TIE) fragment I and
- e) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the non-TIE fragment I is derived from a 5' terminal portion of a non-TIE sequence, and the non-TIE II is derived from a 3' terminal portion of a non-TIE sequence, and

the non-TIE fragment I comprises a ribozyme recognition site I located at its 3' end, the non-TIE fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-TIE fragment I is connected with the non-TIE fragment II to form the non-TIE sequence. In embodiments, the non-TIE sequence is a 3' UTR sequence.

Additionally, the present invention further provides the following items:

1. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment; and
- b) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA through the self-splicing of the circular RNA precursor,

wherein the self-splicing intron is selected from Group I introns and Group II introns, for example, the self-splicing intron is selected from Group I introns.

2. The circular RNA precursor of item 1, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron

(Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At), for example, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

3. The circular RNA precursor of item 1 or 2, wherein the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

4. The circular RNA precursor of any one of items 1-3, wherein when the self-splicing intron is cyanobacterium *Anabaena* Group I Intron such as AnaX, the intron internal split site II and the intron internal split site I are a nucleotide A at position 115 and a nucleotide A at position 116 from 5' end of the AnaX intron (SEQ ID NO:1) respectively, and when connected form a nucleotide sequence of AA.

5. The circular RNA precursor of any one of items 1-4, wherein when the self-splicing intron is *Azoarcus* Group I Intron (Azo) (SEQ ID NO:2),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 24 and a nucleotide A at position 25 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 85 and a nucleotide A at position 86 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P4-5);

3) the intron internal split site II and the intron internal split site I are a nucleotide C at position 71 and a nucleotide C at position 72 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CC (P5);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 108 and a nucleotide U at position 109 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CU (P6);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 154 and a nucleotide A at position 155 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P8); or

6) the intron internal split site II and the intron internal split site I are a nucleotide C at position

198 and a nucleotide A at position 199 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CA (P9).

6. The circular RNA precursor of any one of items 1-4, wherein when the self-splicing intron is *Scytalidium dimidiatum* Group I Intron (Sd) (SEQ ID NO:4),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 88 and a nucleotide G at position 89 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide G at position 174 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of UG (P5);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 204 and a nucleotide G at position 205 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P6);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 252 and a nucleotide G at position 253 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P8); or

5) the intron internal split site II and the intron internal split site I are a nucleotide C at position 370 and a nucleotide G at position 371 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P9).

7. The circular RNA precursor of any one of items 1-4, wherein when the self-splicing intron is *Staphylococcus phage Twort* Group I Intron (Twort) (SEQ ID NO:3),

1) the intron internal split site II and the intron internal split site I are a nucleotide U at position 70 and a nucleotide A at position 71 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P5);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 103 and a nucleotide A at position 104 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P6);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 132 and a nucleotide U at position 133 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AU (P7);

4) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide C at position 174 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UC (P8);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 206 and a nucleotide A at position 207 from 5' end of the Twort intron respectively, and when

connected form a nucleotide sequence of AA (P9);

6) the intron internal split site II and the intron internal split site I are a nucleotide G at position 226 and a nucleotide U at position 227 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of GU (P9.1);

7) the intron internal split site II and the intron internal split site I are a nucleotide U at position 23 and a nucleotide A at position 24 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P2); or

8) the intron internal split site II and the intron internal split site I are a nucleotide A at position 82 and a nucleotide A at position 83 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P4-5).

8. The circular RNA precursor of any one of items 1-4, wherein when the self-splicing intron is *Scytonema-hofmani* tRNA fMet group I intron (Sh) (SEQ ID NO:5),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 129 and a nucleotide A at position 130 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P6);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 60 and a nucleotide A at position 61 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P2);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 109 and a nucleotide A at position 110 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P4-5); or

4) the intron internal split site II and the intron internal split site I are a nucleotide A at position 213 and a nucleotide A at position 214 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of UA (P9).

9. The circular RNA precursor of any one of items 1-4, wherein when the self-splicing intron is *Agrobacterium-tumefaciens* group I intron (At) (SEQ ID NO:6), the intron internal split site II and the intron internal split site I are a nucleotide A at position 107 and a nucleotide A at position 108 from 5' end of the At intron respectively, and when connected form a nucleotide sequence of AA (P6).

10. The circular RNA precursor of any one of items 1-9, wherein the 3' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 3' terminal portion of the self-splicing intron, and the 5' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least

81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 5' terminal portion of the self-splicing intron.

11. The circular RNA precursor of any one of items 1-10, wherein the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

12. The circular RNA precursor of any one of items 1-10, wherein the self-splicing intron is the *Azoarcus* Group I Intron (Azo), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 9, 11, 13, 15, 17 or 19 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 9, 11, 13, 15, 17 or 19, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 10, 12, 14, 16, 18 or 20 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 10, 12, 14, 16, 18 or 20.

13. The circular RNA precursor of any one of items 1-10, wherein the self-splicing intron is the *Scytalidium dimidiatum* Group I Intron (Sd), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 37, 39, 41, 43 or 45 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 37, 39, 41, 43 or 45, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 38, 40, 42, 44 or 46 or a nucleotide

sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 38, 40, 42, 44 or 46.

14. The circular RNA precursor of any one of items 1-10, wherein the self-splicing intron is the *Staphylococcus* phage Twort Group I Intron (Twort), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36.

15. The circular RNA precursor of any one of items 1-10, wherein the self-splicing intron is the *Scytonema-hofmani* tRNA fMet group I intron (Sh), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 47, 49, 51 or 53 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 47, 49, 51 or 53 and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 48, 50, 52 or 54 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 48, 50, 52 or 54.

16. The circular RNA precursor of any one of items 1-10, wherein the self-splicing intron is the *Agrobacterium-tumefaciens* group I intron (At), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 55 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to

SEQ ID NO: 55, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 56 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 56.

17. The circular RNA precursor of any one of items 1-16, which further comprises from 5' to 3' direction: optionally a first residual circularizing element; a nucleotide sequence of interest; and optionally a second residual circularizing element, located between a) and b), wherein the total length of the first residual circularizing element and the second residual circularizing element is about 0 to about 500 nucleotides; or

the circular RNA precursor further comprises a homology arm sequence I upstream of the 3' self-splicing intron fragment and a homology arm sequence II downstream of the 5' self-splicing intron fragment.

18. The circular RNA precursor of any one of items 1-17, wherein the first residual circularizing element comprises or consists of from 5' to 3' direction a 3' exon region and optionally a first spacer, the second residual circularizing element comprises or consists of from 3' to 5' direction a 5' exon region and optionally a second spacer.

19. The circular RNA precursor of item 18, wherein the 3' exon region is derived from the native 3' exon of the self-splicing intron, the 5' exon region is derived from the native 5' exon of the self-splicing intron, and the 3' exon region and 5' exon region can be recognized and/or spliced by the self-splicing intron or a combination of the 3' self-splicing intron fragment and the 5' self-splicing intron fragment.

20. The circular RNA precursor of item 19, wherein 3' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to the native 3' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 5' terminal nucleotide of the native 3' exon,

the 5' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to the native 5' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 3' terminal nucleotide of the native 5' exon.

21. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a first residual circularizing element;
- c) a nucleotide sequence of interest;
- d) a second residual circularizing element; and
- e) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element through the self-splicing of the circular RNA precursor,

wherein the total length of the first residual circularizing element and the second residual circularizing element is about 0 to about 500 nucleotides.

22. The circular RNA precursor of item 21, wherein the self-splicing intron is selected from Group I introns and Group II introns, for example, the self-splicing intron is selected from Group I introns.

23. The circular RNA precursor of item 21 or 22, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At), for example, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

24. The circular RNA precursor of any one of items 21-23, wherein the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

25. The circular RNA precursor of any one of items 21-23, wherein when the self-splicing intron is cyanobacterium *Anabaena* Group I Intron such as AnaX, the intron internal split site II and the intron internal split site I are a nucleotide A at position 115 and a nucleotide A at position 116 from 5' end of the AnaX intron (SEQ ID NO:1) respectively, and when connected form a

nucleotide sequence of AA.

26. The circular RNA precursor of any one of items 21-24, wherein when the self-splicing intron is *Azoarcus* Group I Intron (Azo) (SEQ ID NO:2),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 24 and a nucleotide A at position 25 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 85 and a nucleotide A at position 86 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P4-5);

3) the intron internal split site II and the intron internal split site I are a nucleotide C at position 71 and a nucleotide C at position 72 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CC (P5);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 108 and a nucleotide U at position 109 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CU (P6);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 154 and a nucleotide A at position 155 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P8); or

6) the intron internal split site II and the intron internal split site I are a nucleotide C at position 198 and a nucleotide A at position 199 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CA (P9).

27. The circular RNA precursor of any one of items 21-24, wherein when the self-splicing intron is *Scytalidium dimidiatum* Group I Intron (Sd) (SEQ ID NO:4),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 88 and a nucleotide G at position 89 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide G at position 174 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of UG (P5);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 204 and a nucleotide G at position 205 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P6);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 252 and a nucleotide G at position 253 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P8); or

5) the intron internal split site II and the intron internal split site I are a nucleotide C at position 370 and a nucleotide G at position 371 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P9).

28. The circular RNA precursor of any one of items 21-24, wherein when the self-splicing intron is Staphylococcus phage Twort Group I Intron (Twort) (SEQ ID NO:3),

1) the intron internal split site II and the intron internal split site I are a nucleotide U at position 70 and a nucleotide A at position 71 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P5);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 103 and a nucleotide A at position 104 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P6);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 132 and a nucleotide U at position 133 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AU (P7);

4) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide C at position 174 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UC (P8);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 206 and a nucleotide A at position 207 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P9);

6) the intron internal split site II and the intron internal split site I are a nucleotide G at position 226 and a nucleotide U at position 227 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of GU (P9.1);

7) the intron internal split site II and the intron internal split site I are a nucleotide U at position 23 and a nucleotide A at position 24 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P2); or

8) the intron internal split site II and the intron internal split site I are a nucleotide A at position 82 and a nucleotide A at position 83 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P4-5).

29. The circular RNA precursor of any one of items 21-24, wherein when the self-splicing intron is Scytonema-hofmani tRNA fMet group I intron (Sh) (SEQ ID NO:5),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 129 and a nucleotide A at position 130 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P6);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position

60 and a nucleotide A at position 61 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P2);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 109 and a nucleotide A at position 110 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P4-5); or

4) the intron internal split site II and the intron internal split site I are a nucleotide A at position 213 and a nucleotide A at position 214 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of UA (P9).

30. The circular RNA precursor of any one of items 21-24, wherein when the self-splicing intron is *Agrobacterium-tumefaciens* group I intron (At) (SEQ ID NO:6), the intron internal split site II and the intron internal split site I are a nucleotide A at position 107 and a nucleotide A at position 108 from 5' end of the At intron respectively, and when connected form a nucleotide sequence of AA (P6).

31. The circular RNA precursor of any one of items 21-30, wherein the 3' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 3' terminal portion of the self-splicing intron, and the 5' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 5' terminal portion of the self-splicing intron.

32. The circular RNA precursor of any one of items 21-31, wherein the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

33. The circular RNA precursor of any one of items 21-31, wherein the self-splicing intron is

the Azoarcus Group I Intron (Azo), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 9, 11, 13, 15, 17 or 19 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 9, 11, 13, 15, 17 or 19, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 10, 12, 14, 16, 18 or 20 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 10, 12, 14, 16, 18 or 20.

34. The circular RNA precursor of any one of items 21-31, wherein the self-splicing intron is the Scytalidium dimidiatum Group I Intron (Sd), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 37, 39, 41, 43 or 45 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 37, 39, 41, 43 or 45, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 38, 40, 42, 44 or 46 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 38, 40, 42, 44 or 46.

35. The circular RNA precursor of any one of items 21-31, wherein the self-splicing intron is the Staphylococcus phage Twort Group I Intron (Twort), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 22, 24, 26, 28, 30,

32, 34 or 36.

36. The circular RNA precursor of any one of items 21-31, wherein the self-splicing intron is the *Scytonema-hofmani* tRNA fMet group I intron (Sh), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 47, 49, 51 or 53 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 47, 49, 51 or 53, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 48, 50, 52 or 54 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 48, 50, 52 or 54.

37. The circular RNA precursor of any one of items 21-31, wherein the self-splicing intron is the *Agrobacterium-tumefaciens* group I intron (At), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 55 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 55, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 56 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 56.

38. The circular RNA precursor of any one of items 21-37, wherein the first residual circularizing element and the second residual circularizing element are configured such that the circular RNA comprising them has reduced immunogenicity.

39. The circular RNA precursor of any one of items 21-38, wherein the first residual circularizing element and the second residual circularizing element are configured such that the circular RNA precursor comprising them has comparable or increased circularization efficiency.

40. The circular RNA precursor of any one of items 21-39, wherein the total length of the first residual circularizing element and the second residual circularizing element is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about

100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides.

41. The circular RNA precursor of any one of items 21-40, wherein the first residual circularizing element and the second residual circularizing element are configured to be capable of forming a stem-loop structure, e.g., upon self-splicing for circularization.

42. The circular RNA precursor of item 41, wherein the loop of the stem-loop structure comprises the splicing junction.

43. The circular RNA precursor of any one of items 41-42, wherein the first residual circularizing element comprises the sequence structure of the following formula: 5'-first loop sequence-first pairing sequence-first non-pairing sequence-3'; and the second residual circularizing element comprises the sequence structure of the following formula: 5'-second non-pairing sequence-second pairing sequence-second loop sequence-3',

wherein the first non-pairing sequence or the second non-pairing sequence may be independently present or absent,

the first pairing sequence and the second pairing sequence can complementarily pair to each other to form the stem of the stem-loop structure, wherein the first loop sequence and the second loop sequence can form the loop of the stem-loop structure, e.g., through self-splicing for circularization.

44. The circular RNA precursor of any one of items 41-43, wherein the first residual circularizing element can be selected from various sequence, provided that the first loop sequence of the first residual circularizing element doesn't pair with any sequence of the second residual circularizing element.

45. The circular RNA precursor of item 43 or 44, wherein the first loop sequence comprises or consists of one or more nucleotides which can pair with the P1 region of the corresponding self-splicing intron to form a P10 duplex region during the circularization.

46. The circular RNA precursor of any one of items 43-45, wherein the first loop sequence consist of a nucleotide sequence of (N)_n, wherein N represents any nucleotides (A, G, U, or C), n represents an integer from 1-20, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

47. The circular RNA precursor of any one of items 43-46, wherein the first loop sequence comprises or consists of the sequence of AAAA, AGAA, AA, UUUU, UAAA, CAAA, or GAAA, preferably AAAA.

48. The circular RNA precursor of any one of items 43-47, wherein the second loop sequence comprises or consists of one or more nucleotides which can pair with the internal guide sequence (IGS) of the corresponding self-splicing intron to form a P1 duplex region during the

circularization.

49. The circular RNA precursor of any one of items 43-48, wherein the second loop sequence comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, such as CUU, CAU or CUC, preferably CUU.

50. The circular RNA precursor of any one of items 43-49, wherein the loop of the stem-loop structure has a sequence of CUUAAAA, CUUUUUU, CUUAA, CUUGAAA, CUUUAAA, CUUAGAA, CUUCAAA or CUCAAAA, preferably CUUAAAA.

51. The circular RNA precursor of any one of items 43-50, wherein the stem portion of the stem-loop structure comprises 2-15 or more consecutive matched base pairs, preferably, the stem portion of the stem-loop structure comprises 5, 6 or 7 consecutive matched base pairs.

52. The circular RNA precursor of any one of items 43-51, wherein the stem portion in the stem-loop structure comprises up to 2 base pair mismatches, or the stem portion comprises only 1 base pair mismatch, preferably, the stem portion comprises no base pair mismatches.

53. The circular RNA precursor of any one of items 43-52, wherein the first pairing sequence comprises only G and the second pairing sequence comprises only C, or the first pairing sequence comprises only C and the second pairing sequence comprises only G.

54. The circular RNA precursor of any one of items 43-53, wherein the first pairing sequence comprises only A and the second pairing sequence comprises only U.

55. The circular RNA precursor of any one of items 43-54, wherein the first pairing sequence comprises only U and the second pairing sequence comprises only A.

56. The circular RNA precursor of any one of items 21-55, wherein the first residual circularizing element comprises or consists of from 5' to 3' direction a 3' exon region and optionally a first spacer, the second residual circularizing element comprises or consists of from 3' to 5' direction a 5' exon region and optionally a second spacer.

57. The circular RNA precursor of item 56, wherein the 3' exon region is derived from the native 3' exon of the self-splicing intron, the 5' exon region is derived from the native 5' exon of the self-splicing intron, and the 3' exon region and 5' exon region can be recognized and/or spliced by the self-splicing intron or a combination of the 3' self-splicing intron fragment and the 5' self-splicing intron fragment.

58. The circular RNA precursor of item 56 or 57, wherein the 3' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the native 3' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 5' terminal nucleotide of the native 3' exon,

the 5' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the native 5' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 3' terminal nucleotide of the native 5' exon.

59. The circular RNA precursor of any one of items 21-58, wherein the first residual circularizing element comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69, preferably the first residual circularizing element is selected from a nucleotide sequence of AAAAUCCGUUGA or AAAACCCCCCCA.

60. The circular RNA precursor of any one of items 21-59, wherein the second residual circularizing element comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70, preferably the second residual circularizing element is selected from a nucleotide sequence of AGACGCUACGGACUU or AGACGGGGGGGGC UU.

61. The circular RNA precursor of any one of items 21-60, wherein the first residual circularizing element comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69 and the second residual circularizing element comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70.

62. The circular RNA precursor of any one of items 21-61, wherein,

1) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 57 and the second residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 58;

2) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 59 and the second residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 60;

3) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 61 and the second residual circularizing element comprises or consists of a nucleotide

sequence of SEQ ID NO: 62;

4) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 63 and the second residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 64;

5) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 65 and the second residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 66;

6) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 67 and the second residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 68; or

7) the first residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 69 and the second residual circularizing element comprises or consists of a nucleotide sequence of SEQ ID NO: 70;

preferably, the first residual circularizing element is selected from a nucleotide sequence of AAAAUCCGUUGA and the second residual circularizing element is selected from a nucleotide sequence of AGACGCUACGGACUU, or the first residual circularizing element is selected from a nucleotide sequence of AAAACCCCCCA and the second residual circularizing element is selected from a nucleotide sequence of AGACGGGGGGGCUU.

63. The circular RNA precursor of any one of items 21-62, wherein the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

64. The circular RNA precursor of item 63, wherein the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

65. The circular RNA precursor of any one of items 61-62, wherein the homology arm sequence I or the homology arm sequence II is about 5-500, about 5-450, about 5-400, about 5-350, about 5-300, about 5-250, about 5-200, about 5-150, about 5-100, about 5-80, about 5-60, about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-15 or about 5-10 nucleotides in length, preferably, about 20, 25, 40, 50, 100, or 300 nucleotides in length.

66. The circular RNA precursor of any one of items 61-65, wherein the two homology arm sequences may be polyA and polyU, respectively, or polyG and polyC, respectively.

67. The circular RNA precursor of any one of items 61-66, wherein one of the homology arm sequence I and the homology arm sequence II comprises a nucleotide sequence of any one of SEQ ID NO: 71-94 or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%,

at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 71-94, while the other homology arm sequence comprises the corresponding complementary sequence.

68. The circular RNA precursor of any one of items 21-67, wherein the nucleotide sequence of interest comprises at least one protein-coding sequence and a translation initiation element such as an internal ribosome entry site (IRES) operably linked thereto.

69. The circular RNA precursor of item 68, wherein the translation initiation element, such as an IRES, is located upstream to the 5' end of the at least one protein-coding sequence.

70. The circular RNA precursor of item 68, wherein the translation initiation element, such as an IRES, is located downstream to the 3' end of the at least one protein-coding sequence.

71. The circular RNA precursor of any one of items 68-70, wherein the translation initiation element is an internal ribosome entry site (IRES) which is selected from the following IRES sequences: Taura syndrome virus, blood-sucking bug virus, Tyler's encephalomyelitis virus, simian virus 40, red fire ant virus 1, cereal constrictor virus, reticulovirus Endothelial hyperplasia virus, Forman poliovirus 1, soybean inchworm virus, Kashmir bee virus, human rhinovirus 2, glass leafhopper virus-1, human immunodeficiency virus type 1, glass leafhopper virus-1, lice P virus , Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinovirus, Tea inchworm-like virus, Encephalomyocarditis virus (EMCV), Drosophila C virus, Cruciferous tobacco Virus, cricket paralysis virus, bovine viral diarrhea virus 1, black queen cell virus, aphid lethal paralysis virus, avian encephalomyelitis virus, acute bee paralysis virus, hibiscus yellow ring spot virus, swine fever virus, human FGF2, human SFTPA1, Human AML1/RUNX1, Drosophila Antennae, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 α , Human n.myc, mouse Gtx, human p27kipl, human PDGF2/c-sis, human p53, human Pim-1, mouse Rbm3, Drosophila reaper, canine Scamper, Drosophila Ubx, human UNR, mouse UtrA, human VEGF-A, human XIAP, Drosophila hairless, Saccharomyces cerevisiae TFIID, Saccharomyces cerevisiae YAP1, human c-src, human FGF-1, simian picornavirus, turnip crepe disease virus, eIF4G aptamer, Coxsackie Virus B3 (CVB3) or Coxsackie virus A (CVA1/2), preferably, the IRES is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2.

72. The circular RNA precursor of any one of items 68-71, wherein the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to one of SEQ ID NOs: 120-526.

73. The circular RNA precursor of any one of items 68-72, wherein the at least one protein-coding sequence encodes a protein of eukaryotic, prokaryotic or viral origin, for example, a protein for therapeutic or diagnostic use;

preferably, the at least one protein-coding sequence encodes IL-2, POLR2A, or an aquaporin, such as AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

74. The circular RNA precursor of any one of items 68-73, wherein the at least one protein-coding sequence comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

75. The circular RNA precursor of any one of items 68-74, wherein the nucleotide sequence of interest is a non-coding sequence, for example, the non-coding sequence can be antisense RNA, aptamer, guide RNA, or non-protein-coding RNA existing in any organism.

76. The circular RNA precursor of any one of items 21-75, which comprises a nucleotide sequence selected from any one of SEQ ID NO: 527, 532 and 537-542, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 527, 532 and 537-542.

77. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a circularizing sequence fragment II comprising a ribozyme recognition site II at its 5' end and a circularizing sequence fragment I comprising a ribozyme recognition site I at its 3' end,

wherein the circularizing sequence fragment I is derived from a 5' terminal portion of a circularizing sequence, and the circularizing sequence fragment II is derived from a 3' terminal portion of a circularizing sequence, and

wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the circularizing sequence fragment I is connected with the circularizing sequence fragment II to form the circularizing sequence.

78. The circular RNA precursor of item 77, wherein the nucleotide sequence of interest comprises one or more sequences selected from the following group between the circularizing sequence fragment II and the circularizing sequence fragment I: at least one protein coding sequences, a non-coding sequence, translation initiation element (TIE), a non-TIE functional element or any combination thereof.

79. The circular RNA precursor of item 77 or 78, wherein the circularizing sequence is selected from at least one protein-coding sequence, a non-coding sequence, a translation initiation element (TIE), or a non-TIE functional element.

80. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES), and a protein-coding sequence fragment I;

wherein the protein-coding sequence fragment I is derived from a 5' terminal portion of at least one protein-coding sequence, and the protein-coding sequence fragment II is derived from a 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence;

optionally, the at least one protein-coding sequence comprises at least two protein-coding sequences, and a further translation initiation element, 2A sequence or spacer may be linked

between any two adjacent protein-coding sequences, and the translation initiation element comprises: IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof.

81. The circular RNA precursor of any one of items 77-80, wherein the self-splicing intron is selected from the Group I introns and Group II introns, for example, the self-splicing intron is selected from Group I introns.

82. The circular RNA precursor of any one of items 77-81, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At), for example, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

83. The circular RNA precursor of any one of items 77-82, wherein the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

84. The circular RNA precursor of any one of items 77-83, wherein when the self-splicing intron is cyanobacterium *Anabaena* Group I Intron such as AnaX, the intron internal split site II and the intron internal split site I are a nucleotide A at position 115 and a nucleotide A at position 116 from 5' end of the AnaX intron (SEQ ID NO:1) respectively, and when connected form a nucleotide sequence of AA.

85. The circular RNA precursor of any one of items 77-83, wherein when the self-splicing intron is *Azoarcus* Group I Intron (Azo) (SEQ ID NO:2),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 24 and a nucleotide A at position 25 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 85 and a nucleotide A at position 86 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P4-5);

3) the intron internal split site II and the intron internal split site I are a nucleotide C at position

71 and a nucleotide C at position 72 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CC (P5);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 108 and a nucleotide U at position 109 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CU (P6);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 154 and a nucleotide A at position 155 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P8); or

6) the intron internal split site II and the intron internal split site I are a nucleotide C at position 198 and a nucleotide A at position 199 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CA (P9).

86. The circular RNA precursor of any one of items 77-83, wherein when the self-splicing intron is *Scytalidium dimidiatum* Group I Intron (Sd) (SEQ ID NO:4),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 88 and a nucleotide G at position 89 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide G at position 174 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of UG (P5);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 204 and a nucleotide G at position 205 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P6);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 252 and a nucleotide G at position 253 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P8); or

5) the intron internal split site II and the intron internal split site I are a nucleotide C at position 370 and a nucleotide G at position 371 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P9).

87. The circular RNA precursor of any one of items 77-83, wherein when the self-splicing intron is *Staphylococcus phage Twort* Group I Intron (Twort) (SEQ ID NO:3),

1) the intron internal split site II and the intron internal split site I are a nucleotide U at position 70 and a nucleotide A at position 71 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P5);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 103 and a nucleotide A at position 104 from 5' end of the Twort intron respectively, and when

connected form a nucleotide sequence of UA (P6);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 132 and a nucleotide U at position 133 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AU (P7);

4) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide C at position 174 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UC (P8);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 206 and a nucleotide A at position 207 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P9);

6) the intron internal split site II and the intron internal split site I are a nucleotide G at position 226 and a nucleotide U at position 227 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of GU (P9.1);

7) the intron internal split site II and the intron internal split site I are a nucleotide U at position 23 and a nucleotide A at position 24 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P2); or

8) the intron internal split site II and the intron internal split site I are a nucleotide A at position 82 and a nucleotide A at position 83 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P4-5).

88. The circular RNA precursor of any one of items 77-83, wherein when the self-splicing intron is *Scytonema-hofmani* tRNA fMet group I intron (Sh) (SEQ ID NO:5),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 129 and a nucleotide A at position 130 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P6);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 60 and a nucleotide A at position 61 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P2);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 109 and a nucleotide A at position 110 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P4-5); or

4) the intron internal split site II and the intron internal split site I are a nucleotide A at position 213 and a nucleotide A at position 214 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of UA (P9).

89. The circular RNA precursor of any one of items 77-83, wherein when the self-splicing intron is *Agrobacterium-tumefaciens* group I intron (At) (SEQ ID NO:6), the intron internal split

site II and the intron internal split site I are a nucleotide A at position 107 and a nucleotide A at position 108 from 5' end of the At intron respectively, and when connected form a nucleotide sequence of AA (P6).

90. The circular RNA precursor of any one of items 77-89, wherein the 3' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 3' terminal portion of the self-splicing intron, and the 5' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 5' terminal portion of the self-splicing intron.

91. The circular RNA precursor of any one of items 77-90, wherein the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

92. The circular RNA precursor of any one of items 77-90, wherein the self-splicing intron is the *Azoarcus* Group I Intron (Azo), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 9, 11, 13, 15, 17 or 19 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 9, 11, 13, 15, 17 or 19, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 10, 12, 14, 16, 18 or 20 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100%

identity to SEQ ID NO: 10, 12, 14, 16, 18 or 20.

93. The circular RNA precursor of any one of items 77-90, wherein the self-splicing intron is the *Scytalidium dimidiatum* Group I Intron (Sd), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 37, 39, 41, 43 or 45 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 37, 39, 41, 43 or 45, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 38, 40, 42, 44 or 46 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 38, 40, 42, 44 or 46.

94. The circular RNA precursor of any one of items 77-90, wherein the self-splicing intron is the *Staphylococcus phage Twort* Group I Intron (Twort), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36.

95. The circular RNA precursor of any one of items 77-90, wherein the self-splicing intron is the *Scytonema-hofmani* tRNA fMet group I intron (Sh), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 47, 49, 51 or 53 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 47, 49, 51 or 53, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 48, 50, 52 or 54 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%,

at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 48, 50, 52 or 54.

96. The circular RNA precursor of any one of items 77-90, wherein the self-splicing intron is the *Agrobacterium-tumefaciens* group I intron (At), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 55 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 55, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 56 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 56.

97. The circular RNA precursor of any one of items 77-96, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA comprising them has reduced immunogenicity.

98. The circular RNA precursor of any one of items 77-97, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA precursor comprising them has comparable or increased circularization efficiency.

99. The circular RNA precursor of any one of items 77-98, wherein the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

100. The circular RNA precursor of any one of items 77-99, wherein the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides,

about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

101. The circular RNA precursor of any one of items 77-100, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured to be capable of forming a stem-loop structure, e.g., upon self-splicing for circularization.

102. The circular RNA precursor of item 101, wherein the loop of the stem-loop structure comprises the splicing junction.

103. The circular RNA precursor of any one of items 77-102, wherein the ribozyme recognition site I comprises the sequence structure of the following formula: 5'-third non-pairing sequence-third pairing sequence-third loop sequence-3'; and the ribozyme recognition site II comprises the sequence structure of the following formula: 5'-fourth loop sequence-fourth pairing sequence-fourth non-pairing sequence-3',

wherein the third non-pairing sequence or the fourth non-pairing sequence may be independently present or absent,

the third pairing sequence and the fourth pairing sequence can complementarily pair to each other to form the stem of the stem-loop structure, wherein the third loop sequence and the fourth loop sequence can form the loop of the stem-loop structure, e.g., through self-splicing for circularization.

104. The circular RNA precursor of any one of items 77-103, wherein the ribozyme recognition site II can be selected from various sequence.

105. The circular RNA precursor of item 103 or 104, wherein the fourth loop sequence comprises or consists of one or more nucleotides which can pair with the P1 region of the corresponding self-splicing intron to form a P10 duplex region during the circularization.

106. The circular RNA precursor of any one of items 103-105, wherein the fourth loop sequence consist of a nucleotide sequence of (N)_n, wherein N represents any nucleotides (A, G, U, or C), n represents an integer from 1-20, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

107. The circular RNA precursor of any one of items 103-106, wherein the fourth loop sequence comprises or consists of the sequence of AAAA, AGAA, AA, UUUU, UAAA, CAAA, or GAAA, preferably AAAA.

108. The circular RNA precursor of any one of items 103-107, wherein the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

109. The circular RNA precursor of any one of items 103-108, wherein the third loop sequence comprises or consists of one or more nucleotides which can pair with the internal guide sequence (IGS) of the corresponding self-splicing intron to form a P1 duplex region during the circularization.

110. The circular RNA precursor of any one of items 103-109, wherein the third loop sequence comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, such as CUU, CAU or CUC, preferably CUU.

111. The circular RNA precursor of any one of items 103-110, wherein the loop of the stem-loop structure has a sequence of CUUAAAA, CUUUUUU, CUUAA, CUUGAAA, CUUUAAA, CUUAGAA, CUUCAA or CUCAAAA, preferably CUUAAAA.

112. The circular RNA precursor of any one of items 103-111, wherein the stem portion of the stem-loop structure comprises 2-15 or more consecutive matched base pairs, preferably, the stem portion of the stem-loop structure comprises 5, 6 or 7 consecutive matched base pairs.

113. The circular RNA precursor of any one of items 103-112, wherein the stem portion in the stem-loop structure comprises up to 2 base pair mismatches, or the stem portion comprises only 1 base pair mismatch, preferably, the stem portion comprises no base pair mismatches.

114. The circular RNA precursor of any one of items 103-113, wherein the third pairing sequence comprises only G and the fourth pairing sequence comprises only C, or the third pairing sequence comprises only C and the fourth pairing sequence comprises only G.

115. The circular RNA precursor of any one of items 103-114, wherein the third pairing sequence comprises only A and the fourth pairing sequence comprises only U, or the third pairing sequence comprises only U and the fourth pairing sequence comprises only A.

116. The circular RNA precursor of any one of items 77-115, wherein the ribozyme recognition site II comprises or consists of a 3' exon region, and the ribozyme recognition site I comprises or consists of a 5' exon region.

117. The circular RNA precursor of item 116, wherein the 3' exon region is derived from the native 3' exon of the self-splicing intron, the 5' exon region is derived from the native 5' exon of the self-splicing intron, and the 3' exon region and 5' exon region can be recognized and/or spliced by the self-splicing intron or a combination of the 3' self-splicing intron fragment and the 5' self-

splicing intron fragment.

118. The circular RNA precursor of item 116 or 117, wherein the 3' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the native 3' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 5' terminal nucleotide of the native 3' exon, and

the 5' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the native 5' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 3' terminal nucleotide of the native 5' exon.

119. The circular RNA precursor of any one of items 77-118, wherein the ribozyme recognition site I comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70; preferably, the ribozyme recognition site I comprises a nucleotide sequence of GAGGUCUU, AGAGACUU, UCCAGCUU or UCUGGCAU.

120. The circular RNA precursor of any one of items 77-119, wherein the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisted of any of SEQ ID NO: 95-116, in particular the nucleotide motif consists of CUU, CUC or CAU.

121. The circular RNA precursor of any one of items 77-120, wherein the ribozyme recognition site II comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69; preferably, the ribozyme recognition site II comprises a nucleotide sequence of AAACCUC, AAUCUCU, AACUGGA or AACCAGC.

122. The circular RNA precursor of any one of items 77-121, wherein the ribozyme recognition site I comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70 and the ribozyme recognition site II comprises or consists of a nucleotide

sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69.

123. The circular RNA precursor of any one of items 77-122, wherein,

1) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 58 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 57;

2) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 60 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 59;

3) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 62 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 61;

4) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 64 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 63;

5) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 66 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 65;

6) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 68 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 67; or

7) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 69;

preferably, the ribozyme recognition site I is selected from a nucleotide sequence of UGUAACUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AGAAGUAAC;

the ribozyme recognition site I is selected from a nucleotide sequence of UGUAACUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AAAAUUACA;

the ribozyme recognition site I is selected from a nucleotide sequence of GAGGUCUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AAACCUC;

the ribozyme recognition site I is selected from a nucleotide sequence of AGAGACUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AAUCUCU;

the ribozyme recognition site I is selected from a nucleotide sequence of UCCAGCUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AACUGGA; or

the ribozyme recognition site I is selected from a nucleotide sequence of UCUGGCAU, and

the ribozyme recognition site II is selected from a nucleotide sequence of AACCAGC.

124. The circular RNA precursor of any one of items 77-123, wherein the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

125. The circular RNA precursor of item 124, wherein the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

126. The circular RNA precursor of any one of items 124-125, wherein the homology arm sequence I or the homology arm sequence II is about 5-500, about 5-450, about 5-400, about 5-350, about 5-300, about 5-250, about 5-200, about 5-150, about 5-100, about 5-80, about 5-60, about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-15 or about 5-10 nucleotides in length, preferably, about 20, 25, 40, 50, 100, or 300 nucleotides in length.

127. The circular RNA precursor of any one of items 124-126, wherein the two homology arm sequences may be polyA and polyU, respectively, or polyG and polyC, respectively.

128. The circular RNA precursor of any one of items 124-127, wherein one of the homology arm sequence I and the homology arm sequence II comprises a nucleotide sequence of any one of SEQ ID NO: 71-94 or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 71-94, while the other homology arm sequence comprises the corresponding complementary sequence.

129. The circular RNA precursor of any one of items 77-128, wherein the translation initiation element is an internal ribosome entry site (IRES) which is selected from the following IRES sequences: Taura syndrome virus, blood-sucking bug virus, Tyler's encephalomyelitis virus, simian virus 40, red fire ant virus 1, cereal constrictor virus, reticulovirus Endothelial hyperplasia virus, Forman poliovirus 1, soybean inchworm virus, Kashmir bee virus, human rhinovirus 2, glass leafhopper virus-1, human immunodeficiency virus type 1, glass leafhopper virus-1, lice P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinovirus, Tea inchworm-like virus, Encephalomyocarditis virus (EMCV), Drosophila C virus, Cruciferous tobacco Virus, cricket paralysis virus, bovine viral diarrhea virus 1, black queen cell virus, aphid lethal paralysis virus, avian encephalomyelitis virus, acute bee paralysis virus, hibiscus yellow ring spot virus, swine fever virus, human FGF2, human SFTPA1, Human AML1/RUNX1, Drosophila Antennae, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L,

Human LEF1, Mouse HIF1 α , Human n.myc, mouse Gtx, human p27kipl, human PDGF2/c-sis, human p53, human Pim-1, mouse Rbm3, Drosophila reaper, canine Scamper, Drosophila Ubx, human UNR, mouse UtrA, human VEGF-A, human XIAP, Drosophila hairless, Saccharomyces cerevisiae TFIID, Saccharomyces cerevisiae YAP1, human c-src, human FGF-1, simian picornavirus, turnip crepe disease virus, eIF4G aptamer, Coxsackie Virus B3 (CVB3) or Coxsackie virus A (CVA1/2), preferably, the IRES is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2.

130. The circular RNA precursor of any one of items 77-129, wherein the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to one of SEQ ID NOs: 120-526.

131. The circular RNA precursor of any one of items 77-130, wherein the at least one protein-coding sequence encodes a protein of eukaryotic, prokaryotic or viral origin, for example, a protein for therapeutic or diagnostic use.

132. The circular RNA precursor of any one of items 77-131, wherein the at least one protein-coding sequence encodes IL-2, POLR2A, or an aquaporin.

133. The circular RNA precursor of item 132, wherein the aquaporin is AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

134. The circular RNA precursor of any one of items 77-133, wherein the at least one protein-coding sequence comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

135. The circular RNA precursor of any one of items 77-134, which comprises a nucleotide sequence selected from any one of SEQ ID NO: 530, 531 and 533-534, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 530, 531 and 533-534.

136. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a translation initiation element fragment II, at least one protein-coding sequence, and a translation initiation element fragment I;

wherein the translation initiation element fragment I is derived from a 5' terminal portion of a translation initiation element, and the translation initiation element fragment II is derived from a 3' terminal portion of the translation initiation element, and

the translation initiation element fragment I comprises a ribozyme recognition site I located at its 3' end, the translation initiation element fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the translation initiation element fragment I is connected with the translation initiation element fragment II to form the translation initiation element which can recruit ribosome to initiate translation;

optionally, the at least one protein-coding sequence comprises at least two protein-coding sequences, and a further translation initiation element, 2A sequence or spacer may be linked between any two adjacent protein-coding sequences, and the translation initiation element comprises: IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof.

137. The circular RNA precursor of item 136, wherein the self-splicing intron is selected from Group I introns and Group II introns, for example, the self-splicing intron is selected from Group I introns.

138. The circular RNA precursor of item 136 or 137, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At), for example, the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

139. The circular RNA precursor of any one of items 136-138, wherein the 3' self-splicing intron fragment is derived from a 3' terminal portion of a self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is derived from a 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

140. The circular RNA precursor of any one of items 136-139, wherein when the self-splicing intron is cyanobacterium *Anabaena* Group I Intron such as AnaX, the intron internal split site II and the intron internal split site I are a nucleotide A at position 115 and a nucleotide A at position 116 from 5' end of the AnaX Intron (SEQ ID NO:1) respectively, and when connected form a nucleotide sequence of AA.

141. The circular RNA precursor of any one of items 136-139, wherein when the self-splicing intron is *Azoarcus* Group I Intron (Azo) (SEQ ID NO:2),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 24 and a nucleotide A at position 25 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 85 and a nucleotide A at position 86 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P4-5);

3) the intron internal split site II and the intron internal split site I are a nucleotide C at position 71 and a nucleotide C at position 72 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CC (P5);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 108 and a nucleotide U at position 109 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CU (P6);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 154 and a nucleotide A at position 155 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of AA (P8); or

6) the intron internal split site II and the intron internal split site I are a nucleotide C at position 198 and a nucleotide A at position 199 from 5' end of the Azo intron respectively, and when connected form a nucleotide sequence of CA (P9).

142. The circular RNA precursor of any one of items 136-139, wherein when the self-splicing intron is *Scytalidium dimidiatum* Group I Intron (Sd) (SEQ ID NO:4),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position

88 and a nucleotide G at position 89 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P2);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide G at position 174 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of UG (P5);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 204 and a nucleotide G at position 205 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of AG (P6);

4) the intron internal split site II and the intron internal split site I are a nucleotide C at position 252 and a nucleotide G at position 253 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P8); or

5) the intron internal split site II and the intron internal split site I are a nucleotide C at position 370 and a nucleotide G at position 371 from 5' end of the Sd intron respectively, and when connected form a nucleotide sequence of CG (P9).

143. The circular RNA precursor of any one of items 136-139, wherein when the self-splicing intron is *Staphylococcus phage Twort* Group I Intron (Twort) (SEQ ID NO:3),

1) the intron internal split site II and the intron internal split site I are a nucleotide U at position 70 and a nucleotide A at position 71 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P5);

2) the intron internal split site II and the intron internal split site I are a nucleotide U at position 103 and a nucleotide A at position 104 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P6);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 132 and a nucleotide U at position 133 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AU (P7);

4) the intron internal split site II and the intron internal split site I are a nucleotide U at position 173 and a nucleotide C at position 174 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UC (P8);

5) the intron internal split site II and the intron internal split site I are a nucleotide A at position 206 and a nucleotide A at position 207 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P9);

6) the intron internal split site II and the intron internal split site I are a nucleotide G at position 226 and a nucleotide U at position 227 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of GU (P9.1);

7) the intron internal split site II and the intron internal split site I are a nucleotide U at position

23 and a nucleotide A at position 24 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of UA (P2); or

8) the intron internal split site II and the intron internal split site I are a nucleotide A at position 82 and a nucleotide A at position 83 from 5' end of the Twort intron respectively, and when connected form a nucleotide sequence of AA (P4-5).

144. The circular RNA precursor of any one of items 136-139, wherein when the self-splicing intron is *Scytonema-hofmani* tRNA fMet group I intron (Sh) (SEQ ID NO:5),

1) the intron internal split site II and the intron internal split site I are a nucleotide A at position 129 and a nucleotide A at position 130 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P6);

2) the intron internal split site II and the intron internal split site I are a nucleotide A at position 60 and a nucleotide A at position 61 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P2);

3) the intron internal split site II and the intron internal split site I are a nucleotide A at position 109 and a nucleotide A at position 110 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of AA (P4-5); or

4) the intron internal split site II and the intron internal split site I are a nucleotide A at position 213 and a nucleotide A at position 214 from 5' end of the Sh intron respectively, and when connected form a nucleotide sequence of UA (P9).

145. The circular RNA precursor of any one of items 136-139, wherein when the self-splicing intron is *Agrobacterium-tumefaciens* group I intron (At) (SEQ ID NO:6), the intron internal split site II and the intron internal split site I are a nucleotide A at position 107 and a nucleotide A at position 108 from 5' end of the At intron respectively, and when connected form a nucleotide sequence of AA (P6).

146. The circular RNA precursor of any one of items 136-145, wherein the 3' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 3' terminal portion of the self-splicing intron, and the 5' self-splicing intron fragment is a sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the 5' terminal portion of the self-splicing intron.

147. The circular RNA precursor of any one of items 136-146, wherein the self-splicing intron

is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

148. The circular RNA precursor of any one of items 136-146, wherein the self-splicing intron is the *Azoarcus* Group I Intron (Azo), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 9, 11, 13, 15, 17 or 19 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 9, 11, 13, 15, 17 or 19, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 10, 12, 14, 16, 18 or 20 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 10, 12, 14, 16, 18 or 20.

149. The circular RNA precursor of any one of items 136-146, wherein the self-splicing intron is the *Scytalidium dimidiatum* Group I Intron (Sd), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 37, 39, 41, 43 or 45 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 37, 39, 41, 43 or 45, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 38, 40, 42, 44 or 46 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 38, 40, 42, 44 or 46.

150. The circular RNA precursor of any one of items 136-146, wherein the self-splicing intron

is the Staphylococcus phage Twort Group I Intron (Twort), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36.

151. The circular RNA precursor of any one of items 136-146, wherein the self-splicing intron is the Scytonema-hofmani tRNA fMet group I intron (Sh), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 47, 49, 51 or 53 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 47, 49, 51 or 53, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 48, 50, 52 or 54 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 48, 50, 52 or 54.

152. The circular RNA precursor of any one of items 136-146, wherein the self-splicing intron is the Agrobacterium-tumefaciens group I intron (At), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 55 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 55, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 56 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 56.

153. The circular RNA precursor of any one of items 136-152, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA comprising them has reduced immunogenicity.

154. The circular RNA precursor of any one of items 136-153, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA precursor comprising them has comparable or increased circularization efficiency.

155. The circular RNA precursor of any one of items 136-154, wherein the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

156. The circular RNA precursor of any one of items 136-155, wherein the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

157. The circular RNA precursor of any one of items 136-156, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured to be capable of forming a stem-loop structure, e.g., upon self-splicing for circularization.

158. The circular RNA precursor of item 157, wherein the loop of the stem-loop structure comprises the splicing junction.

159. The circular RNA precursor of any one of items 157-158, wherein the ribozyme recognition site I comprises the sequence structure of the following formula: 5'-third non-pairing

sequence-third pairing sequence-third loop sequence-3'; and the ribozyme recognition site II comprises the sequence structure of the following formula: 5'-fourth loop sequence-fourth pairing sequence-fourth non-pairing sequence-3',

wherein the third non-pairing sequence or the fourth non-pairing sequence may be independently present or absent,

the third pairing sequence and the fourth pairing sequence can complementarily pair to each other to form the stem of the stem-loop structure, wherein the third loop sequence and the fourth loop sequence can form the loop of the stem-loop structure, e.g., through self-splicing for circularization.

160. The circular RNA precursor of any one of items 136-159, wherein the ribozyme recognition site II can be selected from a various sequence.

161. The circular RNA precursor of item 159 or 160, wherein the fourth loop sequence comprises or consists of one or more nucleotides which can pair with the P1 region of the corresponding self-splicing intron to form a P10 duplex region during the circularization.

162. The circular RNA precursor of any one of items 159-161, wherein the fourth loop sequence consist of a nucleotide sequence of (N)_n, wherein N represents any nucleotides (A, G, U, or C), n represents an integer from 1-20, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

163. The circular RNA precursor of any one of items 159-162, wherein the fourth loop sequence comprises or consists of the sequence of AAAA, AGAA, AA, UUUU, UAAA, CAAA, or GAAA, preferably AAAA.

164. The circular RNA precursor of any one of items 136-163, wherein the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

165. The circular RNA precursor of any one of items 136-164, wherein the third loop sequence comprises or consists of one or more nucleotides which can pair with the internal guide sequence (IGS) of the corresponding self-splicing intron to form a P1 duplex region during the circularization.

166. The circular RNA precursor of any one of items 159-165, wherein the third loop sequence comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, such as CUU, CAU or CUC, preferably CUU.

167. The circular RNA precursor of any one of items 159-166, wherein the loop of the stem-

loop structure has a sequence of CUUAAAA, CUUUUUU, CUUAA, CUUGAAA, CUUUAAA, CUUAGAA, CUUCAAA or CUCAAAA, preferably CUUAAAA.

168. The circular RNA precursor of any one of items 159-167, wherein the stem portion of the stem-loop structure comprises 2-15 or more consecutive matched base pairs, preferably, the stem portion of the stem-loop structure comprises 5, 6 or 7 consecutive matched base pairs.

169. The circular RNA precursor of any one of items 159-168, wherein the stem portion in the stem-loop structure comprises up to 2 base pair mismatches, or the stem portion comprises only 1 base pair mismatch, preferably, the stem portion comprises no base pair mismatches.

170. The circular RNA precursor of any one of items 159-169, wherein the third pairing sequence comprises only G and the fourth pairing sequence comprises only C, or the third pairing sequence comprises only C and the fourth pairing sequence comprises only G.

171. The circular RNA precursor of any one of items 159-169, wherein the third pairing sequence comprises only A and the fourth pairing sequence comprises only U, or the third pairing sequence comprises only U and the fourth pairing sequence comprises only A.

172. The circular RNA precursor of any one of items 159-171, wherein the ribozyme recognition site II comprises or consists of a 3' exon region, and the ribozyme recognition site I comprises or consists of a 5' exon region.

173. The circular RNA precursor of item 172, wherein the 3' exon region is derived from the native 3' exon of the self-splicing intron, the 5' exon region is derived from the native 5' exon of the self-splicing intron, and the 3' exon region and 5' exon region can be recognized and/or spliced by the self-splicing intron or a combination of the 3' self-splicing intron fragment and the 5' self-splicing intron fragment.

174. The circular RNA precursor of item 172 or 173, wherein the 3' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the native 3' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 5' terminal nucleotide of the native 3' exon, and the 5' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to the native 5' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 3' terminal nucleotide of the native 5' exon.

175. The circular RNA precursor of any one of items 136-174, wherein the ribozyme

recognition site I comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70; preferably, the ribozyme recognition site I comprises a nucleotide sequence of UGUAACUU.

176. The circular RNA precursor of any one of items 136-175, wherein the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisted of any of SEQ ID NO: 95-116, in particular the nucleotide motif consists of CUU, CUC or CAU.

177. The circular RNA precursor of any one of items 136-176, wherein the ribozyme recognition site II comprises or consists of a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence selected from SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69; preferably, the ribozyme recognition site II comprises a nucleotide sequence of AGAAGUAAAC or AAAAUUACA.

178. The circular RNA precursor of any one of items 136-177, wherein the ribozyme recognition site I comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 58, 60, 62, 64, 66, 68 or 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 57, 59, 61, 63, 65, 67 or 69.

179. The circular RNA precursor of any one of items 136-178, wherein,

1) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 58 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 57;

2) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 60 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 59;

3) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 62 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 61;

4) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 64 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 63;

5) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID

NO: 66 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 65;

6) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 68 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 67; or

7) the ribozyme recognition site I comprises or consists of a nucleotide sequence of SEQ ID NO: 70 and the ribozyme recognition site II comprises or consists of a nucleotide sequence of SEQ ID NO: 69;

preferably, the ribozyme recognition site I is selected from a nucleotide sequence of UGUAACUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AGAAGUAAC;

the ribozyme recognition site I is selected from a nucleotide sequence of UGUAACUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AAAAUUACA;

the ribozyme recognition site I is selected from a nucleotide sequence of GAGGUCUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AAACCUC;

the ribozyme recognition site I is selected from a nucleotide sequence of AGAGACUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AAUCUCU;

the ribozyme recognition site I is selected from a nucleotide sequence of UCCAGCUU, and the ribozyme recognition site II is selected from a nucleotide sequence of AACUGGA; or

the ribozyme recognition site I is selected from a nucleotide sequence of UCUGGCAU, and the ribozyme recognition site II is selected from a nucleotide sequence of AACCAGC.

180. The circular RNA precursor of any one of items 136-179, wherein the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

181. The circular RNA precursor of item 180, wherein the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

182. The circular RNA precursor of any one of items 180-181, wherein the homology arm sequence I or the homology arm sequence II is about 5-500, about 5-450, about 5-400, about 5-350, about 5-300, about 5-250, about 5-200, about 5-150, about 5-100, about 5-80, about 5-60, about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-15 or about 5-10 nucleotides in length, preferably, about 20, 25, 40, 50, 100, or 300 nucleotides in length.

183. The circular RNA precursor of any one of items 180-182, wherein the two homology arm sequences may be polyA and polyU, respectively, or polyG and polyC, respectively.

184. The circular RNA precursor of any one of items 180-183, wherein one of the homology

arm sequence I and the homology arm sequence II has the nucleotide sequence of any one of SEQ ID NO: 71-94 or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 71-94, while the other homology arm sequence has the corresponding complementary sequence.

185. The circular RNA precursor of any one of items 136-184, wherein the translation initiation element is an internal ribosome entry site (IRES) which is selected from the following IRES sequences: Taura syndrome virus, blood-sucking bug virus, Tyler's encephalomyelitis virus, simian virus 40, red fire ant virus 1, cereal constrictor virus, reticulovirus Endothelial hyperplasia virus, Forman poliovirus 1, soybean inchworm virus, Kashmir bee virus, human rhinovirus 2, glass leafhopper virus-1, human immunodeficiency virus type 1, glass leafhopper virus-1, lice P virus , Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinovirus, Tea inchworm-like virus, Encephalomyocarditis virus (EMCV), Drosophila C virus, Cruciferous tobacco Virus, cricket paralysis virus, bovine viral diarrhea virus 1, black queen cell virus, aphid lethal paralysis virus, avian encephalomyelitis virus, acute bee paralysis virus, hibiscus yellow ring spot virus, swine fever virus, human FGF2, human SFTPA1, Human AML1/RUNX1, Drosophila Antennae, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 α , Human n.myc, mouse Gtx, human p27kipl, human PDGF2/c-sis, human p53, human Pim-1, mouse Rbm3, Drosophila reaper, canine Scamper, Drosophila Ubx, human UNR, mouse UtrA, human VEGF-A, human XIAP, Drosophila hairless, Saccharomyces cerevisiae TFIID, Saccharomyces cerevisiae YAP1, human c-src, human FGF-1, simian picornavirus, turnip crepe disease virus, eIF4G aptamer, Coxsackie Virus B3 (CVB3) or Coxsackie virus A (CVA1/2), preferably, the IRES is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2.

186. The circular RNA precursor of any one of items 136-185, wherein the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprise a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to one of SEQ ID NOs: 120-526.

187. The circular RNA precursor of any one of items 136-186, wherein the at least one protein-coding sequence encodes a protein of eukaryotic, prokaryotic or viral origin, for example,

a protein for therapeutic or diagnostic use.

188. The circular RNA precursor of any one of items 136-187, wherein the at least one protein-coding sequence encodes IL-2, POLR2A, or an aquaporin, such as AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

189. The circular RNA precursor of any one of items 136-188, wherein the at least one protein-coding sequence comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

190. The circular RNA precursor of any one of items 136-189, which comprises a nucleotide sequence selected from any one of SEQ ID NO: 528, 529 and 535-536, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 528, 529 and 535-536.

191. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest;
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a non-coding sequence fragment II, and a non-coding sequence fragment I;

wherein the non-coding sequence fragment I is derived from a 5' terminal portion of a non-coding sequence, and the non-coding sequence fragment II is derived from a 3' terminal portion of the non-coding sequence, and the non-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the non-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-coding

sequence fragment I is connected with the non-coding sequence fragment II to form the complete non-coding sequence.

192. The circular RNA precursor of item 191, wherein the non-coding sequence is an endogenous circRNA.

193. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment;
- b) a nucleotide sequence of interest; and
- c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a non-TIE functional element fragment II, a translation initiation element, and at least one protein-coding sequence, and a non-TIE functional element fragment I;

wherein the non-TIE functional element fragment I is derived from a 5' terminal portion of a non-TIE functional element, and the non-TIE translation initiation element fragment II is derived from a 3' terminal portion of the non-TIE functional element, and

the non-TIE functional element fragment I comprises a ribozyme recognition site I located at its 3' end, the non-TIE functional fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the non-TIE functional element fragment I is connected with the non-TIE functional element fragment II to form the non-TIE functional element.

194. The circular RNA precursor of item 193, wherein the non-TIE functional element is selected from 3' UTR or a replicon.

195. The circular RNA precursor of any one of items 191-194, wherein the self-splicing intron, the 3' self-splicing intron fragment, or the 5' self-splicing intron fragment is defined as in any one of items 136-190.

196. The circular RNA precursor of any one of items 191-195, wherein the ribozyme recognition site I is as defined in any one of items 136-190 and/or the ribozyme recognition site II is as defined in any one of items 136-190.

197. The circular RNA precursor of any one of items 191-196, wherein the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II, which are as defined in any one of items 180-184.

198. The circular RNA precursor of any one of items 191-197, wherein the translation

initiation element is as defined in any one of items 185-186.

199. The circular RNA precursor of any one of items 193-198, wherein the at least one protein-coding sequence is as defined in any one of items 136-190.

200. A circular RNA precursor comprising from 5' to 3' direction:

a) a 3' self-splicing intron fragment;

b) a nucleotide sequence of interest, comprising from 5' to 3' direction: protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES) and a protein-coding sequence fragment I; and

c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising the protein-coding sequence fragment II, the translation initiation element, and the protein-coding sequence fragment I through the self-splicing of the circular RNA precursor;

wherein the protein-coding sequence fragment I is derived from a 5' terminal portion of at least one protein-coding sequence, and the protein-coding sequence fragment II is derived from a 3' terminal portion of at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence;

wherein the at least one protein-coding sequence encodes an aquaporin, such as AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

201. The circular RNA precursor of item 200, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At).

202. The circular RNA precursor of item 200 or 201, wherein the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

203. The circular RNA precursor of any one of items 200-202, wherein the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1 and 3, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 1 and 3.

204. The circular RNA precursor of any one of items 200-203, wherein the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

205. The circular RNA precursor of any one of items 200-203, wherein the self-splicing intron is the *Staphylococcus* phage Twort Group I Intron (Twort), the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 27 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 27, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 28 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 28.

206. The circular RNA precursor of any one of items 200-205, wherein the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 35 nucleotides, such as 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, or 35 nucleotides in length.

207. The circular RNA precursor of any one of items 200-206, wherein the total length of the

ribozyme recognition site I is 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, or 35 nucleotides in length.

208. The circular RNA precursor of any one of items 200-207, wherein the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 58, 60 or 64.

209. The circular RNA precursor of any one of items 200-208, wherein the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisted of any of SEQ ID NO: 95-116, in particular the nucleotide motif consists of CUU, CUC or CAU.

210. The circular RNA precursor of any one of items 200-209, wherein the ribozyme recognition site I comprises a nucleotide sequence of UCCAGCUU or UCUGGCAU.

211. The circular RNA precursor of any one of items 200-210, wherein the ribozyme recognition site II comprises a nucleotide sequence of AACUGGA or AACCAGC.

212. The circular RNA precursor of any one of items 200-211, wherein the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

213. The circular RNA precursor of item 212, wherein the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

214. The circular RNA precursor of any one of items 212-213, wherein the homology arm sequence I or the homology arm sequence II is about 5-50 nucleotides in length, preferably, about 10, 20, 30 or 40 nucleotides in length.

215. The circular RNA precursor of any one of items 212-214, wherein one of the homology arm sequence I and the homology arm sequence II comprises a nucleotide sequence of any one of SEQ ID NO: 71-94 such as 83-84, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 71-94 such as 83-84, while the other homology arm sequence comprises the corresponding complementary sequence.

216. The circular RNA precursor of any one of items 200-215, wherein the translation initiation element is an internal ribosome entry site (IRES) which is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2, preferably CVB3.

217. The circular RNA precursor of any one of items 200-216, wherein the IRES comprises a nucleotide sequence of SEQ ID NO: 121, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to SEQ ID NO: 121.

218. The circular RNA precursor of any one of items 200-217, wherein the at least one protein-coding sequence comprises an AQP1-coding sequence, such as a nucleotide sequence of SEQ ID NO: 118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 118.

219. The circular RNA precursor of any one of items 200-218, which comprises a nucleotide sequence selected from any one of SEQ ID NO: 533-534, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 533-534.

220. A circular RNA precursor comprising from 5' to 3' direction:

a) a 3' self-splicing intron fragment;

b) a nucleotide sequence of interest, comprising from 5' to 3' direction: a translation initiation element fragment II, at least one protein-coding sequence encoding an aquaporin and a translation initiation element fragment I; and

c) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA comprising a translation initiation element fragment II, at least one protein-coding sequence, and a translation initiation element fragment I through the self-splicing of the circular RNA precursor;

wherein the translation initiation element fragment I is derived from a 5' terminal portion of a translation initiation element, and the translation initiation element fragment II is derived from a 3' terminal portion of the translation initiation element, and

the translation initiation element fragment I comprises a ribozyme recognition site I located at its 3' end, the translation initiation element fragment II comprises a ribozyme recognition site II

located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the translation initiation element fragment I is connected with the translation initiation element fragment II to form the translation initiation element which can recruit ribosome to initiate translation;

wherein the aquaporin is preferably AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

221. The circular RNA precursor of item 220, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus phage Twort* Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At).

222. The circular RNA precursor of item 220 or 221, wherein the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

223. The circular RNA precursor of any one of items 220-222, wherein the self-splicing intron is selected from a nucleotide sequence of any one of SEQ ID NO: 1, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 1.

224. The circular RNA precursor of any one of items 220-223, wherein the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

225. The circular RNA precursor of any one of items 220-224, wherein the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 35 nucleotides, such as 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, or 35 nucleotides in length.

226. The circular RNA precursor of any one of items 220-225, wherein the total length of the ribozyme recognition site I is 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, or 35 nucleotides in length.

227. The circular RNA precursor of any one of items 220-226, wherein the ribozyme recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 58 or 60.

228. The circular RNA precursor of any one of items 220-227, wherein the ribozyme recognition site I comprises or consists of from 5' to 3' direction a nucleotide motif consisted of any of SEQ ID NO: 95-116, in particular the nucleotide motif consists of CUU, CUC or CAU.

229. The circular RNA precursor of any one of items 220-228, wherein the ribozyme recognition site I comprises a nucleotide sequence of UGUAACUU.

230. The circular RNA precursor of any one of items 220-229, wherein the ribozyme recognition site II comprises a nucleotide sequence of AGAAGUAAC or AAAAUUACA.

231. The circular RNA precursor of any one of items 220-230, wherein the circular RNA precursor further comprises a homology arm sequence I and a homology arm sequence II capable of complementary pairing to each other to form a homology arm double-stranded region.

232. The circular RNA precursor of item 231, wherein the homology arm sequence I is upstream of the 3' self-splicing intron fragment and the homology arm sequence II is downstream of the 5' self-splicing intron fragment.

233. The circular RNA precursor of any one of items 231-232, wherein the homology arm sequence I or the homology arm sequence II is about 5-500 nucleotides in length, preferably, about 10, 20, 30, 40 or 50 nucleotides in length.

234. The circular RNA precursor of any one of items 231-233, wherein one of the homology arm sequence I and the homology arm sequence II comprises a nucleotide sequence of any one of SEQ ID NO: 71-94 such as 83-84, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence

identity to any one of SEQ ID NO: 71-94 such as 83-84, while the other homology arm sequence comprises the corresponding complementary sequence.

235. The circular RNA precursor of any one of items 220-234, wherein the translation initiation element is an internal ribosome entry site (IRES) which is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2, preferably CVB3.

236. The circular RNA precursor of any one of items 220-235, wherein the IRES comprises a nucleotide sequence of SEQ ID NO: 121, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to SEQ ID NO: 121.

237. The circular RNA precursor of any one of items 220-236, wherein the at least one protein-coding sequence comprises a AQP-1-coding sequence, such as a nucleotide sequence of SEQ ID NO: 118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 118.

238. The circular RNA precursor of any one of items 220-237, which comprises a nucleotide sequence selected from any one of SEQ ID NO: 535-536, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 535-536.

239. A nucleic acid vector for generating a circular RNA molecule, said vector comprises a coding sequence of the circular RNA precursor of any one of items 1-238.

240. The nucleic acid vector of item 239, which further comprises an RNA polymerase promoter sequence operably linked to the coding sequence of the circular RNA precursor.

241. The nucleic acid vector of item 240, wherein the promoter is a T7 RNA polymerase promoter, a T6 viral RNA polymerase promoter, a SP6 viral RNA polymerase promoter, a T3 viral RNA polymerase promoter or a T4 viral RNA polymerase promoter, preferably a T7 RNA polymerase promoter.

242. A method for preparing a circular RNA, comprising:

1) providing the circular RNA precursor of any one of items 1-238 or obtaining a circular RNA precursor by transcribing from the nucleic acid vector of any one of items 239-241;

2) incubating the circular RNA precursor in the presence of a divalent metal cation at a temperature at which RNA circularization occurs; and

3) harvesting the circular RNA obtained in step 2).

243. The method of item 242, wherein the divalent metal cation is Mg^{2+} and/or Mn^{2+} .

244. The method of item 242 or 243, wherein the concentration of the divalent metal cation is about 5 mM to about 550 mM.

245. A method for preparing a circular RNA, comprising

a) providing a nucleic acid vector comprising a self-splicing intron-based RNA circularizing elements as a transcription template; and

b) incubating the nucleic acid vector in an in vitro transcription system comprising a divalent metal cation and an RNA polymerase for a first time period during which the linear RNA produced by in vitro transcription is self-circularized under the action of the RNA circularizing elements to produce a circular RNA.

246. The method of item 245, wherein the nucleic acid vector is the nucleic acid vector of any one of items 239-241.

247. The method of item 245 or 246, wherein the divalent metal cation in the in vitro transcription system is Mg^{2+} .

248. The method of any one of items 245-247, wherein the in vitro transcription system further comprises a monovalent metal cation and/or a monovalent anion.

249. The method of item 248, wherein the monovalent metal cation is Na^{+} or K^{+} .

250. The method of item 248, wherein the monovalent anion is Cl^{-} or CH_3COO^{-} (OAc).

251. The method of any one of items 245-250, wherein the concentration of the divalent metal cation in the system during the first time period is from about 5 mM to about 50 mM.

252. The method of any one of items 248-251, wherein the concentration of the monovalent metal cation in the system during the first time period is from about 5 mM to about 100 mM.

253. The method of any one of items 248-252, wherein the concentration of the monovalent anion in the system during the first time period is from about 5 mM to about 150 mM.

254. The method of any one of items 245-253, wherein the method does not comprise a step of isolating and/or purifying the linear RNA produced by the in vitro transcription.

255. The method of any one of items 245-254, wherein the buffer of the in vitro transcription system is selected from Tris-HCl buffer, HEPES buffer, MES buffer, citrate buffer, or phosphate buffer, preferably, the buffer of the in vitro transcription system is HEPES buffer.

256. The method of any one of items 245-255, wherein the in vitro transcription system has a pH of about 5- about 8, preferably, the in vitro transcription system has a pH of about 7.5.

257. The method of any one of items 245-256, wherein the RNA polymerase is selected from

a T7 RNA polymerase, a T6 viral RNA polymerase, a SP6 viral RNA polymerase, a T3 viral RNA polymerase, or a T4 viral RNA polymerase, preferably, the RNA polymerase is T7 RNA polymerase.

258. The method of any one of items 245-257, wherein the first time period is about 0.5 hours to about 24 hours, preferably, the first time period is 3 hours.

259. The method of any one of items 245-258, wherein the incubation for the first time period is carried out at about 16°C to about 60°C, preferably, the incubation for the first time period is carried out at about 37°C.

260. The method of any one of items 245-259, wherein after incubation for the first time period in step b), the method further comprises step c):

adding an additional amount of metal cation to the in vitro transcription system and incubating for a second time period; or

changing the buffer of the system, adding a metal cation and incubating for a second time period.

261. The method of item 260, wherein the metal cation added for the incubation for the second time period is a divalent metal cation, such as Mg^{2+} or Mn^{2+} .

262. The method of item 260 or 261, wherein during the incubation for the second time period, the metal cation is added to a final concentration of about 5 mM to about 550 mM.

263. The method of any one of items 260-262, wherein the buffer of the system is changed to Tris-HCl buffer, HEPES buffer, MES buffer, citrate buffer, or phosphate buffer during the second time period.

264. The method of any one of items 260-263, wherein the pH of the system during the second time period is 5-8.

265. The method of any one of items 260-264, wherein the second time period is about 5 minutes to about 8 hours, about 5 minutes to about 6 hours, about 5 minutes to about 4 hours, or about 5 minutes to about 2 hours.

266. The method of any one of items 260-265, wherein the incubation for the second time period is carried out at about 25°C to about 75°C, such as about 37°C.

267. The method of any one of items 260-266, wherein the method further comprises a step of recovering or purifying the circular RNA as produced.

268. A method for purifying a circular RNA, comprising:

a) contacting a mixture comprising the circular RNA and the circular RNA precursor with a circular RNA-specific probe under a condition that allows the circular RNA-specific probe to specifically bind to and form a complex with the circular RNA;

b) separating the complex from one or more components in the mixture that are not bound to

the circular RNA-specific probe; and

c) releasing the circular RNA from the complex.

269. The method of item 268, wherein the circular RNA is prepared by circularizing a circular RNA precursor.

270. The method of item 268 or 269, wherein the circular RNA is prepared by ligating both ends of a linear circular RNA precursor with an RNA ligase, or, the circular RNA is prepared by the self-splicing ribozyme activity of self-splicing intron-based circularizing elements contained in the circular RNA precursor.

271. The method of any one of items 268-270, wherein the circular RNA-specific probe is a single-stranded DNA or RNA probe.

272. The method of any one of items 268-271, wherein the circular RNA-specific probe specifically hybridizes to a region flanking the circularization junction of the circular RNA.

273. The method of any one of items 268-272, wherein the circular RNA-specific probe is about 10 nucleotides-about 35 nucleotides, about 10 nucleotides- about 50 nucleotides, about 10 nucleotides- about 80 nucleotides, or about 10 nucleotides- about 100 nucleotides in length or longer.

274. The method of any one of items 268-273, wherein the circular RNA-specific probe is immobilized on a support such as a solid support, e.g., the circular RNA-specific probe is immobilized on the support after binding to the circular RNA or the circular RNA-specific probe is pre-immobilized on the support.

275. The method of any one of items 268-274, wherein the condition in step a) include denaturing the RNA at between about 60°C and about 95°C for about 2 minutes to about 10 minutes, then gradually reducing the temperature to below about 40°C to allow the circular RNA annealing to the circular RNA-specific probe.

276. The method of item 275, wherein step c) is performed by increasing the temperature to about 60°C to about 95°C (e.g., about 60°C, about 62°C, about 64°C, about 66°C, about 68°C, about 70°C, about 75°C, about 80°C, about 85°C, about 90°C, about 95°C) to release the circular RNA.

277. The method of any one of items 268-276, wherein the condition in step a) comprises a high salt concentration range of 0.25M-2M.

278. The method of item 277, wherein the salt is NaCl or a guanidine salt.

279. The method of item 277 or 278, wherein in step c) the circular RNA is released by elution with an elution buffer, such as a low salt buffer.

280. The method of item 279, wherein the elution buffer is Tris-EDTA buffer (TE buffer) or water.

281. The method of any one of items 268-280, wherein in step b), the one or more components are removed by washing the complex with a washing buffer.

282. The method of any one of items 268-281, which further comprises the following steps:

i) contacting the mixture comprising the circular RNA and the circular RNA precursor to a linear circular RNA precursor-specific probe under a condition that allows the linear circular RNA precursor-specific probe to specifically bind to and form a complex with the linear circular RNA precursor;

ii) removing the complex formed by the linear circular RNA precursor-specific probe with the linear circular RNA precursor from the mixture, and

iii) collecting the circular RNA-containing mixture obtained in step ii).

283. The method of item 282, wherein steps i)-iii) are performed before step a), for example, steps i)-iii) may be performed multiple times before step a), e.g., 2, 3, 4 or more times.

284. The method of item 283, wherein steps i)-iii) are performed concurrently with steps a)-c).

285. A method for purifying a circular RNA, comprising:

i) contacting a mixture comprising the circular RNA and the circular RNA precursor to a circular RNA precursor-specific probe under a condition that allows the circular RNA precursor-specific probe to specifically bind to and form a complex with the circular RNA precursor; and

ii) removing the complex formed by the circular RNA precursor-specific probe with the circular RNA precursor from the mixture,

iii) collecting the circular RNA-containing mixture obtained in step ii), and

optionally, steps i)-iii) are performed multiple times, e.g., 2 times, 3 times, 4 times or more.

286. The method of item 285, wherein the circular RNA precursor-specific probe specifically binds to the circular RNA precursor and does not substantially bind to the circular RNA.

287. The method of item 285 or 286, wherein the circular RNA precursor-specific probe is immobilized on a support, such as a solid support, for example, the circular RNA precursor-specific probe is immobilized on the support after binding to the circular RNA precursor, or the circular RNA precursor-specific probe is pre-immobilized on the support.

288. The method of any one of items 285-287, wherein the circular RNA precursor is selected from the circular RNA precursor of any one of items 21-76,

wherein the circular RNA precursor is capable of removing the 3' self-splicing intron fragment and the 5' self-splicing intron fragment by self-splicing, generating a circular RNA comprising the first residual circularizing element, the nucleotide sequence of interest and the second residual circularizing element.

289. The method of item 288, wherein the circular RNA precursor-specific probe specifically

hybridizes to at least a portion of the first residual circularizing element or the second residual circularizing element.

290. The method of item 288 or 289, wherein the circular RNA precursor-specific probe hybridizes to a portion of the circular RNA precursor outside the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element.

291. The method of any one of items 288-290, wherein the circular RNA precursor-specific probe hybridizes to the 3' self-splicing intron fragment or a portion thereof or a 5' flanking sequence thereof, or the 5' self-splicing intron fragment or a portion thereof or a 3' flanking sequence thereof.

292. The method of any one of items 288-291, wherein the circular RNA precursor contains a sequence selected from SEQ ID NOs: 600-610 or a complementary sequence thereof, preferably, SEQ ID NO:606 or a complementary sequence thereof outside the first residual circularizing element, the nucleotide sequence of interest, and the second residual circularizing element, to which the circular RNA precursor-specific probe specifically hybridizes.

293. The method of any one of items 285-287, wherein the circular RNA precursor is selected from the circular RNA precursor of any one of items 77-238,

wherein the circular RNA precursor is capable of removing the 3' self-splicing intron fragment and the 5' self-splicing intron fragment by self-splicing, generating a circular RNA comprising the nucleotide sequence of interest.

294. The method of item 293, wherein the circular RNA precursor-specific probe hybridizes to a portion of the circular RNA precursor outside the nucleotide sequence of interest.

295. The method of any one of items 293-294, wherein the circular RNA precursor-specific probe hybridizes to the 3' self-splicing intron fragment or a portion thereof or a 5' flanking sequence thereof, or the 5' self-splicing intron fragment or a portion thereof or a 3' flanking sequence thereof.

296. The method of any one of items 293-295, wherein the circular RNA precursor contains a sequence selected from SEQ ID NOs: 600-610 or a complementary sequence thereof, preferably, SEQ ID NO:606 or a complementary sequence thereof outside the nucleotide sequence of interest, to which the linear circular RNA precursor-specific probe specifically hybridizes.

297. The method of any one of items 285-296, wherein the molar ratio of the probe to the RNA molecules in the mixture is from about 1:1 to about 200,000:1, from about 1:1 to about 150,000:1, from about 1:1 to about 100,000:1, from about 1:1 to about 90,000:1, from about 1:1 to about 80,000:1, from about 1:1 to about 70,000:1, from about 1:1 to about 60,000:1, from about 1:1 to about 50,000:1, from about 1:1 to about 40,000:1, from about 1:1 to about 30,000:1, from about 1:1 to about 20,000:1, from about 1:1 to about 10,000:1.

298. A method for purifying a circular RNA, comprising:

i) adding a linear RNA-specific tag to a mixture comprising circular RNA and linear RNA, wherein the linear RNA-specific tag can be linked to the linear RNA;

ii) contacting the mixture comprising circular RNA and linear RNA with a linear RNA probe that specifically binds to the tag under a condition that allows the probe to specifically bind to and form a complex with the linear RNA; and

iii) removing the complex formed by the linear RNA probe with the linear RNA from the mixture,

iv) collecting the circular RNA-containing mixture obtained in step iii),

optionally, steps ii)-iv) are performed multiple times, e.g., 2 times, 3 times, 4 times or more.

299. The method of item 298, wherein the tag comprises a polyA, polyG, polyU, or polyC sequence.

300. The method of item 298 or 299, wherein the tag is about 10-200nt in length or the probe is about 10-200nt in length.

301. The method of any one of items 298-300, wherein the tag is added to the linear RNA by adding a PolyA/T/C/G polymerase or a ligase to the mixture.

302. The method of any one of items 298-301, wherein the linear RNA probe specifically binds to the added tag without substantially binding to the circular RNA.

303. The method of any one of items 298-302, wherein the linear RNA probe is a single-stranded DNA probe, or a single-stranded RNA probe.

304. The method of any one of items 298-303, wherein the linear RNA probe is immobilized on a support such as a solid support, for example, the linear RNA probe is immobilized on the support after binding to the linear RNA, or, the linear RNA probe is pre-immobilized on the support.

305. The method of any one of items 242-304, wherein the circular RNA comprises a nucleotide sequence of any one of SEQ ID NO: 543-558, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 543-558.

306. A circular RNA, which is prepared from the circular RNA precursor of any one of items 1-238, prepared by the method of any one of items 242-267, purified by the method of any one of items 268-305 or generated by the nucleic acid vector of any one of items 239-241.

307. The circular RNA of item 306, which comprises a nucleotide sequence of any one of SEQ ID NO: 543-558, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least

87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 543-558.

308. Use of the circular RNA precursor of any one of items 1-238 and/or the circular RNA of any one of items 306-307 as an expression vector.

309. A pharmaceutical composition comprising the circular RNA precursor of any one of items 1-238 and/or the nucleic acid vector of any one of items 239-241 and/or the circular RNA of any one of items 306-307, and a pharmaceutically acceptable carrier.

310. A product expressed by the circular RNA of any one of items 306-307.

311. The product of item 310, which comprises IL-2, POLR2A, or an aquaporin, such as AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

312. A method for preventing or treating a disease or condition in a subject, comprising administering to a subject the circular RNA of any one of items 306-307, or the pharmaceutical composition of item 309.

313. The method of item 312, wherein the disease or condition comprises but not limited to xerostomia, xerophthalmia, Sjogren's syndrome or the like.

314. Use of the circular RNA of any one of items 306-307 or the pharmaceutical composition of item 309 in the manufacture of a medicament for preventing or treating a disease or condition in a subject.

315. The use of item 314, wherein the disease or condition comprises but not limited to xerostomia, xerophthalmia, Sjogren's syndrome or the like.

316. A method of screening a target open reading frame (ORF) sequence encoding a target polypeptide for a nucleotide motif, comprising steps of:

1) extracting amino acid units in the direction of N terminal to C terminal in the target polypeptide containing m amino acids, and each group of amino acid units includes an amino acid with the number of n; with a step of 1 amino acid, slide along the N terminal to the C terminal, n is an integer and $n \geq 2$, m is an integer and ≥ 1 ;

2) listing all nucleotide sequences corresponding to each group of amino acid units according to their degenerate codon; and

3) identifying nucleotide sequences obtained from step 2) that matches the nucleotide motif; wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU, and

4) optionally, selecting the target ORF as the circularizing sequence of a circular RNA

precursor.

317. A system for screening a target open reading frame (ORF) for a nucleotide motif, wherein the target open reading frame (ORF) is a nucleotide sequence encoding a target polypeptide, and the system comprises:

a) an amino acid extraction module, which is used for extracting amino acid units in the direction of N terminal to C terminal in the target polypeptide containing m amino acids, wherein each group of amino acid units comprise an amino acid with a number of n, with a step of 1 amino acid, sliding along the N terminal to the C terminal, and n is an integer and $n \geq 2$, m is an integer and ≥ 1 ;

b) a conversion module, which is used for determining the set of codon sequences of each amino acid unit, each set of codon sequences includes all degenerate codon sequence corresponding to each set of amino acid units; and

c) a target codon sequence filtering module, which is used for determining whether the target codon sequence contain a nucleotide sequence that exactly matches the motif;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

318. A method for screening a target internal ribosome entry site (IRES) for a nucleotide motif, comprising steps of:

1) identifying regions of non-conservative sequences of the target IRES by comparing the target IRES with the other IRES sequences from the same viral subgenus, through multiple sequence alignment tools, for example: Clustal Omega, or MUSCLE;

2) based on the results of multiple sequence alignment, determining a low-structure region of the target IRES through secondary structure prediction software, such as RNAfold, Mfold, RNAstructure or RNAalifold; and

3) editing 0, 1, or 2 bases in the low-structure region in the target IRES to match the nucleotide motif without affecting the translation initiation activity of said IRES;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

319. A system for screening a target internal ribosome entry site (IRES) for a nucleotide motif, wherein the system comprises:

a) a multi-sequence alignment module, which is used for comparing the target IRES with the other IRES sequences from the same viral subgenus to identify regions of non-conserved sequence;

b) a low-structure prediction module, which is used for determining a low-structure region within the target IRES according to the result of the multiple sequence alignments of the target IRES; and

c) a base editing module, which is used for editing the bases at any one or more positions in the region to obtain an edited region matching the nucleotide motif,

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

320. A method for screening a target non-TIE functional element for a nucleotide motif, comprising a step of:

a) editing 0, 1, or 2 bases on the non-TIE functional element so that the non-TIE functional element sequence (such as a 3'UTR or replicon) comprises an edited region that matches the nucleotide motif whilst maintaining the non-TIE function;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

321. A system for screening a target region sequence in a non-TIE functional element sequence (such as a 3'UTR or replicon) for a nucleotide motif, wherein the system comprises:

a) a base editing module, which is used for editing the bases at any one or more positions in the non-TIE functional element sequence to obtain an edited region matching the nucleotide motif;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

322. A method for screening a target non-coding sequence for a nucleotide motif, comprising a step of:

1) editing 0, 1, or 2 bases on the non-coding sequence so that the non-coding sequence comprises an edited region that matches the nucleotide motif whilst maintaining the non-coding sequence's function;

wherein the nucleotide motif consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU; and

2) optionally, selecting the target non-coding sequence as the circularizing sequence of a circular RNA precursor, wherein the target non-coding sequence is an endogenous circRNA.

323. A system for screening a target non-coding sequence for a nucleotide motif, wherein the target non-coding sequence is an endogenous circRNA, and the system comprises:

a) a base editing module, which is used for editing the bases at any one or more positions in the target non-coding sequence to obtain an edited region matching the nucleotide motif,

wherein the nucleotide motif consists of any one of SEQ ID NO: 95-116, preferably CUU, CUC or CAU.

Examples

The following examples have been included to provide guidance to one of ordinary skill in

the art for practicing representative embodiments of the presently disclosed subject matter. In light of the present disclosure and the general level of skill in the art, those of skill can appreciate that the following examples are intended to be exemplary only and that numerous changes, modifications, and alterations can be employed without departing from the scope of the presently disclosed subject matter. The synthetic descriptions and specific examples that follow are only intended for the purposes of illustration, and are not to be construed as limiting in any manner the present invention.

Example 1. The permutation intron exon (PIE) design of Group I Intron

Permuted forms of group I introns in circular RNA precursors, which are interrupted by exon sequence, would fold into an active structure capable of splicing via the normal pathway. If the permuted intron-exon (PIE) sequence is active, the prediction is that a circular form of the ligated exon would be generated. Nonetheless, only a few permuted Group I Introns are used in the art that enables the self-splicing activities of Introns in the PIE system. Whether or not the Group I Introns will maintain the self-splicing activities after permuted at any site within the intron is unknown. Group I introns also possess at their core a conserved architecture formed by 10 paired segments (P1–P10) organized in three domains: a structural domain (P4–P6) that folds independently in the intron, a substrate domain (P1, P10), and a catalytic domain (P3, P7, and P8). These secondary structures create even more uncertainty for the selection of permutation sites within the self-splicing introns used for the PIE system. Which particular site is suitable as a site for PIE design is unknown.

We designed a variety of permutation sites at loop area of P2, P4-5, P5, P6, P8, and P9 of *Azoarcus* Group I Intron (Fig. 1), at loop area of P2, P5, P6, P8, and P9 of *Scytalidium dimidiatum* Group I Intron (Fig. 3), and at loop area of P2, P4-5, P5, P6, P7, P8, P9, P9.1 of *Staphylococcus* phage Twort Group I Intron (Fig. 5), and P2, P4-5, P6 and P9 of *Scytonema hofmannii* Group I intron (Fig. 7).

Example 2. Circularization efficiency of *Azoarcus* Group I Intron with different permutation sites

The basic principle is to connect exons E1- E2 end to end by molecular cloning to generate a continuous circular plasmid. The introns are cleaved by restrictive endonucleases to obtain linearized plasmids. Then, *in vitro* transcription is performed through the promoter upstream of the inverted 3' intron to obtain a linear RNA containing the structure of 3' intron-E2-E1-5' intron. The conserved sequence of specific cleavage site of exon E1 is cleaved by the nucleophilic attack of the free 3' hydroxyl of guanosine, exon E1 produces an exposed 3' hydroxyl, and guanosine

binds to the cleaved 5' intron. Subsequently, the exposed 3' hydroxyl of exon E1 attacks the conserved sequence between the 3' intron and exon E2, the 3' intron is excised, and exon E2 and E1 get connected. The reaction results in circular E1-E2 RNA (Fig. 14).

Taking circPOLR2A as the example, the homology arms, 3' permuted intron-exon (PIE) construct of *Azoarcus* sp Group I Intron (3'intron and Exon2), circPOLR2A, 5' permuted intron-exon (PIE) construct of *Azoarcus* sp Group I Intron (Exon1 and 5'intron) and homology arms were cloned into the plasmid containing a T7 RNA polymerase promoter by Gibson assembly. The circPOLR2A RNAs were produced by *in vitro* transcription (IVT) from the linearized plasmid template using the T7 RNA polymerase according to manufacturer's instruction, and then an additional step of incubation with 30mM Mg²⁺ at 55°C for 15min. Circularization efficiency of constructs was measured by native agarose gel. Circularization efficiency of the 6 main permutation sites (P2, P4-5, P5, P6, P8, P9) of *Azoarcus* Group I Intron (Azop) are shown in Fig.2 and Table 1. The permutation sites are varied from hairpin loop to internal loop. The permutation site of P4-5 and P9 domain which are located in internal loop showed the lower circularization efficiency. On the contrary, when the permutation site in the hairpin loop, such as P5, P6 and P8 domain, the circularization efficiency is relatively high. Although permutation site in P2 domain is a hairpin loop, it involved in forming the catalytic core of the Group I Intron, so disrupting the integrity of the P2 domain affect the splicing efficiency. These results indicate that the permutation site of *Azoarcus* Group I Intron (Azop) is better selected in hairpin loop than in functional domain or internal loop. Circularization efficiency is defined as the grayscale value for each of circRNA/precursor RNA or circRNA/(circRNA+precursor RNA). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 1

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
P2	Functional domain	43%
P4-5	Internal loop	14%
P5	Hairpin loop	52%
P6	Hairpin loop	56%
P8	Hairpin loop	52%
P9	Internal loop	42%

Example 3. Circularization efficiency of *S. dimidiatum* Group I Intron with different permutation sites

Taking circPOLR2A as the example, the homology arms, 3' permuted intron-exon (PIE) construct of Scytalidium dimidiatum Group I Intron (3'intron and Exon2), circPOLR2A, 5' permuted intron-exon (PIE) construct of Scytalidium dimidiatum Group I Intron (Exon1 and 5'intron) and homology arms were cloned into the plasmid containing a T7 RNA polymerase promoter by Gibson assembly. The *circPOLR2A* RNAs were produced by *in vitro* transcription (IVT) from the linearized plasmid template using the T7 RNA polymerase according to manufacturer's instruction, and then an additional step of incubation with 30mM Mg²⁺ at 55°C for 15min. Circularization efficiency of constructs was measured by native agarose gel. Circularization efficiency of the 5 main permutation sites (P2, P5, P6, P8, P9) are shown in Fig.4 and Table 2. The results show that the RNA circularization is inhibited when the permutation site is in function domain, such as P2 domain. On the contrary, when the permutation site is in the hairpin loop, such as P5, P6, P8 and P9 domain, the circularization efficiency is relatively high. It indicates that the permutation site of Scytalidium dimidiatum Group I Intron is better selected in hairpin loop than in functional domain. Circularization efficiency is defined as the grayscale value for each of circRNA/precursor RNA or circRNA/(circRNA+precursor RNA). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 2

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
P2	Functional domain	10%
P5	Hairpin loop	29%
P6	Hairpin loop	30%
P8	Hairpin loop	28%
P9	Hairpin loop	28%

Example 4. Circularization efficiency of Staphylococcus phage Twort Group I Intron with different permutation sites

Taking circPOLR2A as the example, the homology arms, 3' permuted intron-exon (PIE) construct of Staphylococcus phage Twort Group I Intron (3'intron and Exon2), circPOLR2A, 5' permuted intron-exon (PIE) construct of Staphylococcus phage Twort Group I Intron (Exon1 and 5'intron) and homology arms were cloned into the plasmid containing a T7 RNA polymerase promoter by Gibson assembly. The *circPOLR2A* RNAs were produced by *in vitro* transcription (IVT) from the linearized plasmid template using the T7 RNA polymerase according to manufacturer's instruction, and then an additional step of incubation with 30mM Mg²⁺ at 55°C for

15min. Circularization efficiency of constructs was measured by native agarose gel. Circularization efficiency of the 8 main permutation sites (P2, P4-5, P5, P6, P7, P8, P9, P9.1) are shown in Fig.6 and Table 3. The permutation sites are varied from hairpin loop to internal loop. The results show that the RNA circularization is inhibited when the permutation site is in internal loop, such as P4-5 domain. On the contrary, when the permutation site is in the hairpin loop, such as P2, P5, P6, P7, P8, P9 and P9.1 domain, the circularization efficiency is relatively high. It is indicated that the permutation site of Staphylococcus phage Twort Group I Intron is better selected in hairpin loop than in internal loop. And there's no significant difference whether the PIE site is symmetrical or asymmetrical from "stem". Circularization efficiency is defined as the grayscale value for each of circRNA/precursor RNA or circRNA/(circRNA+precursor RNA). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 3

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
P2	Hairpin loop	52%
P4-5	Internal loop	18%
P5	Hairpin loop	61%
P6	Hairpin loop	71%
P7	Hairpin loop	68%
P8	Hairpin loop	67%
P9	Hairpin loop	66%
P9.1	Hairpin loop	70%

Example 5. Circularization efficiency of *Scytonema-hofmani* tRNA fMet Group I Intron with different permutation sites

Taking circPOLR2A as the example, the homology arms, 3' permuted intron-exon (PIE) construct of *Scytonema hofmanii* tRNA fMet Group I Intron (3'intron and Exon2), circPOLR2A, 5' permuted intron-exon (PIE) construct of *Scytonema hofmanii* tRNA fMet Group I Intron (Exon1 and 5'intron) and homology arms were cloned into the plasmid containing a T7 RNA polymerase promoter by Gibson assembly. The *circPOLR2A* RNAs were produced by *in vitro* transcription (IVT) from the linearized plasmid template using the T7 RNA polymerase according to manufacturer's instruction, and then an additional step of incubation with 30mM Mg²⁺ at 55°C for 15min. Circularization efficiency of constructs was measured by native agarose gel.

Circularization efficiency of the 4 main permutation sites (P2, P4-5, P6 and P9) of *Scytonema hofmanii* tRNA fMet Group I Intron (Sh) are shown in Fig.8 and Table 4. The permutation sites are varied from hairpin loop to internal loop. The results show that the RNA circularization is inhibited when the permutation site is in functional domain or internal loop, such as P2 and P4-5 domain. On the contrary, when the permutation site is in the hairpin loop, such as P6 and P9 domain, the circularization efficiency is relatively high. It is indicated that the permutation site of *Scytonema hofmanii* tRNA fMet Group I Intron (Sh) is better selected in hairpin loop than in functional domain or internal loop. Circularization efficiency is defined as the grayscale value for each of circRNA/precursor RNA or circRNA/(circRNA+precursor RNA). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 4

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
P2	Functional	27.16%
P4-5	Internal loop	16.71%
P6	Hairpin loop	58.99%
P9	Hairpin loop	49.16%

Example 6. Comparison of the circularization efficiency between Staphylococcus phage Twort Group I Intron and Anabaena Group I Intron

In order to compare the efficiency of Azoarcus Group I Intron, Twort Group I Intron, Sh Group I Intron and Anabaena Group I Intron for formulation of circRNA, we constructed Azoarcus, Twort, Sh and Anabaena plasmids containing T7 promoter, 5' homology arm, 3' permuted intron-exon (PIE) construct of Group I Intron (3'intron and Exon2), a *Salivirus FHB* IRES with *IL2* sequence or a CVB2 IRES with *Luciferase* sequence, 5'permuted intron-exon (PIE) construct of Group I Intron (Exon1 and 5'intron) and 3' homology arm to generate the *circIL2* or *circLuciferase* using the group I intron circularization method that introduced above and the circularization efficiency was visualized on agarose gel (Fig. 9, Table 5 and Table6). The selection of PIE sites of Azoarcus Group I Intron, Twort Group I Intron and Sh Group I Intron (P6 and P8 from Azoarcus intron, P6 and P9-1 from Twort intron, P6 from Sh intron) were based on the cleanest and the highest circularization efficiency PIE sites for circ*POLR2A* cargo. Circularization efficiency is defined as the grayscale value for each of circRNA/precursor RNA.

The result showed that compared with Anabaena Group I Intron (AnaX as set forth in SEQ ID NO:1), the circularization efficiency of Twort Group I Intron (SEQ ID NO:3) and Sh Group I

Intron (SEQ ID NO:5) is significantly improved in the IL-2 (Fig.9A) or Luciferase cargo (Fig.9B), and Azoarcus Group I Intron (SEQ ID NO:2) has slight increase. RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 5

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
AnaX		18%
Azop-P6	Hairpin loop	26%
Azop-P8	Hairpin loop	22%
Twort-P6	Hairpin loop	41%
Twort-P9.1	Hairpin loop	36%

Table 6

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
AnaX		11%
Sh-P6	Hairpin loop	35%

Example 7. Circularization efficiency of Agrobacterium-tumefaciens Group I Intron

We designed permutation sites at loop area of P6 domain of Agrobacterium-tumefaciens Group I Intron to test the circularization efficiency of the novel group I introns. *circPOLR2A* was generated *in vitro* by Agrobacterium-tumefaciens Group I Intron in a self-splicing reaction buffer (50 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.5). The circularization efficiency was visualized on agarose gel (Fig. 10, Table 7). The results show that Agrobacterium-tumefaciens Group I Intron can achieve circRNA formulation. The circularization efficiency is defined as the grayscale value for each of circRNA/precursor RNA or circRNA/(circRNA+precursor RNA). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 7

Code	Predicted Location	Circularization efficiency circRNA/(circRNA+precursor RNA)
P6	Hairpin loop	42%

Example 8. The effect of sequence characteristic of homology arm outside the self-splicing intron fragment on RNA circularization

In order to improve the RNA circularization efficiency of group I intron, we tested the homology arm sequence outside the intron fragment of the AnaX system, including 1) using the endogenous non-complementary sequence (300 bp) from *POLR2A* intron 1 as 5'homology arm and fully complementary sequence (300 bp) of *POLR2A* intron 1 as 3'homology arm; 2) truncating the endogenous non-complementary sequence (300 bp) from *POLR2A* intron 1 to 100 bp; 3) using the natural *AluSz* sequence (285nt) as 5'homology arm and *AluYa5* sequence (174nt) as 3'homology arm, and incomplete complementary pairing of partial *AluSz* and *AluYa5* sequences; 4) modifying the natural *AluYa5* (147nt) to generate an artificial fully complementary *Alu* pair as homology arm; 5) truncating the endogenous non-complementary sequence (300 bp) from *POLR2A* intron 1 to 50 bp which is strong complementary; 6) truncating the endogenous non-complementary sequence (300 bp) from *POLR2A* intron 1 to 50 bp which is weak complementary (Fig. 11).

Next, we examined the effect of the retrofitted homology arm sequence outside the intron fragment on the circularization efficiency of the AnaX system. Taking the circularization of *CVB3-IL2* RNAs as an example, the circularization efficiency was detected by denature PAGE gel. The results show that increasing the length of the complementary base-pairing sequence of the homology arm can improve the RNA circularization efficiency until the circularization efficiency reaches a level close to 100%. In addition, when the homology arm sequence was changed to highly complementary state, RNA circularization efficiency had obvious better effect. On the contrary, the non-complementary sequence reduced the RNA circularization efficiency (Fig. 12). The circularization efficiency is defined as the grayscale value for each of circRNA/(precursor RNA+circRNA) (Fig. 12). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs. Then, we analyzed the average complementary base-pairing strength of each base on the homology arm, and the results show that the lower the minimum free energy of the base, the stronger the pairing strength, and the higher efficiency of circRNA formation (Fig. 13).

Example 9. The Clean PIE design for AnaX Group I Intron

For AnaX group I intron, the 3 nucleotides motif (CUU, CUC or CAU) is the ribozyme recognition site I that pairs with the internal guide sequence (IGS) of the 5' intron to form the P1 double-stranded region of the ribozyme, which can be used for the determination of the 5' splicing site. We designed three types of precursors in which the ribozyme recognition site I is in IRES, cargo and functional element respectively (Fig.14 and 15).

1. Design Clean PIE in IRES

Taking CVB3 as an example, we searched and identified the motif of CUU or CUC or CAU in the corresponding low structure region of CVB3 through RNAalifold, and performed minimal mutations on the flanking nucleotides of the motif to obtain the minimum base pairing numbers required for circularization. The precursor sequence was designed as follows from 5' to 3' direction (SEQ ID NO:528, 529, 535 and 536):

3' self-splicing intron fragment, the sequence from downstream of motif CUU/CUC/CAU (motif not inclusive) to the 3' end of CVB3 (SEQ ID NO:554,556), cargo sequence, the sequence from the 5' end of CVB3 to motif CUU/CUC/CAU (SEQ ID NO:555), and 5' self-splicing intron fragment.

2. Design Clean PIE in cargo

We searched and identified motif CUU or CUC or CAU through degenerate codon mutation method in the cargo (ORF). The precursor sequence was designed as follows from 5' to 3' direction:

3' self-splicing intron fragment, the sequence from downstream of motif CUU/CUC/CAU (motif not inclusive) to the 3' end of the ORF (SEQ ID NO:557, 559, 561), IRES sequence, the sequence from the 5' end of the ORF to motif CUU/CUC/CAU (SEQ ID NO:558, 560, 562), and 5' self-splicing intron fragment. A specific precursor was designed and the ORF of IL2 or AQP1 is used, wherein the corresponding sequences are SEQ ID NO. 530, 531 and 533 as shown in Fig. 16 and 18.

3. Design Clean PIE in functional element

Taking 3'UTR as an example, we searched and identified CUU or CUC or CAU on the 3'UTR sequence. The precursor sequence was designed as follows from 5' to 3' direction:

3' self-splicing intron fragment, the sequence from downstream of motif CUU/CUC/CAU (motif not inclusive) to the 3' end of the UTR, IRES sequence, cargo sequence, the sequence from the 5' end of the UTR to motif CUU/CUC/CAU, and 5' self-splicing intron fragment.

Example 10. The RNA circularization efficiency of CLEAN-PIE

Precursors (SEQ ID NO. 528-531) for circular *CVB3-IL2* RNA were designed according to the CLEAN-PIE strategy using the AnaX Group I Intron and were prepared by *in vitro* transcription from the linearized plasmid template using the T7 RNA polymerase according to manufacturer's instruction. The circularization efficiency was visualized on agarose gel and circular *CVB3-IL2* prepared from traditional PIE method as the positive control (Fig. 17, Table 8). The results show that CLEAN-PIE designed in IRES or Cargo can both achieve circular *CVB3-IL2* RNA, and the circularization efficiency was determined for different CLEAN-PIE designs. The circular RNAs were identified by Sanger Sequencing. Circularization efficiency is defined as

the grayscale value for each of circRNA/(precursor RNA+circRNA). RNase R was used as a tool enzyme for verifying the circular structure of circular RNAs, and it can effectively degrade linear RNAs with free ends, but not circular RNAs.

Table 8

SEQ ID NO.	CLEAN-PIE site	Circularization efficiency
527	NA	60%
528	IRES: CVB3	40%
529	IRES: CVB3	45%
530	Cargo: IL2	59%
531	Cargo: IL2	35%

Example 11. Translation efficiency and immunogenicity of circRNA obtained in CLEAN-PIE

To investigate the translation efficiency and immunogenicity of circRNA obtained from the CLEAN PIE, CLEAN-PIE were designed in CVB3 IRES or AQP1 cargo.

For CLEAN-PIE-IRES, we searched and identified CUU or CUC on CVB3. The precursor sequence was designed as follows from 5' to 3' direction (Fig. 18, SEQ ID NO.535 and 536):

3' self-splicing intron fragment, the sequence from downstream of motif CUU (motif not inclusive) to the 3' end of CVB3 (SEQ ID NO:554,556), cargo sequence, the sequence from the 5' end of CVB3 to motif CUU (SEQ ID NO:555), and 5' self-splicing intron fragment.

For CLEAN-PIE-Cargo, we first searched and identified CUU or CUC through degenerate codon mutation method on AQP1. The precursor sequence was designed as follows from 5' to 3' direction (Fig. 18, SEQ ID NO.533):

3' self-splicing intron fragment, the sequence from downstream of motif CUU (motif not inclusive) to the 3' end of the ORF (SEQ ID NO:561), IRES sequence, the sequence from the 5' end of the ORF to motif CUU (SEQ ID NO:562), and 5' self-splicing intron fragment.

Next, the CLEAN-PIE-circAQP1 was designed by following the above described strategies based on AnaX Group I intron (SEQ ID NO:1) and the precursors thereof were generated by in vitro transcription using the T7 RNA polymerase according to manufacturer's instruction. *circAQP1* prepared from non-CLEAN PIE method was used as the positive control. Results show that the CLEAN-PIE designed in IRES or Cargo can achieve AQP1 RNA circularization with different efficiency (Fig. 19A). Then, we purified *circAQP1* in vitro by RNase R treatment to

achieve high purity, and transfected *circAQP1* (200ng/well) into human A253 cells in 12-well plate using Lipofectamine MessengerMAX™ (Thermo Fisher) RNA transfection reagent. After 24 hours and 72 hours of transfection, the cells were harvested, and the expression level of the AQP1 was determined using Western blot. HSP90 was used as the internal reference. The results show that *circAQP1* produced from CLEAN-PIE had higher translation efficiency (Fig. 19B).

Further, we transfected the CLEAN-PIE-*circAQP1* (200 ng/well) into human A549 cells in 12-well plate using Lipofectamine MessengerMAX™ (Thermo Fisher) RNA transfection reagent to detect the immunogenicity. After 6 hours of transfection, the cells were harvested, and the expression levels of the cytokines, including IFN β , TNF α , IL6 and RIG-I, were determined using quantitative RT-PCR and normalized to 18S rRNA. The results show that compared to *circAQP1* prepared from traditional PIE method, the purified *circAQP1* prepared via CLEAN-PIE circularization methods showed no significant differences in the expression level of IFN β , TNF α , IL6 and RIG-I (Fig. 19C).

These data demonstrated that the *circAQP1* prepared *in vitro* using CLEAN-PIE circularization methods do not cause significant immune responses. IFN β , TNF α , IL6 and RIG-I are commonly used as marker genes for the measurement of cellular immunogenicity, and the positive control Poly (I:C) is a commonly used long double-stranded RNA to mimic virus infection at circular level. mRNA is used as a reference sequence (SEQ ID NO:624).

Example 12. *In vivo* expression of circRNA-hAQP1 injection

To evaluate the *in vivo* expression of AQP1, female wild Sprague Dawley rats of an age of 6 weeks were administrated with 100 μ L of PBS (Sangon Biotech) or LNP encapsulated CLEAN-PIE-*circAQP1* prepared in Example 11 at 0.5 μ g, 1.5 μ g, or 5 μ g into the left submandibular salivary gland via retrograde ductal infusion. On D3 and D16 post-treatment, in each group, both administrated and non-administrated submandibular glands were collected from three animals for immunohistochemistry (IHC). A specific antibody against hAQP1 (Abcam) was used for IHC staining.

As shown in Figures 22 and 23, higher signals were detected in the groups of higher doses (1.5 and 5 μ g) as compared to the 0.5 μ g dose group. On day 16, hAQP1 expression was still detectable in the animals injected with 5 μ g in contrast to the animals treated with lower dose levels.

Retrograde ductal infusion with LNP encapsulated CLEAN-PIE-*circAQP1* increases in situ expression level of hAQP1 in the administrated submandibular glands rather than the non-administrated one. The semi-quantitative analysis revealed a trend for dose responses with higher expression at higher doses. The duration of in situ expression in administrated submandibular

glands lasted from day 3 up to day 16.

Example 13. The CLEAN-PIE design for Twort Group I Intron

We designed CLEAN-PIE precursor using Twort Group I introns. Taking *circCVB3-AQP1* as an example, we searched the CAU through degenerate codon mutation method. The precursor sequence was designed as follows from 5' to 3' direction (Fig. 18, SEQ ID NO.534):

3' self-splicing intron fragment, the sequence from downstream of motif CAU (motif not inclusive) to the 3' end of *AQP1* (SEQ ID NO:563), IRES sequence, the sequence from the 5' end of the ORF of *AQP1* to motif CAU (SEQ ID NO:564), and 5' self-splicing intron fragment.

Next, *circCVB3-AQP1* was generated by in vitro transcription using the T7 RNA polymerase according to manufacturer's instruction and then incubation with 30mM Mg^{2+} at 55°C for 15min for self-splicing circularization. Results show that the CLEAN-PIE designed in cargo can achieve RNA circularization (Fig. 19A). Then, we purified the *circCVB3-AQP1* in vitro by RNase R treatment to achieve high purity, and transfected *circCVB3-AQP1* (200ng/well) into human A253 cells in 12-well plate using Lipofectamine MessengerMAX™ (Thermo Fisher) RNA transfection reagent. After 24 hours and 72 hours of transfection, the cells were harvested, and the expression levels of the AQP1 were determined using western blot and HSP90 was used as the internal reference. The results showed that *circCVB3-AQP1* produced from CLEAN-PIE has higher translation efficiency (Fig. 19B). It indicates that Twort group I intron can be used in CLEAN PIE designs to achieve RNA circularization and circRNA translation.

Further, we transfected the CLEAN-PIE *circCVB3-AQP1* (200 ng/well) into human A549 cells in 12-well plate using Lipofectamine MessengerMAX™ (Thermo Fisher) RNA transfection reagent to detect the immunogenicity. After 6 hours of transfection, the cells were harvested, and the expression levels of the cytokines, including IFN β , TNF α , IL6 and RIG-I, were determined using quantitative RT-PCR and normalized to 18S rRNA. The results show that compared to *circCVB3-AQP1* prepared from traditional PIE method, the purified *circCVB3-AQP1* prepared via CLEAN-PIE circularization methods show no significant differences in the expression level of IFN β , TNF α , IL6 and RIG-I (Fig. 19C). These data demonstrate that the *circCVB3-AQP1* generated by CLEAN-PIE design of Twort group I intron does not cause significant immune responses. IFN β , TNF α , IL6 and RIG-I are commonly used as marker genes for the measurement of cellular immunogenicity, and the positive control Poly (I:C) is a commonly used long double-stranded RNA to mimic virus infection at circular level.

Example 14. The regulation of circRNA expression in cells by miRNA

MicroRNA (miRNA) are small non-coding RNA molecules containing 20 to 23 nucleotides.

The main function of miRNAs is to induce mRNA degradation and translational repression by forming complementary base-pairing with the 3' untranslated region (3'UTR) of target mRNAs. In this example, we demonstrate the regulation of circRNA expression in cells by introducing miRNA binding sites into circRNAs.

We inserted 1 (1x circRNA) or 2 (2x circRNA) tandem miRNA binding sites targeting miRNA-122 or miR-671 respectively into the circRNA constructs that contain the IRES element and luciferase coding sequence. Then circRNAs containing miRNA binding sites were transfected into HEK293 cells overexpressing miRNA-122 or miR-671. The protein expression level and RNA level of each circRNA were detected after transfection 6h, 24h and 48h, respectively.

The results show that the luciferase expression levels are significantly inhibited when miRNA-122 or miR-671 binding sites were inserted into circRNAs (Fig. 20). In addition, the amount of circRNAs containing miRNA-122 or miR-671 binding sites was also significantly decreased compared to the control group (Fig. 20). The above results demonstrate that miRNA can regulate circRNA expression by degrading circular RNA, and thus realize the cell type- or tissue-specific expression of circular RNA.

Example 15. Regulation of circRNA translation by UTRs

Untranslated regions (UTR) flanking the ORF can bind with regulatory proteins and thus affect the protein translational efficiency. In this example, we describe the regulation of circRNA translation by introducing UTRs into the circular RNA. We inserted the UTR element into the 3' end of RBD ORF sequence in the circular RNA and transfected the circular RNA into the THP1 cells, then detected the expression level after transfection 24h and 48h by western blot, respectively. The results demonstrate that UTRs of different sequences may have different effects. UTR-1 enhanced the translation of circRNA, while UTR-2 and UTR-3 showed inhibitory effects (Fig. 21).

While the invention is described in conjunction with the enumerated embodiments, it will be understood that they are not intended to limit the invention to those embodiments. The invention is intended to cover all alternatives, modifications, and equivalents that may be included within the scope of the present invention. One skilled in the art will recognize many methods and materials similar or equivalent to those described herein, which could be used in the practice of the present invention. The present invention is in no way limited to the methods and materials described. In the event that one or more of the incorporated literature, patents, and similar materials differs from or contradicts this application, including but not limited to defined terms, term usage, described techniques, or the like, this application controls.

All references including patents, patent applications and publications cited in the present ap

plication are incorporated herein by reference in their entirety, as if each of them is individually i
ncorporated. Further, it would be appreciated that one skilled in the art could make various chang
es or modifications to the invention without departing from the scope of the invention defined by
the appended claims below. Accordingly, the present invention is not intended to be limited to th
e disclosed embodiments. Rather the present invention is intended to cover the disclosed embodi
ments as well as others falling within the scope and spirit of the invention to the fullest extent per
mitted in view of this disclosure and the inventions defined by the claims appended herein below.

Sequences mentioned herein are listed as below:

S E Q U E N C E S	Sequence	Description
1	AGACGCUACGGACUUAUAAUUGAGCCUUAAGAAGAAUUCUUUAAGUGGAGUCUCUCAAACUCA GGGAAACCUAAUUCUAGUUUAUAGACAAGGCAUCCUGAGCCAAGCCGAGUAGUAAUAGUAAAAAC AAUAGAGUAGACUUACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUACGUCAGAGCA GGGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAAUC CGUUGA	GROUP I INTRON, Anabaena sp. YBS01
2	GAGCAGCGGACUCAUAAUUCGAUGUGCCUUGCGCCGGGAAACCACGCAAGGGAUGGUGUCAAUUCG GCGAAACCUAAGCGCCCCCGGGCGUAUGGCAACGCCGAGCCAAAGCUUCGGCGCCUGCGCCGAUGA AGGUGUAGAGACUAGACGGCACCCACCUAAGGCAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGG CGAAAGUCACACAAACCGGAAUCCGUUGGUGCUG	GROUP I INTRON, Azoarcus olearius BH72
3	ACUACUGAAAGCAUAAUAAUUGUGCCUUUAUACAGUAAUAGUAUACGAAAAUCCUCUAAUUCAGG GAACACCUAAACAAACUAGAUGUAGGCAUCCUGAGCUAAGCUCUUAAGUAAUAGAGAAAGUGCAA CGACUAAUCCGAUAGGAAGUAGGGUACAGUGACUGAAUUGGGGAUUACCCUUCUAGGGUAGUGAUA UAGUCUGAACAUAAUUGGAAACAUAAUAGAGGAUAGGAGUAAACCAUUAUUCGUAAACUAAUUGAA CUUUUAGUUUAU	GROUP I INTRON, Staphylococ us phage Twort
4	AACAGGUUGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUUCCGUGUUUUUUUUUUCUCCGC UAGUCGAUUAUCCGUGACGGCCGGGCAAGCGCCCGAGUACCAGGACCGUCGGGCCCCCGAGAAGGGG GGGCCUAGGAUGCGGCAAGACACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGGC AGGUGCCGUAGCGGGCACCCGUCGUAAACGCGCGGUAGGGCGUCGGUCCCCCUCACCGCGGAGGG GAGGCUAAAGGUACGUGCUAAACCCCAAGCGAUGGGGCCUGUAGGAAAAGCCUGGUACGGCGAAGC CUACGGGACCGCCGAUGGGCGGUCGGAUGCCGGCGGGCCACGGAGCCCCCGGCGUCGAUUGCUGUU A	GROUP I INTRON, Scytalidium dimidiatum
5	GGGCUCAACGAGUAAGAUGACUAAACGCUUAAAAGUUAUUCUUAUUAUGGGCGGUACGUAAAGAAA CUUACGUUAGUUUACCUGUCAAAACUCGGGGAGCCAAUAGCGCGGUAAUCCCGAACCAGCUCCAGA AAUGGAGAAGGUGUAGAGACUGGAAGGCAGGUACCCUAAACAGUAAAGCUGAGGGUAAAGGGACAGUC CAGACCACAAACUGGUUAAUCCAGGCAGUGAAAACUGUAGAUGGUAAAGCAUAAACC	GROUP I INTRON, Scytonema- hofmani tRNA fMet
6	CAGGAUCCCUAAGGUAAAUUGGGGUUGCGCCCGGAAACGACGCAUUCGAUCUGCUAAAGUCGGGG AAAGCUUCGCUUGGUCUCCAGUAUGCCAAUCCCGAGCCAAGCUCGGAGAAAUCCGGUGAAGGUGUA GAGACUGGAUGGGCAGCACCUAAAGGCCGCAAGGUCCACGGUGAAGGGACAGUCCAGACCACGAACG CCCUUAUCCGAUUGAGAUGGGCGGGCGGAAAGCCGAAGCGGUUAUGAAUCCUGGGG	GROUP I INTRON, Agrobacteriu m- tumefaciens
7	AACAAUAGAUGACUUACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAAGA CGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUG	3' self- splicing intron fragment derived from

		Anabaena pre-tRNA-Leu gene GROUP I INTRON
8	AAAUAUUGAGCCUAAAGAAGAAAUUCUUUAAGUGGAUGCUCUCAAACUCAGGGAAACCUAAAUCU AGUUUAUAGACAAGGCAAUCCUGAGCCAAGCCAAGUAGUAAUUAGUAA	5' self- splicing intron fragment derived from Anabaena pre-tRNA-Leu gene GROUP I INTRON
9	AACCACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGCCCGCCGGGCGUAUGGCAACGCCG AGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGACUAGACGGCACCCACCUAAGGCAAACGC UAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACACAAACCGG	3' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-2
1 0	AUUUCGAUGUGCCUUGCGCCGGGA	5' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-2
1 1	ACGCCGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGACUAGACGGCACCCACCUAAGGC AAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACACAAACCGG	3' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE- 4-5
1 2	AUUUCGAUGUGCCUUGCGCCGGGAAACCACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGC CCGCCCGGGCGUAUGGCA	5' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE- 4-5
1 3	CCGGGCGUAUGGCAACGCCGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGACUAGACGG CACCCACCUAAGGCAAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACACAAACCGG	3' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-5
1 4	AUUUCGAUGUGCCUUGCGCCGGGAAACCACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGC CCGC	5' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-5

1 5	UGCGCCGAUGAAGGUGUAGAGACUAGACGGCACCCACCUAAGGCAAACGCUAUGGUGAAGGCAUAGU CCAGGGAGUGGCGAAAGUCACACAAACCGG	3' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-6
1 6	AUUUCGAUGUGCCUUGCGCCGGGAAACCACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGC CCGCCCCGGGCGUAUGGCAACGCCGAGCCAAGCUUCGGCGCC	5' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-6
1 7	ACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACACAAACCGG	3' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-8
1 8	AUUUCGAUGUGCCUUGCGCCGGGAAACCACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGC CCGCCCCGGGCGUAUGGCAACGCCGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGACUAG ACGGCACCCACCUAAGGCAA	5' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-8
1 9	AAACCGG	3' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-9
2 0	AUUUCGAUGUGCCUUGCGCCGGGAAACCACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGC CCGCCCCGGGCGUAUGGCAACGCCGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGACUAG ACGGCACCCACCUAAGGCAAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACAC	5' self- splicing intron fragment derived from Azoarcus olearius BH72 GROUP I INTRON-PIE-9
2 1	AAUGUAUAUCGAAAAUCCUCUAAUUCAGGGAAACACCUAAACAAACUAAGAUGUAGGCAAUCCUGAG CUAAGCUCUUAGUAAUAAGAGAAAGUGCAACGACUAAUUCGGAUAGGAAGUAGGGUCAAGUAGACUGA AAUGGGGAUUACCCUUCUAGGGUAGUGAUUAUGUCUGAACAUUAUGGAAACAUAUAGAAGGAUAGG AGUAACGAACCUAUUCGUAACAUAAUUG	3' self- splicing intron fragment derived from Staphylococ- cus phage Twort GROUP I INTRON- PIE-2
2 2	AAUAAUUGUGCCUUUAUACAGU	5' self- splicing intron fragment derived from

		Staphylococcus phage Twort GROUP I INTRON-PIE-2
2 3	AUCCUGAGCUAAGCUCUUAGUAAUAAGAGAAAGUGCAACGACUAUUCGGAUAGGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUUCUAGGGUAGUGAUUAAGUCUGAACAUUAUGGAAACAUUAAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAUUG	3' self-splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON-PIE-4-5
2 4	AAUAAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAUUCAGGGAACACCUAAACAAACUAAGAUGUAGGCA	5' self-splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON-PIE-4-5
2 5	AAGAUGUAGGCAAUCCUGAGCUAAGCUCUUAGUAAUAAGAGAAAGUGCAACGACUAUUCGGAUAGGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUUCUAGGGUAGUGAUUAAGUCUGAACAUUAUGGAAACAUUAAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAUUG	3' self-splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON-PIE-5
2 6	AAUAAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAUUCAGGGAACACCUAAACAAACU	5' self-splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON-PIE-5
2 7	AAUAAGAGAAAGUGCAACGACUAUUCGGAUAGGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUUCUAGGGUAGUGAUUAAGUCUGAACAUUAUGGAAACAUUAAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAUUG	3' self-splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON-PIE-6
2 8	AAUAAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAUUCAGGGAACACCUAAACAAACUAAGAUGUAGGCAAUCCUGAGCUAAGCUCUUAGU	5' self-splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON-PIE-6
2 9	UAGGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUUCUAGGGUAGUGAUUAAGUCUGAACAUUAUGGAAACAUUAAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAUUG	3' self-splicing intron fragment

		derived from Staphylococcus phage Twort GROUP I INTRON- PIE-7
3 0	AAAUAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAAUUCAGGGAACACCUAAACAA ACUAAGAUGUAGGCAAUCCUGAGCUAAGCUCUUAGUAAUAGAGAAAGUGCAACGACUAUUCGGA	5' self- splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON- PIE-7
3 1	CUAGGGUAGUGAUUAAGUCUGAACAUUAUGGAAACAUAUAGAAGGAUAGGAGUAACGAACCUAUUC GUAACAUAUUG	3' self- splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON- PIE-8
3 2	AAAUAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAAUUCAGGGAACACCUAAACAA ACUAAGAUGUAGGCAAUCCUGAGCUAAGCUCUUAGUAAUAGAGAAAGUGCAACGACUAUUCGGAUA GGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUU	5' self- splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON- PIE-8
3 3	AACAUAUAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAUUG	3' self- splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON- PIE-9
3 4	AAAUAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAAUUCAGGGAACACCUAAACAA ACUAAGAUGUAGGCAAUCCUGAGCUAAGCUCUUAGUAAUAGAGAAAGUGCAACGACUAUUCGGAUA GGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUUCUAGGGUAGUGAUUAAGUCUGAACAUUAU AUGGA	5' self- splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON- PIE-9
3 5	UAACGAACCUAUUCGUAACAUAUUG	3' self- splicing intron fragment derived from Staphylococcus phage Twort GROUP I INTRON- PIE-9.1

3 6	AAAUAAUUGGCCUUUAUACAGUAAUGUAUAUCGAAAAUCCUCUAAUUCAGGGAACACCUAAACAA ACUAAGAUGUAGGCAAUCCUGAGCUAAGCUCUUAAGUAAUAGAGAAAGUGCAACGACUAAUCCGAUA GGAAGUAGGGUCAAGUGACUCGAAAUGGGGAUUACCCUUCUAGGGUAGUGAUUAGUCUGAACAUAU AUGGAACAUAUAGAAGGAUAGGAG	5' self- splicing intron fragment derived from Staphylococ- cus phage Twort GROUP I INTRON- PIE-9.1
3 7	GCGCCCGAGUACCAAGGACCGUCGGGCCCCGAGAAGGGGGGGCCUAGGAUGCGGCAAGACGACCCGG UUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGCGAGGUGCCGUAGCGGGCACCCGUCGUAA GCGCGGUAGGGCGUCGGUCCCCUCCACGCGGAGGGGAGGCUUAAGGUACGUGCUAAACCCCCA GCGAUGGGGCGUGUAGGAAAAGCCUGUACGGGAAGCCUACGGGGACCGCCCGAUGGCGGUCGGGA UGCCGGCGGGCCACGGAGCCCCCGGCGUCGAUUG	3' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-2
3 8	UGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUUCCGUGUUUUUUUUUCCGCUAGUCGA UCAUCCGUGACGGCCCGGCAA	5' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-2
3 9	GCGCUGGCCGAUCCCGAGGCGAGGUGCCGUAGCGGGCACCCGUCGUAAACGCGCGUAGGGCGUCGGU CCCCCUCCACGGCGGAGGGGAGGCUUAAGGUACGUGCUAAACCCCCAGCGAUGGGGCGUGUAGGA AAAGCCUGGUACGGCGAAGCCUACGGGGACCGCCCGAUGGCGGUCGGAUGCCGGCGGGCCACGGAG CCCCGGCGUCGAUUG	3' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-5
4 0	UGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUUCCGUGUUUUUUUUUCCGCUAGUCGA UCAUCCGUGACGGCCGGGCAAGCGCCCGAGUACCAAGGACCGUCGGGCCCCGAGAAGGGGGGGCCUA GGAUGCGGCAAGACGACCCGGUUCGGGGAACGCCACCGU	5' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-5
4 1	GCGGGCACCCGUCGUAAACGCGCGGUAGGGCGUCGGUCCCCUCCACGGCGGAGGGGAGGCUUAAG GUACGUGCUAAACCCCCAGCGAUGGGGCGUGUAGGAAAAGCCUGGUACGGCGAAGCCUACGGGGAC CGCCCGAUGGCGGUCGGAUGCCGGCGGGCCACGGAGCCCCCGGCGUCGAUUG	3' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-6
4 2	UGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUUCCGUGUUUUUUUUUCCGCUAGUCGA UCAUCCGUGACGGCCGGGCAAGCGCCCGAGUACCAAGGACCGUCGGGCCCCGAGAAGGGGGGGCCUA GGAUGCGGCAAGACGACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGCGAGGUGCC GUA	5' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-6
4 3	GGCGGAGGGGAGGCUUAAGGUACGUGCUAAACCCCCAGCGAUGGGGCGUGUAGGAAAAGCCUGGUA CGGCGAAGCCUACGGGGACCGCCCGAUGGCGGUCGGAUGCCGGCGGGCCACGGAGCCCCCGGCGUCG AUUG	3' self- splicing intron fragment derived from

		Scytalidium dimidiatum GROUP I INTRON-PIE-8
4 4	UGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUUCCGUGUUUUUUUUUUUCCGCUAGUCGA UCAUCCGUGACGGCCGGGCAAGCGCCCGAGUACCAGGACCGUCGGGCCCCCGAGAAGGGGGGGCCUA GGAUGCGGCAAGACGACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGCGAGGUGCC GUAGCGGGCACCCGUCGUAACGCGCGGUAGGGCGUCGGUCCCCCUCCAC	5' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-8
4 5	GGAGCCCCCGGCGUCAUUG	3' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-9
4 6	UGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUUCCGUGUUUUUUUUUUUCCGCUAGUCGA UCAUCCGUGACGGCCGGGCAAGCGCCCGAGUACCAGGACCGUCGGGCCCCCGAGAAGGGGGGGCCUA GGAUGCGGCAAGACGACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGCGAGGUGCC GUAGCGGGCACCCGUCGUAACGCGCGGUAGGGCGUCGGUCCCCCUCCACGCGCGGAGGGGAGGCUU AAGGUACGUGCUAAACCCCGAGCGAUGGGGGCUGUAGGAAAAGCCUGGUACGGCGAAGCCUACGGG GACCGCCCGAUGGCGGUCGGAUGCCGGCGGGCCAC	5' self- splicing intron fragment derived from Scytalidium dimidiatum GROUP I INTRON-PIE-9
4 7	AAUGGAGAAGGUGUAGAGACUGGAAGGCAGGUACCCUAAACAGUAAAGCUGAGGGUAAAGGGACAGUC CAGACCACAAACUGGUAAUCCAGGCAGUGAAAACUGUAGAUGGUAAAG	3' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-6
4 8	CAACGAGUAAGAUUGACUAAACGCUUAAAAGUUAUUCUUACUAUGGGCGGUACGUAAAGAAACUUAC GUAUGUUUACCUGUCAAAACUCGGGGAAGCCAAUAGCGCGGUAAUCCCGAACCAAGCUCCAGA	5' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-6
4 9	AACUUACGUUUGUUUACCUGUCAAAACUCGGGGAAGCCAAUAGCGCGGUAAUCCCGAACCAAGCUCCA GAAAUGGAGAAGGUGUAGAGACUGGAAGGCAGGUACCCUAAACAGUAAAGCUGAGGGUAAAGGGACAG UCCAGACCACAAACUGGUAAUCCAGGCAGUGAAAACUGUAGAUGGUAAAG	3' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-2
5 0	CAACGAGUAAGAUUGACUAAACGCUUAAAAGUUAUUCUUACUAUGGGCGGUACGUAAAGA	5' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene

		GROUP I INTRON-PIE-2
5 1	AUCCCGAACCAAGCUCAGAAAUGGAGAAGGUGUAGAGACUGGAAGGCAGGUACCCUACAGUAAAG CUGAGGGUAAAGGGACAGUCCAGACCACAAACUGGUUAAUCCAGGCAGUGAAAACUGUAGAUGGUAA G	3' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE- 4-5
5 2	CAACGAGUAGAUAUGACUAAACGCUUAAAAGUUUUCUUACUAUGGGCGGUACGUAAAGAAACUUAC GUAUGUUUACCUGUCAAAACUCGGGGAAGCCAAUAGCGCGGUA	5' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE- 4-5
5 3	AAUCCAGGCAGUGAAAACUGUAGAUGGUAAG	3' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-9
5 4	CAACGAGUAGAUAUGACUAAACGCUUAAAAGUUUUCUUACUAUGGGCGGUACGUAAAGAAACUUAC GUAUGUUUACCUGUCAAAACUCGGGGAAGCCAAUAGCGCGGUAAUCCGAACCAAGCUCAGAAAUGG AGAAGGUGUAGAGACUGGAAGGCAGGUACCCUACAGUAAAGCUGAGGGUAAAGGGACAGUCCAGAC CACAAACUGGUU	5' self- splicing intron fragment derived from Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-9
5 5	AAUCCGGUGAAGGUGUAGAGACUGGAUGGGCAGCACCUAAGGCCGCAAGGUCCACGGUGAAGGGAC AGUCCAGACCACGAACGCCCUUAUCCGAUUGAGAUGGGCGGCGCGAAAGCCGAAGCGGUAVG	3' self- splicing intron fragment derived from Agrobacteriu m- tumefaciens GROUP I INTRON-PIE-6
5 6	AAGGUAAAUGGGGGUUGCGCCCGGAACGACGCAAUCCGAUCUGCUCAAAGUCGGGGAAAGCUUCGC UGGUCUCCAGUAUGCCAAUCCCGAGCCAAGCUCGGAGA	5' self- splicing intron fragment derived from Agrobacteriu m- tumefaciens GROUP I INTRON-PIE-6
5 7	AAAAUCCGUUGA	3' exon of GROUP I INTRON from Anabaena

		pre-tRNA-Leu gene
5 8	AGACGCUACGGACUU	5' exon of GROUP I INTRON from Anabaena pre-tRNA-Leu gene
5 9	AAAACCCCCCA	3' exon of GROUP I INTRON from Anabaena pre-tRNA-Leu gene
6 0	AGACGGGGGGGCUU	5' exon of GROUP I INTRON from Anabaena pre-tRNA-Leu gene
6 1	AAUCCGUUGGUCUG	3' exon of GROUP I INTRON from Azoarcus olearius BH72
6 2	GAGCAGCGGACUCAU	5' exon of GROUP I INTRON from Azoarcus olearius BH72
6 3	AACUUUUAGUUUAU	3' exon of GROUP I INTRON from Staphylococcus phage Twort
6 4	ACUACUGAAAGCAU	5' exon of GROUP I INTRON from Staphylococcus phage Twort
6 5	CUGUUA	3' exon of GROUP I INTRON from Scytalidium dimidiatum
6 6	UAACAGGU	5' exon of GROUP I INTRON from Scytalidium dimidiatum
6 7	CAUAACCC	3' exon of GROUP I INTRON from Scytonema-hofmani tRNA fMet gene
6 8	GGGCU	5' exon of GROUP I INTRON from Scytonema-hofmani tRNA fMet gene
6 9	AAUCCUGGGG	3' exon of GROUP I INTRON from

		Agrobacterium tumefaciens
70	CAGGAUCCU	5' exon of GROUP I INTRON from Agrobacterium tumefaciens
71	CCGUCGAUUGUCCACUGGUC	5' homology arm-original (20)
72	GACCAGUGGACAAUCGACGG	3' homology arm-original (20)
73	CCGUCGAUUGUCCACUGGUCCCGUCGAUUGUCCACUGGUC	5' homology arm-2×original (40)
74	GACCAGUGGACAAUCGACGGGACCAGUGGACAAUCGACGG	3' homology arm-2×original (40)
75	GGACACCCCGUACCCGGUCUCAGAAAGCCUGAAAGCAGAAUAGUCAUUAAGAGGGUGGUUGGCUU GGUCGGCAUAGACUUGAGCAGAAAGAGGUUGAAAUUGAGCCUGAUUUCUCUAGGCCCCUCUG CAGUGUCUGUUGUGGAGGCCAGAUACGUACUGCUUCCGCUUUUUUGGUCUCAUUCAGGUGAGCA AAUCCCCUUCAGUUUCACACGACAAUGCAGCUGAUGAGGUUCCAGCUUUGCAAUUGAGUCAUC CAUGAGGACUGCUUCCUGAGAUUUCACAGG	5' homology arm-HA1 (300)
76	CCUGAUGAAAUCCAGGAAGACAGUCCUCAUGGAUGACUACAUUUGCAAAGCUGGAACCUCAUCAGC UGCAUUGUCUGGUGAGAAACAUGAAGGGGAUUGCUCACCUUGAAUGAGACCAAAAAAGCGGAAGC AGUUACGUUUCUGGCCUCCACAACAGACACUGCAGAGGGGCCUAGAGAAUCCAGGCUCACAUUUU CAACCUUUUCUGCUCAAAGUCUAUGCCGACCAAGCCAACCCUUCUAAUGACUAAUUCUGCUUU CAGGCUUUCUGAGACCGGGUCAGGGGGUGUCC	3' homology arm-HA1 (300)
77	GGACACCCCGUACCCGGUCUCAGAAAGCCUGAAAGCAGAAUAGUCAUUAAGAGGGUGGUUGGCUU GGUCGGCAUAGACUUGAGCAGAAAGAGGUUGA	5' homology arm-HA2 (100)
78	UCAACCUUUUCUGCUCAAAGUCUAUGCCGACCAAGCCAACCCUUCUAAUGACUAAUUCUGCUU UCAGGCUUUCUGAGACCGGGUCAGGGGGUGUCC	3' homology arm-HA2 (100)
79	AAAGAAGACAGUCUCGCUUAUCGCCCGAGGUGGAGUGUAGAAUGCAGUGCUGAUCUUGGCUCAC UGCAACCUUACCUCCAGGCUCAAGCGAUUCUUCUGCCUCAGCCUCCCGAGUAGCUAGGAUUACAG GCGCGCGCCACCACACCUAGCUUAUUUUGUAUUUUCAGUAGAGAUUGGGUUUACCAUGUUGACCA GGCUGGUCUCAAACUCCUGAGCUUAAGUGACUGCCCGCCUCAGCCUCCCAAAGUGCUUGGAUUUA GGUGUGAGCCACUGCAC	5' homology arm-HA3 (285)
80	CAUCUCUACUGAAAAUACAAAAAUAGCCGGGCUAGUGGCAUGAGCCUGUAGUCCAGCUACUU GGGAGGCUAGGACAGGAGAAUGGCGUGAACUCGGGAGGCAGAGCUUGCAGUGAGCCGAGAUACACC ACUGCACUCCAGCCUGGGCGACAGAGCGAGACUCCGUCUC	3' homology arm-HA3 (174)
81	AGAGAGUCUCGCUUUCGCCAGGCUAGGAGUGGUGAUCUGGCUCACUGCAACUCUCCUCCC GUCAGCAUUCUCUGCCUCAGCCUCCAGUAGCUGGAUACAGGCCGCCACCCGCUAUUUUGUAUU UUCAGUAGAGAUG	5' homology arm-HA4 (147)
82	CAUCUCUACUGAAAAUACAAAAUAGCGGGUGUGGCGGCCUGUAUCCAGCUACUGGGAGGCUGAGGC AGAGAAUGCUGACGGGAGGAGAGUUGCAGUGAGCCAGAUACACCACUCACUCCAGCCUGGGCGAAG AGCGAGACUCUCU	3' homology arm-HA4 (147)
83	GGCCCCUCUGCAGUGUCUGUUGGAGGCCAGAUACGUAACUGCUUCCGC	5' homology arm-HA5 (50)
84	GCGGAAGCAGUACGUUUCUGGCCUCCACAACAGACACUGCAGAGGGGCC	3' homology arm-HA5 (50)
85	UAGACUUUGAGCAGAAAGAGGUUGAAAUUGUUGAGCCUGAUUUCUCUAG	5' homology arm-HA6 (50)
86	CUAAGAGAAUACGGCUCACAUUUUCAACCUUUUCUGCUCAAAGUCUA	3' homology arm-HA6 (50)
87	AUGAGGUUCCAGCUUUGCAAUGUAGUCAUCCAUGAGGACUGUCUUCUG	5' homology arm-HA7 (50)
88	CAGGAAGACAGUCCUCAUGGAUGACUACAUUUGCAAAGCUGGAACCUCAU	3' homology arm-HA7 (50)
89	GGACACCCCGUACCCGGUCUCAGA	5' homology arm-HA8 (25)

9 0	UCUGAGACCGGGUCAGGGGUGUCC	3' homology arm-HA8 (25)
9 1	GAAAGCAGAAUAGUCAUAGAAGG	5' homology arm-HA9 (25)
9 2	CCUUCUAAUGACUAAUUCUGCUUUC	3' homology arm-HA9 (25)
9 3	CGGCAUAGACUUUGAGCAGAAAGAG	5' homology arm-HA10 (25)
9 4	CUCUUUCUGCUCAAAGUCUAUGCCG	3' homology arm-HA10 (25)
9 5	CUU	Nucleotide motif of ribozyme recognition site I of Aaz.b.trnL, Asp- 2.b.trnL, Cpa.y.trnL- 1, Dai.c.trnL- 1, Dai.c.trnL- 2, Dai.c.trnL- 3, Ddu.c.trnL- 1, Ddu.c.trnL- 2, Dla.c.trnL, Dto.c.trnL- 1, Dto.c.trnL- 2,Dto.c.trnL- 3, Fan.c.trnL, Zma.c.trnL, Vfa.c.trnL, Sel.b.trnL- 2, Sbi.c.trnL, Pth.c.trnL, Osa.c.trnL, Omi.c.trnL, Nta.c.trnL, Mpo.c.trnL, Mae.b.trnL- 2, Dto.c.trnL- 4,
9 6	CCU	Nucleotide motif of ribozyme recognition site I of Aha.b.trnR, Atu.b.trnR,
9 7	CAU	Nucleotide motif of ribozyme recognition site I of Asp.b.trnI, Twort.v.ORE1 42,2

9 8	UU	Nucleotide motif of ribozyme recognition site I of Athu.c.trnL-1
9 9	CU	Nucleotide motif of ribozyme recognition site I of Dsp.b.trnM, Sho.b.trnM
1 0 0	GGCU	Nucleotide motif of ribozyme recognition site I of Gvi.b.trnM, Ssp.b.trnM
1 0 1	GCU	Nucleotide motif of ribozyme recognition site I of Pan.m.COX1,16
1 0 2	CGUCAU	Nucleotide motif of ribozyme recognition site I of Ppo.L1949-1.pdf
1 0 3	CGUCAU	Nucleotide motif of ribozyme recognition site I of Ppo.L1949-2.pdf
1 0 4	GGGAU	Nucleotide motif of ribozyme recognition site I of Ppo.L2449-1.pdf, Ppo.L2449-2.pdf
1 0 5	GGAU	Nucleotide motif of ribozyme recognition site I of Ban.b.nrdE
1 0 6	GGAAAU	Nucleotide motif of ribozyme recognition site I of Ban.b.recA
1 0 7	UUGGAGCC	Nucleotide motif of ribozyme recognition site I of

		Beta-22.v.td,1
1 0 8	UGGAGGU	Nucleotide motif of ribozyme recognition site I of Ceu.c.psbC,1
1 0 9	AAGGU	Nucleotide motif of ribozyme recognition site I of Ceu.mL1931.pdf
1 1 0	GGAGGU	Nucleotide motif of ribozyme recognition site I of Cre.c.psbA,3
1 1 1	AGUGUGGUU	Nucleotide motif of ribozyme recognition site I of LL-H.v.terL,1
1 1 2	GCCU	Nucleotide motif of ribozyme recognition site I of Phi-E.v.DP,1, SP82.v.DP,1, SP01.v.DP,1
1 1 3	GCGU	Nucleotide motif of ribozyme recognition site I of RB3.v.nrdB,1, T4.v.nrdB,1
1 1 4	UAUGAGU	Nucleotide motif of ribozyme recognition site I of T4.v.nrdD,1
1 1 5	GGGU	Nucleotide motif of ribozyme recognition site I of T4.v.td,1
1 1 6	GGU	Nucleotide motif of ribozyme recognition site I of Sdi.S1199-1
1 1 7	AUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCCCUGAGCCUGGCCUGGUCACAAACAGCGCUCCUA CAAGCUCUAGCACCAAGAAAACCCAGCUGCAGCUCGAGCACCUGCUGCUGGACCUGCAGAUGAUCCU GAACGGCAUUACAACUACAAGAACCCAAAGCUGACCAGAAUGCUGACAUUUAAAGUUCUACAUGCCU AAGAAGGCCACAGAGCUGAAGCACCUGCAAUGUCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGC UUAAUCUGGCUCAGAGCAAGAACUCCACCUGAGACCUGGGACCUGAUCAGCGACAUAACGUGAU	IL2-N88D

	CGUGCUGGAACUGAAAGGACGCGAGACAACCUUCAUGUGCGAGUACGCCGAUGAGACGCCACCAUC GUGGAGUUCUGAAUCGUGGAUACCUUCUGGCCAGUCCAUCUCCACCCUGACC	
1 1 8	AUGGCCAGCGAGUUCAGAGAAAGCUCUUCUGGAGGGCAGUGGUGGCCGAGUUCUGGCCACGACCC UCUUUGUCUUCAGCAUCGCGUUCUGCCCUGGGCUUCAAUACCCGGUGGGGAACAACAGACGGC GGUCCAGGACAACGUGAAGGUGUCGCGUGGCCUUCGGGCUAGCAUCGCCACGCGUGGCCAGAGUGUG GGCCACAUCAGCGGGGCCACCUCACACCCGGCUGUCACACUGGGGCGUGCUCAGCUGCCAGAUCA GCUUCUCCGUGCCCUCAUGUACAUCUAGCCAGUGCGUGGGGGCCUAGCUGGCCACGCCCAUCCU CUCAGGCAUCACCUCCUGACUGGGAACUCGCUUGGGCGCAAUGACCUGGCGAUGGUGUGAAC UCGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGGACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUA CCGACCGGAGGCGCCUGACCUUGGUGGCUAGCCCCCUUGCCAUCGGCCUCUCUGUAGCCCUUGG ACACCUCCUGGCUAUGACUACUGGCUUGGGGAUUAACCCUGCUGGUGCCUUGGCUCCGGGUG AUCACACACAACUUCAGCAACACUGGAUUCUUGGGUGGGGCAUUAUCGCGGGGAGCCUUGGCG UACUCAUCUACGACUUCUCCUGGCCCCACGAGCAGUGACCUACAGACCGCGUGAAGGUGUGGAC CAGCGGCCAGGUGGAGGAGUAUGACCUGGAUGCCGACGACAUAACUCCAGGGUGGAGUAUGAGCCC AAA	AQP1
1 1 9	GGUGGAACGGCACAUGUGUGAUGGGGACAUGUUAUCUUAACCCGGCAGCCAAACUCUGCACAAAUG UCCAUGAUGGGGCAUCGGGUCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGACAA CUCCGUACAAUGCAGACUUGACGGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCG AGCAGAGAUCAGGAGCUGGCCAUGGUUCCUGCAUGAUUGUCACCCCCAGAGCAAUCGGCCUUC AUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAUUCACCAAGAGAGCGUCUUCUGGAGC GG	POLR2A
1 2 0	GUUAAGAGACAGGUGUUGCCUUGUCUUCGGACUGGCAUCUUGGGACCAACCCCCUUCUCCCCAG CCAUGGGUUAUAAUGGCAUAUAAAGGACGUAACAACUUGUAACCAUAUAGCUUUGUAUUUUGUAACC ACUAAGCUUUGUGCAUAUAGUAACCAUCAAGCUUGUUAUGUCCAGCAGGAGGUUGCAUGCUUGU AGCCGAAUUGGGGCUCCGACCCCCCAUAGUAGGAUACUUGAUUUUGCAUUCUUAUGUGGACCUGCAA CUCUACACAUAAGAGGCUUUGUCUUGCAUCUAAACACCCUGAGUACAGUGUACCUAGACCUAUAUG ACGGGAGGACCGUUGUUCUCAAUAACCCUACAUAUAGGCUAGGUGGGCAUGCCCAUUGGCAA GAUCCAGACUGGGGUGCGGUCUGGGCAGGGUAGAUCCUGUUAUGCUACUGCCUGAUAGGGUGGUG CUCAACCAUGUGUAGUUAUAAUUGAGCUGUUAUAUACC	CrohivirusB
1 2 1	UUAAAACAGCCUGUGGGUUGAUCCACCCACAGGCCAUUGGGCGCUAGCACUCUGGUUAUCACGGUA CCUUGUGCGCCUGUUUUAUACCCCCUCCCCAACUGUAACUUAAGAUAACACACACCGAUCAACA GUCAGCGUGGCACACCAGCCACGUUUGAUCAAGCACUUCUGUUAACCCCGACUGAGUAUCAUAGA CUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAACUACUUCGAAAAACCUAGUAACAC CGUGGAAGUUGCAGAGUGUUCGUCACACUACCCAGUGUAGAUACAGGUCGAUGACUACCCGAU UCCCCACGGGCGACCGUGGGGUGGCGUGCGUUGGGCGGCCUGCCAUUGGGGAACCCAUUGGGACGCUC UAUAACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGUAGUCCUCCGGCCCCUGAAUUGCGGCUA AUCCUAACUGCGGAGCACACACCCUCAAGCCAGAGGGCAGUGUGUCUAACGGGCAACUCUGCAGCG GAACCGACUACUUGGGUGUCCGUGUUAUUAUUAUCCUAUACUGGCGCUUAUGGUGACAUAUUGA GAGAUUGUUAACUAUAGCUAUGGAUUGGCCAUCCGGUGACUAUAUAGAGCUAUAUAUAUACCCUUA GUUGGUUUAUACCAUAGCUUAGCUUAGAGGUUAAACAUAUACUUAUGUUAAGUUAUAGAAUACAG CAA	CVB3
1 2 2	ACAUGGGGGGUCUGCGGACGGCUUCGGCCCCACCCGCGACAAGAAUGCCGUAUCUGUCCCAUUAACC CGUAUUCUCCUCCUCCCGGCAACCAACACGCUUACUCGCGCACGUGUUGAGUGGCACGUGCGUUG UCCAAACAGCUACACCCACACCCUUCGGGGCGGGUUGUCCCGCCUCCGGGUUCCUGCGGAACCCC CCCCUCCUUCUCUUCUUAUCCGCCCUCACUUCCCAUAACUACAGUGCUUUGGUAGGUGAGGAC CCUGACCCCCCGGGAAGCUGCUAACUGGGCAACUGUGGGGAUCCAGGCAGGUUAUCAAGGCACCC GGUCUUCGCGCCUUCAGGAGUAUCUUCGCGGUGAAUUCGGUAGGGCUCUGCUUGGUGCCAACUUC CCCCAAUUGCGCGUCGCGGAGUGCUUUCCCAACUACUUAUAGUAACCUUCUAGUGUGUGCUUG GUCAGCAUAUCGAGGCGACGUUCCGCUUGCCAGACCAGUCCAGCAAUGGACGGGCCAGUGUGCGU AGUCGCUUUCGCUUUCGCGGCAUGUUGGCGAAACGUGAGGUAAAGGUUGGUGUGCCCAACGCC CGUAUUUGGUGAUACCUCAAGACCACCCAGGAUUGCCAGGAGGUACCCACUUCGGUGGGAUCUG ACCCUGGGCUAAUUGUCUACGGUGGUUCUUCUGCUUCCACUUCUUCUUCUUGGCAUG	SalivirusFHB
1 2 3	CCCACAGCAAGAAUGCCAUAUCUGUCCUACCCCCAAUUUCCUUCUUCUCCCGUGCAACCAUUA CGCUUACUCGCAUGUGCAUUGAGUGUGCAUGUGUUGAAACAAACAGCUACACUACAUGGGGGCGGG UUUUCGCGCCUACGGCCUUCGCGAGGCCACCCUUCUCCUCCCUUAUAACUACAGUGGUUUGGU AGGUAAGCAUCCUGAUCCCCCGGGAAGCUGUCACGUGGCAACUGUGGGGACCCAGACAGGUUAUC AAGGCACCCGGCUUUCGCGCUUCAGGAGUAUCCUACUAGUGAAUUCUACGCGGGGCUUCGCUUGG UGCCAACCUCCCCAAUUGCGCGCUGCGGGAGUGCUUCCCCAACUACCCUAGUAUCCUUCUACUG UGUGUGCUUGGUCAGCAUAUCGAGACGAUUGUCCGUGUCCAGACAGUCCAGUAUUGGACGGGC CAGUCGUGUAGGUCUUCGCGCUUGUCCGGGCAUGUUGGUGAACCUGGUGGUAAGGUGUGG UGCCCAACGCGCCGUACUUGGUGACACCUCAAGACCACCCAGGAUUGCCAGGAGGUACCCACUUC ACGGUGGGAUCUGACCCUGGGCUAAUUGUCUACGGUGGUUCUUCUGCUUCCACUUCUUCUUCUGU UCAG	18-675
1 2 4	UUUGAAAAGGGGUGGGGGGCGCUGCGCCCCUACCCUUCUUCGUGGUGUGUCCCGGACAC CGUUAUCUCCAUUCAGCUUCUUCGGAACCUUGUUCGGAGGAUUUAAACGGGCACCCAUACUCCCCAC CCCCUUUUGUAACUAAGUAUGUGUGCUGGUAUCUUGACUCCACGGAACGGACCAUCCGUUGGU GAACAAACAGCUAGGUCACAUCUCCUUCUCCUGGGAGGGCCCCGCCUCCCAUCCUCCUCCCC CAGCCUACGUUAUCACAGGCUUGUGAAGCCCCCGGAAAGCUCUACCGUGGCAUUGUGGGUCC CCUUCUACAAGACACCAGUUCUUCUCCUUAAGGCUAGCCCCGGCGUGUGAAUUCACGUUGGCA	Aichivirus

160

1 3 1	UUA AAAACAGCGGAUGGGUAUCCCAACUCCGGCCACUGGGUGUAGUACUCUGGUACA UUGUACCUU UGUACGCCUGUUUUCCCCUCUUGUACCCGCCCUUACAAGCUCUUGCCCAAGUAACGUUAGAAGUUU GAACA UUGGUACA AUAGGAAGCAUCACA UCCAGUGGUGUACUGUACA AACA CUCUUGUUGCCCCGGA GCGAGGUUAUAGAUUGUCCCCACCGUCA AAAAGCCUUAAACCGUUUACCGCCAAUACAUCGUAAUGG CUAGUAGCACCUUGGAUUAAGUUGGGGUCGUAUCAGGUGGUAACCCCCACUAGUUUGGUCGAUGAG GCUAGGAUUUCCCCACGGGUGACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCAUCCCGUGGGAC GCCAAUUUAUGACAUGGUGUGAAGACCUUGCAUGUGCUU GAUUGUGAGUCCUCCGGCCCCUGAAUGC GGCUAAACCUAACCCCGAGCCUUGCGACACA AUCCAGUGUUGUUAAGGUCGUAUAGAGCAAUUCUG GGAUUGGACCGACUACUUGGGUGUCCGUGUUUCUUAUUUUUCUUGAAUUUUUCUUAUGGUCACAGC AUUAUACA UUAUACUGUGAUC	HRVB92
1 3 2	UUA AAAACAGCCUGUGGGUUGUCCCAACCCGAGGGCCACUGGGCGCUAGCACACUGGU AUCCCGGU ACCCUUGUGCGCCUGUUUUAUUAUACCCUCCCCCUU AUGUAACUAGAAGUAUGAUUCAACCGGUCGA CAGGCGGCUACAGUGCAACACUGAGUCAUGACC AAGCACUUCUGUUACCCCGGACUGAGUAUCAUA AGCUGUUCACACGGCUGAAGGAGAAAACGUUCGUUACCCGGCCAAUUAUUCUGAGAAACCUAGUACC AGGCCGGUGUGCGUUUGUCUAUUAUGUUUAUUUCCACCAUAUUGCCGCUUUUUGGCAUUGGAGGCC AUUCCCCACGGGCGACCGUGGGCGGUGGUGCGUUGGCGGCCUGCCAUUGGGGCAACCCAUGGGACGC UUCAAUACUGACAUGGUGUGAAGAGUCUAUUGAGCUAAUUGGUAUGUCCUCCGGCCCCUGAAUGCGGC UAAUCCCAACUGUGGAGCAGAUACUCACA AACCAAGUGAGCGGUCUGUCGUAACGGGCAACUCCGCAG CGGAACCGACUACUUGGGUGUCCGUGUUUCUUUAUUCUUAACA UUGGUGCUUAUGGUGACAAUU GACAAAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAAACAGAGCUAUUGUUUACUUGU UUGUUGGUUUAUACCAUUAUUAUACAAGGUCUUAAGAACUCUCAACUUUAUUUGACACUCAAUAC AGCAAA	EVB107
1 3 3	UUGAAAGCCGGGGGUGGGAGAUCCGGAUUGCCAGUCUGCUCGAUAUCGCAGGCGUGGUGCCGUGACUA CCCACUCCCCCUUACAACGUGAAGGCUACGAUAGUGCCAGGGCGGGUACUGCCGUAAGUGCCACCCC AAAACAACAACACCCCCUCUCCUCCCCCCCCCUAACGUAUACUGGCCGAAGCCGCUUGGAAUA AGGCCGGUGUGCGUUUGUCUAUUAUGUUUAUUUCCACCAUAUUGCCGCUUUUUGGCAUUGGAGGCC CGGAAACCGUGGCCUGUCUUCUUGACAGCAUUCUAGGGGUCUUUCCCCUCUGCCAAAGGAUUGC AAGGUCUGUUGAAUGUCGUGAAGGAAGCAGUUCUUGGAAGCUUCUUGAAGACAACAACGUCUGU AGCGACCCUUGGACAGGCAAGCGGAACCCCCACUUGGCGACAGGUGCCUUCUGCGGCCAAAAGCCACGU GUUAUAGAUACACCGCAAGGCGGCACAACCCAGUGCCACGUUGUGAGUUGGAUAGUUGGAAA GAGUCAAUUGGUCUCCUCAAGCGUAUUAACAAGGGGUGAAGGAUGCCAGAAAGGUACCCCAUUG UAUGGGAUCUGAUCUGGGGCCUCGGUGCACAUGCUUUAUGUGUUUAGUCGAGGUUAAAAACGUC UAGGGCCCCGAAACACGGGGACGUGGUUUUCCUUGAAAAACACGAUGAUAAU	EMCV-A
1 3 4	UUGCCAGUCUGCUCGAUAUCGCAGGCUUGGUGCCGUGACUACCCACUCCCCCUUACAACGUGAAGGCU ACGAUAGUGCCAGGGCGGGUACUGCCGUAAGUGCCACCCCAAACAACAACAACAACUCCCC CUCCCCCCCCUUAUAUACUGGCCGAAGCCACUUGGAAUAAGGCCGGUGUGCUUUUGUCUAACUGCU AUUUUCUACCCGAUUAACGUCUUAUGGUAUUGUGAGGGUCCAGAACCGUACCCUGUCUUCUUGACGA ACACUCCUAGGGGUCUUUCCCCUCUGACAAGGAGUGUAAGGUCUGUUGAAUGUCGUGAAGGAAGC AGUUCUUCUGGAAGCUUCUUAAGACAACAACGUCUGUAGCGACCCUUUGCAGGCGAGCGGAACCCC CACUCCUGGUGACAGGUGCCUCUGCGGCCAAAAGGCCACGUGUAUAGAUAACCUUGCAAGGCGGCC AACCCACAGUGCCACGUUGUGAGUUGGAUAGUUGUGGAAAGAGUCAAUUGGUCUCCUCAAGCGUAUU CAACAAGGGGUGAAGGAUGCCAGAAAGGUACCCCAUUGUAUGGAUCUGAUCUGGGGCCUCGGUGC ACGUGCUUUAACAGUGUUGAGUCGAGGUGAAAAACGUCUAGGCCCCCCGAAACACGGGGACGUGGU UUUCCUUGAAAAACACGAUUACA AU	EMCV-B
1 3 5	UUGAAAGCCGGGGGUGGGAGAUCCGGAUUGCCAGUCUGCUCGAUAUCGCAGGCGUGGUGCCGUGACUA CCCACUCCCCCUUACAACGUGAAGGCUACGAUAGUGCCAGGGCGGGUACUGCCGUAAGUGCCACCCC AAAUUAACAACAGACCCCCCUAACGUAUACUGGCCGAAGCCGCUUGGAAUAAGGCCGGUGUGCGU UUGUCUAUAUGUUAUUUCCACCAUAUUGCCGUCUUUUGGCAUUGUGAGGGCCCCGAAACCGUGGCC UGUCUUCUUGACGAGCAUUCUAGGGGUCUUUCCCCUCUGCCAAAGGAAUGCAAGGUCUGUUGAAU GUCGUGAAGGAAGCAUUCUUGGAAGCUUCUUGAAGACAACAACGUCUGUAGCGACCCUUGCA GGCAGCGGAACCCCCACUUGCGACAGGUGCCUCUGCGGCCAAAAGCCACGUGUAUAGAUAACCC UGCAAAGGCGGCACAACCCAGUGCCACGUUGAGUUGGAUAGUUGGAAAGAGUCAAAUGGCUC UCCUCAAGCGUAUUAACAAGGGGUGAAGGAUGCCAGAAAGGUACCCCAUUGUAUGGGAUCUGAUC UGGGGCCUCGGUGCACAUGCUUUAUGUGUUUAGUCGAGGUUAAAAACGUCUAGGCCCCCCGAAAC CACGGGGACGUGGUUUUCCUUGAAAAACACGAUGAUAAU	EMCV-pEC9
1 3 6	UUA AAAACUGGGUACAGGUUGUCCCAACCUUAUCACCCACGUGGUGUGGUCUCUUGUAUUCGGUA CACUUGCACCGCCAGUUUGCCACCCCUACCCGUCGUAACUUAAGAGCUAACACUCCGACCAACAGGC GGUGGUAACCAUACCACUACGGUCAAGCACUCCUGUUUCCCCGGUAUGCGAGGAAUAGACUCCUA CAGGGUUGAAGCCUCAAUAUCCGUAUUGGUAUACGCAAGCUUAGUAGUCCUUGAAAG UCCCUUGGUUGGUCUCCGCUAGUUUCCCCUAGUAGACCUGGCAGAUAGGCGAGGACACUCCCAAC UGGCGACAGUGGUCUGCCUGCGUGGCGCCUAGGGGUGCGAAGCCAGUGACAGAC AAGGUGUGAAGAGCCCCGUGUGCUACCAUAGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCAACCCC ACAGCUAUUGCACACAAGCCAGUGUGUAUGUAGUCGUAUAGAGCAAUUGUGGACGGAACCGACUAC UUUGGGUGCCGUGUUUCCUUAUUCUUAUACUUCUGCUUAUGGUGACAAUACUGGAAUAGUGU UGUUACC	HRVC-02
1 3 7	UAAAACUGGAUCCAGGUUGUCCCAACCUUGGAUCUCCAUUUGGGAGUUGUACUUAUUAUCCGGUAA UUUGUAGCGCCAGUUUUUAUCCCCCCUCCCAUUGUAACUUAAGAGGUUAUCAUACGACCAUA GGUGUAGUUAAGCAAACUACAAAGGUCAAGCACUUCUGUUUCCCCGGUCAAAGUUGUAUAGCUCC AACAGGGCAAAAACACUGAGAUUGCUUACCGCAAAGUGCCUACGCAAGCCUAGUAACACCUUUGA	HRV-A21

162

	GAGCUUCCAGGUUGGGAAACCUUGGACUGGGCCUAUACUACCUGAUAGGGUCGCGGCUGGCCGCCUGU AAUAGUAUAGUCAGUUGAAACCCCCC	
1 4 4	UUUCCUCCUUUCGACCGCCUUACGGCAGGCGGGUCCGCGGACGGCUUCGGCCUACCCGCGACAAGAA UGCCGUAUCUGUCCUUUAUACCCAUUAUUCUUUCCUCCCCCGCAACCAUCACGCUUACUCGCGCA CGUGUUGAGUGGCACGUGCGUUGUCCAAACAGUUAACACUCACACCCUUGGGCGGGUUUGUCCCGC CUCGGGUUCCUCGCGGAACCCUCCUUCUCUCUCCUUCUUAUCCGCCUUCACUUAUACUA CAGUGCUUUGUAGGUAAGCAUCCUGACCCCCCGCGGAAGCUGCCAACGUGGCAACUGUGGGGAUCC AGGCAGGUUAUCAAAGGCACCCGCUUUCGCGUUCAGGAGUAUCCUGCCGGUGAAUCCGACAG GGCUUCGUUGGUGCCAACCUCCCCCAAUGCGCGUCGCGGAGUGCUUCCCCAACUACUUAUAG UAACCUUCUAUGUGUGUGCUUGGUCAGCAUAUCUGAGGCGACGUUCCGCUUGCCAGACCAGUCCAG CAAUGGACGGGCCAGUGUGCGUAGUCGUUUCGGUUCUCCGCGCAUGUUUGGCGAAACGUCAGG UAAGGUUGGUGUGCCAAUGCCCGUAUUUGGUGACACCUCAAGACCACCCAGGAUUGCCAGGGAGG UACCCACUUCGGUGGGAUUCUGACCCUGGGCUAAUUGUCUACGGUGGUUCUUCUUGCUUCCACUUCU CUUUUUUCUGGCAUG	Salivirus-A- BN5
1 4 5	UAUGGCAGGCGGGCUUGUGGACGGCUUCGGGCCACCCACAGCAAGAAUGCCAUCAUCUGUCCUACC CCAUUUUUCCCCUUUCCUUCUGCAACCGUUACGCUUACUCGAGGUGCAUUGAGUGGUGCAGC UGUUGAAUAAACAGCUACACUCACAUGGGGGCGGGUUUCCCGCCUUCGCGCCUCUCGCGAGGCCA CCCCUCCCUUCCUCCAUAAUACUAGUGCUUUGGUAAGCAUCCUGAUCCCCCGCGGAAGCUG CUCACGUGGCAACUGUGGGACCCAGACAGGUUAUCAAGGCACCCGUCUUCGCGCUUCAGGAGU AUCCUGCUAGUGAAUUCUAGUAGGGCUCUGCUUGGUGCCAACUCCCCCAAUUGCGCGUCGCGGA UGUCUUCUCCCAACUACCCUAGUAUCCUCUAGUGUGUGCUUGGUCAGCAUAUCUGAGACGAUG UUCCGCUGUCCAGACCAGUCCAGUAUUGGACGGGCCAGUGUGCGUAGUCGUCUUCGCGCUUUCG GCGCAUGUUUGGUGAACCUGGUGGGUAAGGUUGGUGUGCCCAACGCCGUAUUCUGGUGAUACCUCA AGACCACCCAGGAUUGCCAGGGAGGUACCCGCUUCACAGCGGGAUCUGACCUGGGCUAAUUGUCU ACGGUGGUUCUUCUUGCUUCCACUUCUUCUACUGUUC	Salivirus-A- BN2
1 4 6	UAUGGCAGGCGGGCUUGUGGACGGUUUCGGGCCACCCACAGCAAGAAUGCCAUCAUCUGUCCUACC CCCAAUUUUCCUUUCCUUCUGCAUACUACGCUUACUCGCAUUGAGUGGUGCAUUGUG UUGAACAAACAGCUACACUCACAUGGGGGCGGGUUUCCCGCCUACGGCCUUCGCGAGGCCACC CUUCCCUCCCGCUUAUAACUACAGUGCUUUGGACAGGUAAGCAUCCUGAUCCCCCGCGGAAGCUGUC ACGUGGCAACUGUGGGACCCAGACAGGUUAUCAAGGCACCCGUCUUCGCGCUUCAGGAGCAUC CCCAUAGUGAAUUCUAGUGGGGCUUCGCUUGGUGCCAACUCCCCCAAUUGCGCGUCGCGGAGUG CUCUUCUCCCAACCAUCCUAGUAUCCUCUAGUGUGUGCUUGGUCAGCAUAUCUGAGACGACGUUC GGACCCAGACAGGUUAUCAAGGCACCCGUCUUCGCGCUUCAGGAGUAUCCUGCUAGCGAAUUC UAGUAGGGCUCUGCUUGGUGCCAACUUCUCCAAUUGCGCGUCGCGGAGUGCUUCCCCAAUUA CCCCAGUAUCCUCUAGUGUGUGCCUGGUCAGCAUAUCUGAGACGAUGUUCGCGUUCUCCAGACCA GUCUCCGUAUUGGACGGGCCAGUGUGCGUAGUCGUCUCCGCGCUUGUCCGCGCAUGUUUGGUGAAC GGUGGGGUAAGGUUGGUGUGCCCAACGCCGUAUACAGGGGAUACCUCAAGGCACCCAGGAUUGCCA GGGAGGUUAUCCGCCUACAGCGGGAUCUGACCCUGGGGUAUUGUCUGCGGGGGUCCUUCUGGCC CAAUUCUACAGUAUUUCCAGG	Salivirus-A- GUT
1 4 7	UUCUCCUGCAACCAUUAACGUUAUACGCAUGUGCAUUGAGUGGUGCAUGUGUUAACAAACAGCUAC AAUCACAUUGGGGGCGGGUUUCCCGCCCCACGGCUUCUCGCGAGGCCAUCCUCCCUUUUCCUCCA UAACUACAGUGCUUUGGUAAGGUAAGCAUCCCGAUCUCCCGCGGAAGCUGCUCACGUGGCAACUGUG GGACCCAGACAGGUUAUCAAGGCACCCGUCUUCGCGCUUCAGGAGUAUCCUGCUAGCGAAUUC UAGUAGGGCUCUGCUUGGUGCCAACUUCUCCAAUUGCGCGUCGCGGAGUGCUUCCCCAAUUA CCCCAGUAUCCUCUAGUGUGUGCCUGGUCAGCAUAUCUGAGACGAUGUUCGCGUUCUCCAGACCA GUCUCCGUAUUGGACGGGCCAGUGUGCGUAGUCGUCUCCGCGCUUGUCCGCGCAUGUUUGGUGAAC GGUGGGGUAAGGUUGGUGUGCCCAACGCCGUAUACAGGGGAUACCUCAAGGCACCCAGGAUUGCCA GGGAGGUUAUCCGCCUACAGCGGGAUCUGACCCUGGGGUAUUGUCUGCGGGGGUCCUUCUGGCC CAAUUCUACAGUAUUUCCAGG	Salivirus-A- CH
1 4 8	UCUGUCCUACACCCAUUCUCCUUCUUCUCCUACCGUUACGCUUACUCGCAUGUGCAUUGAGUGGU GCACGUGCUUGAACAAACAGCUACACUCACAUGGGGGCGGGUUUCCCGCCUUCGCGCCUCUCGCGA GGCCACCCUCCUCCUUCUCCCAUAAUACAGUGCUUUGGUAAGGAUAGCAUCCUGAUCCCCCGCGG AAGCUGCUCACGUGGCAACUGUGGGACCCAGACAGGUUAUCAAGGCACCCGUCUUCGCGCUUC AGGAGUAUCCUGCUAGUGAAUUCUAGUAGGGCUCUGCUUGGUGCCAACUCCCCCAAUUGCGCGCU GCGGGAGUGCUUUCUCCCAACUACCCUAGUAUCCUCUAGUGUGUGCUUGGUCAGCAUAUCUGAG ACGAUGUUCGCGUGUCCAGACAGUCCAGUAUUGGACGGGCCAGUGUGCGUAGUCGUCUUCGCGCU UGUCCGCGCAUGUUUGGUGAACCGGUGGGGUAAGGUUGGUGUGCCCAACGCCCGUACUUGGUGAU ACCUCAAGACCACCCAGGAUUGCCAGGGAGGUACCCCGCUUCACAGCGGGAUCUGACCCUGGGCUAA UUGUCUACGGUGGUUCUUCUUGCUUCCACUUCUUCUACUGUUCAG	Salivirus-A- SZ1
1 4 9	UUAAAACAGCCUGUGGGUUGUUCUCCACCCACAGGGCCAUUGGGCGUCAGCACCCUGGUUAUCAGGU ACCUUUGUGCGCCUGUUUAUAUCCUUCUCCCAUUGUAACUUAAGAAGAAACACACACCGAUCAAC AGCAAGCGUGGCACACAGCCAUUUUUGGUCAAGCAUUCUGUUACCCCGGACUGAGUAUCAAUAG ACUGCUCACGCGGUAGAAGGAGAAAGCGUCCUUAUCCGCGCAGCAUUCGAGAAACCUAGUAACA CCUAGGAAGUUGCGGAGUGUUUCGCUAGCAUACCCAGUGUAGAUACGGUCGAGUAGUACCCGCU UUCUCCACGGGCGACCGUGGCGGUGCGUUGGCGGCCUGCCUAGGGGGAACCCAUAGGACGCU CUAAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGCUGGUUAUUCUCCGCCCCGUAUUGCGGCU AAUCCUAACUGUGGAGCACAUGCCCUAAUCCAAGGGUAGUGUGUAUAGAGCAAUUCCGCGC GGAACCGACUACUUGGGUGUCCGUGUUUCUUCUUAUUCUUGUACUGGCGCUUAUUGGUGACAAUUG AGAGAUGUUAACCAUUAAGCUAUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUGUGUAUCUUCU UGUUGGAUUUGUACCACUUAUUUGAAAGAAUACAGGACACUACGCUACAUUUUACUAUUGAACACC GCAAA	Echovirus-7

1	UUAAAAAGCCUGUGGGUUGCACCACACAGGGCCACUGGGCGCAAGCACUCUGGCACUUCGGU	EVA71
5	ACCUUUGUGCGCCUGUUUUUAUACCCUCCCCCAAUGAAUUUAGAAGCAGCAAACCCGAUCAAUA	
0	GCAGGCAUAACGCUCAGUUUAUGUCUUAAGCAAGCACUUCUGUUUCCCCGGACUGAGUAUCAAUAGA	
	CUGCUCACGCGGUUGAAGGAGAAACGUUCGUUAUCCGGCUAACUACUUCGGAAAGCCUAGUAACAC	
	CAUGGAAGUUUGCGAGAGUUUCGUUCAGCACUCCCCAGUGUAGAUACGGUCGAUGAGUACCCGCAU	
	UCCCCACGGGCGACCGUGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGGGUAACCCAUGGGACGCUC	
	UAAUACGGACUGGUGUGAAGAGUCUACUGAGCUAGUUAGUAGUCCUCCGGCCCCUGAAUUGCGGCUA	
	AUCCCAACUGCGGAGCACACGCCACAGCCAGUGGGUAGUGUGUCGUAACGGGCAACUCUGCAGCG	
	GAACCGACUACUUUGGGUGUCCGUGUUCCUUUAUUCUUAUGUUGGCGUCUUAUGGUGACAAUUA	
	AGAGUUUUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGUGCAACAGAGCGAUCCUUUACCUAUUU	
	AUUGGUUUUGUACCAUUGACACUGAAGUCUGUGAUCACCCUUAAUUUUUAUCUUAACCCUCAACACAG	
	CCAAAC	
1	CCUCGGUCCUCUUUCCGUGCGCCGCCACGACGUUAAAUGCGGUGUUGUGGUGCUUAGGUGCCACAC	Oscivirus
5	CACUGCUAUUUGGUGCCCCUUUCCCCUAUAUAUGUUUGUUUUUAUUUCAUUUUCUUGAGGAUUG	
1	GCACCUCCUUAUGCCAAAUUCUAAUACUGGGAGGAUCCAGGCUUUCUGGCUUUUAACAGAAACUCCAC	
	GUCCAGGUCUAUAGAAACUGGUUGGUAGGCGUCCUGAGUAGUCCAUUUGCUAGUAGUCCUUGUGAAC	
	AGGGUGGCUCCGUUUAUCUGCUGGUAUUCGGUGUAGGUGCGCAUGGUGGUAACACCAUCCUGCAU	
	UGUGUGUGAACCCAGUACCGCAAGGAUAGCAAGGUAUGAACACUUGUGGACGAAAUGGUUAGUGAUCA	
	UUGAGCUACUAUAGAAUCCUCCGGCCCGUGAAUGCGGCUAAUCCCAACCCUGGAGCAGGUGGUCACA	
	UGAGGGUGGGCUAAGCAGACCCUGGCCAUGUGGCUAAGCUUUUCCCUUAUUGUUUUACUUUGUAACAU	
	UUUAUUGUGGUUAGUGAUUUUGUGUUGUGGCCUUUGUGAGCUAUUCCAGUAUAAGUUCGCAGCUA	
	GAGUUAUUCUUCGACAUCGGCUGUAUUGGAA	
1	GUUAUCCGCUUAUGUACUUCGAGAAGCCAGUACCAUCCUGGAAUUCUUGAUGCGUUGCGUCAGCA	PV-Mahoney
5	CUCAACCCACAGAGUGUAGCUUAGGCUAGAGUCUGGACAUCCUACCCGGUGACGGUGGUCAGGC	
2	UGCGUUGGGCGCCUACCUAUGGCUAACGCCAUGGGAGCGUAGUUGUGAACAAGGUGUGAAGGCCUA	
	UUGAGCUACUAUAGAAUCCUCCGGCCCGUGAAUGCGGCUAAUCCCAACCCUGGAGCAGGUGGUCACA	
	AACCCAGUGAUUGGCGUGUCGUAACGCGCAAGUCCGUGGCGGAACCGACUACUUUGGGUGUCCGUGUU	
	UCCUUUUUAUUUAUUGUGGCGUCUUAUGGUGACAAUCACAGAUUGUUUAUCAUAAAGCGAAUUGGAUU	
	GGCCAUCGCGUGAAGUGAGAUUCAUUAUCUAUCUGUUUGCUGGAUUCGCUCAUUGAGUGUGUUUA	
	CUCUAAGUACAAUUAACAGUUAUUUCAUACAGACAAU	
1	CACCCAUACACCCCAACCCCUUUCUGUAACUCAAGUAUGUGUGUCUGUAUUCUUGACUCCACGG	Aichivirus-A-FSS693
5	AAUGGAUCCAGUCCGUGGAGAACAAACUGCUGAGUCCACAUCCUCCUCCUUGGGAGGACCUCGG	
3	UCCUCCCAUACUCCUCCUCCAGCUCAGCUAUCACAGGCUUGUGUGAAGCCCCCGCGAAAGCUGCUCA	
	CGUGGCAAUUUGGGUCCUCCUUAUCAAGACACCAAGGUCUUCUCCUUAAGGCUAGCCCCGAGU	
	UGUGAAUUCACAUUGGGCAACUAGUGGUGUCACUGUGCGUCCCAUUCUGGCCCGCGAGUGCUGUU	
	CCCCAAGCCAAACCCUUGGCCUUCACUAUGUGCCUGGCAAGCAUACUGAGAAGGUGUUCGCGU	
	GGCUGCCAGCCUGGUAACAGGUGCCCCAGUGUGCGUAACCUUCUUCGUCUCCGGACGGUAGUGAUU	
	GGUUAAGAUUUGGUGUAAGGUUACUGUGCCACGCCUUGUGCGGGAUGAAACCUUACUGCCUAGG	
	AAUGCCAGGCAAGUACCCACCUUGGGUGGGAUCUGAGCCUGGGCUAAUUGUCUACGGGUGAUUUU	
	AUUUCCAAUUCUUUAUGCUGGAGUC	
1	UGGUUACUCCAUUCAGCUUUCGGAACCGUUCGAGGAAUUAAACGGGCACCCAUACUCCCCCCC	Aichivirus-KVGH
5	ACCCCCCUUUGUAACUAAGUAUGUGUGCUCUGACCUUGACUCCACGGAACGGACCGAUCCGUGU	
4	UGGAACAACACAGCUAGGUCCACAUCUCCUUCUCCUUGGAGGGUCCCGCCUCCCAUCCUCCCC	
	CCCAGCCUGACGUGUCACAGGCUUGUGAAGCCCCCGCGAAAGCUGUCACGUGGCAAUUGGGGUC	
	CCCCUUAUCAAGACACAGGUCUUCUCCUUAAGGCUAGCCCCGGCGUGUGAACUCACGUGUGG	
	CAACUAGUGGUGACUGUGCGCUCCCAUUCUGGCCGCGGAGUGCUGUCCCCAAGCCAAACCCU	
	GGCCUUCAGTUAUGUGCCUGGCAAGCACACCUAGAGAAGGUGUUCGUGUGGCGUGCCAGCCUGGUA	
	CAGGUGCCCCACUGUGCGUAACCUUCUUCGUCUUCGGACGGUGUGAUUGGUUAAGAUUUGGUGUA	
	AGGUUCAUGUGCCACGCCUUGUGCGGGAUGAAACCUUACUGCCUAGGAAUGCCAGGCAGGUACC	
	CCACCUUGGGUGGAUCUGAGCCUGGGCUAAUUGUCUACGGGUGGUUCAUUUCCAAUUCUUUCAU	
	GUCGGAGUC	
1	GUUACUCCAUUCAGCUUUCGGAACCGUUCGAGGAAUUAAACGGGCACCCAUACACCCCCACCC	Aichivirus-DV
5	CCUUUUCUGCAACUUAAGUAUGUGUGCUGUAUUCUUGACUCCACGGAACGGAUCGAUCCGUGGA	
	GAACAAACUGCUAGAUCCACAUCUCCUUCUCCUUGGAGGACCCCGGUCCUCCCAUCCUCCCC	
	CAGCCUGACGUAAACACAGGCUUGUGAAGUCCCGCGAAAGCUGUCACGUGGCAAUUGGGUCCC	
	CCCUUACCAAGACACAGGUCUUCUCCUUAAGGCUAGCCCCGAUGUGUGAAUUCACAUUGGGCA	
	ACUAGUGGUGUCACUGUGCGCUCCCAUUCUGGCCGCGGAGUGCUGUCCCCAAGCCAAACCCUGG	
	CCCUUACUAUGUGCCUGGCAAGCAUUCUGAGAAGGUGUUCGCGUGUGGCGCCAGCCUGGUAACA	
	GGUGCCCCAGUGUGCGUAACCUUCUUCGUCUCCGGACGGUAGUGAUUGGUUAAGAUUUGGUGUA	
	GUUCAUGUGCCAACGCCUUGUGCGGGAUGAAACCUUACUGCCUAGGAAUGCCAGGCAGGUACCC	
	ACCUUCGGGUGGGAUCUGAGCCUGGGCUAAUUGUCUACGGGUAGUUUCAUUUCCAAUUCUUUAUGU	
	CGGAGUC	
1	GUAAUUCAGAGUGUGUGCUCGUAUUCUUGACUCCUGCCGAAUUGCCGCCGUGUCAGUGAACAAA	Murine-Kobuvirus-1-aichivirusA3
5	CAGCUAGGCAAGUCCUCCUUCUCCUGUGGUGUUCUACCGGCCACCAUCCUCCUCCAGCCUG	
6	ACGUGUACAGGCUUGUGCAAAGCCCCCGCGAAAGCUGCUCACGUGGCAAUUGUGGGUCCCCUUUG	
	UCAGACACCCAGUCUUUCUCCUUAAGGCUAGCCCGGUCCACGAACGUGGAAUUGGCAACUAGUG	
	GUGUCACUACGCGUCCGACCUCCGACGCGGAGUGCUGUUCUCCCAAGCUGUAACCCUGACCAAGA	
	CUGUGUGCCUGGCAAGCACCGUCUGGGAAGAUUCCGCGUGGCGUCCAAACCGGUAAACAGGUG	
	CCCCAGUGUGUGAUGUCUCCUCCAGUCUCCGACUGGCAGUCUUGUGUAAAGAUAGCAGUGUAAGGU	

	UCAAGUGCCAAAUCCUGGAAGGAGUGACCCUCUACUGCCCUAGGAAUGCUGCAGGUACCCCCAA CUUCGGUUGGGGAUCUGAGCACAGGCUAAUUGUCUACGGGUAGUUUCAUUCCCAUCCUCUUUUU UGGCAUC	
1 5 7	UUUGAAAAGGGGUGGGGGGGCCUCGGCCCCUCACCCUCUUUCCGGUGGCCACCCGCCGGGCCA CCGUUACUCCACUCCACUCCUUCGGGACUGGUUGGAGGAACAUAACAGGGCUUCCCAUCCUCUUU ACCCUUAUCUCACUACCCUCCCCUUGACCAACCCUAUCCACACCCACUGACUGACUCCUUUGGA UCUUGACCUCGGAUUGCCUACUUGACCUCUCCACUUGCCUCUCCUUUCCGGAUUGCCGGUGGGUCU GGCGGAAAAAGCACAAAGUGUGUUGGCUACCAACUCCUACCCGACAAAGGUGCGUGUCCGGUG CUGAGUAAUGGGAUAGGAGAUGCCAAUAACAGGCUCGCCAUGAGUAGAGCAUGGACUCCGGUGCAU GUGACUUCGGUCACAGGGGCAUAGCAUUGCUCACCCUGAAUCAAGUCAUCGAGAUUUCUGACC UCUGAAGUGCACUGUGGUUGCGUGGCGUGGGAUCCACGCUGACCAUGUACUGCUUAGUAGAGUCGC GGCUGGCCGACUCAUGGGUUAAGUCAGUUGACAAGACAC	Porcine- Kobuvirus-K- 30
1 5 8	CCACCGUUAUCUACUCCACUCCUCGGGACUGGUUGGAGGAGCAUAACAGGGCUUCCCAUCCUG UUCACCCUCAUACACCCACCCUUCUCCUCAACCAUCCCUAUCACACCCACUGACUGAUUCCCU UGGAUUUUGACCCUACAAGCGCCUACUUGACCUCUCCACUUGCCUUUCCGGAUUGCCGGUGGU GCCUGGCCGAAAAAGCACAAAGUGUGUUGCAGGCUACCAACUCCUACCCGACAAAGGUACGUGCCG CCUGUCUAGUAUUGGGAUAGGAGAUGCCUACAACAGGCUCGCCAUGAGUAGAGCAUGGACUCCGGU GCAUGUGACUUCGGUCACACGGGCAUAGCAUUGCUCACCCUGAAUCAAGUCAUUGAGAUUCCUCU GACCUCUGAAGUGCACUGUGGUUGCGUGGCGUGGGAUCCACGCUUGACCAUGUACUGCUUAGUAGAG UCGCGGCGUGGCCGACUCAUGGGUUAAGUCAGUUGAUAGACAC	Porcine- Kobuvirus-XX
1 5 9	UUUGAAAAGGGGUGGGGGGGCCUCGGCCCCUCACCCUCUUUCCGGUGGCCACGCCGGGCCAC CGAUACUCCCUUACUCCUUCGGGACUGUUGGGGAGGAACAACAAGGGCUUCCCUUUCUCCCAU UCCUUCUCCCUUUUCCCAACCCCAACCGCGUAUCUGUGGGCGGCAAGACACAGGGUCUUUCCUC UAAAGCACAAUUGUGUGUGUCCAGGUCCUCCUGCGUACGGUGCGGGAGUGCUCCACCCCAACUG UUGUAAGCCUGUCCAACGCGUGUCCUGGCAAGACUAUGACGUCGCAUGUUCCGUGCGGAUGCCGA CCGGGUAAACCGGUUCCCCAGUGUGUGUAGUGCGAUUCCAGGUCCUCCUGUUGGCGUUGUCCAGA AACUGCUUACAGGUAAGUGGGUGUGCCCAAUCCCUACAAGGUUGAUUUCUACACCCUUAAGGAU GCUCCGGAGGUACCCAGCAACAGCUGGGAUCUGACCGGAGGCUAAUUGUCUACGGGUGGUGUUCC UUUUUUUUUACACAACUUCUACUGUGACACUACUGACUUAUCCAUUGCUCUUCUUGGCUUUUC UGCUCUGGUUCAAGUUCUUGAUUGUUUUGACUGCUUUUACUGCUUUUUCUUCACAAUCCUUGC UCAGUUCAAAGUC	Caprina- Kobuvirus- 12Q108- aichivirusC2
1 6 0	UUCUCCUUCUCCCUUUUACGGCCCCGGCCGCCCGGGUCGUCACCCGUUCAGACACCUACGGGUU AUCUGUUCGGGCUAUAUAUUGGGCAUUCUUCUCCUCCUUCUCCUUUGAAGAUAGUGCGCAUAU UCUUGACUCCGCCUGGAUUGGCCGCCAAGGCGUGAACAAGCAGCUAGGCCACCAUGACACUGCGGU GGUGUCCGAACCCCGGGUGGCCUUCACGGGCACCUUGGUUAUGUAGGACUCCACCCUGGUCUUCC UUUCCUCCUCAAUCUUUCCUCCUGGUUCGACUAACGGGACAGUGCUUGGAACCUUGUCCGGUAACGG UAUAGCAGGCCCCCCCGGAGAAACACCCGGUGCUUACCCCUUAAGGCUAGCCCCUUCUUGAAU UGGUUGGGGCAACUAGUGGUGUACAGUUGGCGUGAACCUCUCCGGUUAAGGAGUGCUUUGCCCAU CCUCUGUGUGGCCUUGCAGUAGGGACUGGCAUCCUUCGCGUAGGUGAUCCGCGUGGCCAUGCCAU CCUGGCCAGCAGGAGGCCAGUGUGGCGCAACCUACGUCCUUCUGGGUGCUGCAUUGCAUUAACUUUG GAGUAAGCUUGGUGUGGCCAAACCCCGGGUUUACGUACCAUCUGUGUGUAGGGAUUGGCCGAG GUACCCCAUCCUUGAGGUGGGAUCUGACCGGUGAGCUAAUUGUCUAGCACCAUUCUUCUUCUUUUUC UUUGCUGGUCACG	Rabbit- Kobuvirus
1 6 0	UACGCAUGUAUUCACACUCAUUCUCCUCCUCCACCCUUAAGGUGGUUGUAUCCCAUACCUUACCC UCCUUCUCCACAUGGACGGACAUAUGGAUUGACUCCACGGCAACACAUAUGGUAUGAUUUCGGAU ACACCUUACGGCAGUAGCGUGCGAGCUAUGGAAAAUUCGAAUUGUCGAUAGCCAUGUUAUGAG GGCGUUCGGCGUGUCCUUGGUGUAUUGGGCGACUGGUUACAGGAGAGUAGGCAGUAGCUAUGGC AAACCUUACAGUAUUAUAGAGGGAUUGGCAUUGAGACUUGACGAGCUCUCCUGGAGAUUGU GGCGAUGCUCUUGGCAUUAACUAGUAGGCUUCCAGGUUGGGAACCUUGGACUGGGCCUUAUACUAC CUGAUAGGGUGCCGCGUGGCCCGCUUAACUAGUAUAGUCAGUUGAAACCCCCC	Porcine- Sapeloivirus- JD2011
1 6 2	UUGAAAUGGGUGUGGGGUAUGCGUAUUAACGGUACGCAUAUAUUCACACUCAUUCUCCUCCUCC ACCCUUAAGGUGGUUGUAUCCCAUACCUUACCCUCCUUAUAAACAGAUAGGACAAUUGGAUUGA ACUUAUGGCAAGUGAAUUGGUAUGACUUAUGGAUACUUAUACGGCAGUAGCGUGGCAGCUAUGG AAAAUCCGCAUUGUCGAUAGCCAUUGUAGUACGCGCUUCGGCGUGCUCCUUGGUGUAUUCGGCGA CUGGUUAACAGGAGUAGGCAGUAGCUAUGGGCAACCUUACAGUAUUAUAGAGGGAAUUGUC AAUUGAGACUUGACGAGCGUCUUAAGAGAUGUGGCGCAUGCUUUGGCAUUAACCAUAGUAGCUUC CAGGUUGGGAACCUUGGACUGGGCUUAUACUACCUGAUAGGGUCGCGGUGGCGCCUGUAACUAGU AUAGUCAGUUGAAACCCCC	Porcine- Sapeloivirus- A2
1 3	CUCACACUCAUUCUCCUCCUCCUGCUGUACAGGCAACCCCCACCCUUCUUCUACGGGAGUAGGGC CAAGGAUCUGUUGCAUAGGCGUUGUAUCCCUUACCUUUUACCUUACCAUCCCAUAGGACUGGUU UUCGGUUUUGAUUGAGUAUUGGAUACUGAUUUAUACCUUUAACCAUUCAGGGGAAAAUUGGAGUU UCUUUACUGGACUGACUUGAUUAGACCAAGAGUCAACAUUUGCGUGUUGGCCGUAUGGAUUGCUU UAAGGUUAUUCUUGGAUUAUGACUUCAGGGUUGGCCGCCAGGAUAAAGGCAUUGUGGUAAGU GAUGUUAUGCUAUGUGGUUGAAACCUCCUAAAGACGUCCUAGGUCUACGUGUGCGGGCCGAAGUA AGCUUAGGAUAACAGGGAGUAUGCCAUUUUCUGCUUUCACCCAAACAGACCGUACAGGAAAGAGCU AGAGGCAUUCUUGGGGCAAGGGGAAAGCUUUGCUUAGCCGGAUUGUUAUUGAGUCCUUGACGAAU GCGUCCCGUCUGCCGACGGUGAGGCGUAUGGCGCAUGCUAUGGCAUUAACCAUUGGUGUAUCUG UGAGGGGGGGCCUCCACACUAGUCUAGUCUACUUGACAGGGCCGGCGUGGCGUUGGUGUUGUUAU	Stntan- Sapeloivirus- 1

	GGUUAUACCAGUAGUAAUCCCCAUGGAUUGCUUUAAAUCCCCUCCUCCCUUACCAAGACAUCUCUAAG	
1 6 4	UUUUAAUUGUUAUGACAUUCAAAGGgAAAAUUGUCUUUUUAUUAUGGGACUGACCUGUUUAUGAAC AUGAGCAGCGGCACUGCUCCACGGGCUAUCCGUGUAAGAAUUAUUAUUAUGGAUCAUGAU UUCAGGGUUGGCGGCCAGUCUAAAAGGCAAUUGUGGUAAGCUAUGUAAGUAGUUGGCUUGGAAAG GAGCCAAGUACAUCUAGGUCUACGCUUGCGGGCCGAAGUAAGACUUGGAACAACUCUGAGUAGGC AGUUUUUCUCUUAGCCCCAACACGACCGCAUACUGAAGAGCUAGAGGCACUUUGGGGCAAAGGUAAA AGCAUUGCUUAGACCGAAUGUCAAUGAGACCUUGACGAGUGCUGUCACAGUGUCCCCUGAUGGCAG UAUGGCGCAUGCUCUUGGCAUUAACCAUAUGUGUAUCUAUAGGGGGGGGGCCCCUUAUCUAGUCU AGUGCUACCUGACAGGGCCGCGGUGUCGUGCGUGUGGUAUAACAGUAGUAAUCCCCAUGGA UUGCUUUAACUCCCCUCCUCCUCAACAAAACUUUCUUAAG	Stmnan- Sapellovirus- 2
1 6 5	CCGGGUAUAACCCGAGUUUUGGGGACGGUCCAAGCCCCACAUAGGAACAUAACGAUCCACGGAUCCU GUGUUCUUUUAUGCUUUUAACCUUACCCUUGUAACCAUUAACGCUUUAACGCGCAUGGUGUUUGGC GGCACCAGUAGCUGGACAAGAGGUUACGCCAUUACGAUAUGUACCCUCCCCUUUGGGGAGAGACCG ACCAAUUAUGGUACAGUAUCCAACUGUAUUGUGGUAAGUACUUCUGUUUCCCCGUGAUGCGGGAU AGGCUUUAACCCAGGCCAAAACCGUGUAUCCGUUACCCGACUCACAUUACGAGGAGGCUAGUAAA AGGCAUGAAGUUAAGAGUAUGAUCCAACAGAUCCCCACUGGUAACUAGUGAUGAGGGUUCGCU UCCGAACAUUGGCAACAUUGGGUUCUCCUGCGUUGGCACUAGGCCCUUCCGAGGGGUGUCUGAAGA UGGAUUGUUGAUGAAGACCAAUUGUGCAUGUGUUUAUCCUCCGGCCUUGA	Rabovirus-C
1 6 6	UUGAAAGACCUCGGCAUAUAUCGUUGUCACAACGGUAUAUGUCGAGAUUUUCCCCACCCCCUCC AAUUCUUUUUCCCCCUCUUGCAACUUAAGAGUUUGUACACACAGGGCAAUAGGAUACGUGAUCCAG CCAGGACAGUGAGCUCUACGACUUCUGUUUCCCCGUCUUUACGUAUACGCGGAUUGUAGUAA UUUGUGUGCAUUAUAGUAGGUUGAUCCGGGAUUAACCCCAAUCCAGAAACUGGUGAUGAGCGUU ACCACCCCCCGCGGGCGACCGGAAGGUUUCGUGCGUUGGCACAGGGCUUCGGCACAGAAAAGG UAAAGCAAUAGAAGCGCUACUGUGCUACGAGAAGUUUCCUCCAGGCCUUGAAUGCGGCUAAUCCU AACCAUGAUCCACCGGUGCAAACCAUGUACUAGGUGGUCGUAACGCGCAAGUCGUGCGGAAC GACUACUUUGGGUGUCCUGUGUUCCAUAUUUUUAUUUAUCAAUUUUUUGGUGACAAGAGUAAAGA GAUACAGAUUUGCAGCC	Boosepivirus -A
1 6 7	UUUUUCCCCUCCCCUCCCAACUACCUUUUCCCCCUUUGUAACGCUAGAAGUUUGGCAAAACCGCC UGUAGGGUACUGCAAUCCAGCAGUGCAUAGGCUAAGCUUUUCUUGUUUACCCACCCCAUUAUACU GAGGAGGAUUGUGAAAUUGUGUAGUAGGGUUAUGUAGCGGUGACCCGGGUAACCCCAACCCAGAAA CUCACGGAUGAGAUAACAGGACCCCAUGGUAACGUGUGUUCGUCUGCCCCGCAAGGUGAGGC CGUGAGAGCUUUGCAGCGGAAAACCUUGAAACCCAAAAGUACCUUGAGCUUUCGCUAUUUUGUGU UUCUCCAGGACCCUGAAUGCGGCUAAACCUAACC CGCAUCCGACUAGCAACCCAGCUAGAGUG UGGUCGUAUUGCGCAAGUUGCGGGCGGUACCGACUACUUUGGUGUUCUGUGUUUCCUUUAUUUUUAU UUUGAAUUUUUUGGUGACAACAGCUAGAAAUAAAGAGUGAAC	Boosepivirus -B
1 6 8	UUAAAACUGGGUUUGGGUUGUUCACCCCAACCACCCACGCGGUGUUGUACACUGUUAUUCGGUA ACCUUGUACGCCAGUUUUAUACCUUCCCCCCCCUUGUAACUUAAGAAGCAUGCGAAUCCGACCAUA GCAGGCAUUAACAGAUUGUCACCGGUAAGCACUUCUGUUUCCCCGGCUCUGGUUAUUGCUCC AACAGGGCAAAAACAAUUGGAGUCGUUACCCGCAAGAUGCCUACGCAAAACCUAGUAGCAUUCUCCA AGAUUUUUGGUGUGGUGCGUCAGUUGCUACCCAGCAUUAAGACUUGGCGAGUAGGCUAGAAUACCC CACUGGUGACAGUUGUACGCUUGCGUGGUGCCUGCACACCCACACGGGUGUGAAGCCAAAGAUUG GACAAGGUGUUGCAGAGUCACGUGUGCUCAUCUUGAGUCCUCCGGCCCCUGAAGUGCGCUAACCUUA CCCCGUGACCAUUGCUGCAAUCCAGCGAGUAUUGGUCGUAUAGUAAUUAACGGGAUGGGACCGA CUACUUUGGGUGUCCGUGUUUACCUUUUUAUUAUCAAUUGCUUAUGGUGACAAUAUAUAGAUUA UAUAUUGACACC	HRVA73
1 6 9	ACAUGGGCCAGCCCCACACACCCACUGGGUGUAGUAGUCUGGUUCUAUGGAACCUUUCUACGCCUCU UUUGCUUCCUCCCCCAUUUCUCCUUCGAUUGCUCCACCGUGUAUCUUUGCAACUUAAGAAGAAUAA UGAACCCGCAACAUAAGCGGGCGCUGAGCCACAGCGUCAUUGUGCAAGCAUUCUGUUUCCCCGGAU GGGCCCCAUAGGCGUGUACCCACGGCUGAAAGGGACCGGCCGUUACCCGCCUUGGUUACUGCGAGAAG UUAGUAACUCCCCUGAUAGCUUUAGGCGUUAACGCUACGCCUUUAGAGCCGAAGGGUAGUUCGGGUC GAUAGGCGUCGUUAUCCCCACUGGCGACAGUGUGACUUGCCUGCGUUGGCGGCCCGGGUGGGGGG CAACCCCAUCCACGCCUACUGAAGGACAGGGUGUGAAGGCGCUAUUGCGCUACUAAGGAGUCCUCC GGCCCCUGAAUGCGGCUAACCCGAACCCGAGCCACGGUGGUAACCCGCCACAAGUGGGUGCUAA UGAGUAAUUUGGGGAGGGACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUUCCAUAACGAUGGCU GCUUAUGGUGACAACCAUAGCAAUUGGAUUGGCCAUCCGGUGUUAUAUUGCGAAU	EV-L-1
1 7 0	UUUAAAAGAGCCUGGGGUUGUUCACCCACCCAGGAUGCCACUCGUGGGCAUGUGGUACUGUGGUA UCGCGGUUACACAGUACGCCUUCUUAGCCGCCUGACCCAAAGAAUUCUUUCCUUUAACAGAUUCUU CGUAACAAGGGCCUACGCCUGACGCAAGUGCCUCCAUAUGGAGUCUCUCCAAAGCCUCCGGGGCUUGG AGGGCGCCGACCCUUGCUAGGGGGAGCCACGACACGGCUGGAGUCCAUGGCACACCGCAGCCCA CGAUUCAAGCCAGAAUUGAAAGCGGGAAGCACUUCUGUCUCCCCGGUGUGGAUCAUACGCUUACCC ACGGCGAAAAGUAAGCAUCGUUACCCGACUCGGUACUUCGAGAAGCCAGUACAGUUGGUAUCUC UGCAGGGUAUACGCUACGCGUACCCUACGUAGUUCUUGAGAUGGCGUGAGAGAACACCCACGGG CGACCGUGUCUUCGCGCGUGGCUAAGGCGGGCCUUCAGUGGCUUGGUGCCUUGCAGAGUAGG CCUCCGAACAGCCUAUUGAGCUACCGUUUAGCCUCCGCCUUCUUGAAUGCGGCUAAUCCUAACCAUG GAGCGCCCGCCACAGUCCAGUGGUGAGGCGUCGUAACGCGCAAGUCCUGGCGGAACCGACUACU UUAGAGUGGCGUGUUCCAUAUUUAUCCUUUAUAAAGUUGCUUAUGGUGACACCAAGAGAUCCACG AUUUUUUGUUUUUAUCACUGAGACACAAGUCAUAUUAUCAAUCUUUAUUGCGGAUUUAACUUGGU GCGUCCAACACAUACG	EV-K-2

1	UUA AAAA UAG CCU CAG GGU UGU UCC CAC CCU GAG GGC CCAC GUG GC GUA GU AUA CU CUG GUA UCA AAG GUA	EV-J-1631
7	CCU UUG UAC GCG CUA UUU UUA UUU CCU UCC CCCC CAC AGA UAC UUA GAA GCG UUA UCUCUA UAG UUC AAC AG	
1	UAG GGU CAC UUA CCA AG UGG GCG CAG C GAA CAG CAC UUC UGU UUU CCCC GGU CUC UAG UAC CUG UGA AG	
1	CUG UAC CCA CCG CAG CCG GAG GGG AAA AGA UCG UUA UCC GCG CCCC UAC CUC CGG AAA GCG CUA GUA A CAC	
1	CAU UGA AG CAA UCG AGU GUU GCG CUA CAG CAC AGA UA ACC CCU GUG UAG CUU UGG UUG AUA GAG UCU GGG	
1	CAC UCC CCA CUG GCG AC AG C GCG C CAG GC UGC GUU GCG CGC CAA CCG AC CUG CGG CCA CCG GGU CCG GA	
1	CGC UCG UUU UGG UGA CAG UGU GUA GAG GCG CUA UCC GCG UAG CUC CUC CGC CCCC UAG CUC GUA AG C	
1	GGU AUA UCC UAA CCCC GAG CAC C CAC ACU CAA UCC CAG AGU G CAG GAU GUC GUA A CCG CUA AG GUC UG	
1	GGAC GGA ACC GAC UAC UUU UGG GUG CUC CGU GUU UCC UGU UUA CUA UAC UUU UGG CUG CUU AUG GUG ACA	
1	AUC UAG UGU UGU UAC CAA UAG CUA UUG GAU UGG CCA UCC CGG UGU UUG AAU UGUG UGU UUA UAC UA	
1	AUUC UUU UAC AUA UCC AGA CACA ACC AAAU	
1	UUA AAAA UAG CCU CAG GGU UGU UCC CAC CCU GAG GGC CCAC GUG GC GUA GU AUA CU CUG GUA UUA CCG GUA	EV-J-N125-J108
7	CCU UUG UAC GCG CUA UUU UUA UUU CCU UCC CCCC UUG UAA CUA AAG CAA AG CAA ACC GUA UCA AUA GUA	
2	AGCAACACAA CACC CAG UGU UUG AG CAG AACAAGUACUUCUGU UUC CCG GAG GGU CUA GCG UUA AGCU	
1	GU ACC CAC GCG CUA GUA GU AG C CUA CCG UUA ACC C GCG UAC CUA CU CCG AGA AG UCU AGU AUA UAC CAU	
1	UGA AGU UUU UGU UGG CGU UAC GCU CAA CAC ACU ACC C GUG UGU AG UUU UGG CUG A GUA GUA CAC GCG CA	
1	UU CCCC CAG GCG CAG CUG GCG CUG CGU UGC GCG CCA CCA CAG GGG CCG CAA GCU CUC CUA GAC GU	
1	CAC UUA AC CAG CAC UGG UGU GUA AGA ACC UUA UUG AG CUA GGU AUA GUC CUC CCG CCCC UGA AU GCG GC	
1	UA AUCC UAA CUC CCG AG CACA UUC AG UGA CAA C C CAG CAC AUU UGG UGU GUG UUA UAC CCA AGUC UCG GAG	
1	CGGA ACC GAC UAC UUU UGG GUG CUC CGU GUU UCC UGU UUA C C UUA UUG GCG CUG UUA UGG GUG ACA UUU	
1	UGA UUA UGU UUA CCA UUAUG CUG UUG GAU UGG CCA UUC CGG AUU UUG GAG CAC CCA AAG C UUC UCU	
1	CUC UAC UUC AGA UUA CAG UG CGA AGU UUC C UUU UCA UUA UUA C UUA C UAA UUA UUG AGU ACC CAA AG	
1	UUA AAAA CAG CCUG GGG UGU UCC CAC CCCC CAG GGC CCAC GUG GC GCA AGU AUA CU CUG UUA CCG UUA GA	EV-1
7	AGCCUUGUAG CCG UGU UUU UCC CUC C UUA AACA AUA AAG UUA CCG CAC UUC CAG GGG AGU AG CCG C	
3	ACU CCG CUC C GAU AC GUC GCU C CAG UGC GCU AGU AC GCU AUA GU AC UUU UC CAC GGG GAU AG UGG UAU	
1	CCU CCCC UAG UAA UUU UAG AAG ACUUG AAAA CCG ACC AAU AGG CACC UC C GCA UCC CAG C GGG GUA AAG	
1	GUC AAG CAC UUC UGU UUC C CCG GGU CAG UAG C GAU AG AC UGU GCC CAC GCG UCA AGG UGA AAC CAA C	
1	CGUUA UCC GAC UCU UUG AACU CCG GAG CCG UAGU AC CAC CAA AGA UUG CUG AUA UGC UUG GGG UUA CCG CAG	
1	CAC GACC CUG GUG UAG AUC AGG CCG AUG GAUC ACC CGCAU UCC CUC ACAG GC GAC UGUG GCG GUG GUC G	
1	CGUG GC AGC UUG CCG AUG GGG CAA C CCAU CCG CAG CCAAG CAA UUA G CAG GGG UGU GAAG CAG C CUA C	
1	GAG CUA CAA AGUA UUC CUC CCG CCCC UGA AUUG CCG C UAA UUC CCA CCG CAG GAC AAU UG CCG CAA C	
1	CCAG GU AGUG AAU GUC GUA CCG GGU AACUCUGUGGC GGA ACC GACUACU UUG GGG UGU CCGUGU UUC	
1	CUUUUAAU UUAUC AUAUUGCU GAUAUGGUGACAACUAUAGUGCUAUCUC GAU UUGCAUUA CUAUUGUUG	
1	AGAUUAAAACUUUAUUAUCUUGU GCAU UUA C C C UUA UUGAGUGAUUUUACCUGAAC CAGAUUAAU	
1	UACUCAUC CCGUUAUAUAUAUUAACAGCAGAAUAUCUUGCAAAG	
1	UUUAAAACAG CCUG GGG UGU UCC CAC CCCC CAG GGC CCAC GUG GC GCU AGU AUA CU CUG GUA CCG CUA GU	EV-F1-BEV-261
7	ACCUUUGUAG CCG UGU UUU UCC CCCC CUC UAA UAAU AAG UUA CCG UUA CUA CUG AGGG GAGU AGUC	
4	CGAC CCG CUC CAG CCAU GCG CAC CAG CAG CUG GUA CCG UAC GU A CUA CUU UAC CCG GAG UAG UGU	
1	AU CCCC UUA CCCC GGA ACC UAGA GAUUG CAC ACA AACC GAC CAAU AGG CCG CACC GCAUCC AG CCGUG	
1	CAG CCG UCA AG CACUUCUGUCU CCCC GGU CUG UAA GAU CUG UUA UCC GCG C GACC CACUAC C GAA AAG	
1	CCUAGU AACUGGCCAAGUGAAGCGGAAGUUGCGCUC CCGCCAC CAC C C CAG UGGUAG CUCUGGAAGAU	
1	GGGGCUCGCAC CAC CCGUGAAGUAACG GUGUUAACGGUUGUC CCGCGCUGUC CCGUGUC CUGUC CUGUG	
1	CCG UUCACUUC AACUUC CAG CCA CCG CCAAGAGCCUAUUGUGUCUGGGACGGU UUUCC CCG GGG CCG	
1	GUGAAUGG CUCUAAUCCCAAC CUC CCGAG CCGUGUGCGCACA UCC CAGUUGUCUACGUCGU AACCGGU	
1	AAGUUGGAGCGG AAGACAG CACUUCUUGCUA C CCGUAC CCG CUGU UCC CUGU UUAUUAUUAUUAUUAUG	
1	GUGACAAUUGUUGAGAUUUGCGCUCUUGCAACGUUGCCAUUGAAUUAUUGGCUUAUACUAUUAUGGUUG	
1	CCUUUACAA AACCUCUGAUUAUAC CAGUUCUUA CUAUUGAUCUGCUUGU UUUUCUCAAUUAUGAGUA	
1	UAGACUACAAUAGCAAA	
1	UUAAAACAG CCUG GGG UGU UCC CAC CCCC CAG GGC CCAC GUG GC GCG CUA GU AUA CU CUG GUA UAC CCG	EV-D94
7	UACCUUUGUAG CCG CUG UUU UUAU AUC C C C UCC CCG CAA CUA GAA GAAA ACAAU CAA GUUCACUA	
5	GGAGGGGGUGACAAACCAGUAC CAC CAG AACAAG CACUUCU GUU UUU CCCC GGU GAGUCUGUAUAGACU	
1	GUAAGCAGGUGUAGAAACGAUUGAAC CCGUUAUCCGCUACUUGUACUUGUAGAAAGCCAGUAUACCUU	
1	GGAAUCUUGCAGUGCGUUGCGCUCAGCACUCAAC C C CAGAGUGUAGCUUAGGUCGAGUAGUCUGGACA	
1	CUCUCCACCCGCGACGUGUGUC CAGGCUGCGUUGGCGGGCCUAC CUGUGGUC CCAAGGCCACAGGACGC	
1	UAGUUGUGAACAAGGUGUGAAGAGCCUAUUGAGCUACAGAAGAUUC CUCGCGCCUUGAAUGUUCGCGCU	
1	AAUCCUAAACCAGGAGCAAGGGUACACAAACCAGUGUAUUCUUGGCUAACCGCGCAAGUCUGUGGC	
1	GGAAACCACUACUUGGGUGUC CUGUUC C UUUUGU UUUUAUAGUGGUCGUUAUGGUGACAAUCU	
1	AAGAUGUUAUCAUAUAGCUGUUGGAUUGGCCAUCCGGUAAUUAUUGAGAUUUGAGCAUUGGCUUG	
1	UUUCUUAACCAAUUAUCGAUAUUAUUGCAUUAUCGACUGCAA	
1	UUAAAACAG CUCUG GGG UGU UCC CAC CCCC CAG GGC CCAC GUG GC GCG CUA GU AUA CU CUG GUA UAC CCG	PV3
7	UACCUUUGUAG CCG CUG UUU UUAU AUC C C C UCC CCG CAA CUA GAA GAAUACA UUA CAG CUC AUA AG	
6	GAGGGGGUGCAG CCG CAG CCG CUC CCG GGG CAA GCA CUA CUG UUU UUU CCCC GGU GAGGCGUA GAC CUG	
1	UU CCCC CCG UUG AAG UGG CCGAUCCCGUUAUCCGCUCAUGUACUUCGAGAAGCCUAGUAUCG CUCUG	
1	GAAUCUUCGACGCGUUGCGCUCAGCACUCAAC C C CAGAGUGUAGCUUGGGCCGAGUAGUCUGGACAG	
1	UCC C CAGUGGC CAGUGGUC CAGGCUGCGUGGCGGCC CAC CUGUGCC CCAAGCCAC CAGGACCGCU	
1	AGUUGUGAACAGGGUGUGAGAGCCUAUUGAGCUACAUGAGAGUCCUGGCGCCUGAAGUGCGGCUA	</

	UUGGAUCCACUCCCGAAACGUUUACUCCUUUACCUUUAUGAAAUUGUUUGAAGACAGGAUUUCAGUG UCACA	
1	UUAAAAAGCUCUGGGUUUGUCCACCCACAGAGGCCACGUGGCGGCUAGUACUUGGUUACACGG	EV-C102
7	UACCUUUGUAGCCUGUUUUAUACCCUCCUCCACGUACUUUAGAAGCAAUUAACAAAGUUCAAU	
7	AGAGGGGUAACAACACAGUAUACACCAGCAACAAGCACUUCUGUUUCCCCGGUGAUUUUACUAAGCU	
	GUGCCACCGCUGAAAGUAGUAUACCCGCUUACCCGCUAGUACUUCGAAAGCCUAGUAUCCGUU	
	GGGAUUCUCCGACCGUUGCGCUAGCACUUAUACCCGAGUGUAGUAGUAGUAGUAGUAGUAGUAGU	
	UCCCCAUCGCGCAGCAUGGCCAGGCGUGCGUUGGCGGCCUACCCAUGGCUUACGCCAUGGGACGCUA	
	GUUGUGAACAAGGUGUGAAGAGCCUUAUUGAGCUACUCGAGAGUCCUCCGCGCCUGAAUGCGGCUAA	
	UCCCAACCCAGCAUCGUGCCUCCAAACCCAGGAGGUGGCCUGUCUUAACCGCGAUGGUGCGG	
	AACCGACUACUUGGGUGUCCGUGUUUCCUUUUAUCUUUUAAGUGGUGCUUAGUGGACAAUUAU	
	GAUUGUUUAUCAUAAAGCGAAUUGGAUUGGCCAUCCGGUGAAUACAACACAUUAUUUACUUGUUUG	
	UUGGAUUUACUCCGCUACACAGCUUACUCCUAGAUAUAUUUAUUGUAUUGCUGGUAAGGAGACA	
	CUAUUAUA	
1	UUAAAAAGCUCUGGGUUUGUCCACCCACAGGGCCACUGGGCGCCAGCACACUGGUUACUGUGGU	EV-30
7	ACCUUUGUGCGCCUGUUUUUAUACCCUACCCUACUAGUACUCUAGAAGCAAGGCAACUUGACCA	
8	AUAGUAGGUGUGGCACACACCGCGCAUUUUGGUCACAGCACUUCUGUUUCCCCGACCGGAUAUCAAU	
	AAGCUGUCUACCGCGGCUAGAGGAAACCCGUCUUAUACCCGACAGCUACUUGGAGAAACUAGUAA	
	CACUAUGAACGUUGCGGAGUGUUUCCGUUACGACUUCUCCCGUGUAGAUACAGGUCGAUGAGUACCG	
	CAUUCUACCGGUGAGCCUGGCGGUGGCGUGCGUUGGCGGCCUGCCUACGGGUGUCCGCCUGAGGACG	
	CUCUAUACCCAGCAUGGUGUAGGUCAUCCUUAUGAGCUAGGUAUCCUCCGCCUUGAUGCGG	
	CUAAUCCUAACUGCGGAGCAGGUGUCACAGACCAGUGAGUAGCCUGUCUUAACGGGCAACUCUGCA	
	CGGGAACCGCAUCUUGGGUGUUUUUCCUUUUUUCUUAUUAUUGGCGCUUAGUGGUAACAUAU	
	UAAAGAAUUGUUUACCAUAUAGCUUAUGGAUUGGCAUCCGGUGAGCAGAGCGAUGGUGUACUCU	
	UUUGUUGGAUUUGUACCUUUGAACCAACAAGUCUUGAAUACCAUUCUUCUUAUUUAAAGUUAACU	
	CAGCUAAAGAGAA	
1	UUAAAAUAGCCUGGGUUUGUCCACCCACAGGGCCACAGUGGCGUAGUACUUGGUUUAUCCGGUAC	SA5- enterovirusB 114
7	CUUUGUACACCUUAUUUACAACCCUACCCUUGUAACCUUAGAAGCAAUUUAUUUAAACCGCUACUAG	
9	GGGUGUGCUAUUCCAAGCAUCAUAGAGCAAGCACUUCUGUCUCCCGGGAGGGGCUAAUGGUACGC	
	UGUGCCACCGCGGAAUUGAGCCUACCCGUUAACCCGGAGUCUACUUCGGGAAGCCGAAUACUACA	
	UUGAAACUUUGAGCGCUUACACGCUACGACCAUAACCCAAUGUGUAGUUCUGGUCGUAAGCCUUGG	
	AUCCCCACAGGCGACUGUGGCCAAGGUCGCGUUGGCGGCCAGCCUGCGGACCAAAAGUCCGUAGGA	
	CGCCUAAUUGUGGACAUUGGUGUGAAGAGCCUACUGAGCUAGACUGUAGUCCUCCGCCCCUGAAUUG	
	GGCUAAUCCUAACCCUGGAGCAUCCCGGUGCAACCCAGUACGUAGGUGGUGCUUAUUGCGUAAGUCUG	
	GGUUGGAAACCGACUACUUGGGUGUGCGUGUUUCUUGUUUUAUACUUGGCGCUUAGUGUUUACA	
	CUAAUUGUGUAUUAUUAUUGGAGUGCGCGUGACACGCGAUUAUUGUAUUAUCCAUUUGUUGGAUA	
	CUCCAAUAGUGUCAUCUUAUACACAACAUUUUACCACUGAUCAGAAUAAA	
1	UUAAAAAGCCUGGGUUUGUCCACCCACAGGGCCACUGGGCGGACGACACUCUGGUUUCUCCGGA	EV-A114
8	ACCUUUGUGCGCCUGUUUUGAAACCCUCCUCCCAACUCGAAACGUAAGUAUUGUACACUACUGA	
0	UCAGUAGCAGGCGUGGGCCACAGCCAUUGUCUGAUACAGCACUUCUGUUUCCCCGGACUGAGUAUC	
	AUAGACUGUCACCGCGUGUGAAGGUGAAACCGCCGUUACCCCGCUAACUACUUCGAGAAACUAG	
	UAGCAACUAGAAACUGCGAGUGUUUCCGUCGACACUUCUCCCGUGUAGUACUGGUGAGUGAUCU	
	CUGCAUUCUCCACGGGUGACGUGGCAUGGCGUGCGUUGGCGGCCUGCCUUAUGGGGCAACCAUAGG	
	ACGCUCUAAGGUGGACAUUGGUGAGAAGUCUUAUUGAGCUAGUAGUAGUCCUCCGCCCCUGAAUG	
	CGCUUCCUUAACUUGAGUGGAGCAUAUCUCCAAACAGGAGGAGUAGGUGGUCUUAAGGGCAACUCC	
	GCAGCGGAACCGACUACUUGGGUGUCCGUGUUCCUUAUUUUAUCCUAUACUGGCGUCUUAUGGUGAC	
	AAUUGAGAGAUUGUUAACCAUAUAGCUUAUUGGAUUGGCCAUCCAGUGUGUAUAGAGCAUUAUUAU	
	CAAUUUGUGGAUUUACUCCAUUAACCCACACGUCUCUCAACACACUACAUAUUUACUUAUACUGA	
	ACACUAGAAA	
1	CUCCCCCCCCUCCCUUCCCGAGUAGGAGAUUGGCAUGUAUGUCUACAUUGCCCGAUUCUCUCUUG	Burpengary- virus
8	CUCACUCUCUUAUAAUCCUGGUGGCGGUCUCGGAUUAACAAUUAUUGUCUACUGGGUAGUCUUAU	
1	UUGGUGGUAUUCUCCUUGUACGGGACUCCGACUGCCGUAUAAAGUCCUACAGAGGGCA	
	AUGUACAAGGAAGUCAUUAUACGCUAAUUAAGUAUUUGAUGAAGUACUAGUGUGACAGGGCUGAGGA	
	ACUCCCCCGGGUAACCGGUGGUCACGCGUCCGAAAGACACGUGUAUAGGAUACCCUUGUAUUAUC	
	CAGCACUGUAUUAUAGUCAAAUCCUAGUUAUUGUAGGAUGUAUAAUUGGCUAUGUAUAAUUAU	
	GGCGUAGAGAUUGUCCGAGGUACCCCGUUUACGGGAUCUGAUCGGAGGCUAAUUAACCAUUGCG	
	CCCUAAAUAAUCUACAUAUAUUUUCUUUUUUAUUCAAA	
1	AUGGGGAAGGUAUGACGUGGCCCCUCCUUCUUGGAGAACUCCGCUAGUGGUCUUUCCACUUCUG	GFTV
8	GAAAAGAGUGAGGUCAGUUAUAGAGACGUCUGAAGACGACAAUUAUCCUGGUGUCUUAUUCUAAGA	
2	CGUUUCACAGCUGUAGCGACCCUCAGUAGCAGCGGAAGCCCCUCCUGGUGACAGGAGCCUCUGCG	
	GCCAAAAGCCACUGGUAUAGAUUCCACUGUGAGGGCGGUGCAGCCUAGCACCCUGUAGUGCAUAC	
	UAGUGUACGUGGCGCGGACUAUUGGUGUCUUAUAGACACUUGAAGAGACCAUUAUUGGUGG	
	AGGUACCCCGGUGCGGGAUCUGACAGGAGACCAUUGCCCAAUGCUUUAACACGGGUCUAUGGUUU	
	AAAAACUGUCGACUUCUCCAAACCAAGUGGUCUUGGUUUUCAAUUAUUAUUGAAUUAUUCACU	
1	AAGACCAUAUUAUUAUCCUUGAAUGCGUGGCUAAUACAGAGGAAAGUCAGCAUUUCCGGCCACGC	SAFV-V13C
8	UAAGAGAAACAUUAAGUUAGAUUUGAUUUAAUCCUUGAGGAAUUCGAAACAGAGAUUCCUCCU	
3	CCUUGCGCGUUGUUCACACCCACUUAUUAUUAUUCGGCCUUUUGUGACAAGCCCCUCGGUGAAAGAA	
	CCUUCUCUUAUUCGACUGGUGGGAUUGCCAUUUGCCAGCAAAGCUUAUACUAGCCUCCGCCGA	
	UUUAGUGAUGUUUUUCCGCGGUGGCGGAGCAUUCUUGGUGUAGGAAACUUGGUAUUUUUUAU	

	UGGAACCUUGGUUCCCCCGGUCUAGCCGCUUGGAAUAUGACAGGGUUAUUUUUUGAUCUUAUUUC UACUUUUGCGGGUUCUUAUCCGUAAAAAGGGUACGUGCUGCCCCUUCUUCUGGAGAAUUCACACG GCGGUCUUUCCGUCUCUACAAGUGUGAAUGCAGCAUGCCGAAACGGUGAAGAAAACAGUUUUUCU GUGGAAUUUAGAGUGCACAUCGAAACAGCUGUAGCGACCUCACAGUAGCAGCGGACUCCCCUCUUG GCGACAAGAGCCUCUGCGGCCAAAAGCCCCGUGGAUAAGAUCCACUGCUGUGAGCGGUGCAACCCCA GCACCCUGGUUUGAUGAUCAUUCUCUUGGAACAGAAAUGGUUUUCUACAGCCUCCGGUAGAGA AGCCAAAGAUUGCCUGAAGGUACCCCGCGUGCGGGAUCUGAUCAGGAGACCAAUUGCGCGUGCUUUA CAGUGCACUUUGGUUAAAAAUGUCACAGCUUCUCCAAACCAAGUGGUCUUGGUUUUCCAAUUUU GUUGAAUGGCAAU	
1 8 4	GGGAUCUCCCCCGCGGAGCUGGUUACAGCUGUCGCGGACGGAGAUCAUCCCCCGGCCACCCCCUU UCGACCGGGUACUGCGAUAGUGCCACCCAGUCUUUCCUACUCCCGACUCCCGACUCUAACCCAGG UUCCUUGGAACAGGAACACCAAUAUACUAUCCCCUGGAUGCUGACUAAUCAGAGGAACGUCAGCAU UUUCCGCGCCAGGCUAAGAGAAGUAGUAAGUUAGAUUCCAAUUGAUUUUAUCAUCCCCUUGACGAA UUCGCGUUGGAAUAGCACCUCUCACUUGCCGUCUUCACACCAUUAACUUGAUUUGCGCCUCUGUGU UGAGCCCCUUGUGAAGUGCUUCCUCCCAUCGUGACGUGGUUGGAGAUCAAGUCAACCGACUCCGA CGAAACUACCAUCAUGCCUCCCCGAUUAUGUGAUGCUUUCUGCCUGCUGGGUGGAGCAUCCUCGGG UUGAGAAAACCUUCUUCUUUUUCCUUGGACCCCGUCCCGGUCUAGCCGCUUGGAAUAAGACA GGGUUAUCUACCCUUCUUCUUCUACUUAUAGUGUUCUAUACUAUGAAAGGGUAUGUGUCGCC CCUUCUUCUUGGAGAACACGCGCGGCGGUUUUCCGUCUCUGAAAAGCGCGUGUGCGACAUGCA GAGAACCUGAGAAAGCAGUUUGCGGACUAGCUUUAUGGCCACAAGAAAACAGCUGUAGCGACCA CACAAGGCGAGCCGCCUCCUGGCAACAGGAGCCUCUGCGGCCAAAAGCCAGUGGAUAAGA UCCACCUUUGUGUGCGGCACAACCCAGUGCCUGGUUUUUGUGACACUUCAGUGAAAACGCAAA UGGCGAUCUGAAGCGCCUCUGUAGGAAAGCCAAGAAUGUCCAGGAGGUACCCUUCUCCUGGGAAGG GAUCUGACCUGGAGACACAUCAUGUGCUUACACUGUGCUUGUGUUUAAAAAUGUGACAGCUU UCCCAACCAAGUGGUCUUGGUUUUACUCUUUAAACUGAUUUACU	SAV-P-113
1 8 5	GACCCUAAACCCAGGUUCCUGGGAUCAGGAACACAAUUAUCUACUCCCGUGAUGCUGACUAAUCAG AGGAAAGUCAGCAUUUUCGCGCCAGGCUAAGAGAAGUAGAUAAAGUAGAUUCCAAUUGACCUACC AUCCCAUGACGAAUUCACAUUGGAAUACACCUCUCACUUGCUGCUUACACCCAUCAAUUAU UCCGCCCACUGUGUGAGCCUCCGUUGAAGUGUUUCCUCCUUGGACGUGGUUGGAGAUCCAAAG UCAUCCGACUCCGACGAAACUACCAUCAUGUCUCCCGAUUAUGUGAUGCUUUCUGCCUGCUGGGU GGAGCAUCCUGGGUUGAGAAUUCUUCUUCUUCUUGGACCUCGGUCCCGGUCUAAAGCCG CUGGAAUUGUGCGCCUUCUUCUUGGAGAACGUGCGUGCGGUCUUCUCCGUCUUCGAAACGUG CGUGCGACAUAGCAGAUAAACGCAAGAAAGCAGUUCUUGGUCUAGCUCUGGUGCCACAAGAAAACA GCUGUAGCGACACACAAGGCGAGCGAAACCCUCCUGGUAACAGGAGCCUCUGCGGCCAAAAGC CAGUGGAUAAGAUCCACCUUUGUGUGCGGUGCAACCCAGCACCUGGUUUUUGGUGACACCUUA GUGAACCCUCGAAUGGCAUUCUACAGCGCCUCUGUAGGAAAGCCAAGAAUGUCCAGGAGGUACCCU UCUCUAGGAGGGAUUCUGACCUGGAGACACAUACACGUGCUUACACUUGUGCUUGUGUUAAAAA UUGUCACAGGCUUCCCAAACCAAGUGGUCUUGGUUUUCCUUAACUUGGAAAGUCACU	VHEV- cardiovirusB 2
1 8 6	CACCCCUUUGGACGCGGGCCUGCGAUGAGGCCACCCCGUCCUCCUUAUCCAAUACCCGACC CUAACCAGGGUCCAGGAACUGGAUCAUAACAAGUCAUCCCGUAAUGCUGGCUAAUCAGAGGAAA GUCAGCAUUUCCGCGCCAGGCUAAGAGAAACACAAUAAGUUAAGAUUAAUUAUCCCUUGACG AAUUCGCAACGAUAAAGUCCUCCUCCUUGCCGUCGUUACACCCAGAACUAACAAUUCGCCCCU UGUGACGAGCCCCUUGGUAAGGACCUCUUAUUGCGACGUGGUUGGAGAUUAAACCGACUCCGA CGAAAGUGCUUUAUGCCUCCCGAUUAUGUGAUGUUUUCUGCCUGUGGGCGGAGCAUUCUCGGG UUGAUACACCUUGAAUCCUUAUCCUUGGACCUCAGGUCCCCGGUUAAGCCGCUUGGAAUACCUA AGGGUUAUUUCCAAUUCUUCUUCUACUUAUUGAGUCCUUAUUAUGAAAGGGUUCUGUGCUGC CCUUCUUCUUGGAGAAUCUGCGCGCGGUCUUCUCCGUCUCUGAAAAGCGCAGAUAGCAGUGCU GGAACCGGUGAAGAAACAGUUCUUUGGGAACUUAAGAGCAGACAUUGAAACAGCUGUAGUGACCU CACAGUAGCAGCGGAACCCUCCUGGUAACAGGAGCCUUGCGGCCAAAAGCCCGUGGAUAAGAU CCACUGCUGUGAGCGGUGCAACCCAGCACCUGGUUCGAUGGUUUGUCUUGUGGAACAGAGAAU GGUCUUUCUACAGCCUCCAGUAGAGAAGCCAAGAAUGUCCUGAAGGUACCCCGCAUGCGGGAUCUG AUCAGGAGACCAUUCGUCAGUGCUUACACUGGCGUUUGGUUAAAAACUGUCACAGCUUCUCCAA ACCAAGUGGUCUUGGUUUUACUUAUUAACUUAACUGUUUC	TRV-NGS910- cardiovirusB 4
1 8 7	AAAUACUGGUCGAAACCGCUUGGGAUAGACCGGGUUUGUUAUGUCUCAUGUUAUUCUCCACCC AAUUGACGUCUUUUGUCAAUUGGAGGCGAGUAAACCUUGCCUUCUUGCAGAGGAUUCACAG UGGUCUUUCCGUCUCGACAAGGGAUUAUGAUCCACCAAAAGUUGUGAAGAGAGCAGGUCCCAUG GAAGCUUUCUGACGACUGAUGACUGUAGCGACCCUUGCAGGACGCGGACCCCCACCUUGGUG ACAGGUGCCUCUGCGGCCAAAAGCCACUGUUAACAGACACCUGCAAGGCGGCAACCCAGUG CCUCAUCAAAAGUCUGAUGACUGUGGAAUAGUCAACCGGCUUUUUAAGCAAAUUGGUGUCGGG GCUGAAGGAUGCCGGAAGGUACACACUGGUUGAUGAUCUGGCGGCCACAGUACAUGUGCUUU ACACAUGUAGCUGCGGUUAAAAACGUCUAGGCCCCCCGAACACGGGACGUGGUUUUCUUGGAA AACCACGAUUAACAU	EMCV2- RD1338- cardiovirusA 2
1 8 8	UUGAAAGCCGCGGGGUGGAGAUCCGGAUUGCCAGUCUGCUGCAUUCGACGGCUGGGUCCGUGACUA CCACUCCCCUUAACGUGAAGGCUACGAUAGUGCCAGGCGGGUACUGCCGUAAAGUGCCACCCC AAACAAACAAACACCCCCCUAACAUUACUGGCCAGCGCGCUUGGAAUAAAGCCGGUGUGCGUUU GUCUAUUGUUUAUUUCCACCAUUGCCGUCUUUUGGCAUUGUGGCGCCGGAACCUUGGCCUG UCUUUUGACGAGCAUUCUAGGGGUCUUCUCCUUCUGCCAAAGGAUUGCAAGGUCUGUUGAAUGU CGUGAAGGAAGCAGUUCUUGAAGCUUUGAAGACAACACGUCUGUAGCGGCCUUGCAGG	EMCV1-JZ1203

	CAGCGGAACCCCCACCCUGGCGACAGGUGCCUCUGCGGCCAAAAGCCACGUGUAUAGAUAACCCUG CAAAGGCGGACAAACCCAGUGCCAGCUUGUGAGUUGGAUAGUUGUGGAAAGAGUCAAAUGGCUCUC CUCAAGCGUAUUAACAAGGGGCGUGAAGGAUGCCAGAAAGUACCCCAUUGUAUGGGAUCUGAUCUG GGCGUCGGGCGACAUUGCUUUAACUGUGUUUAGGAGGUAUAAAACAGCUCAGGCCCCCGAACCA CGGGGACGUGGUUUUCCUUUGAAAAACACGAUGAUAAU	
1	AUGUGUCGAGCCACUUGGAAUAGAACCAGGCGUGCGCUUGUCUUAUGUUAACUUCACCAUAUUC	EMCVI-AnrB-3741
8	GCUCUUUUGCAAUGUGAGGGCCCGGAACCCUGGCCCGUGUCUUUUGAGCAGCAUUCUAGGGGACUU	
9	UCCCCUCUCGCCAAAGGAAUGUAAGGUCUGUUGAAUUGCGUGAAGGAAGCAGUUCUCUGGAAGCUU CUUGAAGACAAACAGCGUCUGUAGGACCCUUGCAGGCAGCGGAACCCCCACCCUGGUAACAGGUG CCUCUGCGGCGGAAAGCCAGUGUAUAGAUAACCCUGCAAAGCGGCGACAACCCAGUGCCACGUU GUGCUUUGGUAUGUUGUAGGAAGAGCAAUUGGCUUUCGCCAUGAGUAUUCAAACGAGGGCGUGAAGG AUGCCCAAGAAAGUACCCACUGGUUGGGAUCUGAUCUGGGGCGUCGGUGCAGGUGCUUUAACCCUGU UGAGUCGAGGUUAAAAACGUCUAGGCCCCCCGAACCAGGGGACGUGGUUUUCCUAGAAAAACCAU AGACAAU	
1	CUCUCUCUUUGACCCUGGCUACUGGGGUUCAGUAGCAUCUUUCUUUUUCUUAACUUAUCUUACUUUC	Malagasiviru s-B- Lesavirus-2
9	GACUUCUUUCUUUAUCUUAUCUUUACUUUUCUAUCUUCUUUUUAUUCUUAACUUUAUUAUCAAAC	
0	UUUUUCUAAUUCUUUUUUUUUUUAUAGUAAACUUCUUCUUUUUAGUUAUUUUUGCCUACUUGUGA GCUUUAUGCGGACCAUCUUAUUGUAGACAACCCCAACUUUGCAUGAGUAUACCAACCAUUAACG AUUACUUUUUAACCGUCUGACCUUUUGAUAAACACUGAAGUUAAGGCGUGAAACAUUCAUUUAACCA AAGUAGCCCCCAUUUCCCAACUACCGUGGGGGGGUACCCUACUGGCUUUGGAACUUGAGCCAUUA UGUGUUGCCGCGUUUCAGGACUACCAACACACAGUUCUUCUACAAGUUAUUGGGUGAGAUG CAGUGACAUGAACAGUAUCUAGUAGUACAUAAGACUACAAGCCUAGUUGCCUGCGGAACAACAUUGG UAACACUAGGCCAGGGUCCAAAGACAAGGGUUAACAGCCCCUUCUAGGUGUCUGUGUGUAGAAGAA UACUUUAUGAGUGUUAUAGUACCUUACCUUUAAGUACAGAAGUAGGUGCUGUGGUAAGGAUGUC CUACAGGUACCCAUAUAUAGGAUCUGAGUAGGAGACCAUGUGUGGUGCUUUAACGCCAGGUGAGUG GUUUAAAAAGCGUCUAGCCAAGCCAACGACACUAGGGAUAGUGCUUUCUAUUAUUUAUUAUUUCAGU GUAUUGGUGACAA	
1	CCCCUUUACGUAACGCAACUUAAGAGUACCCUACUGCAUUGGAUGUGUGGUAACUUUUACGCAC	Tottorivirus
9	ACAUUUGUAGUAGUGUUAGUUAUGUUCUACCUAAUGAGUAUGCAUGCACCUGCGAAACACGCUUGU	
1	GAUAAGAUAAGGUGAUCCAUGUGACUAAUCUUAUAGAUAUUAAGCACCUCUACACGCACGGCAC GUCUGUGUCUUCUGUGCGGGCCGGGACAACAGCGGCCUAAAUUCUUCUAGGUGACCAACUGCUUUU GGGACUAUGGCACCACUGUGGACGUGAGUACCUGGCAGUAAGUCUGUGAAAAGAUGGAAGGUGUCC AAGCUAUGGGGCGUAUGCAUAUAGCCUGCGGAACAAACAACGGCGACGUUGUCCCCAGGGGCCAAAA GGCACUGGUAAGAAGUACCUUAUUGUUUACCCCAUGUGUAAGUCACUGGAAGUCCUAGUAUUG AUGUCUGAGUAUAGGCUCAAGGGUAGGGGGAAGGAUAGCCGGAAGGUAACCGUAGGUAACCUUAAG AGACUAUGGAUCUGAUCUGGGGACCGGAUGGGCGCAUACCAUAGCUGUGAGGCGCGGUUAAAAAAC GUCUAAGCCCGCAACCAACUAGGGGACUAGGUUUUCCUUUUUAUCAUGUAUGACGU	
1	UUUCCAGCCGCUUUAAGGACGGGGGUGUGGACGGCGUUGCGGCCACCACAGCAAGAAGGCCAUCA	Salivirus-A-02394-A105-675
9	UCUGUCCUACACCCCAUUCUCCCCUCCUUCUCCUGCAACCAUUAAGCUUACUCGCAUGUGCAUUGA	
2	GUGGUGCAGUGUUGAACAACAGCUACACUCAGUGGGGGCGGGUUUUCUCCGCCUUCGCGCUCUC CGGAGGCCACCCUUCUCCUCCUCCCAUUAACUACAGUGCUUUGUGAGUAAGCAUCCUGUACCCUCC CGCGAAGCUGUCUGCGGCGGCAACUGUGGGGACCAAGGUAUUAUCAAAGCGACCCGCUUUCUCCG CUCCAGGAGUAUCCUGCUGUAGUAAUUCUAGUGGGGCGUCUGCUGGUGGCAACUCCUCCCCAAUGCG CGCUGCGGGAGUGCUUUCUCCCAACUACCCUAGUAUCCUUCUACUGUGUGUGCUGUGGACGCAUUC UGAGACGAUGUUCUGCGUCCUCCAGGACUCCAGUCCAGUCCAGGCGGACGUGUGGCGUAGUGUGCU GGCUUGUCCGGCGCAUGUUUGGUAACCGGUGGGGUAAGGUUGGUGGCCAACGCCCGUACUUGG UGACAACUCAAAGCACCACAGGAUUGCCAGGGGAGUACCCCGCCUACAGCGCGGGAUCUGAGCCUGGG CUAAUUGUCUACGGUGUUCUUCUGUCCAUUUUCCUUCUGUGUC	
1	UUA AAAACAGCCUGUGGGUUUAACCCACCAACAGGGCCUACUGGGCGUAGCACACUGGUUUUACGGU	CVA3-A115-675
9	ACCUUUGUGCGCCUGUUUUUAACCCCCCCAACUUCGAAACUUAAGAAGUAAGCAAAACCCGAUCAAU	
3	AGCAGUGGCGGCGCACAGGUGCAUCUUGAUCAAGCAUUCUGUAUACCCCGGACCGAGUAUCAUAG ACUGUCUCCGGGUUGAAGGAAGAAACGUUCGUUAACCCGCGUUAACUUCUAGUAGCAUCCAGUACA UCAUGAAGUUGCAGAGUGUUUCGCUACAGCAUACCCCGUGUAGAUACAGGCCGAGUAGUACCCGCA CUUCCCCACGGGCGACCGUGGCGUGGCGUGCGUUGGCGGCGUCCUAUGGGGCAACCCUAGGACGC UUAUAUACGAGCAUGGUGCGAAGAGUCUAUUGAGCUAGUUAAGUAGUUCUCCGGCCCGUAAGUGCGG UAAUUCUUAUCCGGAGACAUACCCUUAACAAAGGGCAGUGUGUCUUAACCGGUAACCUUGCAG CGGAACCGACUACUUGGUGUGCCGUGUUUCCUUUUAUUUUUACUGGCGUGCUUAUGGUGACAAUUG AGGAAUUGUUGCCAUUAGCUUAUUGGAUUGGCCAUCCCGGUGACUAAACAGAGCUUAUUGGUUCCAAU UGUUGGAUUUAACCCGCUACACUCGACAGUGAAGAACCCUUCUAUUAUGGUAUUUUCUACACUCA AGAA	
1	UCACCCUUCUUUCGGGUGGUGUGGUCGGGACCAACCCUUAUCUACUUCAGCUUCUUCGGAACUGU	Kobuvirus-16317
9	CGGAGGAUUUAACGGGCAACCAUACCCCCUACCCUUCUGCAUUAAGUAUGUGGUGCUGU	
4	AUCUUGACUCCCAAGGAUUGGAUCGUAACCGCUGGAGAAACAAACUGCUAGAUCCACAUCCUCCUCC CCUGGAGGACCUUGGUUCCUCCACAUCCUCCCCCAGCCUGACGUACCAACAGGCGUGUGAAGGCC CCGCGAAGAGCUGUACUGGCGAAUUGUGGGUCCCCCUUCAACAGACACAGGCUUUCUCCUCCU AAGGCUAGCCCGUAGUGUGAAUUCACAUUUGGGAACUAGUGGUGGACUGGUGGCUCCCAUUCUG GCCGCGAGUGUCUGUCCCCAAGCCAACCCUUGGCCUUCACUAUUGGUCUGGCAAGCAUACCUGA GAAGGUGUUCUCCUGUGGUGGCGAGCCUGGUAACAGGUGGCCAGUGUGGUAACCUUUCUCCGUCU UCGCGAGGUAUGUAUUGUUAAGAUUUGGUGUAAGGUUACUUGGCAACGCGUACGCGGCGGAAGAA	

	CCUCUACUGCCCCUAGGAAUGCCAGGCAGGUACCCCCACCCCGGGUGGGAUUCUGAGCCUGGGCUAAUU GUCUACGGGUAGUUUCAUUUCCAAUUCUUUUUUGUCGGAGUC	
1 9 5	UCACCCUCUUUCCCGGUGUCUGGUCUCCGGACACCGUUACUCCAUCAGCUUCUUCGGAACCGUUC CGGAGGAAUUAACGGGCACCCAUAUCCCCAUCCCCUUUCUGUAACUUAAGUAUGUGUGCUUGUA AUCUUGACUCCACGGAUUGGAUCCGUGGAGAACAAACUGCUAGAUCCACAUCUCCCUUCC CCUGGGAGGACCUUGGUCUCCACAUCUCCCCCAGCCUGACGUACCACAGCCUGUGUGAAGCCC CCGCGAAAGCUGUCUACGUGGCAUUGUGGGUCCCCCUUCAUCAAAGACACAGGUCUUUCCUCCUU AAGGCUAGCCCCGAUGUGGAAUUCACAUUGGGCAACUAGUGGUGACUGUGCGUCCCAUUCUG GCCGCGGAGUGUCUUCUCCCAAGCCAAACCCUUGGCCUUCACUAGUGGCCUGGCAAGCAUUCUGA GAGGUGUUCUCCUGUGGUGCCAGCCUGGUAACAGGUGCCCCAGUGUGCGUAACCUUUCUCCGUCU CCGGACGGUAGUGAUUGGUUAAGAUUGGUGUAAGGUUCAUGUGCCAACGCCUUGUGCGGGAUGAAA UCUCUACUGCCCCUAGGAAUGCCAGGCAGGUACCCCACCCUGGGUGGGAUUCUGAGCCUGGGCUAAUU GUCUACGGGUAGUUUCAUUUCCAAUCCUUUUUUGUCGGAGUC	Aichivirus- Chshc7
1 9 6	GGCUUCGGCCCCUACCCUCUUUCCGGUGUCUGGUCUCCGGACACCGUUAUCCAUCUACGCUUC UUCGGAACCGUUCUGGAGGAAUUAACGGGCACCCAUAUACCCCCAUCCCCUUUUUUGCAACUUAAG UAUGUGUGUCUGUAUUCUAGACUCCACGGAUUGGAUCCGUGGAGAACAAACUGCUAGAUCC ACAUCUCCUCCUCCUCCUUGGGAGGACUCCGUGUCCUCCACAUCUCCUCCAGCCUGACGUUACACA GGCUGUGUGAAGCCCCCGGAAAGCUGUCUACGUGGCAUUGUGGGUCCCCCUUCAUCAAAGACACC AGGUCUUUCCUCCUUAAGCUAGUCCGUAUGUGGAAUUCACAUCCGGCAACUAGUGGUGUCACUGU GCCUCCCAAUUCUGGCCCGGAGUGUCUUCUCCCAAGCCAAACCCUUGGCCUUCACUAGUGGCUU GGCAAGCAUUCUGAGAAGGCGUUCGUGUGGCCAGCCUGGUAACAGGUGCCCCAGUGUGCGU AACCUUCUCCUGUCCCGACGUGAGUGAUUGGUUAAGAUUGGCUAAGGUUCAUGUGCCAACGCC CUGUGCGGGGAUGAAACCUUACUGCCCUAGGAAUGCCAGGCAGGUACCCCACCUUCCGGUGGGAUUCU GAGCCUGGGCUAAUUGUCUACGGGUAGUUUCAUUCUUAUUCUUCUUGUGCCGAGUC	Aichivirus- Goiania
1 9 7	ACCCUCUUUCCGGUGGUCUGGUCUCCGGACACCGUUAUCCAUCAGCUUCUUCGGAACCGUUCG GAGGAAUUAACGGGCACCCAUAUACCCCCACCCCUUUUUGCAACUUAAGUAUGUGUGCUUGUAU CUUGACUCCACGGAUUGGAUCCGUGGAGAACAAACUGCUAGAUCCACAUCUCCCUUCCCU UGGGAGGACUCCGUGUCCUCCACAUCUCCUCCUCCAGCCUGACGUACCACAGGCUGUGUGAAGCCCC CGGAAAGCCGUCACGUGGCAUUGUGGGUCCUCCUUCUUAAGACACAGGUCUUUCCUCCUUA GGCUAGUCCCGAUGUGUGAAUUCACAUUGGGCAACUAGUGGUGUCACUGUGCGUCCCAUUCUGGC CGCGGAGUGUGUUCUCCCAAGCCAAACCCUUGGCCUUCACUAGUGGCCUGGCAAGCAUUCUGAGA AGGUGUUCGCGUGUGGCCAGCCUGGUAACAGGUGCCCCAGUGUGCGUAACCUUUCUCCGUCUUC GGACGGUAGUGAUUGGUUAAGAUUGGCGUAAGGUUCAUGUGCCAACGCCUUGUGCGGGAUGAAAC UCUACUACCCUAGGAAUGCCAGGCAGGUACCCCACCCUUGGGUGGGAUUCUGAGCCUGGGCUAAUUGU CUACGGGUAGUUUCAUUCUCCAAUUCUUCUUAUGUCGGAGUC	Aichivirus- ETHP4
1 9 8	GUUACUCCAUCUACGCUUCUUCGGAACCGUUCGAGGAAUUAACGGGCACCCAUAUACCCCCACCC CCUUUUCUGCAACUUAAGUAUGUGUGUCUGUAUUCUAGACUCCACGGAUUGGAUCCGUGGUA GAACAAACUGCUAGAUCCACAUCUCCUCCUCCUUGGGAGGACCCCGUCCUCCACAUCUCCUCC CAGCCUGACGUUAUCACAGGCUUGUGAAGUCCCGGAAAGCUGUCUACGUGGCAUUGUGGGUCCC CCCUUACUCAAAGACACAGGUCUUUCCUUAAGGCUAGCCCCGAGUGUGAUAUCCACAUUGGGA ACUAGUGGUGUCACUGUGCGUCCCAUUCUGGCCCGGAGUGUGUCCCAAGCCAAACCCUUGG CCCUUACUAGUGCCUGGCAAGCAUUCUGAGAAGGUGUUCGUGUGGCCAGCCUGGUAACA UGUGCCCGAGUGUGGCUAACCUCUCCGUCUCCGGACGUGAGUGAUUGGUUAAGAUUGGUGUAAG GUUCAUGUGCCAACGCCUUGUGCGGGAUGAAACCUUACUGCCCUAGGAAUGCCAGGCAGGUACCCC ACCUUCGGUGGGAUUCUGAGCCUGGGCUAAUUGUCUACGGGUAGUUUCAUUCUCCAAUUCUUUUGU CGGAGUC	Aichivirus- DVI2169
1 9 9	GGGUGGGGGGUGUCGGCCCCUACCCUCUUUCCGGUGGUCUGGACCCGGACACCGUUAUCCA UUCAGCUUCUUCGGAACCGUUCGAGGAAUUAACGGGCACCCAUUUCUGUCCUCCUCCUUCU CUGUAACUCCAAGUGUGUGUCUGUAUUCUAGACUCCCGCGGAUUGACCCGUCUCCGUGGUAACAAAC UGCUAGGUAUCUCCUCCUCCUCCUUGGGGUCUCCUCCGGGCGUCCACACCCUCCUCCUCCAGCCUGAC GUGUCACAGGCUGUACAAGACCCCGGAAAGCUGCUAACGUGGCAUUGUGGGUCCCCCUUUGUA AAGGAACCGAGUCUUUCCUUAAGGCUAGACCCUUGUGAAUUCACAGGUGGCAACUAGUGGU CCACUGCAUGUCCGACUCCGGCCCGGAGUGUCUUCUCCCAAGUCGUAACACUGACCUUACUUA UGUGCCUGGCAAGCAUUCUGAGAAGAUGUCCCGUGUGGUGCCAAACCGGUAACAGGUGCCCCA GUGUGCGUAGUCUUCUCCGUCUUCGGACGGUAGGUGUUAAGUAAGGUGGCGUAAGGUUCAAGU GCCAACGCCUUGGAAGGGAUGACCCUUCUACUGCCCUAGGAAUGCCGCGCAGGUACCCCAGGUUCCG CUGGGAUCUGAGCGCGGCUAAUUGUCUACGGGUAGUUUCAUUCUCCUUCUCCACUGGCAUC	Aichivirus- rat08
2 0 0	GGGUGGGGGGUCUGGCCUCCUACCCUCUUUCCGGUGGUCUGGACCCGGACACCGUUAUCCA UUCAGCUUCUUCGGAACCGUUCGAGGAAUUAACGGGCACCCAUUUCUGUCCUCCUCCUUCU CUGCAACUCAAAGUGUGUGUCGUAAUCCUGACUCCACGGGUUGACCGGCCUUGGUGAACAACA GCUAGGUAUUCUCCUCCUACCCUUGGGCGCCAUUUCAGUGGCGUUAUCCUCCUCCUCCAGCCUGA CGUGUCACAGGCUUGGCAAGUCCCGCGAAAGCUGUCUACGUGGCAUUGUGGGUCCCCCUUUGU GAAGGAACCGAGUCUUUCCUUAAGGCUAGACCCUUGUGUGAACUACAGGUGGCAACUAGUGGU UCCACUGCAUGUCCCGACUCCGCGCGGAGUGUCUUCUCCCAAGUCGUGACACUGACCUCCAU AUGUGCCUGGCAAGCAUUCUGAGAAGAUGUCCCGUGUGGUGCCAAACCGGUAACAGGUGCCCC AGUGCGUGUAGUCUUCUCCGUCUCCGGACGGUAAGUGUGUGGUAAGAUUGCGGCUAAGGUUCAAG UGCCAAACGCCUUGGAAGGGAUGACCCUUCUACUGCCCUAGGAAUGCCGCGCAGGUACCCCAGGUUCC CCUGGGAUCUGAGCGCGGCUAAUUGUCUACGGGUAGUUUCAUUCUCCUUCUUCUACUGGCAUC	Aichivirus- Rt386

2 0 1	UUUGAAAAGGGGGUGGGGGGCCUCGGCCCCUCACCCUCUUUCCGGUGGCCAUUCCGCCGGGCCA CCGUUACUCCACUCCACUCCUCCGGGACUGGUUUGGAGGAACACAACAGGGCUUCCCAUCCUGUUU ACCCUUUAUUCACUCCUCCUCCCAAGUUUACCCUACUCCACACCCACUGACUGACUCCUUUGGA UUUUGACCUCAGAAUGCCUAAUUGACUCCCAACUCGCCUCUCCUUUCCGGAUUGCCGGUGGUGCCU GGCGGAAAAAGCACAAAGUGUGUUGCAGGCUACCAAACUCCUACCCGACAAAGGUGCGUGUCCGCCGUG CUGAGUAAUGGGAUAGGAGUCCUACAACAGGCUCCGCCAUGAGUAGAGCAUGGACUGCGGUGCAU GUGACUUCGGUACACACGGGCAUAGCAUUGCUACCCGUGAAUCAAAGUACUCCGAGAUUUUCUGACC UCUGAAGUGCACUGUGGUUGCGUGGCGGGAUCCACGCUUGACCAUGUACUGCUUGAUAGAGUCGC GGCUGGCCGACUCAUGGGUUAAGUCAGUUGACAAGACAC	Porcine- Kobuvirus- GS2
2 0 2	UUUCACACCCUCUUUCCGGUGGUGCCGACCCAGACCACCGUUACUCCAUAUACGUACUUCGGUACC UGUUCGGAGGAAUUAACGGGCACCCUACCCAAGGGUUAUAGGGACCAUAUUCUCCUCCUCCUGUA ACUUUAAGUUUUGGCCCGUAUUCUUGACUCCAGGCGGAUGUUUGUGCCCGUCCUGUGAACAAAC AGCUAGACACUUUCUCCUCCUCCUUGGGGUGUCUCCGGAGUCCACUCCUCCUCCCGAGCGUAACA UGCCCGCGUGAGUGAUGCACCUGGAAGUCUGGACGUGGGUUAGUAACUUCGGUGAAAAACCCACUA UAUGACAACUUGGUUGACCCCAACUCAAAGGACUCGAGUCUUUCCUCCUUAAGGCUAGCCCGCC ACAUGAAUUUGCAGCUGGCAACUAGUGAGUCCACCAUGUCCCGCAACUCCGGCUGCGGAGUGCGUU CCCCAAGCGUAUGCCUCCUUCUGUAAGAGUGCGCCUGGCAAGCACAUUCUGAGAAGUCGUUCGCGUG CGUCGUGCCAACCCUGGCGACAGGUGACCCAGUGUGCGUAGACUUCUCCGGAUUCUGCCGGCUCUUC UCUAGGAAACAUUGCGUAAAGGUUCAUGUGCCAAAGCCUUGCGCGCGGUGUUCUUCUACUGCCUAG GAUUGGCCGAGGUACCCUACUUCGGUAGGGAUCUGAGCGGUAAGCUAAUUGUCUACGGGUAGUUU CAUUCUACUUCUUCUACAGGUCGACAUC	Kobuvirus- SZAL6- aichivirusA5
2 0 3	UUUGAAAAGGGGGUGGGGGGCCUUCGGCCCCUCACCCUCUUUCCGGCGGCCAAGUUCGCGUCGU CGUAAGCCUGGACACCCAUUGUUCGCAUCGGUCUCCGGGCAAAUACAAGAACCCAGACCUUCUGGUA CUUCUUCGCCUGGGUCACAAAAGCGAAGAACCUGCCUCUUAACGCCAGACGAGCGGCAUUAACUU GAACUUCUGGCACUUCACUUCUCCUUUCCUUGUCCUUCUCCUACUGCGUCUCAAGGUCGCGC AAUUCUGGGACUAGCCAGUUUAAAAGUUCUGGCACCCUUGCCUUCUAGGCCCUAAGGUAGG AACUGACCUUGUGUGUGAUCUGGUGCGGGAGUGCUACACGUAUGCAUCGUAAGCCUCGUUUGUG GUUCUGCCUGGCAAGGCUACAGAGUACCGUUGUCCGUGUGGAUGCCAUCGGGUAACCGGACCCC CAGUGUGUGUAGCGGUAUGUUCACGGUCCGCGGUGUACCAAGAUUCCUGACUUGGCUUUGCUAGAA AUGGUGUGUGCCAAUCCUUGUGACCAUUAUACAUCACCUAGGAUUGCUAGGAAGGUACCCC AGUUCUGAGCUGGGAUUCGAUCCUAGGCUAAUUGUCUACGGUGAUGCUCCUUUAUUUUCUACAC UGCUAUUGACUGAUGUUGCUAUCUUGGUCUUCUUGGUCUUCUUGGUCUAGUUCUACAGGUVU CUCUUUGUCCAAUCCUUUGGUUCUUCUUGUCCACUUGCCACUGCCAACGCUUGUC	Kobuvirus- sheep-TB3
2 0 4	GACAGGUGUUUGGAGGGCGGAUGACGAUAUCUGGCGGGCCACAGGGAAUACGGCAAAUGUCUGA UCAUACGGUUCACAAGUCUACCGGCGAUGAGUGGUACAACACCAUGUGUAGCAGGGAUUCUUGCGU GUGAAGCGGACAGUGC	Bat- sapelovirus
2 0 5	GAAAGGGGGUGGUAGGGGCGGUACGGUUAUGCCGUGCGGUUCCGCCACCCUAGGGGGCCACACGGU CCUGCCGUGUGGUUCCCGUGGUUGACAGUACGCAUUGGGGGCCGUACGGUCCUGCCGUGCGGUU UCCUUUGCUUGCUGGCAUUGCGGGAGUACACCCCUUUAACGUGGGUACUACGAAAGUGCCCUUC GCUCCGAGGUUAAAGGAGAACCCTCCUUCUUAACCCACUCAGCUCGCCUUCAGUGCGGCGCUA GCCUUUCCACUUGCAGCUUCUGUUGAUGAUGCUUGCACCUGAUGUGGUGCGCUUCUUGCUUAGUC GCUUGGCUUCUUAUCUUCUGACGAUUCAGUUCUUAACGCCAGUUGUUGACCGGCCAAGGGGUA UUGCGGUUAAGUAUCCUACCGCAUUAUCCUUCUCCGUUCGUAUGGUUUGGAUUCUUGGAUUCU CUCUCCUUCUUCUCCCGCUCUCAAUUUAGCUUCGUAUUGAAGCAUCACUGUCUUAUUAUUUA UGUCGGACUGACGAUUGAGUACGUUCAGAUUGUUGUUGGGAGGCCAAGGGAUUCGAUGGACAACACU UCGAAAGAGUACUUGUACCGCUCUUCUCCUACCCUAGCAACUUGGAUUGCUACGUGGAG UUCGAGUUCUGUACUUAACUUGACUUGCUUUCUUAACUUGCUUAUUCUGCAGUGGAUUGGUU GUAGACUUAUUCACGUUCGUUAGAUUGCUGACGUGGAGUACGAUCGUGUACAUUCACUACUGCC AAUUAUCUCCCUUCCCGUUGCUCUCCUUAUAAGGAGAGCCUUCUUGCAAAGGUGAAGCCUUC ACCCCGGUCGAAAGCCGUUGGAUAAGACAGGGUUAUUUUCUCCUCCUGCGCGCUUGCCUUCU UAAGCUGAAUAGGUUCUAUUAUUCAGGCGGAUGGUCUGGUCGUCCUUCUUGGACAGAGUGUGUA UCUGGGUUUCCGGAUUCGACACACUACACAGAGCUCAGGAGUGAUUAAGUCAAGGCCGGAUC UGCGGCGAAAAGGAAAUGAAGUAUUUGCAGCUGUAGCGACCUUCUAAAGGCCGCGGAUUCUCCAC CUGGUGACAGGUGCCUCUGGGGCCAAAAGCCACGUGUUAUAGCACCCUUGAGAGCGGUGGUACCCC ACCACCCUGCAAAUUAUGGAUUUGACUUAAGUAAUAAAAGAUUGACUUGGCAUACCUCAACUAGC GGCGGCUAAGGAUGCCUGAAGGUACCCGUGUGAAUUCGUUCGCGGACCAUGGAUCUGAUCAGG GCCUGCCUGGAGUGGUUCUAUCCACACAGCGUAGGGUUAAAAAACGUCUAACCGCCCCCAAAGA CCCCGGCAGGAUGCCGGUUCCUUUUACCAAUUCUUGACACU	Bovine- cardiovirus- 1
2 0 6	UUGGAUCUGAGCAGGGGCCCCUUGGGUUGCUUUAACAACUACACUGGGGGUAAAAACGUCUAAAC CCGACACGCGCAGAGGAUCUGGUUCCUUUAUUUUAUUCUUAACUACACUAGCAGAUUGACGAU AAACGUUGUUGUUGUGACUAUUGACUUGAUCUGCUUCUACGGGUUAUUCUACUGUUAUACUUCU UGCUUUGUUGGUGUUCACUGUACUUGUCUCCUUCUACAUUUCACA	Bovine- picornavirus -isolate- TCH6
2 0 7	UUUAAGUGUUGGCCAAUCUCUUGACUCCUGCGGAACACCGACAGUAGUGUCCAAAAGGCCAG GUGGAAAAUCCUCCUCCUUGGGCUUAUGCCCGGAUCCUCCCGCAGCCUGACGUGCCACA GGCUGUGCAAAGACCCCGGAAAGCUGCCAAAAGUGGCAUUGGGUGCCCCCUUUGUCAAGGCGU CGAGUCUUCUCCCUUAAGGCUAGUCCUGUAGUGAAUCUGUCGGGCAACUAGUGACGCGCACUGCA UCGACCGCAGCUGCGCGCGGAGUGUCUCCCGCAAGUCAUGCCCGUACCAAGAUUGUGUGUCU GGCAAAAUUGUCUGAGAAUUGUCCCGUGUGGCGCCAGCCUGGUAACAGGCGUCCCGCAGUGUG CGUAAUUCUCAUCCAGACUUCGGUCUGGCAACUUGCUGUUAAGACAUGGCGUAAGGGCGUGUGCCA	Canine- kobuvirus

	ACGCCUUGAAGCAGUGUCCACUCUAAUACCCGAGGAAUGCUACGCAGGUACCCUUGGCUGCCAG GGUUCUGAGCGUAGGCUAAUUGUCUAAAGGUAUUUUCUAAUUCACCCUUCUUCUUGUUAUA	
2 0 8	UUAAAAUAGCCUGAGGGUUGUCCACCCACAGGGCCACGUGGCGCAGUACUCUGGUUACAGGUA CCUUUGCAGCCUUAUUUUAUACCCUUCUCCCAUCGUAACUUAGAAGCAACAAACAAACUGCCAAUA GCAGCAACACCCAGUUGUUAAGGGCAAGCACUUCUGUUCUCCCGAAGGGUCUGACGGUAUGC UGUACCCACGGCAGAGUAUGACCUACCGUUAACCGGCCAUGUACUUCGAGAAGCCUAGUACCAUUA UGAAGUUGAUGUAGUUAACGCUCCCAAGCAACCCAGCUGGUAGUUCUGGUCGAGUAGUCUCGGCA UUCUCCACGGGCGACCGUGGCCAGGCGUGGUGGCGGCCAGCUACACCAUACGGUGUAGGACGUC AAGAUACUGACAUGGUGUAGAGGCCUUAUAGGCUACGUGGUAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAAUCUCCGGAGCAUCCGCCAGUAAGCCACUGGAAGGGUGUCGUAUUGCGAAAGUCUGGAGC GGACCAGUACUUCUUGGUGUCCGUGUUUCCUGUUUACUUAUUGUUUGGUGUCUUAUGGUGACAAAC UUUAUAGUUAUCAUCAUAAGCUACUUGGUCUUGCCACCCGAGAAUUAUUUGGUUAUUUGUUGGUUC AUAACCUACAGUCGUAUUUUCUGUCUUAUUAUUGUUCUCAAUUUAACAACA	Enterovirus- sp-isolate- CPML
2 0 9	UUAAAAACAGCCUGGGGUGUUGUCCACCCCGGGGCCACGUGGCGCCAGUACUCUGGCACGUUAGU GCCUUUGUACGCCUGUUUUCUCCUUAACAAUUAAGAUUGCCACUACUGAGGGGAGUAGUCCG ACUCCGCUCCAGUACCCGUGCACCAGUGAGCUGGUACGCUAGUACCUUCGCACGGAGUAGUAGGCAU CCCCACCCGUAAUCUUAAGAGCAAGUACACAUUCUGGCCAAUAGUGGCGUGCAUCCAGCCGCGCA ACGGUCAAGCACUUCUGUUUCCCGUCCGCAAGGGUGCUUAUCCGCCACAGUCCACUACGGAAAGCC UACUAAACCAUAGAGCUAUCGAGAGGUUGCGUCUGGCCACGACCCCGUGGUAGCUCUGAGUGAGG GGCUCGCAAAACCCCGUGGUAAACAGGAUGCUUGCCCGCGCGUGCACUCGGGUUACGCUUAUUGG UUGUUCACCUCAACAUGUGUAAUUGGCCAAGAGCCUACUGUGCUGGAUUGGUUUUCCUCCGGAGCC UUGAAUCGUGUAAUCCCAACUCCGAGCGUGUGCGCAUUAUCCAGUGUUGCUACGUCGUAACGCGU AAGUUGGAGGCGGAACAGACUACUUCGGUACUCCGUGUUCUUAUUGUUUAUUUUGAAUUUAUGG UGACAAUUGCUGAGAUUUGCGAAUUAAGCAGCUCUACCGCUGAACAUUGCCUGUACUACCUAAUCGC AUUUCACAAAACCCUACAGAUACCAAGCUCUUAUUAUGAUCUGCUGUUUCCUGAAUCUCAAUAU AAAUUGGAACAAGCAAA	Enterovirus- AN12-F7
2 1 0	UUAAAAACUGGGUGUGGGUUGUCCACCCACACCACCAUAGGGUGUUGUACUCUGUUAUUCGGUA ACUUUGUACGCCAGUUUUCUCCUCCUCCCAUCCUUUACGUAACUUAGAAGUUUAAUACAAAG ACCAAUAGUAGGCAACUCCAGGUUGUCUAAAGGUAAGCAUUCUGUUUCCCGGUUGAUGUUGAU AUGCUCACACAGGGCAAAACAACAGAUACCGUUAUCCGCAAGUGCCUACACAGAGCUUAGUAGGA UUCUGAAAGAUUUUGGUUGGUGUUCAGCUGCAUACCCAGCAGUAGACCUUGCAGUAGAGGCGGGA CAUUCUCCACUGGUAACAGUGGUCCAGCCUGCGUGGCGUGCCUGCGCACCUUCUAGAGGUGUGAAGC CACUAAACCAUAGAGGUGUGAAGAGCCGCGUGUGUCACUUAUUGAGUCCUCCGGCCCCUGAUGCGGC UAACCUUAAACCCUGCAGCCAUUGGCUCAUAAGCCAAUGAGUUUAUGGUCGUAACGAGUAAUUGCGGGA UGGGACCGACUACUUGGGUGUCCGUGUUCACUUAUUCUUUAUUAUUGCUUAUGGUGACAAUAU AUUAUUGAUUAUUAUUGGCAUC	Human- rhinovirus-1
2 1 1	ACCUGCGACAGUGCCAACCCACAUUCUUCUCCACUUAUUAUUAUCCGGUUCUCCGCCUUAAGGUCCA GAAAAGACCAACAUCUACUCCUUGGAUGCUGGCUAAUCAGAGGAAAGUCAGCAUUUCCGGGCCA GGCUAAGAGAAACAUAUAGUUAAGAAUUAUUAUCCUUGACGAAUUCGGAACGAGAAAGUUCUC CCUCCUUGCCGCUUGUUCACACCAUUAUUAUUCGGCCUUAUUGAGACAAGCCCCUGGUGAAAG GACCUUCUCUUCUCCGACGUGGUUGGAUUAACAUUUUCCGACGAAAGUGCUUAUUAUGCCUCCCC GAUUGUGUGAUGCUUUCUGCCUGCUGGGCGGAGCGUCCUGGGUUGAGAAACCUUGAAUCUUUCC UUUGGAGCCUUGGCUCCCGGCUUAAGCCGCUUGGAUAUAGACAGGGUUAUUUCCAAACUUCUUA UUUCUACUUUAUGGGUUCUAUCCAUGAAAAGGGUAUGUGUUGCCCCUCCUUCUUGGAGAAUCUG CGCGGCGGUCUUCUCCGUCUCUACAGCGGUGGAUGCAACAUCCGGAACGGUAGAGAAACAGU UUCUGGGAUUAUUAAGAGUGGACAUCAAGAACAGCUGUAGCGACCUACAGUAGCAGCGGAUUCUCCU CUUGGCGACAAAGAGCCUUCUGCGGCCAAAGCCCGUGGAUAAGAUCCACUGCUGUGAGCGGUGCAAC CCAGACCCUUGGUUCGAGUGGCAUUCUUAUGGAACCAAGAAUUGGUUUUCUAAAGCCUCCGGUA GAGAAGCCAAGAAUGUCCUGAAGGUACCCCGCGCGCGGGAUCUGAUCAGGAGACCAUUGGCAGUGC UUUACGCGCCAUUUGGUUUAAAAACUGUCACAGCUUCUCCAAACAGUGGUCUUGGUUUUCCAA UUUUGUUGACUGACAAU	Human-TMEV- like- cardiovirus
2 1 2	UUGAAAAGGGUGUGGGGGGGAUACCCUCCUCCUCCACACUCUUUUCGACGGCGCGCCUCGCGAAGG CCGCGAUACCGCACUGUCUACUUCGGUACCAUCUCGCGGGAUGACAAUUGCCCGUGUAUAGU UUGCCUGUUUCUGCACCGUUAACCGUCUGUUCGGGAUUGGAAACUGGCACCCUCCUUCUCCUAC CACCUCUUCUCCUUGCCUCCAUUAUUAUUAACACGCGCACACAGCGCGCGCCGCCAAGGGCUA GCCUGGCGGUUAUAAAGGAACUUGGUCUUCUCCUUAAGCCAAAGGUAGGUUCCUUGUGUCC CUGAAUGCUCGGUGAGGAUUGCUGCACCGUAACGCUUUGUGAAGUUGUUGCAAGUUCUGGCCCGCA AGCCUACAGAGUGCUGUAUCCGUGCGGACGCCAUCCUGGUAACAGGACCCCAUGUGCGCAACA GUAUUGUUCAGACUUCGGUUGUUCACUUGCUUUAUGGACCAUUGCGCGAAAGUGCGUGCGCAUAU CCUGUACUUCAGGUGUGUUCUUGGACCCUAGGAUUGCUGCGAAGUACCCGUUUCGGCGGGAU CUGAUCGCGAGCUAAUUGUCUAUGGUUUCAGUUUCCUUUUAUUAUCCACAAUUGACUGCUUAAAC UGACUUCGGAUCUUGGCUUCCACUGCUUUCUUGCUCUCAAACGGCUUCACUUAACCAUCUCAC CUUUCGACCAACACCAUUAACACACUAAUUAUUCGACUCUUCGACUCCUGGCUUGGUGAAGAC	Kagoshima-2- 24-KoV
2 1 3	CUUCUUUGACCCUGGCCACUGGGGUACAGUGGCACUACAUUCUCUCUUCUUUAUCUUAUCUUAUCGU UUCUUUACUUUACUUUACUUAUUAUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU AUUAUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU AUAAAGCAAGUUUAUGCCUGCAACUGCUUAUUAUGGACACCGCUUCGCGGACCCUUGUUGAG UUUGUAUGUUUUUAAGUAAUUAUUGCAACCAUACGAUUAUUUAAGCCGCCUCUCUUAUUAUGAUCUUG UUUAUGUGGACGUGAAACAUUGGUUUUUCUACACACGUCUCCGGGUGACCCGGCGUGUGUUCUCCG	Lesavirus-1

[illegible]

2	UUUAAAACAGCCUGGGGUGUUCACCCACCCAGGGGCCACUGGGCGUUAGUACUCUGGUUACGCGG	Human- enterovirus- 109
2	UACCUUAGUAGCCUGUUUUUAGUCUCCUUCUCCCCGCAACUUAGAAGUAAUCAAAGUUUAGGCUCA	
0	ACAGUCGCCACACCCCAAGUGUGUUCGAGCAAGCACUUCUGUUCUCCCGUUGCGUCUUUAUGCU GUGUGAACGGCAGAAAGGACAAUUCGUUAUCCGCUCAACUACUGGGAAGCCUAGUACCACCAU GGAUUGACCUGAAAGUUGCGUUCAGCGCACCCCGAGCGAGCUAGGCCGAUGAGGCUCCGAAUACC CCACGGGCGACCGUGUCGAGCCUGCGUUGGCGGCCUGCCACAGUUGCAAAACUGGGAGCGUCUUU UCAUGACAUGGUGCGAAGAGCCUACUGUGCUAGUUGAGAGUCCUCCGGCCCCUGAAUGUGGAUAAUC CUAACCACUGAACCACGGGCGCAAAACAGCGUCUGGUUAGGCCGUAACGCGCAAGUCGGUGGCGGAA CCAACUACUUUGGGUGUCCGUGUUUCCUUUUUACUUUUUAGAUUUUAGUGGACAAUUGUGUGUA CAGUUGUUAACAUAGUUUGCAUUCAGAAAUAACCUAACACUUUCCAAUUAUUUGUUACA	
2	UUAAAACAGCUCUGGGGUGUACCCACCCAGAGGCCACGUGGCGGCUAGCACUCCGGUUAUCACGG	Human- poliovirus- 2-strain- NIE0811460
2	UACCCUUGUGCGCCUGUUUUUAUUCUCCCCCGCAACUUAGAAGCAGAAACCAAGUUCAAUAGAA	
1	GGGGUACAAACCAGUACCACAGCAACAGCACUUCUGUUUCCCCGGUGACAUUGCAUAGACUGCU CACGCGGUGUAAAGUGAUCCAUCGUUACCCGCUUGUGUACUUCGAAAGCCUAGUAUUGCCUUGGA AUCUUCACCGGUGUGCGUCAGCACCCGACCCGGGGUGUAGCUUAGGCUAGUAGUCUGGACAUUC CUCACCGGUGACGGUGGUCAGGCGUGCGUUGGCGGCCUACCUAUGGCUAAUGCCAUGGACGCUAGA UGUGAACCAAGUGUGAAGAGCCUUAUAGGCUACAUAAGAGUCCUCCGGCCCCUGAAUGCGGCUAAGC CUAACACCGGAGCAGGCGUGCGCAACACAGUACUAGCUUGUCGUAACGCGCAAGUCUGUGGCGGAA CCGACUACUUUGGGUGUCCGUGUUUCCUGUUUUUUUAUUGGCGUCUUAUGGUGACAAUCAGAGA UUUUUAUCAUAAAGCGAAUUGGAUUGGCCAUCCGGUGAGUGUUGUGCAGGUAUACAACUGUUUGU GGAACACUGUGUAGUUUAACCUUCUUCUUAACCAAUUAAGUCAAAAACAUAACGAAGAUAAGACAA CAUACUACA	
2	GGGGUUGUUCACCCACCCAGAGGCCACGUGGCGGCUAGUACUCCGGUUAUACGGUACCCUUGUACGC	Human- poliovirus- 1-strain- EQG1419328
2	CUGUUUUUAUACUCCCCUCCCCGUAACUUAGAAGAAACAAAUAGUUCAAUAGGAGGGGUAACAAAC	
2	CAGUACCACACGAAACAAGCACUUCUGUUCUCCCGUGACAUUGCAUAGACUGUCCCCACGGUUGAA AGCAAUUGAUCCGUUACCCGCUUCUUGUACUUCGAGAAGCCUAGUACCAUCUUGGAUUAUCGAUGCG UUGCGUCCACACUACAGUCCAGAGUGUAGCUUAGGCGUAGAGUCUGGACAUCUCCUACCGGCGAC GGUGGUCCAGGCGUGCGUUGGCGGCCUACCUUGUGGCCAAAGCCACAGGACGCUAGUUGAACAAGG UGUGAAGAGCCUAUUGAGCUAUAAGAGAGUCCUCCGGCCCCUGAAUGCGGCUAUAUCCCAACACGGA UCAAGGGUGCACGAACACAGUUAUACCUUGUCGUAACGCGCAAGUCCGUGGCGGAACCGACUACUUU GGGUGACCGUGUUUCCUUUAUUAUUUCAAUUGGCGUCUUAUGGUGACAAUCAUUGAUUGUUAUCAUA AAGCGAAUUGGAUUGGCCAUCCGGUGAAGUGAAACAUAUUGUUUGCCUUCUGUUGGUCUACUUC AACCAUUCUUUACUUAACAUCUUAACCAUACAGUUUUGCUGGUUAGAAGUGUGUUUACG	
2	CUGGGGUGUACCCACCCAGAGGCCACGUGGCGGCUAGCACUCCGGUUAUACGGUACCCUUGUGC	Human- poliovirus- 2-isolate- IS061
2	GCCUGUUUUUAUACUCCCCUCCCCGUAACUUAGAAGCAGAAACCAAGUUCAAUAGAAGGGGUAACAAA	
3	CCAGUAACCAUACGGAACAAGCACUUCUGUUUCCCCGGUGACAUUGCAUAGACUGUCUACCGCGUUGA AAGUGAUCGAUCCGUUACCCGCUUGUGUACUUCGAAAAGCCUAGUAUCGCCUUGGAAUUCUACGCG GUUGCGUCACGACCCGACCCCGGGGUGUAGCUUAGGCCGAUGAGUCUGGACAUCUCCUACCGGUGA CGUGGUGCCAGGCGUGCGUUGGCGGCCUACCUAUGGCUAACGCCAUAGGACGUUAGAUGUGAACAAGG UGUGAAGAGCCUAUUGAGCUACAUAAGAGUCCUCCGGCCCCUGAAUGCGGCUAUAUCCUAAACCGGA GCAGGCGGUCGCAACACAGUAGCUGGCUUGUCGUAACGCGCAAGUCUGUGGCGGAACCGACUACUUU GGGUGUCCGUGUUUCCUGUUUUUUUAUCAUGGCGUCUUAUGGUGACAAUCAGAGAUUGUUUAUCAUA AAGCGAAUUGGAUUGGCCAUCCGGUGAGUGUUGUGUCAGGUAUACAACUGUUUGUUGGAACCAUCGU GUUAGCUUUGCUUCUUAUUAACCAUUAUCAAACAAACAAUACGAGGAUAAAACAACAUAUACA	
2	UUAAAACAGCCUGUGGGUUGUACCCACCCAGGGGCCACUGGGCGCUAGCACUCCGGUUAUCACGGU	Coxsackievir us-B5
2	ACCUUUGUGCGCCUGUUUUUAGCCCCUCCCCAAUUGAAACUUAGAAGUUAACACACACCGAUCAA	
4	CAGCGGGCGUGGCAUACCGCCGCGUCUUGAUCAAGCACUCCUGUUUCCCCGGACCGAGUAUCAUA GACUGUCACGCGGUGAAGGAGAAAACGUUCGUUACCCGGCUAACUACUUCGAGAAACCUAGUAGC AUCAUGAAAGUUGCGAAGCGUUCGCUACGACAUCCCCAGUGUAGAUACAGGUCGAGUAGUACCCGC AUUCCCCACGCGCGACCCGUGGCGUGGUGCGUUGGCGGCCUGCCUACGGGGCAACCCGUAAGGACGC UUCAAUACAGACAUGGUGCGAAGAGUGCAUUGAGCUAGUUAGUAGUCCUCCGGCCCCUGAAUCCGGC UAAUCCUAACUGCGAGCACAUACCCUCAAACCCAGGGGCAUUGUGUCGUAACGGGUAACUCUGCAG CGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUUUAUUCUUAUAAUGGCGUCUUAUGGUGACAAU GAAAGAUGUUACCAUAUAGCUAUGGAUUGGCCAUCCGGUGUCUAAACAGAGCUAUUAUUAUACCUU UUUGUUGAUUUUGUACCAUUGAUUAAGGAAGUCAAGACACUACAUAUUAUACAUAUUGAACAC AGCAA	
2	CUGUGGGUUGUUCACCCACCCAGGGGCCACUGGGCGCCAGCACUCUGAUUCCACGGAAUCUUUGUGC	Coxsackievir us-A10
2	GCCUGUUUUAACAACCCUCCCCAACUUGUAACGUAAGUAUAACACACUACUGAUCAAUAGCAGGC	
5	AUGGCGCGCCAGUCAUGUCUCGAUCAAGCACUUCUGUUCUCCCCGGACUGAGUAUCAUAGACUGCUC ACGCGGUGUAAGGAGAAAACGUUCGUUACCCGGCUAACUACUUCGAGAAACCUAGUAGCACCAUAGA AGCUGCAGAGUGUUCGCUACGACAUUCCCCGUGUAGAUACGGCUGAUGAGUCACUGCAAUCCCA CGGGUGACCGUGGACUGGCUUGCGUUGGCGGCCUGCCUAGGGGCAACCCAUAGGACGCUCAUUGU GGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUAUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUA ACUGCGGAGCAUAGCCUUAACCCAGAAGGUAGUGUGCGUAACGGGCAACUCUGCAGCGGAACCG ACUACUUUGGGUGUCCGUGUUUCCUUUUUAUCCUAUAUUGGCGUCUUAUGGUGACAAUCACGGAAU GUUGCCAUUAAGCUAUGGAUUGGCCAUCCGGUGUCUAAUAGAGCUAUGUGUACCUAUUUUGUGGA UUUACUCCGCUAUCACAUAAAUCUCUGAACACUUUGUGCUUUUAUUAUUGAACUUAACACCCGAAA	

2	AUUAAGGUUUUAUACCUUCCAGGUAACAAACCAACUUCUUGAUCUCUUGUAGAUUCUUCUCU	Human-XIAP
2	AAACGAACUUUAAAAUCUGUGGGCUGUCACUCGGCUGCAUGCUUAGUGCACUCACGCAGUAUAAU	
6	AAUAAACUAAUUAUCUGUCUUGACAGGACACGAGUAACUCUGUCUUAUCUUCUGCAGGCUGCUACGGUU	
	UCGUCGUGUUGCAGCCGAUCAUCAGCAUCUAGGUUUCGUCGGGUGUGACCGAAAGGUAAG	
2	UUAAAACAGCCUGGGGUUGUACCCACCCUGGGGCCACGCGGCCAGUACUCUGGUACGCUAGU	iBEV1
2	ACCUUUGUACGCCUGUUUCCCCACCCUAAAUAUAAUUAAGAUUACCACUGCUGUGGGAGUAGUCC	
7	GACUCCGCACCGAUACGUCGCACAGUGAACUGGUUCGCUUAGAACCUUUGCAGGAGCAGCUGGUA	
	UCCCCCCCCGUAACUUAGAAGCCUGGACAAACCGACCAUAAGAGCGUUGUAGCCAGCAGCGCAACG	
	GUCAAGCACUUCUGUUUCCCGGGGCCAUGCGUCGUUACCCGCCCGGCCUACUGCGGGAAGCCUAGU	
	AGGACCCCAUUCGGAGCGCGGUUGCGUUCAGCCACAACCCAGUGGUAGCUCUGAGGAUUGGGACU	
	CGCCUACCCCCACAGCAUUGUGGUGGCUUUGCCGCGUGUGUCCUGGGUUCGCUUGUAGGCGAUC	
	ACCGCAACUCUGAGUAAGGUCUCAAGAGCCUACUGCGCUAGGUCGGUUCUCCCGGAGCCGUGAAU	
	GGUGCUAAUCCCAACUCCGAGCGUGUGCGCACACCCAGUGUUGCUACGUCGUAUUGCGUAAGUUG	
	GAGGCGGAACAGACUACUUCGGUACUCCGUGUUUCCUAAUUUUUGUUAUUUAUACUUGGUGACA	
	UUGACUGACAUUUGGAAUUGGCCCGCUGCCUUGAACAUAGCUCUCUUAUACUUGAUUGCAUUC	
	ACAAAACAACCCAACUCACUGUUGAACUUGUUGACUUCGCAGUCUACUUGAACUACAGUACAUA	
	CAUUCACA	
2	CUUGUAAAAACAACUUUGAUGCCUUGAAUUAUAAUGAUUCAUUAUAAUUAUGCAUAGAUUUUA	iCovid19
2	AUAAUCUGCAUUAUUUAUGCUUUUAUGUUUUUCCUAAUUAUAGAUUUGACAUUGGUUAAUAAUUA	
8	UAUAUUCUGCAUCACAGUUUACAUUUUAUGUAAAAUAAAGCAUUAUAAAAUUAUAGUUUUUUCUG	
	CCUGCUUAAUUAUACUUCUCAAAGAGAGAAAACAAAAUAGCAUUAUUAUUAUAGACUUGAA	
	UGAUGUGGUAAUUGCGAAACUCUAGUAUUUAGAAUUAAGAUUUCUAGCGGUCGUGUAGUUAUUU	
	UAUGUCAUAAUGGUAUUAUUGUAGCUCUUAUAAAGUCUGUUGCUUGUGUUUACAUUUUGG	
	AUUUCCUAAUUAUUGUUCUUUUUAGAAAGGUGGACAGUCCUUAUUUUAAGAGA	
2	UUUAAUAAAGUUGUGCAGAUUAUUCUGCACAGCACUAGCCACAAGUUCAGUGGUCCCAAUUAUGG	iCrPV
2	GAUUAUAAACUCAAUGGUCGACGAGUGACCUUGGAAUUGCUCCCGUGAGAAACCUUGUUUAAACGC	
9	ACUUGGACCUUCUCGGCCUCAGCGCCAGGCAUAGCUGGAUCGUCAUCCCAAGUUGGACUUAAGAC	
	UUGCUUAUGUUAACAAUUGGGAGGAGCGUGACUUAACAUUGGUCGUUGAUGACGUAAUUAUCCAAACA	
	AAACAACAACAACAACUUAAGUUUAAUAAUUCGGUCUUUAAGCAUCAGGUAAGACCGAUGAGGAA	
	UAAAGAACCCUUAUAGACUGAUUCGAUACCAAGAGCUGGUGUCUGAGGGAUUAACGACUCGCGUGGUC	
	GGACAGCACAUUGUAGUCAUGUAUUUAGGGUGUUUCCCAACAGUAUCUUUGCUUAGUAUUGAACUU	
	UAGUGAUGAGUGUACUUGGCAUGUACACUGAUUAGUUAGCCUUGUGGAUACGCCACUGCCGCUCA	
	GUACAGUAGGACAUACCCCAAUAUUGGGUUCACUCUGUAGAGAAGAAAAUUAUGGGCGAAACGUA	
	CGGUUUAUGUCAAGGAGAGGUGAUUUUAUUGCACCUAGAAGCAACCCUAGGUCUUCUUUGAUG	
	UAGGCUAAUAAAGGAGUAACCGAUUCUUAACAUGUGAUC	
2	AGCGCCCAAUACGCAACACCGCCUCCCGCGCGUUGCGGAUUAUUAUGCAGCUGGCACGACAG	iEMCV
3	GUUUCCCGACUGGAAAGCGGGCAGUGAGCGCAACGCAUUAUUGAGUUAAGCUCACUUAUAGGCA	
0	CCCCAGGCUUACACUUAUGCUUCCGGCUCGUAGUUGUGUGGAAUUGUGAGCGGAUAAACAUUUC	
	ACACAGGAAACAGCUAUGACCAUGAUUACGCCAAGCUUCCAGGCACUCCUUCAAGACCUGAAGGU	
	CGAUUAGCUGCAAGAAUUCUUCUGUUUAAACUUAUCCAUUUUGCAAAAGCUUUUUGCAAAAGC	
	CUAGGCCUCCAAAAAGCCUCCUACUACUUCUGGAUUAAGCUCAGAGGCCGAGGCGGCCUCGGCCUC	
	UGCAUAAUAAAAAAUUAUGUCAGCCAUUGGGCGGAGAAUGGGCGGAACUGGGCGGAGUUAAGGGGC	
	UUGCGCUUAAAGGAGCCCUUCGCCUCUGCUUAGUUGAGGCCUGGCCUGGGCGUGGGGCCGCCGC	
	UGCCUGCUGGGGAGCCUGCGUGAGGCUCCGUGCCGUCAGUGGGCAGAGCGCACUAGCCACAG	
	UCCCCGAGAAUUGGGGGGAGGGGUCGGCAUUGAACCGGUGCCUAGAGAAGGUGGCGCGGGGUA	
	CUGGGAAGUGAUGUCUGUACUGGCUCCGCCUUUUCCGAGGGUGGGGAGAACCUGUAUUAAGU	
	GCAGUAGUCGCCGUGAAGCUUUCUUUUCGCAACGGGUUUGCCGCCAGAACACAGGUUAGUCCGUGU	
	UGGUUCCCGCGGGCUGGCCUUCUUACGGGUUAUGGCCUUGCGUGCCUUGAAUUAUUCACGCC	
	CCUGGCGCAGUACGUGAUUCUUGAUUCCGAGCUUCGGGUUGGAAGUGGGUGGGAGAGUUCGAGGCC	
	UUGCGCUUAAAGGAGCCCUUCGCCUCUGCUUAGUUGAGGCCUGGCCUGGGCGUGGGGCCGCCGC	
	GUGCGAAUUCUGGUGGCACCUUCGCGCCUGUCUCGCGUUCUGAUUAGUUCUAGCCAUUUAAAAUU	
	UUUGAUGACCUGCUGCGACGUUUUUUUCGCAAGAUAGUCUUGUAAAUGCGGGCCAAAGAUUCGA	
	CACUGGUUAUUCGGUUUUUGGGGCCGCGGGCGGCGACGGGGCCGUGCGUCCAGCGCACAUUUUCG	
	GCGAGGCGGGGCCUGCGAGCGCGGCCACCGAGAAUCGAGCGGGGUAGUCUCAAGCUGGCCGCCUG	
	CUUGUGUCCUGGCCUCGCGCCCGCGUUAUCGCCCGCCCGUGGCGGCAAGGCUGCCCGGUGCGGC	
	ACCAGUUGCGUGAGCGGAAAGAUUGGCCGUUCCCGGCCUGCUGCAGGGAGCUCAAAAUGGAGACG	
	CGGCGCUCGGGAGCGGGCGGUGAGUACCCACACAAGGAAAGGGCCUUCUCCGUCCACGCC	
	UCGCUUACUGUGACUCCACGGAGUACCGGGCGCCGUCAGGCACUUCGAUUAUUGUUCGAGCUUUUG	
	GAGUACGUCGCUUUAAGGUUGGGGGAGGGGUUUUUGCGAUGGAGUUUCCACACUGAGUGGGUG	
	GAGACUGAAGUUUAGGCCAGCUUGGCACUUGAUGUAUUCUCCUUGGAUUUGCCCUUUUUGAGUUUG	
	GAUCUUGGUUAUUCUACAAGCCUCAGACAGUGGUUCAAAGUUUUUUUUCUUAUUAUAGGUGUCGUG	
	AGGAAUUCUUCUAGGUCGACCCCGGGGAUCCUUAAGACUAGUGCUAGCAUGCUGCAGUCGAAGGUACC	
	GCGGGCCCGGGAUCCGCCCUUCUCCUCCCCCCCCCUAACGUUACUGGCCGAAGCGCGCUUGGAAUA	
	AGGCGCGGUGCGUUUGUUAUAUUGUUAUUUCCACCAUUAUUGCCGCUUUUUGGCAUUGGAGGGCC	
	CGGAAACUGGCGCCUGUCUUCUACAGCAUUCUAGGGGUCUUUCCCUUCUGCCAAAGGAUUGC	
	AAGGUCUGUUAAUUGCUGAAGGAAGCAGUUCUUCUGGAAGCUUCUUGAAGACAACAACGUCUGU	
	AGCGACCUUUGCAGGCGAGCGGACCCCCACCUGCGGACAGGUGCCUCUGCGGCCAAAAGCCACGU	
	GUUAUAGAUACACUGCAAAGGCGGCACAACCCAGUGCCACGUUGUGAGUUGGAUAGUUGGGAAA	
	GAGUCAAAUGGCUCUCCUACAAGCGUAUUAACAAGGGGCGAAGGAUGCCAGAGGUAACCCAUUG	

177

	GGCGACUGUGGCCAGGUCGCGUGGCGGCCGCCACCCCCUGGGUGGACGCCUUGAUAAUGACA AGGUGGGAAGAGCCUUAUUGGGCUAGCUGGUUUCUCCGGCCUCCUGAAUGCGGCUAACCUUAACCCC AGAGCAUUGGUAGCAACCCAGCUACUAGUAUGUCAUAAUGCGUAAGUCUGGGAGGGACCGACUAC UUUGGAGAGUCCGUGUUUCUUAUUGUUUCUUAAUCAUUCUUAUGGUGACAAUUUAUGUGCCUGAG UAUUGAUUGGUUGGUUGCUUUUGACAAUUAUUGAGACAUCACAUAGACAUA	
2 3 7	UUAAAAUAGCUGCUGGGUUGUCCCAUCCAGCAGGCCACUGGGCGUGAUACUCUGGUUUGCGGUA CCUUUGUUAUGCCUUAUUUACCUCCUCCCCUGAAACUUAAGAAGAAAGAAACACCGCUCACUA GGUACACUGCAUCCAGCAGUGUAUAGCAAGCACUUCUGUUUCCCCGGAAGGGAUUAUUGGUACGC UGUGCAAACGGCGGAAAUUAUCCUACCGUUAACCGCCCAUCUACUCCGAGAAGCCUAGUACCUAAU UGAACUUAUCAUUGGAGUUGCGCUCAGCAGGUGACCCUGACCUGCCAGCUCUCCGGCUGAUGACCUGG GCAUCCCCCACAGGCGACUGUGGCCAGGUCGCGUGGGCGGCCGCCACUCUUUUGAGUGGGACGCC UUGAUAAUGACAAGGUGGGAAGAGCCUUAUUGGGCUAGCUGGUUUCUCCGGCCUCCUGAAUGCGGCU AACCCUAACCCCGGAGCAUUGGUGGCAACCCAGCCACUGGUUUGUCGUAACGCGUAAGUCUGGGAU GGGACCGACUACUUUGGAGAGUCCGUGUUUCCAUUAUUUUCUUUAUUGAAUCUUAUGGUGACAAAUU GGAGUUAUCUGGGUGACGAUUGAUUAUACUUCUGAAACGUAAGAGAUUAAGAAUCAACACA	iSimianV4
2 3 8	UUAAAAACAGCCUGUGGGUUGUCCCAACCCACAGGGCCACUGGGCGCUAGCACACUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUGACUUAACCCUCCCCAAACGCAACUUAAGAAGCACACUUAUAAUGGUCAAU AGGCGGCUAGUAUGCCAACUGAGUCUCGAUCAAGCACUUCUGUUACCCCGGACUGAGUACCAUAG GCUGCUCACGCGCUGAAGGGGAAACCGUUCGUUACCCGACUAAUCUCCGAGAACCCUAGUACCA CCAUGAAAGUUGCGCAGCGUUUCGCUCCGCACAACCCAGUGUAUGAUCAGGUCGAUGAGUACCCGCA AACCCACAGCGGACCGUGGCGGUGGCGUGCGUGGCGGCCUGGCCAUGGGGCAACUAGGGACGCU UCAUAUACUGCAUUGGUGGGAAGAGCUUAUUGAGCUAGUUGGUUAGUCCUCCGGCCCGUAAUGCGGCU AAUCCUAACUGCGGAGCAGAUACCCACGCACCAGUGGGCAGUCUGUCGUAUUGGGCAACUCUGCAGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUUUGUCUUAUACUGGCUACUUAUGGUGACAAUUG AGAGAUUGUAACCAUAUUGCUUAUUGGAUUGGCCACCCGGCGACGAAUAGAACAGUUGCUUACCGUU UGUUAAGUCUGUAUCACUGAACUACAAUCCUUAACACCCUUAUUAUUGGCUAUAACGCUCAAUAC GUUAAA	iSwineVesicu lar
2 3 9	CCCCUCUCCUCCCCCCCCCUAACGUUACUGGCCGAAGCCGCUUGGAAUAAAGCCGGUGUGCGUUU GUCUUAUUGUUAUUUCCACCAUAUUGCCGUCUUUUGGCAUUGUGAGGGCCCGGAAACCUUGGCCUUG UCUUCUUGACGAGCAUUCUAGGGGUCUUUCCCCUCUGCCAAAGGAUUGCAAGGUCUGUUGAAUGU CGUGAAGGAAGCAGUUCUUGAAGCUCUUGAAGACAACAACGUCUGUAGCGACCCUUUGCAGG CAGCGGAACCCCCACCUGGCGACAGUGGCCUCUGCGGCCAAAGGCCACGUGUAUAGAUACACCUUC GAGGAACGUCACCAUUCUCCGGCCAGGCUAAGAGAAGUAGAUAGUUGGAAAGAGUCAAUUGCCUUC CUCAAGCGUAUUAACAAGGGGUGAAGGAUGCCCAAGAGGUACCCCAUUGUAUGGGAUCUGAUCUG GGGCCUGGUGCACAUGCUUUACAUGUGUUAGUCGAGGUUAAAAACGUCUAGGCCCCCCGAACCA CGGGACGUGGUUUUCCUUUGAAAAACACGAUGAUA	cardiovirusA 1;encephalom yocarditisvi rus1
2 4 0	UUGAAAGGGGGCCCGGGGAUCUCCCCCGCGUAACUGGUCACAGUUGCCGCGGACGGAGAUCAUCC CCCGGUUACCCCUUUUCGACGCGGGUACUGCGAUAGUGCCACCCAGUCCUCCUACUCCCGACUCC CGACCCCAACCCAGGUUCCUUGGAACAGGAACACCAAUUAUUAUCCUUGGAUGCUGACUAAUUA GAGGAACGUCACCAUUCUCCGGCCAGGCUAAGAGAAGUAGAUAGUUAAGUAAUCAAUUAUUAUUC AUCCCCUUGACGAUUCGCGUUGGAAAGCACCUCUCACUUGCCGCUUCUACACCCAUCAUUCUUA UUCGGCCCCUGUGUUGAGCCCCUUGUUAAGUGUUUCCUCCAUCGCGACGUGGUUGGAGAUCAAG UUAAACGAGUCCGACGAAACUACCAUAGCUGCCUCCCGAUUAUGUGAUGCUUUUGCCUUGCGUGGU GGAGCAUCCUGGGUUGAGAAAUUUUUCUUAUACUUGGACUCCGGUCCCCCGGUCUUAAGCCG CUUGGAUUAAGACAGGGUUAUCUUCACUCCUUCUUCUUAUACUUCACAGUUGUUAUGCUGUGAAA CGUUGGCGGACUAGCAGAGUCUCGGAAGAAAGCAGUUCUGGUUAAGUUGCCCAAGAAAGAAA CAGCUGUAGCGACCACAAAGGCAGCGGAACCCCCUCCUGGUAACAGGAGCCUCUGCGCCAAAA GCCACGUGGAUUAAGAUCCACCUUGUGUGCGGUGCAACCCAGCACCUGUUUUCUUGGUGACACUC UAGUGAACCCUGAAUGGCGAUUUAAGCGCCUCUGUAGGGAAGCCAGGAUUGUCCAGGAGUACCC CUUCCGUCUGGAAGGGAUCUGACCUGGAGACAUACAUGUGCUUUAACCCUGUGCUUGUGUUUA AAAUUGUCGACGUUCCCAACCAAGUGGUUUGUUUUCUUAUUAUUAUUGUCAU	cardiovirusB 1;Theiler'sm urineencepha lomyelititsvi rus-GD
2 4 1	UUGAAAGGGGGCCCGGGGAUCUCCCCCGCGCAACUGGUUACAGUUGUGCGGACGGAGAUAUCC CCCGGUACCCCUUUUCGACGCGGGUACUGCGAUAGUGCCACCCAGUCCUCCUACUCCCGACUCC GACCCUAACCCAGGUUCCUGGUUAUAGGAACACCAUUAUUAUCCUUGGAUGCUGACUAAUCAG AGGAACGUCAGCAUUCUCCGGCCAGGCUAAGAGAAGUAGAUAGUUAACAUAUCCAAUUAUUAUUA AUCCCUUGACGAUUCGCGUUGGAAAGCACCUCACUUGCCGCUUCUACACCCAUCAUUCUUA UUCGGCCUUGUGUUGAGCCCCUUGUUAAGUGUUUCUCCAUCGCGACGUGGUUGGAGAUCAAG UCAACCGACUCCGACGAAACUACCAUAGUCUCCCGAUUAUGUAGUCUUUCUGCCUUGCGGUG GGAGCAUCCUGGGUUGAGAAAUCCUUCUUCUUAACCUUGGACCCCGGUCUCCCGGCUAAGCCG CUUGGAUUAAGACAGGUUAUUAUCCUUCUUCGAGAACACGCGUGGCGUUUUUCCGUCUCUGACAAGCG CGUUGGCGGACUAGCAGAGUCUCGGAAGAAAGCAGUUCUGGUUAAGUUGCCCAAGAAAGAAA CAGCUGUAGCGACCACAAAGGCAGCGGAACCCCCUCCUGGUAACAGGAGCCUCUGCGCCAAAA GCCACGUGGAUUAAGAUCCACCUUGUGUGCGGUGCAACCCAGCACCUGUUUUCUUGGUGACACUC UAGUGAACCCUGAAUGGCGAUUUAAGCGCCUCUGUAGGGAAGCCAGGAUUGUCCAGGAGUACCC CUUCCGUCUGGAAGGGAUCUGACCUGGAGACAUACAUGUGCUUUAACCCUGUGCUUGUGUUUA AAAUUGUCGACGUUCCCAACCAAGUGGUUUGUUUUCUUAUUAUUAUUGUCAU	cardiovirusB 1;Theiler'sm urineencepha lomyelititsvi rus-DA

2	UUUAAAGGGGGCCCGGGGAUCUCCCCCGGGUAACUGGUGACAGUUGUCGCGGACGGAGAUCAUCC	cardiovirusB 1;Theiler'sm urineencepha lomyeltitsvi rus
4	CCCGGUUACCCCUUUCGACGCGGGUACUGCGAUAGUGCCACCCAGUCUUUCCUACUCCGACUCC	
2	CGACCCUAACCCAGGUUCCUCGGAACAGGAACACCAAUUUACUACUCCCUUGGAUGCUGACUAAUCA	
2	GAGGAACGUCAGCAUUUCCGGCCAGGCUAAGAGAAGUAGAUAGUAGAAUCCAAAUUGAUUUUAU	cardiovirusB 3
4	CAUCCCCUUGACGAAUUCGCUUGGAAAAACCCUCUCACUUGCCGCUUUCACACCCAUUAAUUUA	
2	AUUCGGCCUCUGUGUAGCCCCUUGUAGAGUUGUUCUCCUCCUACGCGACGUGGUUGGAGAUCUAA	
2	GUCAACCGACUCCGACGAAACUACCAUACUGCCUCCCCGAUUAUGUGAUGCUUUCUGCCUUGCGG	cardiovirusB 5
4	UGGAGCACCCUCGGGUUGAGAAAACCUUCCUUCUUCUUGGACUCCGGUCCCGCGUCUAAAGCC	
2	GCUCGGAUAUAGACAGGUUAUUUACCCUUCUUCUACUCCACAGUGUUCUAAUACUGUGGAAG	
2	GGUAUGUGUUGCCCUUCCUUCUUGGAGAACUGCGCGCGGUCUUCUCCGUCUCGACAAAGCGGC	cardiovirusC 2;Boonecardi ovirus2
4	GUGCAACAUACAGAUAAACGCGAAGAAAGCAGUUCUGGUCUAGCUCUAGUGCCACAAGAAAACAG	
2	CUGUAGCGACACACAAAGGACGCGGAACCCCUCCUGGUAACAGGAGCCUCUGCGGCCAAAAGCC	
2	ACGUGGAUAAAGAUCCACCUUUGUGGCGGUGCAACCCAGCACCCUGGUUUCUUGGUGACACUCUAG	cardiovirusC 3
4	UGAACCCUGAAUGGCAUUCUAAAGCGCCUCUGUAGGGAAGCCAAAGAUUCCAGGAGGUACCCCU	
2	CCUUCUGGAAGGGAUCUGACCGGAGACACAUACACAGUCUUAACCCUGGCUUGGUUUAAAAA	
2	UUUUUACAGCUUCCCCGAACCAAGUGGUCUUGGUUUACUUUUUAUCACACUGUCAAU	cardiovirusC 3
4	CGCAGUACCCGCGUUGAAAGAGGAGGGCCGGGGAUGAUUCCACCGCGACACUGUAGCCAGUUGC	
2	CGCGGAGAGAUCCCGGGGCCCCUUUACAGCGGGUACUGCGAUAGUGCCACCCAGCCUUAUCC	
2	UUUCCCCGACUCCCGACCUAAACCCAGGUUCCUCGGAACAGGAACACAAAUUACUACUCCCGUGAUG	cardiovirusC 3
4	CGUGACUAAUCAGAGAAAGUUGCAUUUCCGGCCAGGCUUAGAGAAGUAGAUAGUUAGAAUCUA	
2	AAUUGAUUGACUCCUCCUUGGCGAAUCCGCGUGCAGGGCGCCACCUUCACUUGCCGUCCACACC	
2	CAUCCUUAUUUGGCGCCUCUGUGUUGGACCCCUUGUUGAAGUGGCUUCCUUCUGCGCGCGCGUUGG	cardiovirusC 3
4	AGAAUUUUGUCAUCCGACUCCGACGAAACUACCAUACUGCCUCCCGAUUGUGUGAUGCCUCCGCC	
2	CUGCGUGGUGAGCAUCCUCCGGUUGAGAAACCUUCUCCUUAACCUUGGACCCCGUCCCGGUC	
2	UAAGCCGACAGGGAUAAAGAUCCACUCUUGAGUGCGGUGCGGCCCGGCACCCCGUUCGUGGUUG	cardiovirusC 3
4	CCUGAAGGGUAGUGUUGCCCCUCGCUUCUUGCGGAACGCGCACGGUGGUCUUCACCUCUCGAA	
2	AAGUGUGCGUGCAACAUAGCAGAUACUGAAGAAAGCAGUUUGUGAACUAGUCCUAGAGCGAACAC	
2	GAAUCCGGCUGUAGCGACACUUAAGAGCAGCGGAACCCCUCCGGCGACGGGAGCCUCUGCGGC	cardiovirusC 3
4	CAAAAGCCACGUGGAUAAAGAUCCACUCUUGAGUGCGGUGCGGCCCGGCACCCCGUUCGUGGUUG	
2	UUUUUAGCAGAACCCACAAUUGGCGAUCUUAAGCGCCUCUGUAGAGAAGCCAAAGAUUCCAGAGG	
2	UACCCUUAUCAGGGAUCUGAUCUGGAGACGUGUACUUAUGCUUUAACAGAGUGUCGCGGUUAAA	cardiovirusC 3
4	AACUGUCACAGCUUCCCAAACCAAGUGGUCUUGGUUUUCCUUCUAAUUCUGUGUCAU	
2	GGAGAUAUCCCCCGGUUACCCCUUUCGACGCGGUACUGCGAUAGUGCCACCCAGUCUUUCCUA	
2	CUCCCGACUCCCGACCCUAACCCAGGUUCCUCGGAACAGGAACACAAAUUACUACUCCCGUGAUG	cardiovirusC 3
4	CGUGACUAAUCAGAGGAACGUCAGCAUUUCCGGCCAGGCUUAGAGAAGUAGAUAGUUAGAAUCUA	
2	AAUUGAUUGACUCCUCCUUGAGCAUUUACGUUGGAAUUGCACCUCACUUGCCGCUUUCACAC	
2	CCACUAAUUGAUUCGGCCUACUGUGUUGAGCCCCUUGUUGAAGUGCUUCCUCCUUGUGACGUGGU	cardiovirusC 3
4	UGGAGAAAUUCUUGUACCCGACUCCGACGAAACUACCAUAGUCCUCCCGAUUAUGUGAUGCUUUC	
2	UGCCUUGCUGGGUGGAGUAUCCUCGGUUGAGAAUCCUUCUCCUUAUACCUUGGACCUUGGUC	
2	CCGGUCUAAGCCGCUUGGAAUAAAGACAGGGUUAUUUACCCUUCUUCUUCUACUUCUAGGUGCUC	cardiovirusC 3
4	UAUACCAUGAAAGGUUAGUGUGCCCCUUCUUCUUGGAGAACUCACGCGGCGGUCUUCGUGUCU	
2	CAAAAAGCGCGAGUGCGACAUGCAGAGUAACGCGAAGAAAGCAGUUCUUGGCCUAGCUCUAGUGCCC	
2	ACAGAAACCGCGUGUAGCGACACACAAAGCAGCGGAACUCCCUCCUGGUAACAGGAGCCUCUG	cardiovirusC 3
4	CGGCCAAAAGCCACGUGGAUUAAGAUCCACCUUUGUGGCGUGCAACCCAGCACCCCGUUCUUG	
2	UUGACACUCUAGUGAAUCCUUGAAUGGCAUUCUACAGCGCCUCUGUAGGAAAGCCAAAGAUUGCCAG	
2	UUGGUAACCCCUCCUGCGGAAGGGAUCUGACCUGGAGACACAUACACGUGCUUUAACUUGGCU	cardiovirusC 3
4	UGUGUUUAAAAUUGUACAGCUUUCCAAACCAAGUGGUCUUGGUUUUCCUUUUUUACCUACUGU	
2	CAAU	
2	GGCAUACCUCUACCCUGAGCGGCGGCUAAGGAUGCCUGAAGGUACCCUAGUAGAAUCCGUCUGGCG	cardiovirusC 3
4	ACCAUGGAUCUGAUUAGGGGCCUUGCCUGGAGUGGAUCUUAUCCACACAGCGUAGGGUUAUAAAAACG	
2	UCGAACCGCCCCACAAUGACCCCGGAGGGAUGCCGUUUUUCUUAUCCAAUUCUGACACU	
2	GGCCUAGUCGCGUGCGGUUCCCUUUGUUUUUGCUAUGCAAUCGGGGAUGACACCCCUUUCACGUG	cardiovirusC 3
4	GGUACUACGAAAGUGCCCCUCGCUCCGAGGUUAAAGGAGAACCCCUUUCUUAACCCACUACAGCU	
2	CGCCCUACAGUGUGGGCGUAGCCUUCACUUGCAGCUUCUGCUUGUAGAUGCUUGCACCCGUGAUC	
2	GGUGCGCUUCUUGCUUUUAGUCGCUUGUGCUUUAUCGUUUCUGACGAUUCAGUUCUUAACGCCAGUG	cardiovirusC 3
4	UUUCGACGGCCCAAGGGGUAGUUGCGGUCAGUUAUCCUACCCGAUUAUCCCUUUCGCCGUCGUA	
2	CGUGGUUUGGAUCUUGGAUCUCUCCUUCUCCCGUCUUAUUAUAGCUUCUGAUGAUGAAGCA	
2	UCUCACUGUCUCUAGUUAUUUAGUGCAACUGACGAUUGAGUACGUUUCAGAUUGUGUUUGGAGGCCC	cardiovirusC 3
4	AAGGGACCGAUGGACAACACUUCGAAAGAGUCACUUGUCCACCCGCUUUCUCCCUACCCUAGCAAC	
2	UCUGGAUUUGUCUACGUGGAGUUCGAGGUCUAUACUUUAACUCUGACUUGCUUUUUAACCUUGCUA	
2	UCUUGCUGACGUGGAUUGGUUGUAGACUGGUUACGUUCCCGUUAAGUUGCUGACGUGGAGUACGAUC	cardiovirusC 3
4	GCUGUACAUUCCACUACUGCCAAUUAAGCUCCCUUCCCGUUGCUCCCUUUAUAGGAGAGCCUUC	
2	UUUUGCAAAGGUGAAGCCUACCCCGGUCGAAGCCGCUUGGAAUAAACAGGCGUUAUUUUCUCCU	
2	CUUCUGCGCGCUUGCCUUUUCUAAAGCUGAAUAGGUCCUUAUCUUAUACGGCGGAUGGUCUGGCGUU	cardiovirusC 3
4	CCUCCUUGGACAGAGUGUUAUCUGGGUUUUCGGAUUCGACCGCACUCACACAGAGCUCAGGAG	
2	UAGCUAAGUCAAGGCCCGAUCUGCGGCGAAAGGAAUAGAAUUAUUUGCAGCUGUAGCGACCUUC	
2	UAGGCCAGCGGAUCUCCCAACCUUGUAACAGGUGCCUCUGGGGCCAAAAGCCACGUGUUAUAGCAC	cardiovirusC 3
4	CCUUGAGAGCGGUGGUUACCCACCCUGCAAACUAGGAUUUGACUUAAGUAAACAAAAGAUUGAC	
2	UUGGCAUACCUCUACCCUGAGCGGCGGCUAAGGAUGCCUGAAGGUACCCGUGUUGAAUCCGCUCCG	

	CGACCAUGGAUCUGAUCAGGGGCCUGCCUGGAGUGGUUUCUGUCCACACAGCGUAGGGUUAACAAACUUGACACU	
2	UUUCAAGGGGGCCUGGGGUCUUUGUCGCGGAGAUAGCGUAGCAUCUUCGCGGUAAAGAACACCC	cardiovirusD 1;Saffoldvir us1
4	AGGGGUAACCCUUUGAACGCGGGACUGCGAUAGUGCCACCCACGUCUUUCCACUCCUAACCCGA	
7	CUCCACCUCGCGGUCCAAUGAAAGGACCAAUAAUUAUCCUUGAAUGUUGGCUAAUCAGAGGAAA	
	GUCAGCAUUUCCGCGCCAGGCUAAGAGAAACAUAUAAGUUAGAUUUGAUUCAAUCCUUGACGAA	
	UUCGAAACGAGAUUUUCCUCCUCCUUGCCGCUUGUUCACACCAUACAUUUAUUCGGCCUCCUGU	
	GACAAGCCCCUGGUGAAAGAACCCUCUCUUUUCGACGUGGUUGGAAUUAACAUAUUUCCGACGA	
	AAGUGCUAUAUGCCUCCCGAUUAUGUGAUUUUUCGCCCUGCUGGGCGGAGCAUUCUGGGUUG	
	AGAAACUUGAAUUCUUUCCUUGGAACCUUGGUUCCCGGUCUAAGCCGCUUGGAAUUGACAGG	
	GUUAUUUUUCUUUAUUCUUAUUCUACUUUACGGGUUCUAUCCGUGAAAAGGGUACGUGUUGCCCUU	
	CCUUCUUCGAGAAUUCACACGCGGCUUUUCCGUCUCUCGACAAGUGUGAAUGCAACAUGCCGAA	
	ACGGUGAAGAAACAGUUUUCGCGGAAGUCUAGAGUGCGCAUCGAAACAGCUGUAGCGACCCUACA	
	GUAGCAGCGGACUCCCUUUGGCGACAAGAGCCUUCGCGGCCAAAAGCCCCGUGGAUAAGAUCAC	
	UGCUGUGAGCGGUGCAACCCAGCACCUGAUUCGUAUGAACGUUUCUUAACGGAACAGAGGAUGUU	
	UUCUUAAGCCUCCGGUAGAGAAGCCAAAGAUUGCCUGAAGGUACCCCGCGUGCGGGAUCUGAUCAG	
	GAGACCAAUUGACUGUCUUUACACCGUCACUUUGGUUUAACACUGUCACAGCUUCCAAACCAA	
	GUGGUCUUGGUUUCCAUAUUUAUUAACUGGCAU	
2	UUUGAACGCGGACUGCGAAAGUGCCAACCCACGUCUUUCCACUCCUAACCCGACUCCACCUCG	Saffoldvirus 3
4	GGUCCAAUGAAAGGACCAUAUAUUAUCCUUGAAUGCUGGCUAAUCAGAGGAAAGUCAGCAUUUU	
8	CCGCGCCAGGCUAAGAGAAACACAUAAGUUAGAUUUGACUCAAAUCCUUGACGAAUUCGAAACGAG	
	UAUUCUCCUCCUCCUUGCCGCUUGCUCACACCCACACAUAUUAUUCGCGCCUUCUGGACAGCCCU	
	UGGUGAAAGAACCCUCUCUUUUGACGUGGUUGGAAUUAACAUAUUCGACGAAAGUGCUAUA	
	UGCCUCCCGAUUGUGUGAUUUUUCGCCCUGCUGGGCGGAGCAUUCUGGGUUGAGAAACCUUGA	
	AUCUUUCCUUGGAACCUUGGUUCCCGGUCUAAGCCGCUUGGAAUAUGACAGGGUUAUUUCCU	
	UAUCUUACUUCUACUUUCACGGGUUCUAUCCGUGAAAAGGGUACGUGUUGCCCUUCCUUCUCCGA	
	GAUUCACACGCGGUCUUUCCGUCUCUCGACAGGUGUGAAUGCAACUGCCGGAACGGUGAAGAA	
	AACAGUUUUCGCGGAAGUCUUGAGUGCGCAUCGAAACAGCUGUAGCGACCUACAGUAGCAGCGGA	
	UCCCCUCUUGGUGACAAGAGCCUCUGCGGCCAAAAGCCCCGUGGAUAAGAUAUCCACUGCUGAGCG	
	GUGCAACCCAGCACCCUGAUUCGAUGAACAUUCUUUACGGAACAGAGAUGGUUUUCUUAAGCCC	
	UCCGGUAGAGAGCCAAAGAUUGCCUGAAGGUACCCCGCGUGCGGGAUCUGAUCAGGAGACCAUUG	
	ACUGUGCUUAACAGUCACUUUGGUUUAACACUGUCACAGCUUCCAAACCAAGUGGUUUGGU	
	UUUCCAUAUUUAACUACUGGCAU	
2	UUUCAAGGGGGCCUGGGGUCUUCGUCGCGGAGGUGCGUAGCACCUUCGCGGCGAAGAACACCC	Saffoldvirus 4
4	AGGGGUACCCCUUUGAACGCGGGACUGCGAGAGUGCCAAACCCACGUCUUUCCCAUUUUACCC	
9	GGUUCACCUCGCGGUCCAAUGAAAGGACCAAGCAAUUCAUCCUUGAAUGUUGGCUAAUCAGAGGA	
	AAGUCAACAUUUCCGCGCCAGGCUAAGAGAAACAUAUAAGUUAGAUUUGAUUCAAUCCUUGACG	
	AAUUCGGAACGAGAUUUCUCCUCCUUGCCGCUUGUUCACACCAUGUUUAGAUUCCGCGCUUU	
	GUGACAAGCCCUUGGUGAAAGAACCCUCUCUUUUCGACGUGGUUGGAAUUAACAUAUUCGAC	
	GAAAGUGCUAUAUGCCUCCCGAUUGUGUGAUUUUCUGCCUUGCGGCGGAACAUUCGCGGU	
	UGAGAAACCUUGAAUUCUUUCCUUGGAACCUUGGUUCCCGGUCUAAGCCGCUUGGAAUAUGACA	
	GGGUUAUUUUAUAUUAUUAUCCACUUUUGGGUUCUAUCCACAAAAGGGUACGUGUGCCCC	
	UUCUUUCUUUGGAGAACUCGCAAGCGGCUUUUCCGUCUCUAACAAGUGUGAGUGCAGCAUGCCG	
	AAACGGUGAAGAAAACAGUUUCUUGGGAAGUCUAGAGUCACAUCCGAAACAGCUGUAGCGACCUCA	
	CAGUAGCAGCGGACUCCCUUCUUGGCGACAAGAGCCUCUGCGGCCAAAAGCCCCGUGGAUAAGAUC	
	ACUGUGAGAGCGGUGCAACCCAGCACCCUGAUUCGAUGGCUAUUCUUAACGGAACAGGAUAUG	
	UUUUUUAAGCCUCCGGCAGAGAAGCCAAAGAUUGCCUGAAGGUACCCCGCGUGCGGGAUCUGAUC	
	AGGAGACCAUUCGACAGUCUUUACACUGUCGUUUUGGUUUAACACUGUCACAGCUUCCAAACC	
	AAGUGGUCUUGGUUUUUAUUUUUGUUGAAUGGCAU	
2	UCCCUUGCCGCUUGUUCACACCAUAUAUUAUUAUUCGCGCCUUGUGACAAGCCCUUGGUGAAAGA	Saffoldvirus 5
5	ACCUCUCUUUUCGACGUGGUUGGAAUUAACAUAUUUCCGACGAAAGUGCUAUCAGCCUCCCG	
0	AUUAUGUGAUUUUUCUGCCUUGCGGCGGAGCAUUCGCGGUUGAGAAACCUUGAAUUCUUUCCU	
	UUGGAACCUUGGUUCCCGGUCUAAGCCGCUUGGAAUAUGACAGGGUUAUUUUCUUUAUCUUAUU	
	CUACUUUACGGGUUCUAUCCGUGAAAGGGUAUGUGUUGCCCUUCCUUCUUCGGAGAUAUCACAC	
	GGCGGUCUUUCCGUCUCUCGACAAGUGUGAAUGCAUAUCCGGAACCGUGAAGAAACAGUUUUC	
	UGCGGAAGUCAGAUUGGCAUCCAAACAGCUGUAGCGACCUACAGUAGCAGCGGACUCCCUU	
	AGCCUUGGUUCCCGGUCUAAGCCCGUUGGAAUAAGAUCCACUGCUGUGAGCGGUGCAACCC	
	AGCACCCUGAUUCGAGAACAUUCUACCGAACCAGAGGAUGGUUUUCUUAAGCCCUCCGGUAGAG	
	AAGCCAAGAAUGUCCUGAAGGUACCCCGGUGCGGGAUCUGAUCAGGAGACCAUUGA	
2	GAGAAACAUAAGUUAGAUUCGAUCCAAUCCUUGACGAAUUCGAAUUGAGAUGUUUCCUCC	Saffoldvirus 6
5	UUGCCGCUUGUUCACACCAUAUUUGUGAUUCGGCCUUCUGUGACAAGCCCUUGGUGAAAGACCU	
1	CUUCUUUUCGACGUGGUUGGAAUUAACAUCUUUCCGACGAAAGUGCUAUAUGCCUCCCGAUUA	
	UGUGAUGCUUUCUGCCUUGCGGCGGAGCAUUCGCGGUUGAGAAACCUUGAAUUCUUUCCUUGG	
	AACCUUGGUUCCCGGUCUAAGCCCGUUGGAAUAUGACAGGGUUAUUUUCUUUAUCUUAUUUCUAC	
	UUUCACGGGUUCUAUCCGUGAAAGGGUACGUGUUGCCCUUCCUUCUUCGGAGAUAUCACAGCGG	
	GUUUUCCGUCUCUCGACAAGUGUGAAUGCAUUGCCGGAACGGUGAAGAAACAGUUUUCGCG	
	GAGUCUAGAGUGCAUCAGAAACGGCUGUAGCGACCUACAGUAGCAGUGGACUCCCUUUGGCG	
	ACAAGAGCCUCUGCGGCCAAAAGCCCGUGGUAAGAUCCACUGCUGAGCGGUGCAUCCCGAGCA	

	CCUGAUUCCGAUGAGCAUUCUCUACGGAACCAAGAGAAUGGUUUUCUCAAGCCCUCCGGUAGAGAAGC CAAGAAUGUCCUGAAGGUACCCCGGUGCGGGAUCUGAUCAGGAGACCAAUUGA	
2 5 5	UUUCAAGGGGGCCCUUGGGGUCUUUGUCGCGGAGAUAGCAUCUUCGCGGUGAAGAACACCCC AGGGGUACCCCUUUGGACGCGGGACCUUGCGAGAGUGCCAACCCACGUCUUUCCCAUCCCUACCCC GACACUCACCCUGGGUCCAUGAAAGGACCAAAUUAUUCAUCCCUUGAAUGCUGGCUAAUCAGAGGA AAGUCAGCAUUUCCGGCCAGGCUAAGAGAAACAUAUAAGUUAGAUUUGACUUAAUCCCUUGACG AAUUCGGAACGAGAUUUCUCCUCCUUGCCGCUUGUUCACACCCAUACUUCAAUUGACCCCUU GUGACAAGCCCUUGGUGAAAGAACCUUCUCUUUUCGACGUGGUUGGAAUUAACACCAUUUCCGAC GAAAGUGUUAUCAUGCCUCCCGAUUAUGUGAUGCUUUCUGCCUUGGCGGAGCACUCUGGGU UGAGAAACCUUGAGCCUUUUUUGGAACCUUGGUUCCCGGUCUAGCCGCUUGGAAUUAUGACA GGGUUAUCUUUUUAUCUUUAUUUCUACUUCACGGGGUCUUAUCCGUGAAAAGGUGACGUGUUGCCCC UUCCUUUCUGGAGAAUUCACACGGCGGUCUUCCGUCUCUGACAAGUGUGAAUGCAACAUCCCG AACCGGUGAAGAAAACAGUUCUUUGUGGAAGUCUAGAGUACAUCAAGAACAGCUGUAGCGACCUCA CAGUAGCAGCGGACUCCCUUCUUGGCGACAAGAGCCUCUGCGGCCAAAAGCCCGUGGAUAGAUC ACUGCUGAGCGGUGCAACCCAGCACCUGAUUCGAUGAACAUCUUUACGGAACCAAGAGGAUGG UUUUUCAAGCCCUCCGUGAGAGAAGCCAAGAAUGUCCUGAAGGUACCCCGGUGCGGGAUCUGAUC AGGAGACCAAUUGACUGUCUUUACACCGUCACUUUGGUUAAAAAUUGUCACAGCUUCUCCAAACC AAGUGGUCUUGGUUUUCCCAUUUUGUGACUGGCAAU	Saffoldvirus 7
2 5 3	UUUCAAGGGGGCCCUUGGGGUCUUUGUCGCGGAGAUAGCAUCUUCGCGGCGAAGAACACCCC AGGGGUACCCCUUUGAAGCGGGACCUUGCGACAGUGCCAACCCACGUCUUUCCCAUCCCAUCCC GGCUCACCCUUGAGUCCAUAGAAAGGACCAAGUACAUCUCCCUUGGAUGCUGGCUAAUCAGAGGA AAGUCAGCAUUUCCGGCCAGGCUAAGAGAAACAUAUAAGUUAGAAUAGUACACAUCCCUUGACG AAUUCAGAACGAGAUUUCUCCUCCUUGCCGCUUGUUCACACCCACUUAUAAUUCGGCCUUUU GUGACAAGCCCUUGGUGAAAGGACCUUCUCUUUCGACGUGGUUGGAAUUGACAUUUUCCGAC GAAAGUGCUAUCGUGCCUCCCGAUUAUGUGAUGUUUUCUGCCUUGGCGGAGAGCAUUCUGGGU UGAGAAACCUUGAACCUUCUCCUUGGAACCUUGGUUCCCGGUCUAGCCGCUUGGAAUUAUGACA GGGUUAUUUCCCUUAUCUUUAUUUCUACUUUCAUGGGUUCUUAUCCAUAGAAAGGGAUGUGUUGCCCC UUCCUUCUUGGAGAAUUCACACGGCGGUCUUCCGUCUCUGACAAGUGUGAAUGCAACAUCCCG AAACGGUGAAGAAAACAGUUCUCUGUGGAAGUCUAGAGCGGCUAUCGAAACAGCUGUAGCGACCUCA CAGUAGCAGCGGACUCCCUUCUUGGCGACAAGAGCCCUUGCGGCCAAAAGCCCGUGGAUAGAUC ACUGCUGAGCGGUGCAACCCAGCACCUGAUUCGAUGAAGUUCUUUACGGAACCAAGAGGAUGG UUUUUCUAGCCCUCCGUGAGAGAAGCCAAGAAUGUCCUGAAGGUACCCCGGUGCGGGAUCUGAUC AGGAGACCAAUUGACGUGUCUUUACACCGUCACUUUGGUUAAAAACUGUCACAGCUUCUCCAAACC AAGUGGUCUUGGUUUUUCUACUUUAUUGACUGGCAAU	Saffoldvirus 8
2 5 4	UUCAAAGGGGGCCCUUGGGGUCUUUGUCGCGGAGAUAGCAUCUUCGCGGCAAGAACACCCCA GGGUAACCCCUUUGAAGCGGGGUCUUGCGGAGAGGCCAACCCAGUCUUUCCCAUCUUUCCCGA CUCCACCCUUGGUCUAUUGAAAGACCAAAUUAUUCAUCCCUUGAAUGCUGGCUAAUCAGAGGAA GUCAGCAUUUCCGGCCAGGCUAAGAGAAACAUAUAAGUUAGAUUUGAUUUAUCCCUUGACGAU UCGAAACGAGAGCUUUUCCUCCUUGCGCUUGUUCACACCCACUUAUUAAUUCGGCCUUUUUG ACAAGCCCUCCGUGAAAGAACCUUCUCUUUUGACGUGGUUGGAAUUAACAUUCUCCGACGAA AGUGCUAUCAGCCUCCCGAUUAUGUGAUGUUUUCUGCCUUGGCGGAGCAUUCUGGGUUGA GAAACCUUGAAUCUUUCCUUGGAACCUUGGUUCCCGGUCUAGCCGCUUGGAAUUAUGACAGGG UUUUUUUCUUAUCUUUAUUUCUACUUUACGGGUCUUAUCCGUAAGAAAGGUAUGUGUCCCUUC CUUCUCUGGAGAAUCUACACGGCGGUCUUUCCGUCUUUAACAAGUGUGAUGCAGCAUGCCGAAA CGGUGAAGAAACAGUUUUCUGUGGAAGUUUAGAGUGCAUCCGAAACAGCUGUAGCGACCUACAG UACGAGCGGACUCCCUUCUUGGCGACAGAGCCUCUGCGGCCAAAAGCCCGUGGAUAGAUCACU GCUGUGAGCGGUGCAACCCAGCACCCUGAUUCGAUGGUCAUUCUACGGAACAGAAAUGGUUU UCUCAAGCCCUCCGUGAGAGAAGCCAAGAAUCUCCUGAAGGUACCCCGGUGCGGGAUCUGAUCAGG AGACCAAUUGACGGUGCUUUACACUGUCACUUUGGUUAAAAUUGUCACAGCUUCUCCAAACCAAG UGGUCUUGGUUUUCCAAUUUUGUGAGUGGCAAU	Saffoldvirus 9
2 5 5	UUUCAAGGGGGCCCUUGGGGUCUUUGUCGCGGAGAUAGCAUCUUCGCGGUAAGAACACCCC CAGGGGACCCCUUUGAACGCGGGACCUUGCGAAAGUGCCAACCCACGUCUUUCCCAUCUUUCCC GAUUCACCCUUGGUCUUAAGAAAGGACCAAAUUAUUUUAUCCCUUGGAAUGCUGGCUAAUCAGAGGA AAGUCAGCAUUUCCGGCCAGGCUAAGAGAAACAUAUAAGUUAGAUUUGGUUCAAUCCCUUGACG AAUUGGAAACGAGAUUUCUCCUCCUUGCCGCUUGUUCACACCCACAUUUAAUUCGGCCUUUU GUGACAAGCCCUUGGUGAAAGAACCUUCUCUUUCCACGUGGUUGGAAUUAACAUCAUUUCCGAC GAAAGUUGUUAUCAUCCUCCCGAUUAUGUGAUGCUUUCUGCCUUGGCGGAGCAUUCUGGGU CGAGAAACCUUGAAUCUUUUCUACUGGAACCUUGGUUCCCGGUCUAGCCGCUUGGAAUUAAGACA GGGUUAUUUUCUUUAUCUUGUUUUAUUUUAUGGUUUCUUAUAAAAAGGGAUGUGUGGCCCC UUCCUUCUUGGAGAAUUCACACGGCGGUCUUCCGUCUCUGACAGGUGUGAAUGCAGCAUGCCGG AAACGGUGAAGAAAACAGUUUUCUGUGGAAGUCUAGAGUGUACAUCAAGAACAGCUGUAGCGACCUCA CAGUAGCAGCGGACUCCCUUCUUGGUAAACAAGAGCCUCUGCGGCCAAAAGCCCGUGGAUCAGAUCC ACUGCUGAGCGGUGCAACCCCGGACCCUGAUUCGAUGGACAUUCUACGGAACAGAAAUGGUUU UUUUUCUAGCCCUCCGUGAGAGAAGCCAAGAAUCUCCUGAAGGUACCCCGGUGCGGGAUCUGAUCAGG AGACCAAUUGACGGUGCUUUACACUGUCACUUUGGUUAAAAUUGUCACAGCUUCUCCAAACCAAG AAGUGGUCUUGGUUUUUCUAAUUUUAUUGAUGGCAAU	Saffoldvirus 10
2 5 5	UCAAGGGGGCCCUUGGGGUCUUUGUCGCGGAGAUAGCAUCUUCGCGGUGAAGAACACCCAG GGGUACCCCUUUGAACGCGGGACCUUGCGAGAGUGCCAACCCACGUCUUAUCCCAUCUUUACCCGAC CCCCACCUUGGUCUUAAGAAAGGACCAAAUUAUUUUAUCCCUUGGAAUGCUGGCUAAUCAGAGGAA AGUGGUCUUGGUUUUUCUAAUUAUUAAGUUAGAUUUGGUUCAAUCCCUUGACG AAUUGGAAACGAGAUUUCUCCUCCUUGCCGCUUGUUCACACCCACAUUUAAUUCGGCCUUUU GUGACAAGCCCUUGGUGAAAGAACCUUCUCUUUCCACGUGGUUGGAAUUAACAUCAUUUCCGAC GAAAGUUGUUAUCAUCCUCCCGAUUAUGUGAUGCUUUCUGCCUUGGCGGAGCAUUCUGGGU CGAGAAACCUUGAAUCUUUUCUACUGGAACCUUGGUUCCCGGUCUAGCCGCUUGGAAUUAAGACA GGGUUAUUUUCUUUAUCUUGUUUUAUUUUAUGGUUUCUUAUAAAAAGGGAUGUGUGGCCCC UUCCUUCUUGGAGAAUUCACACGGCGGUCUUCCGUCUCUGACAGGUGUGAAUGCAGCAUGCCGG AAACGGUGAAGAAAACAGUUUUCUGUGGAAGUCUAGAGUGUACAUCAAGAACAGCUGUAGCGACCUCA CAGUAGCAGCGGACUCCCUUCUUGGUAAACAAGAGCCUCUGCGGCCAAAAGCCCGUGGAUCAGAUCC ACUGCUGAGCGGUGCAACCCCGGACCCUGAUUCGAUGGACAUUCUACGGAACCAAGAGAAUGG UUUUUCUAGCCCUCCGUGAGAGAAGCCAAGAAUCUCCUGAAGGUACCCCGGUGCGGGAUCUGAUCAGG AGGAGACCAAAUAGCAGUGCUUUACCGUUAUUGGUUAAAAUUGUCACAGCUUCUCCAAACCA AAGUGGUCUUGGUUUUUCUAAUUUUAUUGAUGGCAAU	Saffoldvirus 11

1	UCAGCAUUUUCCGGCCAGGCUAAGAGAAACAUUAUAGUUAAGAUUCAAUCCUUGACGAAU UCGAAACGAGAUUUUCUCCUCCUUGCGCUUGUUCACACCCAUACACUUAUUCGGCCUUCUGUG ACAAGCCCCUCGGUGAAGAACCCUCUCUUCUUCGACGUGGUUGAAUUAACAUUCCCGAGAA AGUGCUUAUUGCCUCCCGAUUGUGUGAUGUUUUCGCCUGCGGCGAGCAUUCGCGGUUGA GAAACCUUGAAUUCUUUCCUUGGAAACCUUGGUUCCCCCGGUCUAAGCCGCUUGGAAUUGACAGGG UUUUUUUCUUUAUCUUUUUCUACUUUACCGGCUUCUACCCGUGAAAGGGUAUGUGUGUCCCCUUC CUUCUUCGCGAAGAUUCACACGCGGCUUGUCCGCAAGUGAAGUCCGCAACUUGCCGAGAA CGGUGAAGAAAAAGUUUUCUGCGGAAGUCUAGAGUCGCAUCGAAACCGGUGUAGCGACCCACAG UAGCAGCGGACUCCUCCUUGUGUACAGAGCCUUCUGCGGCCAAAAGCCCCGUGAUUAGAUCCACU GCUUGAGCGGUGCAACCCAGCACCUGAUUCGAUGAACUUCUACCGAACAGAGGAGUGUUU UCUUUAAAGCCUCCGGAAGGAGGACCAAGAAUGUCCUGAAGGUUCCCGCGUGCGGGAUCUGAAGG AGACCAUUGACUGUGCUUUAACACCGUCACUUGGUUUAAAAACUGUCACAGCUUCUCCAAACCAAG UGGUCUUGGUUUUCCAUUUUUGUGACUGGCAU	
2	UUUCAAAGCGCGGCGUGGGUGUGGGUGCGGCAACAGUCAGUGUGUGCGCGGUCCACUCCCUCA 5 GCAACCCGCUUUGGACGUGGGUACUACGAUAGUGCCAUCCAUUUCCCUUACCUUGUUAACCUCCG 7 AGCGACUCCAACACCCGAGGUGUCACACCCUUGUAAGGAAAAUUGAAUUCGUCUGUGCGGUC UAAACGAUUAAGAUUAGCAUUUUCCGGCCAGGCUAAGAACCCUUAAGAUAGUCGAUAAACUU AAUUUGGCGAACUGUAACAAAGAACGUGUCCUACCUUCCGCCACACACCAACCCUAAAAUUUG GCCUUUAGUGUGUGGCCCCUUGUGUAUGACCUCACUCUUGCGUACACGCGGAUUUAACAUUUUC CGCUAAAGGAGUUUUCGUGUCCUCCCGAUUAUGUGAUGCUCGUCUUUUGCGGGGAGCAACGCA CGAAACAGGCCUUUUUCUUAACUUAAGAAUCCAUAUCCCGCGUCUAGCGCGUUAAGAACAG GGUUUUUUCACCCUUCUUCUACUCUACAGAGCUCUAUUCUGUGAGAAGGGAUUGUUGCGCC CCUCCUUCUUGGGAGUCGUCUGUGCGGUGUUCUCCGUCUCUGUGCAAGAGCGAGUGCGCAUUGUGU GAAGGUGAAGAAAGAAUUAUUUGGGAUUAACGUUUGAAGCGAACGUCUACGAGCCACUACG AAAGGCAGCGGAACCCCCUCCUGUGACAGGAGCCUCUGCGGCCAAAAGGCCAGUGGUUAGCAGCC ACCUUUGUGAGCGCGGUGAACCCGACGCCACGGUUUCGUGUGUGACACUACUACCCCAUUCGGU CUUUUUUAAAGCCUUAACACAGUCCACCAAGAAUGUCCUGAAGGUACCCUUCGGAUGAUCUGAUCAG GAGCUUACGCGGUGCUUUAACCGGUCGGAAGUUAAAAACUAGCAGUGGACCAAAUUCUAGA UGGUUCUAGAUAUUCUUAUGGUUUUAAAAUUUCAUU	cardiovirusE 1;red- backedvoleca rdiovirus
2	CUAGACUUUCUACCCUUCUUGAAUGGCGGAAUUCAGAGUAAAGUCACCUUCCUUGCCGCUUAACAC 5 ACCCAACCAUUCUUGCGGCCAGGUGUAUAGCCCUUUGGUAAGUACCUUUAAGCUUGGAGCG 8 GGAUUUAAGACCUUGAAUUAUCCGUGAAGAUUAACACCAUGUUUGCCCCGAUUAUGUGAUGUUUCGC UCUUGCUGGGCGGAACUCCGUCGGGAUGCCAAAGCAAUUGUGUCCCCUUCUUUUUUGGACUUCG GUCCCCGUGCGAAGCCGCUUGGAUAAGACAGGGUUAUUUUCCUCCUUCUUCUUCUACUUCUUG UAAGCCUUAUUUAAACAAAGGGAUGUUCUGGCCUCCUUCUUGCGGAUUGUGCGGCGGUCUUU CCGUCUCUCGAAAAGCGUGCAUGCCAGAUCCAAGAAAGUGAAGAAAGUAGUUGUGUCUAACGUU AAGACAGGAACGAUUGUGUCUGUAGCGACCCUUCUCAAAGGCGCGGUCUCCCCACCCUGGUAACAGGUG CCGACGCGGCCAAAAGCAGUGGUUAACAGCCACCUUUGAGAGCGGUGUACCGGCAACCCACACU UUCGUUGUGUUUUCUACUAACCAACACUCGCUUCUCUGAAGCCCUUUGUAGAGUUGCCAAGAAU GUCCAGAAAGUACCCCGUGACCGGGAUGAUCUGGAGCUCUACGCGUGUCUCUACACAGGUUCG GGAGUUAAAAACUGUCACCAACUCAAACUAGGUGUCCUAGUUUACCUUUUUUUGUCAAGUC AUUAGUGUAGGCUUAAAGGAGAACGACA	cardiovirusF 1;grey- backedvoleca rdiovirus
2	UUUAAAACAGCCUGUGGGUUGCUCCACCCACAGGGGCCACUGGGCGCUAGCACUGAAUUUACGG 5 AAUUCUUUGUGCGCCUGUUUUUAUCCUCCUCCCCAACAGAAUUUAAGAAGUUUAUGCAUUUAUGG CCAGUAGCGGCAUGAGCCACACAGUACUGUCUUGGCUAGCAUUGUUUCCCGGACUGAGUAC AAUAAACUGCUCACGCGGUGAAGGAGAAAAAGUUCGUUACCCCGCUAACUACUUCGAGAAACCCAG UAACACCAUGAAAGUUGCGGAGUGUUUUCGUCACGACUUCUCCCGUGUAGUACAGGUCGAGUAGUCA CUGUAAACCCACGGGCGACCGUACAGUGGCGUGCGUUGCGCCUUGCCAGUGGACAACCCAGUGG ACGCUCUAAUACAGACAUGGUGUGAAGAGUCUAUUGAGCUAGUAGUAGUCCUCCGGCCCCUGAAUG CGGCUAAUCCUUAACUGCGGAGCACACACCCUCAAUCCAGGGGGCAGUGUGUCUAAACGGGCAACUCU CGAGCGGAACCGACUACUUUUGGGUGUCCGUGUUUCCUUUUUAUUCUUUUGUGGCGCUUUAUGGUGAC AAUUAAAGAAUUGUUACCAUUAAGCUUAUGGAUUGGAAUUGGCUUACACAGACGAUUAUGUGAU UUUUUUGCUGGAUUUGUCCAUUAUCAUACACAUUUUUUAUACUUUAUUAUAAAAUUGA ACACAAGAAA	coxsackievir usA4
2	UUAAAACAGCCUGUGGGUUGCACCCACCCACAGGGGCCACUGGGCGCAGCACACUGAUUUCGCGGU 6 ACCUUUGUGCGCCUGUUUUUAGACCCUCCUCCCAACCGAAACUUAGAAGUUAACACAUACCGAUCAAUA 0 GUGGGCAUAGCAGCGCCAGCUGUGUCUUGAUAACGACAUUCUGUUUCCCGGACCGAGUAUCAAUAGA CUGCUCACGCGGUGUAGGAGGAAGCGUUCGUUAUCCCGCUAACUACUUCGAGAAACCUAGUAACAC CAUGAAAGUUGCGGAGUGUUUCGUAUUAUCCUCCAGUGUAGUACAGUCUACGUGACUACGCGGU UCCCCACGGGCGACCGUGGCGUGGCGUCGCGUGCGGCCUGCCUAUGGGAGUAGCCCAUAGGACGCUU AACACAGACAGAGUGUGAAGAGUCUAUUGAGCUAGUUAUAGUUCUCCGGCCCCUGAAUUGCGGCUA UUCUUAACUGCGGAGACGUGCCUCAAAGCCAGAGGGUGGCGUGUAGCUAACCGGCAACUUGGACG GAACCGACUACUUUGGGAGUCCGUGUUUCCUUUUAUUCUUAACUUGGCGUCUUUAGGUGACAAUUGA GAGAUUGUUAACAUUAGCUAUUGGAUUGGCCAUCCGGUGAGCAACAGAGCGAUUGUAUACUCUUUU GUGGUUUUGGCCAUAAAACUUUACUGUGAUUACCACACUGAAACUACUUCUGCAGUUUAUACAG CAAA	coxsackievir usA5
2	UUAAAACAGCCUGUGGGUUGCACCCACUCACAGGGGCCACUGGGCGCUAGCACACUGAUUUCGCGGA 6 AUUCCUUGUGCGCCUGUUUUUAUCCUCCUCCCCAUGCGCAACUUAAGAAGCAUUCACACUCCUAGCA 1 AAGUACAGGCGUGGCGCGCGACGCAAGUCUUAAGUACUACCAUUCUCCCGGACCGAGUAGUACAA ACUACAGGCGGUGUAGGAGGAGUUAUUGAGCUAGUUAUAGUUCUCCGGCCCCUGAAUUGCGGCUA	coxsackievir usA6

	UAAACUGUCACGCGGUUGAAGGAGAAAAUGUUCGUUACCGGCUAACUACUUCGAGAAACCUAGUAGCACCAGUAAAAUUGCAGAGCGUUUCGUCAGCGCUUCCCCCGCUAGAUACGGCUGAUGAGUCACUGCAUUCUCACGCGGCGACCGUGGCGAGUGGCGUGGCGGCGUGCCCAUGGGGUAACCCAUUGGGACGCUCUAAUUGGACAUUGGUGAAGAGUCUUAUUGAGCUAGUUAGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUUCUAAACUGCGGAGCACAUAACCCCAACACAGGGGCGGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUUGGUGUGCCGUUUUCCUUUAUUCUUAUAGUGGUGCUUAGUGAGACAAUUGAAAGAUUGUACCAUAGCUAUUGGAUUGGCCAUCCGGUGAAUAAACAGAGCCUUGAUUACCUUUUGUAGGGUUUAACCAUACUCUUCGCGUUGUAGACUCUAAAGUACAUCUAAUCUUGAACACUAGAAA	
2 6 6	UUUAAAACAGCCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGCGUAGCACUCUGGUUUUACGGUACCUUUUGUGCGCCUGUUUUUAUACCCCUCCCCAACUGUAAAUAAGAAGCAACACACGUCGAUCAUAGUAGGGCGUGACGCACCCAGUCACGUCUUGAUCAAGCACUUCUGUUUCCCCGACCGAGUAUCAUAGACUGUCACGCGGUUGAAGGAGAAAACGUUCGUUACCCGACCAACUACUUCGAAAAACCUAGUAGCAACAUGAAGGUUGCGGAGUGUUUCGUCGGCACUCCCCAGUGUAGAUACAGGUCGAUGAGUCACCCGCAUUCGCCACGGCGACCGGAGCGUGGCGUGGCGUGGCGGCGUGCCUAGGGGUAACCCAUAGGACGUCUAAUACGGACAUUGGUGCGAAGAGUCUUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACUGCGGAGCGCGCGCCCCAUUCCAGGAGGUAGCGGUGCUAACGGGCAACUCUGCAGCGGAACCGACUACUUUGGUGUGCCGUGUUUCCUUUAUCUUUAUUCUGGCGUCUUAUGGUGACAAUUGACAGAUGUUUACCAUAGCUAUUGGAUUGGCCAUCCGGUCACAAAGCAGAGCGAUUAUCUACCAAUUUAUGGAUUUAUCCACUCGACACUCUGGAAUUAACACAUUACUGCGUAUCAGUAGUUUGGACAAAGCAA	coxsackievirusA7
2 6 3	UUAAAACAGCCUGUGGGUUGCACCCACCCACAGGGCCCAUUGGGCGUCAGCACUCUGGUUACCGAGGUACCUUUUGUGCGCCUGUUUAUUCUCCCUUCCCCUGAUGCAACUUAAGAAGCUCCGAACUUAUGAUCAAUAGUAGGUGUGGACGCGACGCCAGCCACAUCUUGAUCAAGCACUUCUGUUUACCCGGACCGAGUAUCAUUAAGCUGCGCAAGCGGUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAACUACUUCGAGAAGCUUAGUACCAUGAAGACGUGGACAGAGUGUUUCGUCAGCACAAACCCCGGUGUAGAUACAGGCCGAUGAGUCACCCGCUUCCCCAUGGGCGACCAUGGGCGUGGCGUGGUUGGCGGCGUGCCAUUGGAGCAAUCCAUGGGACGCUUUAUACUGACAUUGGUGCGAAGAGCCUUAUUGAGCUAGUUGGUGAUCCUCCGGCCCCUGAAUGCGGCGUUAUCCUAAACUGCGGAGCAUGCGCCCCACAAGCCAGUGGGUGGUGUGUCGUAACGGGUAACUCUGCAGCGGAACCGACUACUUUGGUGUGCCGUGUUUCCUUUAUCUUUAUACAUUGGCGUCUUAUGGUGACAAUUGAAGAAUUGUUAACCAUAGCUAUUGGAUUGGCCAUCCGGUGUGCAUAGAGCGAUUAUUAUACCUAUUUGUUGGAUUUGUCCAUUGACAUUAAGAUUCUUAAACACUCUACAACACAUUCUUGAUUCUUGAACACGAGAAA	coxsackievirusA8
2 6 4	UUAAAACAGCCUGUGGGUUGCACCCACUCACAGGGCCACUGGGCGCUAGCACUCUGGUUACCGUGGUACCUUUUGUGCGCCUGUUUAUACCCCUUCCCCUGUUAACAUUAAGAAGUAAACGCACCUCGAUCAGUAGCAGCGCGCGCACCCAGCGUGUCUUGAUCAAGCACUUCUGUUUCCCCGACCGAGUAUCAUAGACUGUCACGCGGUUGAAGGAGAAAACGUUCGUUAUCCGGCUAACUACUUCGAGAAACCUAGUAGCAACAUGGAAGUUGCAGAGUGUUUCGUCAGCACUCCCCCCCCAGUGUAGAUACGGCUGAUGGAUACCCGCUUCCUACGCGGUGACCGUGGCGGUGGCGUGGCGGCGUGCCAUUGGGCAACCCAUUGGGAGGCUUUAUUAUGGACAUUGGUGUGAGUUAUUGAGCUAGUUAUUGAGUCCUCCGGCCCCUGAAUGCGGCUUUAUUCUAAACUGCGGAGCACGUGUCCCCAACCCAGGGGAUAGCGUGUCGUAACGGGUAACUCUGCAGCGGAACCGACUACUUUGGUGUGCCGUGUUUCCUUUAUCUUUAUUCUGGCGUCUUAUGGUGACAAUUGAGAGAUUGUUAACCAUAGCUAUUGGAUUGGCCAUCCGGUGUCUAAACAGAGCUAAUAAUUAUUGCUUUUGUUGGGUUCGUACCCCUCAAUUUUGAGGUGUUCUCACAUUAAGUACAUCUUGGUGUAAUUAUCAGAAAG	coxsackievirusA10
2 6 5	UUAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUCUGGUUACUACGGUACCUUUUGUGCGCCUGUUUAUAGCCCCUACCCCCACUCGUAACUUAAGAAGCUUCUCACACUCGUAUCAAUAGUAGGUGUGGACGCGCAGUCACACCGUGAUCAAGCACUUCUGUUUACCCGGUCUGAGUACCAAUAAGCUGCUAACCGCGGUGAAGGGGAAAACGAUCGUUAUCCGGCUAACUACUUCGAGAAACCCAGUAACACCAUGAACGUGUAGCGGUGUUUCGUCGCGCACAAACCCAGUGUAGAUACAGGUCGAUGAGUCACCUGAUUCCCCACGGGCGACCGUGGCGGUGGCGUGGUGGCGGCGUGCCAUUGGGUGACCCAUUGGGAGCGUCUAAUACUGACAUUGGUGCGAAGAGUCUUAUUGAGCUAGUUAUUGAUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACUGCGGAGCACAUACCCUUAUUAUCCAAAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUUGGUGUGCCGUGUUUCCUUUAUUCUUAUACAUUGGCGUCUUAUGGUGACAAUUGAAAAGUUGUUAACCAUAGCUAUUGGAUUGGCCAUCCGGUGACAAAUAAGAGCUAUUGUAUUAUCUUUGUUGGUUACGUACCCCUAAUUUAACAAGUGGUUUCAACUUGAAAUAACUUAACACUAAAUUGUAGAAA	coxsackievirusA12
2 6 6	UUAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUGAUUUCGCGGUAUCUUUGUGCGCCUGUUUAUUAUUCUCCCCAACUAGGUAAUUAAGAAGUUAAGCACUUGUAUUGACCAUAGAGGCGUGACGCGCCAGUCAUGUCUUGGUCAAGCACUUCUGUUUCCCCGACUGAGUAUCAUUAAGCUGCUUUGCGCGGUGAAGGAGAAAACGUUCGUUACCCGGCUAACUACUUCGAGAAACCUAGUAAGUAACCAUGAAAGUUGCGGAGUGUUUCGUCAGCACUUCUCCCCGUGUAGAUACAGGUCGAUGAGUCACUGCAAACCCACGGGUGACCGUGGCGAGUGGCGUGCGUUGGCGGCGUGCCAUUGGGUGACCCAUUGGGAGGACGCGUCUAAUACUGACAUUGGUGCGAAGAGUCUUAUUGAGCUAGUUAUUGAUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACUGCGGAGCACACGCUUGCAAACACAGUGAGUAGUUGUCCUUAUUGGUGACAAUUGAAAAGUUGUUAACCAUAGCUAUUGGAUUGGCCAUCCGGUGACAAAUAAGAGCUAUUGUAUUAUCUUUGUUGGUUACGUACCCCUAAUUUAACAAGUGGUUUCAACUUGAAAUAACUUAACACUAAAUUGUAGAAA	coxsackievirusA14

5	UUA AAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUCUGAUUUCUACGGA	coxsackievirus A16	
6	AUCCUUGUGCGCCUGUUUUUAUGUCCCUUCCCCCAAUCAGUAACUUAAGAAGCAUUGCACCUCUUUCGA		
7	CCGUUAGCAGGCGUGGCGGCACACCGCAUGUCUUGGUCACAGCACUUCUGUUUCCCCGGACCGAGUAUC		
	AUAAGACUGCUCACGCGGUUGAGGGAGAAACGUCGCUUACCCGGCUUACUUCUUGACUUCGAGCAAGCUAG		
	UAGCACCAUGAAAGUUGCAGAGUGUUUCGUCACGACUUCUCCCGUGUAGAUACAGGUCGAUGAGUCA		
	CUGCGAUCCCCACGGGCGACCGUGGCGAGUGGCGUGCGUUGGCGGCCUGGCCUGUGGGGUAAACCCACAGG		
	ACGCUUAUAUAGGACAUGGUGCAAAAGAGUCUAUUGAGCUAGUUAUGUAGUCCUGGCCGCCUGAAUGCGAAUG		
	CGGCUAAUCCUAAUCUGCGGAGCACAUACCCUCGACCCAGGGGGCAGUGUGUCGUAAACGGGCAACUCU		
	GCAGCGGAACCGACUACUUGGGUGUCCGUGUUUCUUUUUAUUCUUAUACUGGCUGCUUAUGGUGAC		
	AAUUGAAAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGUGCAACAGAGCUAUUAUUUAC		
	CUAUUUGUUGGUAUAUACCACUCACAUCCAGAAAAACCCUCGACACACUAGUAUACAUUCUUUACU		
	UGAAUUCUAGAAA		
2	UUA AAAACAGCUGUGGGUUGUACCCACCCACAGGGUCCACUGGGCGCUAGUACACUGGUUACUCGCU		enterovirus 71
6	ACCUUUGUAGCCGCUUUUUUAUACCCUCCUCCUGAUUUGCAACUUAAGAAGCAACGCAACGAGUACA		
8	UAGUAGGUGUGACAUACCAGUCGCAUCUUGAUCAAGCACUUCUGUAUACCCCGGACCGAGUAUCAUA		
	GACUGUGCACACGGUUGAAGAGAGAAACGUCGCUUACCCGGCUAACUACUUCGAGAAGCCUAGUAAC		
	GCCAUUGAAGUUGCAGAGUUGUUCGUCACGACUUCUCCCGUGUAGAUACAGGUCGAUGAGUCACCGC		
	AUUCUCCACGGGCGACCGUGGCGUGGUGGUCGUGUUGGGCGGCCUCUAUGGGGUAGAACUAGAGACG		
	UCUAUACGGACAUGGGCGUGAAGAGUCUAUUGAGCUAGUUAUGUAGUCCUCCGGCCCCUGAAUGCGGC		
	UAAUCCUAAUCUGCGGAGCAUAACCCUUAUUAUCCAAAGGGCAGUGUUGGCUAACGGGCAACUCUGCAG		
	CGGAACCGACUACUUGGGUGUGCGUGUUUUUUUAUUCUUGUAUGUUGGUGCUUAUGGUGACAAU		
	AAAGAAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCAGUGUCAACACAGAGCUAUUGUAUAUCUCU		
	UUGUUGGAUUCACACCUCUCACUCUUGAAACGUUACACACCUCUCAAUUAUAUUAUCUGCUGAACAC		
	GAA GCG		
2	UUA AAAACAGCCUGUGGGUUGUACCCACCCACAGGGGCCACUGGGCGCUAGCACACCCGUUUAUACGG	enterovirus 76	
6	UACCCUUGUGCGCCUGUUUUUAUUAACCUUCCCCUUGUAACUUAAGAAGAAACAAUAAACGAUCAU		
9	AGUAGGCUUAGCGACCAAGCUAAGUCUUGAUCAAGCACUUCUGUUAACCCCGGACUGAGUAUCAAUA		
	GCUGUACCCACGGCUGAAAGAGAAACGUUCGUUAUCCGGCUAUCUACUUCGGGAAACCCUAGUACCA		
	UCAUGAAGAUUGCAAAGCGUUUCGUCACGACAAACCCAGUGUAGAUACAGGUCGAUGAGUCACUGCA		
	UUCUCCACAGGCGACUGUGGCGAGUGGUGCGGUUGGCGGCCUCCUACCGGGGAAACCCUGAGGACGCU		
	UUAUAUUAACUAGUGGUGCGAAGAGUCUACUGAGCUAGGUAGUAGUCCUGGCCGCCUGAAUGCGGCU		
	AAUCCUAAACCAGGAGCAGGUGUCACACUCCAGUGAGUGGCCUGUCGUAACGC GCAAGUCUGUGGC		
	GGAAACCGACUACUUGGGUGUCCGUGUUCUUUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUCU		
	AAGAUGUUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGAACUCUAGAGAUUAUUAUUUACUUGUU		
	GUUGGUUUUAUCAAGUUAUAUACCAAGUUAACCAACUCCUUAUUAUGUACUACUGAAGCAAUA		
	GAAAGAGAAA		
2	UUA AAAACAGCCUGUGGGUUGUACCCACCCACAGGGGCCACUGGGCGCUAGCACACUGGUUACUCGCU		enterovirus 89
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCUUGUAACUUAAGAAGAAACAAACGAUCAU		
10	AGUAGGCUUAGCACUCCAGCUAAGUCUCGAUCAAGCACUUCUGUUAACCCCGGACCGAGUAUCAUA		
	GCUGUACCCACGGCUGAAAGAGAAACGUUCGUUAUCCGGCUAUCUACUUCGGGAAACCCAGUACCA		
	UCAUGAAGAUUGCAAAGCGUUUCGUCACGACAAACCCAGUGUAGAUACAGGUCGAUGAGUCACUGCA		
	UUCUCCACAGGCGACUGUGGCGAGUGGUGCGGUUGGCGGCCUCCUACCGGGGAAACCCUGAGGACGCU		
	UUAUAUUAUGACUUGGUGCGAAGAGUCUACUGAGCUAGGUAGUAGUCCUCCGGCCCCUGAAUGCGGCU		
	AAUCCCAACCCAGGAGCGGCGCACACUCCAGUGAGUAGCUUGUCGUAAGUCGAAAGUCUGUGGC		
	GGAAACCGACUACUUGGGUGUCCGUGUUCUUUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUCU		
	AAGAUGUUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGAAUUCUAGAGAUUAUUAUUUACUUGUU		
	GUUGGUUUUAUCAAGUUAUAUACCAAGUUAACCAACUCCUUAUUAUGUACUACUGAAGCAAUA		
	GAAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 90	
7	GUACCUUUGUGCGCCUGUUUUUAUACCCUCCUCCCAUGUAACUUAAGAAGCUAUUAUAUAAAGUUA		
1	AUAGAUGGGGUACAACACAGUACCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUACAUA		
	ACUGUUAUCCACGGUUGAAGUAGGCGAUCGCGUUAUCCGCUUUGGUACUCCGAGAAACGUAUCAU		
	CUUGGAAUCUUCGAUGCGUUGCGUUCAGCACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGG		
	ACGAUUCUACUGGCGACAGUGGUCCAGGCGUGCGUUGGCGGCCUACUUGGUGCGAAAGCUACAGGAC		
	GCUAUGUUGAACAAGGUGUGAAGAGUCUAUUGAGCUAUCAAAAGAGUCCUCCGGCCCCUGAAUGCGCG		
	CUAUUCCCAACCCAGGAGCAAGUGUCUCAAACCCAGUGAGUGGCUUGUCGUAACCGCGCAAGUCUGG		
	GCGGAACCGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	GUACCUUUGUGCGCCUGUUUUUAUACCCUCCUCCCAUGUAACUUAAGAAGCUAUUAUAUAAAGUUA		
1	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG	enterovirus 91	
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCGCUUACCCGCUUUGGUACUUCGAGAAACCUAGUAUACCUU		
	GGAAUCUUCGAGUGCGUUGCGUCACGACUCAACCCAGAGUGUAGCUUAGGUCGAUGAGUCUGGACG		
	AUCCUACUCUGGUGACGUGGUCGAGGCGUGCGUUGGCGGCCUACUUGGUGGAGAAACACAGGACGCU		
	AGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUACCAAAGAGUCCUCCGGCCCCUGAAUGCGGCU		
	AUCCUAAACCCAGGAGCAAGUGCCUUAACAAACAGUAGGUAGCUUGUGUAACCGCGCAAGUCUGGCG		
	GAAACGACUACUUGGGUGGCCGUGUUUCCUUUAUUAUUAUCAUGGCUGCUUAUGGUGACAUA		
	CUAAGAUUGUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAUAACAGAGAUACAUAUAUUAU		
	UUGUUGUUUUUUAACAUCAGACAACUACAUCUUAACUUCUUCUUAUUAUUAUUAUUAUUAUUAUUA		
	UAGAAGAGAAA		
2	UUA AAAACAGCUCUUAAGGUUGUUCUCCACCCUAGAGGCCACUGGCGGCUAGCACUCUGGUUAUCAG		enterovirus 91
7	ACCUUUGUGCGCCUGUUUUUAUUAACCCUCCUCCCAUGUAACUUAAGAAGUUAUUAACAAGUUA		
2	GAAGGGGUGUAACAACAGUGCCACCACGAACAAGCACUUCUGUUCUCCCGGUGAAGCUGUAUAGACU		
	GUCUCCACGGUUGAAAAGUGGCAUCCG		

	GAUUGUUUAUCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAACAGAGAUUCAUAUACCUUUUUGU UGGUUUUACUAAAUAUUAUUAUACCAGUUAAGACACUUCUCCUUAUCAUAACAUAUUAUAGUGAG AAGAGAAA	
2 7 3	UUUAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACGCGGCGUUAGUACUCUGGUUACACGG UACCUUUGUAUGCCUGUUUUUACUCCCUUCCCCGUAACGUUAGAAGAAUGAACAAAAGACAAUUAAG CAAAUCCACCAACCCAGGUGGGUUUUCGUCUAGCAGUUCUGUUUCCCCGACCUAGUACCAUAAAGC UGUGCCACGGCUGAAGGGGAAAACGUUCGUUAUCCGGCUACCUUACUCCGAGAAGCCUAGUACCAAC AAUGAUUGUCUCAAGUGUUAACGUCAGCACAACCCAGUGUAGUUCAGGUCGAUGAGUCACCGAAUU CCCCACGGGCGACCGUGUGCGGUGGCGUUGGCGGCCUGCCUGCGAGUUUACUCGCGAGGACGCUU GAUUCUAGACAUUGGUGCGAAGAGCCUUAUUGAGCUAGGCGAGUAGUCCUCCGGCCCCUGAAUGCGGCUA AUCCUAACCCAGGAGCCAUUGCGUCACAAACCCAGUGGCGUCAUGGUGCUAACGAGCAAUUCUGUGGCG GAACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUUAUUUACCUUGGCGUUAUGGUGACAAUCUA GAGUUGUCUACUAUAGCGAUUGGAUUGGCCAUCCGGAUUAUACUUUGUGUUAUUAUCUUCUGGUC GUAAGCAAUUUUAAUAGUUUAUAGCAUUAU	enterovirusA 92
2 7 4	UUAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCAGACUGAUUUACACGGA AUCCUUGUGCGCCUGUUUUACACCCCUUCCCCAACUCAGUAACUUAGAAGUUGUAAACACACGUG GCCAACAGCGGGCGUAGCGCACCAGCUAUGUCUUGGUCAGCAGUUCUGUUUCCCCGGACAGAGUUA CAAUAAGCUGCUUGCGCGGCGUAGGAGAGAAAUGUUCGUUAUCCGGCUACUACUUCGAGAAACCUA GUAGCAUCAAGAAUGUUGCGAGGCGUUUUCGUCAGCAGUUCUCCCGUGUAGUACAGGUCGAUGAGUC ACUGCAGUCCCCACGGGCGACCGUGGCGAGUGGCGUGGUUGGCGGCCUGCCUAGGGGUGACCCAUUG GACGCUUCAAUACGGACAUGGUGUGAAGAGUCUAUUGAGCUAGUUAUUAUCCUCCGGCCCCUGAAU UGAGCAUAUCCUAAACUGCGGAGCACAUAACCCUCAAACCCAGGGGGCGGUGUGUGUAACGGGCAACUC UGCAGCGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUAUACUUUAUACUGGCGUCUUAUGGUGA CAAUUGAAGAAUUGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAACAGAGCUAUUAUAUA UUUAUUUGUGGAUUUGUUCUUAUACAUAACAUUUUCCAGCAGUUAUACCCUUAUUUAACAAUAC AAUACAGAAA	enterovirusA 120
2 7 5	UUUAAAACAGCUCUGGGUUGUUCACCCUAGAGGCCACUGGGCGCUAGCAGCUCUGGUUAUACG GUACCUUUGUGCGCCUGUUUUUAUUAUCCCUUCCCCACGCAACUUAAGAAGUUAUUAACAAGUUCAA UAGAAGGGGGUACAAACCAGUACCCACGACAAAGCAGUUCUGUUUCCCGUGAGCUACAUAAGA CUGUUCACCAGGUUAGAAUGGCGAGAUCCGUUAUCCGCUUUGGUACUUCGAGAAACCUAGUACACC UUGGAUUCUUGCAUGCGUUGCGUCAGCAGUUAUCCAGAGUUAUAGCUUAAGGCGUAGAGUCUGGA CGAUCCUACUGGCGACAGUGGUCCAGGCGUGCGUUGGCGGCCUACCUUGGUGAAAGCCACAGGACG CUAGUUGUGAACAAGGUGUGAAGAGUCUAUUGAGCUAACAAAGAGUCCUCCGGCCCCUGAAUGCGCG UAUCCCAACACCGAGCAGCGCUCACAAACCCAGUGAGUAGCUUGGUGUAACGCGCAAGUCUGUGG CGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUAUUAUUAUAGGCGUCUUAUGGUGACAAUC UAGAUAUGUUAUUAUAGCUUUUGGAUUGGCCAUUCGGUGACUAAUAGAGACCUUGUGUAUCUAU UGUUGGUUUUACCAACUAAUUAUUAUACACUcAAAACACUCCUUCUUAUUAUACAACUUAUAGU AGAAAGAGAAA	enterovirusA 121
2 7 6	UUAAAACAGCCUGUGGGUUGUACCCACCCGAGGGCCACUGGGCGCUAGUACACUGAUUACACGGU ACCUUUGUACGCCUGUUUUUAUUAUCCCUUCCCCAUCGUAACUUAGAAGCUUAACAUAACAGAUCAAU AGCCUCUGCACCAAACCAGGUGUGGAUUAUUAAGCAUUAUCCUGUUAUCCCGGACCUAGUACCAUAG GCUGUGUCAACCGCUGAAGGGGAAAAGUUCGUUAUCCCGGCCACCUACUUCGAGAAGCCUAGUAGC ACUACUGAAGCUGUGAAGCGUUAACGUAUCAGCACAACCCAGUGUAGUUAUAGGUCGAUGAGUCACUGC AUUCCCCACAGGCGACUGUAGCAGUGGCGUGGUUGGCGGCCUGCCUUGGUCUCUGACCAAGGGACG CUUUGGAGUGGAUUAUGGUGCGAAGAGCCUUAUUGAGCUAGGUGGUAUCCUCCGGCCCCUGAAUGCGG CUUAUCCUAACCCCGGAGCUGGUGCACACAACCCAGUGUGGCGAGUCGUAUAGCAAUUUCUGG AUGGAACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUACCUUACCUUAGGCGUCUUAUGGUGACAAU UGUAUAGUUGUCAUCAUAUAGCUAUUGGAUUGGCCAUCCGGAUUAUUAUUCUUGUGUUUAUUGUGUA AUUUGAUUAUUAUAGAACCUGUAUAGA	enterovirusA 122; simianvi rus19
2 7 7	UUAAAACAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACUGGGCGCUAGUACUCUGGUUAUACGGU ACCUUUGUACGCCUGUUUUUAUUAUCCCUUCCCCAUGUAACUUAGAAGCAACAAUAAACGACCAUA GCACACUCACCAUCCAGAUAGUGUAGGUAAGCAGUUCUGUUUACCCCGGCUUAGUACCAUAGA CUGCCAUCGCGGUUGAAGGGGAAAACGUCUGUUAUCCGGCUCUACUGCGAGAAGCCUAGUAAACGC CUGUAGAAUUGCUAAGCGUUGCGUCAGCACAACCCAGUGUAGCUCAGGUCGAUGAGUCACUGCAU UCCCCACGGGCGACCGUGGCGAGUGGCGUUGGCGGCCUGCCUACAGCUCUACGUGUAGGACGCU UGUGUAUGGACAUGGUGUGAAGAGCCAUUGGCGUAGGGAGUAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAAACCCCGGAGCUGGUGCUUGUAAACCAACAGUGGCCAGUAGUAGGGCAACUCUGGGAU GGAACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUAUUAUACCUUGGCGUCUUAUGGUGACAAUCU AGUGUUGUGGCAUAUAGCGAUUGGAUUGGCCAUCCGGAUUAUUGUUUAGACACUUGUGUUUGUGAG UAGAUCUCUUCUGACAUAGGAUUAACA	enterovirusA 123; simianvi rus43
2 7 8	UUAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCAGCUCUGGUUAUACGGU ACCUUUGUGCGCCUGUUUUUAUUAUCCCUUCCCCAUGUAACUUAGAAGAAACAAAACUGGAUCAUA GCCAUACACCAAACAGGGUGGUGAGCAUAAGCAGUUCUGUUUCCCCGGAUCUAGUUAUCAAUAGA CUGUAGACACCGUUGAAGGAGAAAGCAUUCGUUAUCCGGCUCUACUUGCGAGAAGCCUAGUAAACGC CUGUAGAAUUGCUAAGCGUUGCGUCAGCACAACCCAGUGUAGCUCAGGUCGAUGAGUCACUGCAU UCCCCACGGGCGACCGUGGCGAGUGGCGUUGGCGGCCUGCCUACAGCUCUACGUGUAGGACGCU UGUGUAUGGACAUGGUGUGAAGAGCCAUUGGCGUAGGGAGUAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAAACCCCGGAGCUGGUGCUUGUAAACCAACAGUGGCCAGUAGUAGGGCAACUCUGGGAU GGAACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUAUUAUACCUUGGCGUCUUAUGGUGACAAUCU AGUGUUGUGGCAUAUAGCGAUUGGAUUGGCCAUCCGGAUUAUUGUUUAGACACUUGUGUUUGUGAG UAGAUCUCUUCUGACAUAGGAUUAACA	enterovirusA 124; simianvi rus46

	GGAUUGUUGCCAUAUAGCUAUUGGAUUGGCCAUCCGGAAUUAUAAACACACAAUACUGUUAUUUUUUC CUUUUUGGGAAUUCUCCCUUGAGUUUUACUAAUUUUUAUCUUGUUAUAGUUUUAAAGUUGAAU ACAUU	
2 7 9	UUAAAAACAGCCUGUGGGUUGUCCACUCACAGGGCCCAAGUGGCGUAGAACUCUGGUUAUCACGGUA CCUUUUGUUGCGCUAUUUUAAAAACCCCCCCCCCGUAACUUAGAAGAAUGAACAAAAACAAACUCAGUA GUAGGUAUAGCAUCCAGCAUUAUCACAGAUUAAGCAUUCUGUUUCCCCGGUGAUGGAGAUAGGCUG UACCCACGGCUGAAACUCUAUGAUCCGUUAUCCGGUAUGGUACUGCGAGAGUCUAGUACCGUCUAG AAAGUUUCUAAGCGUUAACGCUCCCGAGUAACCCCAACUGGUAGUUUUGGCGUAGAGUCUGGGCACU CCCCACAGGCGACUGUGGCCCAGGCGGCUUGGCGGCCAGCUCUCCUCAUUGAGGGUGAGACGCUU AAAUACAGACAUUGGUGAGAAGAACCUAUUGCGCUACCAUUAUAGUCCUCCGGCCCCUGAAUGCGGCUA AUCCCAACCCUGGAGCAUUAACUGCAAACAGCAGUGAGAGUGGCUAACGCGAAAGUCCGAGGCG GAACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUUAACUUUUUGAUGGCGUUUAUGGUGACAAUUU AGAGUAUUCAGAUUGUCAUCAUAAACUACUAGGUCUUGCCAACCGGUCUUAUAAUUAUUUUUGUU GGAUUUGUCCACUCACGCCAAUUUUUACAAAUUUUUGCUUAGAUUCCUACUCGCAACCUAAGAU AUCAAAG	enterovirusA 125;baboonen terovirusA13
2 8 0	UUAAAAACAGCCUGUGGGUUGAUCCACCCACAGGGCCCAUUGGCGCGUAGCACUCUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUUAUACCCCUUCCCCAACUGUAACUUAGAAGUUAACACACACCGAUCAAC AGUCACGGUGGCGACACACGACCGUUAUUGAUCAAGCACUUCUGUUAACCCGGACUGAGUAUCAUAG ACUGCUACCGCGGUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAAACCUAGUAACA CCGUGGAAGUUGCAGAGUUUUCGCUACAGCAUACCCAGUGUAGAUAGGUCGAUGAGUACCCGCA UUCGCCACGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUGCCCAUGGGGAAACCCAUGGGACGCU UAUAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUUGGUAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAACUGCGGAGCACACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGC GGAACCGACUACUUUGGGUGUCCGUGUUCAUUAUUAUCCUAUACUGGCGUUAUGGUGACAAUUG AGAGAUUGUUAACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGACCAUAGAGCUAUUAUUAUUCUUCU UGUUGGUUUUAUACCAUAGCUUAGCUUAGAAAGAGGUUAAACAUUAACAUAUUGUUAAGUUGAAUACA GCAA	coxsackievir usB3
2 8 1	UUAAAAACAGCCUGUGGGUUGUCCACCCACAGGGCCCAUUGGCGCGUAGCACUCUGGUUAUCACGGUA CCUUUUGUGCGCCUGUUUUAUACCCCUUCCCCAAUUGUAUUUAAGAAGUUUACACACCGAUCAUUA GCAAGCGUGGCGACACACGCAUGUUUUGAUCAAGCACUUCUGUUAACCCGGACUGAGUAUCAUAGA CCGCUAACCGCGGUUGAAGGAGAAACGUUCGUUAACCCGGCCAAUACUUCGAAAAACCUAGUAACAC CAUGGAAGUUGCGGAGUGUUUCGCUACAGCACUACCCAGUGUAGAUACGGUCGAUGAGUCACCGCGU UCCCCACGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUGCCUACGGGGAACCCGUAAGGACGCU UAUAUACAGACAUUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUUAUCCUCCGGCCCCUGAAUGCGGCUA AUCCUAACUGCGGAGCACAUACCCUCAAAACAGGGGGCAGUGUGUCGUAACGGGCAACUCUGCAGCG GAACCGACUACUUUGGGUGUCCGUGUUUCAUUAUUAUCCUAUACUGGCGUUAUGGUGACAAUUGA CAGGUUGUUACCAUUAUAGUUUAUUGGAUUGGCCAUCCGGUGACUAAACAGAGCAAUUAUUAUUCUUCU GUUGGGUUUAUACCAUUAAGCUUAGCUUAGAAAGAGGUUAAACACUACUUCUACUUAUAAACUAAUACA CAAA	coxsackievir usB1
2 8 2	UUAAAAACAGCCUGUGGGUUGAUCCACCCACAGGGCCCAUUGGCGCGUAGCACACUGGUUAUACGGU ACCUUUGUGCGCCUGUUUUAUACCCUUGCCCCAUGCAACAUAGAAGCUUAGUCCAAAAGGUCUAAU AGGCAGCCACAGUCACCAACUGGGUCCAGACCAAGCACUUCUGUUAACCCGGACCAAGUAUCAGUAG GUGUCUACCGCGGUGAAGGAGAAACGUUCGUUAACCCGGCCAAUACUUCGAGAAACCUAGUAGCAC CAUGAAGGUUGCGCAGUGUUUCGCUCCGACAGCCCGAGUGUAGAUACGGUCGAUGAGCCACCGCAU UCCCCACGGGUGACCGUGGCGGUGGCGUGGCGGCGCCUGCCUAGGGGUAACCCAUGGGACGCG UUCAUAUCUGACAUUGGUGAAGAGUCUAUUGAGCUAGUUUGGUAGUCCUCCGGCCCCUGAAUGCGGCG UAAUCCUAACUGCGGAGCGUGCGUUGCAACCCAGUGAGUAGCAGCGUUAUAGGGUAACUCUGCAG CGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUGUCUGGCGUUAUGGUGACAAUUG AGAGAUUGUUGCCAUAUAGCUAUUGGAUUGGCCAUCCGGUAUCUAAACAGGCGGAUCAUUUACCGUU UGUUGGUUUAUACCCUUGAAUCUUGAGGUUUAUAAUACCCUACUCUUAUAAUAAACUCAAUACA UCAAG	coxsackievir usB2
2 8 3	UUAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCAUUGGCGCGUAGCACACUGGUUAUCCGGU ACCUUUGUGCGCCUGUUUUAUACCCCCCCAGUUCGCAACUUAGAAGCAAAGAAACAAUAGGUCAA UUACUGACCGAGCAACCCAGCUGUGUUUUGGCCAAGUACUUCUGUCCCGGACUGAGUAUCAUUA AGCUGCUUGCGCGGCUAAGGAGAAACCGUUCGUUAACCCGGCCAAUACUUCGAGAGCCUAGUAAC GCCAUACCAACGUUGAGGAGUGUUUCGCUACGACUUCUCCCGUGUAGUUAAGGUCGAUGAGUCACCGC GUUCCCAACGGGUGACCCGUGGCGGUGGCGUGGCGGCGCCUGCCUUGGGGCAACCCGCGAGGCGC UCUGAUACAGACAUGGUGUGAAGAGCCUAUUGAGCUAGUUUGGUAGUCCUCCGGCCCCUGAAUGCGGCG UAAUCCUAACUGCGGAGCAGCUGUUCGCAACCCAGUGAGUAGCAGCGUUAUAGGGUAACUCUGCAG CGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUGUCUGGCGUUAUGGUGACAAUUG AGAGAUUGUUGCCAUAUAGCUAUUGGAUUGGCCAUCCGGUAUCUAAACAGGCGGAUCAUUUACCGUU UGUUGGUUUAUACCCUUGAAUCUUGAGGUUUAUAAUACCCUACUCUUAUAAUAAACUCAAUACA UCAAG	coxsackievir usB4
2 8 4	UUAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCAUUGGCGCGUAGCACACUGGUUAUCCGGU ACCUUUGUGCGCCUGUUUUAUACCCCCCCAGUUCGCAACUUAGAAGCAAAGAAACAAUAGGUCAA UUACUGACCGAGCAACCCAGCUGUGUUUUGGCCAAGUACUUCUGUCCCGGACUGAGUAUCAUUA AGCUGCUUGCGCGGCUAAGGAGAAACCGUUCGUUAACCCGGCCAAUACUUCGAGAGCCUAGUAAC GCCAUACCAACGUUGAGGAGUGUUUCGCUACGACUUCUCCCGUGUAGUUAAGGUCGAUGAGUCACCGC GUUCCCAACGGGUGACCCGUGGCGGUGGCGUGGCGGCGCCUGCCUUGGGGCAACCCGCGAGGCGC UCUGAUACAGACAUGGUGUGAAGAGCCUAUUGAGCUAGUUUGGUAGUCCUCCGGCCCCUGAAUGCGGCG UAAUCCUAACUGCGGAGCAGCUGUUCGCAACCCAGCAGGAGUGUGUUGCGUUAACGGGCAACUCUGCAG CGGAACCGAGUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUGUCUGGCGUUAUGGUGACAAUUG GAAGAUGUUAACCAUUAAGCUAUUGGAUUGGCCAUCCAGUGUCAAAUAGAGCAAUCAUUAUUCUGU UUGUUGGUUUCGUUCCUUGGACUACAGAAUUCUAAAACUCUUAUUAUUAUUGAGACUCAUAC GAUAAA	coxsackievir usB5

	UUCAAUAUGGACAUGGUGUGAAGAGUCUAUUGAGCUAGUUAAGUAGUCCUCCGGCCCCUGAAUGCGGC UAAUCCUAACUGCGGAGCAGUGCCUCCAUUCCAGGGGGUGGCGUGUCGUAACGGGCAACUCUGCAG CGGAACCGACTACUUUGGGUGUCCGUGUUCUUUAAUUUUUAUACUGGCUGCUUAUGGUGACAAUUG AAAGAUGUUGCCAUUAUAGCUAUUGGAUUGGCCAUCCGGUAUCCAGAGCAAUUGUGUACCUUUU UGUUGGAUUUGUACCACUUAACAGAAACAAGUUUUCUAUACACUGUGUUACAUAUUAUAGACUAAACACA GAAAA	
2 8 5	UUAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUCUGGUUAUUGCGGU ACCUUUUGUGCGCCUGUUUUUACCCACCCUCCCCAACGUAACUUAGAAGCCUGACAUUAACGGUCACA AGCCAGCCAGUACGCCAACUGGGUACCGACCAAGCACUUCUGUUAACCCCGGACUGAGUAUCAUAA GCUGCCAAGAGCGGUGAAGGAGAAAACGUUCGUUACCCGGCCAAUUAUUCGAGAAACCCAGUAAACA CCAUGAAAGUUGCGUAGCGUUUCAAUUCAGCACAAUCCAGUUGAGUACAGGUCGAUGAGUCACCGCA UUCCCCACAGCGACUGUGGGCGUGGCGUGGCGGCGCCUAGGGGCAAAACCAUUGGGACGC UUCAAUACUGACAUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCCUGAAUGCGGC UAAUCCUAACUGCGGAGCAGGCACUCGACAGCAGCGAGCAGCUUGUCGUAUUGGGCAACUCCGCAG CGGAACCGACTACUUUGGGUGUCCGUGUUCUUAUUCUUUAACUGGCUGCUUAUGGUGACAAU GAGAAAUGUUACCAUAUAGCUAUUGGAGUGGCCAUCCGGUGACUACAGAGCAAUAUAUUAUUGU UUGUUGGCUUUGUGCCACUUAUUGAAGGUUGUUAAGACGCUCAUUCUGAUACUACUUCUAAUAA AGCAAA	coxsackievir usB6
2 8 6	UUUAAAAACAGCCUGUGGGUUGUUCACCCACAGGGCCACCGGGCGUAGCACACUGGUUAUCACGG UACCCUUGUGCGCCUGUUUUUAACCCACCCCGAGUAAACCUUAAGAAGCAUAGCACCUCUGGUCAA UAGUAGGUGUGACACACCAGUCACAUCUGACCAAGCACUUCUGUCUCCCCGGACUGAGUAUCAUA GGCUGCUCGCGCGGUGAAGGAGAAAGCGUUCGUUACCCGGCCAGCUACUUCGAGAAAGCCUAGUAAC ACCAUGAAGGUUGCAGAGUGUUCGUCUAGCACUUCUCCUGUGUAGUACAGGCCGAUGAGUACCCGC GUUCCUACAGGGCGACCGUGGGCGUGCGUGCGUGGUGGCCUGCCUAGGGGCAACCCUAGGAGCGC UCUAAUGCUGACAUGGUGCGAAGAGUCUAUUGAGCUAGCUGGUAGUCCUCCGGCCCCUGAAUGCGGC UAAUCCCAACUGCGGAGCAGCACCCUCAAACAGGGGGCAGCGUGUCGUAACGGGCAACUCUGCAG CGGAACCGACTACUUUGGGUGUCCGUGUUCUUUUUAUUCUUAUUGGCUGCUUAUGGUGACAAU GAGAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGAGCAACAGAGCUAUUGUGUAUCUGU UUAACUGGUUUAACCCUUAUUAACAAAGAGUCAAAACCCUACUUAUUGUUAUUGCAUUAAC AACA	coxsackievir usA9
2 8 7	UUAAAAACAGCCUGUGGGUUGUUCACCCACAGGGCCACUGGGCGUAGCACACUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUUUAUACUCCCCCCCCUAGGAACUUUAAGAAGCAUAGCACCUCUGGUCAA UAGUGGGUUAUGGCACACCAGUCUAUUCUUAUUAAGCACUUCUGUUCUCCCCGGACUUAUAGUACCAUA GACUGCUCAAGCGGUGAAGGGGAAACGUUCGUUACCCGGCCAAUACUUCGAGAAACCUAGUAGC ACCAUGAAGUUGCGGAGUGUUCGUCUAGCACUUCUCCCGUGUAGUACAGCGUGAUGAGUACCCGU AUUCCCCACGGCGACCGUGACGUGGCGUGCGUUGGGCGCCUGCCUAGGGGUAACCCUAGGGACGC UCUAAAAACAGACACGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGC CAUCCUAACUGCGGAGCACAUAUCCCAUCCAGGGAGCAGUGUGUCUUAUUGGUUAACUCUGCAG CGGAACCGACTACUUAUGGGUGUCCGUGUUCUUUUUAUUCUACAUAGCUGCUUAUGGUGACAAU UAAGAGUUGUUAACCAUAUAGCUAUUGGAUUGGCUAUCGGUGAGCAUAGAGCUAUAUUGUUAUUAU UUGUUGGAUUGUACCAUCUACUUAUUCUGUUAUGAGAACUACAACUACAUCUUAUCUGCUAAACAC AUCAA	echovirus1
2 8 8	UUAAAAACAGCCUGUGGGUUGAUCACCCACAGGGCCAUUGGGCGCUAGCACUCUGGUUAUACGGU ACCCUUGUGCGCCUGUUUUACACGCCCCUCCCCAAGCAACUUAGAAGUCUGUCACAAUUGGUCAAUA GACAGCUCAACACGCGCAGUUGAGUCUAGACCAAGCACUUCUGUCGCCCCGGACCGAGUAUCAUAGA CUGCUAGCGCGGUGAAGGAGAAAACGUUCGUUACCCGGCCAAUACUUCGGGAAACCUAGUACCAU CAUGAAAGUUGCGCAGUGUUCGUCUCCGACAAACCCAGUGUAGUACAGGUCGAGUAGGCCACCGCAU UCCCCACAGGGCAGUGUGCGGUGGCGUGCGUUGGGCGCCUGCCUAGGUACAAACCAUUGGACGCUU CAAUACUGACAUGGUGCGAAGAGUCUACUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUA AUCCUAACUGCGGAGCAGUACCCACGAGCCAGUGGGCAGCCUGUCGUAUUGGGCAACUCUGCAGCG GAACCGACUACUUAUGGGUGUCCGUGUUCUUAUUAUACUUGGCUGCUUAUGGUGACAAUAC GGAAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCAGUGACUAAACAGAGCAUUAUAUUCUUAU GUGUUGGUUUGUACCAUCUACCGUCACGAGCUUAACACGCUAAACUACAUAUUGUACAUAUAGACG CAAG	echovirus2
2 8 9	UUAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCCAGCACUCUGGUUAUACGGU ACCCUUGUGCGCCUGUUUUUAUACUCCCCUCCCCAAGGUAACGUUAAGAAGCAUAGCACCUCUCCGAUCAA UAGAGGUGGCGCGCGCCAGUCGUAUCUGAUAAGCACUUCUGUUCUCCCCGGACCGAGUAUCAUAGA GACUGCUAGCGCGGUGAAGGAGAAAACGUUCGUUACCCGGCCAAUACUUCGAGAAAGCCAGUAGC GCCAUGAAAGUUGCAGAGUGUUCGUCUAGCACUUCUCCCGUGUAGUACAGGCCAGUAGUACCCGC AAUCCCCACGGCGACCGUGGGCGUGGCGUGCGUUGGGCGCCUGCCUAGGUACAAACCAUUGGACGCUU CAAUACUGACAUGGUGCGAAGAGUCUACUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUA AUCCUAACUGCGGAGCAGUACCCACGAGCCAGUGGGCAGCCUGUCGUAUUGGGCAACUCUGCAGCG GAACCGACUACUUAUGGGUGUCCGUGUUCUUAUUAUACUUGGCUGCUUAUGGUGACAAUAC GGAAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCAGUGACUAAACAGAGCAUUAUAUUCUUAU GUGUUGGUUUGUACCAUCUACCGUCACGAGCUUAACACGCUAAACUACAUAUUGUACAUAUAGACG CAAG	echovirus3
2 9 8	UUAAAAACAGCCUGUGGGUUGCACCACCCACAGGGCCAUUGGGCGCUAGCACCCUGGUUAUACGGU ACCUUUGUGCGCCUGUUUUUAUACUCCCCUCCCCAAGGUAACGUUAAGAAGCAUUAACACUAGUAAU AGGAGGCGCGGCACGCCAGCCAGUACCAAGCAUUCUGUUCUCCCCGGACCGAGUAUCAUAG GAGGAGGCGCGGCACGCCAGCCAGUACCAAGCAUUCUGUUCUCCCCGGACCGAGUAUCAUAG	echovirus4

	ACUGCUUGCGCGGUUGAAGGAGAAAAACGUCGUUACCCGCCAACUACUUCGAGAAACCUAGUACCA CCAUAGAAAGUUGCGGAGUGUUUCGCGUCAGCACUACCCCGAGUCUAGAUACAGGCCGAUGAGUCACCCGG UUCCCCACGGGUGACCGUGGCGGGUGGCGUCGCGUGGCGGCCUGCCUUAUGGGGCAACCCUAGGACGCU CUAAUACAGACAGUGGUGGAGAGUCUUAUUGAGCUAGUUGGUGAUCCCGGCCCGUUAUUGCGCGCU AAUCCUAAACUUGCGGAGCACACGCGUCACAAAGCCAGUGAGUGGUGUGUCGUAAUUGGGUAACUCGCGAGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUAUACUUAUUGGCGUCUUAUGGUGACAAUUA AGAAAUUGUUACCAUAGCUUAUUGGAUUGGCCAUCGCGGAGCUAGAGACUUAUUAUACUCUGUU UGUUGGCUUUGUACCACUAAACUUAUAAAGUCCUUAAGAACUCUUGAUUUUAUACUAAUUAUUGAAUAG GCAA	
2 9 1	UAAAAACAGCCUGUGGGUUGCACCACCCACAGGGCCACUGGGCGCUGUACACUGGUUAUUAAGGU ACCCUUGUACGCCUGUUUUUAUACCCUCCCCCGAUUUUUAACUUAAGAUUAAAGCAACCCGAUCAAA UAGUAGGUGUGGCAAAACAGCCACAUUCUGAUCAAGCACUUCUGUUUCCCCGGACCGAGUAUCAUA GACUGUGACGCGGUUGAAGGAGAAACGUCGUUACCCGGCCCAACUACUUCGGGAAACCUAGUACC ACCAUGAAGUUGCGAGUGUUUCGUCUCCGCACAAACCCAGUAGCAUAGGUCGAUGAGCCACCGC AUUCCCCACAGGCGACUGUGGCGGUGGCGUGGUGGCGGCGCCUGCCAUUGGGGUAAACCAUGGGACGC UUAUUAUCUGACUAGUGGUGCGAAGAGUCCACUGAGCUAGUUGGUUAUUCUCCCGGCCCUUGAAUGCGCG AAUCUUAACUGCGGAGCAGGCACCCACAGGCCAGUGGGCAGCCUUGUGUAAUUGGUUAACUCUGCAG CGGAACCGACUACUUGGGUGUCCGUGUUUCCAUUUUUUUUUGGUGUACUUAUGGUGACAUUA CGGAUUGUUUACCAUUAAGCUUAUUGGAUUGGCCAUCCAGUGACUAAACAGAGCAGUAAUUAUCCU UGUUGGUUUUGUACCACUACCCGUCACGAGCCUCAGAACACUCGUUAUUCUUAACCCUUAACAAG AUAAA	echovirus5
2 9 2	UAAAAACAGCCUGUGGGUUGCCCCCACCAGGGCCCAUUGGGCGCUGACACACUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUUUAUACCCUCCCCCAAGUGCAACGUAAGACAGACACAAAGGUCUAA AGAUAGCCACGACGCUAAUUGGUGUUUUGACCAAGCACUUCUGUCCCGGACCGAGUAUCAUAA GCUGCUACGCGGCGUGAAGGAGAAAGUGUUCGUUACCCGGCUAGUUAUUCGAGAAACCUAGUACCA CCAUAGAGGUUGCGCAGCGUUUCGCUCCGCACAAACCCAGUGUAGAUACAGGUCGAUGAGUCACCCGG UUCCCCACGCGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCCAUUGGGGCAACCCUAGGGACGCU UCAUACUGUACUAGGUGCGAAGAGUCUUAUUGAGCUAAUAGUAGUCCUGGCCCCUGAAUGCGCGAU AAUCUUAACUGCGGAGCAGGUGUCACAAUCCAGUGGGUGGCGUGCGUAAACGGGCAACUCUGCAGC GGAAACCGACUACUUGGGUGUCCGUGUUUCCUUAUUCUUAUUGGUGUCUUAUGGUGACAAUUG ACGAUUGUUGCCAUUAAGCUUAUUGGAUUGGCCAUCCGGUGAGCAGCAUUGAUUUUACCAUUG UGUUGGAUUUGUACCACUUGAUACAUAGUCUUUAGAACACUUGUCUUUAUUCUAAACUUAACAAG GCAAG	echovirus6
2 3	UAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCAUUGGGCGCUGACACCCUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUUUAUACCCUCCCCCAUUGUAACUUAAGAAAGAACACACACCGAUCAAC AGCAAGCGUGGCACACCAGCCAUUGUUUGGUCAAGCACUUCUGUUAACCCGGACUGAGUAUCAUAG ACUGUCUACGCGGUUGAAGGAGAAACGUGCCGUUAUCGCGCCAGCUACUUCGAGAAACCUAGUAACA CCAUUGGAUUGCGGAGUGUUUCGUCACGACUACCCAGUGUAGAUACAGGUCGAUGAGUCACCCGU UUCCCCACGCGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUUAUGGGGGAACCCAUAGGACGCU CUAAUACAGACAUUGGUGCGAAGAGUCUUAUUGAGCUAGCUGGUUAUUCUCCCGGCCCCUGAAUGCGCGU AAUCCUAAACUGUGGAGCAUAGCCCCUAAUCCAAGGGGUGAGUGUGCGUAAUAGGACAAUCCCGCAGC GGAAACCGACUACUUGGGUGUCCGUGUUUCCUUAUUCUUAUUGGUGUACUUGGUGACAAUUG AGAGAUGUUAACCAUUAAGCUUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUUAUUGUGUAUCUCUU UGUUGGAUUUGUACCACUUAUUUGAAAGAAAUACAGGACACUACGCUACAUUUUACUUAUUGAACACC GCAA	echovirus7
2 9 4	UAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCAUGGGCGCCAGCACUCUGGUUAUCGCGGU ACCUUUGUGCGCCUGUUUUUGUUUCCUUAUACCCCGAAUGGAACUUAAGAGUAAUUGCACACUGCAUCAA UAGCAGGCGUGGCACACGACCGAUGAUCAAGCAGCACUACUUCUCCCGGACCGAGUAUCAUAGA GUCUGUUGCGCGGUUGAAGGAGAAACGUGUCGUUAUCCGGCUAAUACUUCGAGAAACCCAGUAAAC CAUGGAAGUUGCGGAGUGUUUCGUCACGACUACCCAGUGUAGAUACAGGUCGAUGAGUCACCCGU UUCCCCACGCGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUUAUGGGGGAACCCAUAGGACGCU CUAAUACAGACAUUGGUGCGAAGAGUCUUAUUGAGCUAGCUGGUUAUUCUCCCGGCCCCUGAAUGCGCGU AAUCCUAAACUGUGGAGCAUAGCCCCUAAUCCAAGGGGUGAGUGUGCGUAAACGGGCAACUCCGACGCGGA ACCGACUACUUGGGUGUCCGUGUUUCCUUAUUCUUAUUGGUGUACUUGGUGACAAUUG AGAGAUGUUAACCAUUAAGCUUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUUAUUGUGUAUCUCUU UGUUGGAUUUGUACCACUUAUUUGAAAGAAAUACAGGACACUACGCUACAUUUUACUUAUUGAACACC GCAA	echovirus9
2 9 5	UAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCAUGGGCGCUGACACACUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUUUAUACCCUCCCGCAACCGCAAAUUAAGAGCAAGACUAAACCGAUUC GAUAGCGGAUGCGCAUGCCAGCCGAUUAUUGAUCAAGUACUUCUGUUUCCCCGGACCGAGUAUCAU AGACUGUCUACGCGGUUGAAGGAGAAACGUCGUUAUCCGGCUAAUACUUCGAGAAACCCAGUAA CAUCAUGAAUUGUUGCAGGCGGUUUUCAUACGACGACGACCGUGUGUAAACGAGCUUAGUGUACCG CAUCCCCACGGGUGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUUAUGGGGUGACCCAUAGGACG CUCUAAUACGGAUUGGUGCGAAGAGUCUUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGG UUAUACCUAACUGCGGACCAUACUACCCUAAUCCAAGGGGUGAGUGUGCGUAAACGGGCAACUCCGACGCGGA ACCGACUACUUGGGUGUCCGUGUUUCCUUAUUCUUAUUGGUGUACUUGGUGACAAUUG GAUUGUUGCCAUUAAGCUUAUUGGAUUGGCCAUCCGGUGUCUAAUAGAGCGUAUUAACUACAGUUGU UGGAUUUGUACCACUACUUAUUCAACACUCAAGACACUACUACUAGCAGCUACACACAGUA AA	echovirus11

2	UUAAAAAGCCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUUAUCCCGGU	echovirus12
9	ACCUUUGUGCGCCUGUUUUUAUUAUACCCUCCCCUCAGUAACCUAGAAGUUAUACACAAUAGUCAAUA	
6	GUUAGCUCACAAACACAGUUGAGCCUAGAUCAAGCACUUCUGUUAACCCCGGGCUGAGUAUCAAUAAG	
	CUGUUGACACGGCGUGAAGGAGAAACGCCGUUACCCGACCAGCUACUUCGGAGAACCUAGUAUCAC	
	CAUAGAGGUUUGCGUAGCGUUCGCUCCGCAACCCACAGUGUAGAUACGGUCGAUGAGUACCCGCGU	
	UCCCCACAGGCGACUGUGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGGGUUACCCAUGGGACGCUU	
	CAAUACUGACAUUGGUGUAGAGUUGACUGAGCUAGCUGGUAGUCCUCCGGCCCCUGAAUUGCGGCUA	
	AUCCUAACUGUGGAGCAAGUGCCCACACCCAGUGGGUGGCUUGUCGUAUUGGGCAACUCUGCAGCG	
	GAACCGACUACUUGGGUGACCGUGUUUCUUAUUCUUAUUAUUGGCGUCUUAUGGUGACAAUCUC	
	AGAGUUGUUACCAUAUAGCUAUUGGUUUGGCCAACAGUGACUACAGAGCAAUUAUUAUUCUUUC	
	AUUGGAUUUAUACCUAUAAACAACACAGUUAUACCAUUGUGUUAUUAUUGGCUUGAACUCGA	
	AAAG	
2	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCCAGCACACUGGUUAUCACGGU	echovirus13
9	ACCCUUGUGCGCCUGUUUUUAUUAUACCCUCCCCGAGUAACUUAAGAAGCUCUAAAUAAAGCGGUCAGC	
7	AGCCAGCCACACACACAGCUGGGUUUUGACCAAGUACUUCUGUUUACCCCGGACCGAGUAUCAAUAG	
	GCUGCACACGGCGUGAAGGAGAAAACGUUCGUUACCCGGCCAUUACUUCGAGAAACCUAGUACCA	
	CUAUGAAAGUUGCGAAGCGUUUCGCUCCACACAACCCACAGUGUAGAUACGGCUGAUGAGUACCCGCA	
	UUCCTCACAGGUGACUGUGGCGGUGGCGUGGUGGCGGCCUGCCCAUGGGGCAACCAUGGGACGCU	
	UCAAUACCGACAUUGGUGCGAAGAGUCUAUUGAGCUAGAUAGGUAGUCCUCCGGCCCCUGAAUUGCGGCU	
	AAUCCUAACUGCGGAGCAGGUGCACAAACACAGUGAGUAGCUGUGCUAAUUGGGCAACUCUGCAGC	
	GGAACCGACUACUUGGGUGUCCGUGUUUCUUUAUUCUUAUUAUUGGCGUCUUAUGGUGACAAUUG	
	AAAGAUUGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAGCAGACUGAUUAUACCUAUU	
	UGUGGGCUUUGUACCAUUAUUGUACAGCUACACCAUUGCUAUUAUUGUCCAACUGAACGCA	
	AGAAG	
2	UUAAAAAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACAGGGCGCAAGCACACUGGUUAUCACGGU	echovirus14
9	ACCUUUGUGCGCCUGUUUUUAUUAUACCCUCCCCCAAAGUAACGUAGAAGCCUGAACAAAGUGGUCACUA	
8	GUUAGCUCGGAUAGCCAACCGAGGUUAGACCAAGCACUUCUGUCUCCCCGGACUGAGUAUCAAUAGG	
	CUGCUAACCGCGGUGAAGGAGAAAACGUUCGUUACCCGACCAGCUACUUCGAGAAACCUAGUACAC	
	CAUGGAUGUUGCGCGCGCGUUCGUUACACAACCCACAGUGUAGAUACGGUCGAUGAGUACCCGCU	
	UCCCCACGGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCCAUGGGGCAACCAUGGGACGCUU	
	CAAUACUGACAUUGGUGCGAAGAGUCUAUUGAGCUAGCUGGUAGUCCUCCGGCCCCUGAAUUGCGGCUA	
	AUCCUAACUGUGGAGCAGGCGCUCGCAACCCAGCGAGUAACCUUGCUUACGGGCAACUCCGACGCG	
	GAACCGACUACUUGGGUGUCCGUGUUUCUUUAUUCUACCUUGGCGUCUUAUGGUGACAAUUGA	
	GAGAUUGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUAAGCAACAGAGCAGUAUUUAUUAUUGUUC	
	GUUGGUUUAUACACGUGACUUAUAAAGUCUUAACAACUAAACCUCAUUAUUAUUAUUAUAGUG	
	CAAA	
2	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUCUGGUUAUACGGU	echovirus15
9	ACCCUUGUGCGCCUGUUUUUAUUAUACCCGCCCCAAGCAAACCGUAGAUGUAACCGCACUUAUGAUCAA	
9	UAGCAGGCGGGCACUCCAGCCACGUUAUGAUCAAGCACUUCUGUCUCCCCGGACCGAGUAUCAAUA	
	GACUGUCACCGCGGUCGAAAGGAGAAAACGUUCGUUACCCGACCAGCUACUUCGAGAAACCUAGUAAC	
	UCCAUGGAGGUUGCAGAGUGUUUCGUUCAGCACUUCUCCCCGUGUAGAUACGGCUGAUGAGUACCCGC	
	GUUCCUACCGGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUUAUGGGUAUACCCAUAGGACGC	
	UCUAAUACUGACAUUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUUGCGG	
	UAAUCCUAACUGCGGAGCACAUACCCCAACCCAGGGGGCAGUGUGUCGUAACGGGUAACUCUGCAG	
	CGGAACCGACUACUUGGGUGUCCGUGUUUCUUUAUUCUUAUUAUUGGCGUCUUAUGGUGACAAUUG	
	GAGAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCAGUGACUAAACAGAGCUAUUAUUAUUAUUAU	
	UUGUUGGCUUCGUACUACUUGGUUUUAAAGAGGUUAGUACUUUAUUAUUGCAUUAUUAUUAUUAUUAU	
	GAGAAA	
3	UUAAAAAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACUGGGCGCAAGCACACUGGUUAUACGGU	echovirus16
0	ACCUUUGUGCGCCUGUUUUUAUUAUACCCUCCCCUCCCAUUAUAACUUAAGAAGCACAUACAAAACGACCAU	
0	AGGUGGUCCAGCACUCCAGCUGGGCUCUGGUCAAGUACUUCUGUUAACCCGGACUAAAGUAUCAAUA	
	GCGGCUAACCGCGCUGAAGGAGAAAGCGUCCGUUACCCGGCCAAUACUUCGGGAAUUAUUAUUAACA	
	CCAUGAAGGUUGCGCAGUGUUUCGCUCCGCAACAACCCAGUGUAGAUACGGUCGAUGAGUACCCGCA	
	UUCCTCACAGGCGACUGUGGCGGUGGCGUGGUGGCGGCCUGCCCAUGGGGCAACCAUGGGACGCU	
	UCAAUACUGACAUUGGUGCGAAGAAUUAUACUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUUGCGGCU	
	AAUCCUAACUGCGGAGCACAUACCCUUAACCAAGGGGCGAGUGUGUCGUAACGGGCAACUCUGCAG	
	GGAAACCGACUACUUGGGUGUCCGUGUUUCUUUAUUCUUAUUAUACUGGCGUCUUAUGGUGACAAUUG	
	AGAAUUGUUAACCAUAUAGCUUUUGGAUUGGCCAUCCAGUGACUAAACAGAGCUAUUAUUAUUAUUAU	
	UGUUGGCUUUGUACCAUACCUUAGAGUUUAAGGACUCUCAAUUGGUUCUGCUAUCGAAUACA	
	ACAAA	
3	UUAAAAAGCCUGUGGGUUGUUCACCCACCCACAGGGCCAUUGGGCGCUAGCACUCUGGUUAUUGCGGU	echovirus17
0	ACCUUUGUGCGCCUGUUUUUAUUAUACCCUCCCCUGACCGUAACUUAAGAAGUAACACACACCGAUCAA	
1	CAGUAAGCAUGGCACACACAGCCAUUUUGAUCAAGCACUUCUGUUAUCCCGGACUGAGUAUCAAUA	
	GACUGUCACCGCGGUUGAAGGAGAAAACGUUCGUUACCCGGCCAAUACUUCGAAUUAUUAUUAUUA	
	ACCAUGGAAGUUGCAGAGUUGUUCGCUACGACUACCCAGUGUAGAUACGGUCGAUGAGUACCCGC	
	ATUCCCCACGGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUUAUGGGGAAACCCAUGGAGCGC	
	UCUAAUACAGACAUUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUUGCGG	
	UAAUCCUAACUGCGGAGCACACACCCUAAAGCCAGGGGGCAGUGUGUCGUAACGGGCAACUCUGCAG	
	CGGAACCGACUACUUGGGUGUCCGUGUUUCCCCUACUCUUGUAUUGGCGUCUUAUGGUGACAAUUA	

190

	UAAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUAUAGUAGUCCUCCGGCCCCUGAAUUGCGGAUA AUCCUAAACUGUGGAGCAGAUACCCACGAACACAGUGGGCAGUCUGUCGUAACGGGCAACUCCGCAGCG GAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCCAAUCUGGUGUCUUAUGGUGACAAUUGA GAGAUUGUUGCCAUAUAGCUAUUGGAUUGGCCAUCCGGUGAAUAAUAGAGCGAUAAUUAUUAUUGUUU GUUGGAUUCGUGCCACUUAAGUCUGAAAGUUUUGAGAACACUACUACGUAUUUAUUGCUGAAUAGUG CAAG	
3 0 8	UUAAAAAGCCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGCGCUAGCACACCGGUUAUCACGGU ACCCUUGUGCGCCUGUUUAUACCCACCCCGUUGUAACUUAAGAAGCUAAUACAAUUGGUGCUAA GCCAGCCCGAUACGCCAAUUGGGUUUUGACCAAGCACUUCUGUUAACCCCGGACUGAGUACCAUAGG CUGCCAAAGCGGCUGAAGGGGAAACGUCGUUACCCACCAACUACUUCGAGAAACCUAGUACCCAC CAUGAAGAUUGCGCAGCGUUUCGUCUAGCACAACCCAGUGUAGAUACAGGUCGAUGAGUCACCGCAU UCCCCACAGGGCAGUCUGGCGGUGGCGUUGGCGGCCUGCCAUUGGGGCUACCCAUGGGACACUU CAAUUAUCGACAUUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUUGCGGAUA AUCCUAAACUGCGGAGCAGGACCCACAAGCCAGUGGGCAGCCUUGUCGUAACGGGCAACUCUGCAGCG GAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUAACUAGGUGCUUAUGGUGACAAUUGA AGAAUUGUUACCAUAUAGCUAUUGGAUUGGCCAUCCAGCGACUAAACAGAGCUAAUUAUCUUCUGUUU GUGGGUACGUACCACUCAAUUCACAGUAGUAAGGACUCUCAAUUCAUUUUGUGCUUGAAUAGG CGAA	echovirus26
3 0 9	UUAAAAAGCCUGUGGGUUGUACCCACCUACAGGGCCACUGGGCGCUAGCACACUGGUUAUCCCGU ACCUUUGUGCGCCUGUUUAUUAUCCCCACCCACGUGUAACUUAAGAAGCAUUGCAAUUUGCGUCA UGAGUAGACGCGGCAACCCAGUCGUGUUUCGACCAAGCACUUCUGUAACCCCGGACCGAGUAUCAAU AAGCUGUCUGCGCGCUGAAGGAGAAACCGUUCGUUACCCGGCUAACUACUUCGAGAAAGCUACAA CACCAUGAACGUUGCAGAGUGUUUACUCUAGCAGCUUCCCCCGUGUAGAUACAGGUCGAUGAGUACCCG CGAUCCCCACGGGCAGCCUGGCGGUGGCGUGCGUUGGCGGCCUGCCUUGUGGGCAACCCACAGGACG CUCUAAUACAGACAUGGUGCGAAGAGCCUAUUGAGCUAGUUAUAGUAGUCCUCCGGCCCCUGAAUUGCGG CUAAUCCUAAUCUGCGGAGCACAUUGCCACAACACAGUGGGUGGUGUGUCGUAACGGGCAACUCUGCA GCGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUACUACUGGUGCUUAUGGUGACAAU UGAAAGAUUGUUGCCAUUAGCUAUUGGAUUGGCCAUCCGGUGUCAACAGAGCAUUAUUUACCUA UUCGUUGGUUUUGCACCCUUGAAUUUAAGGUUCUUAACACUUAUUCUUAUUAUAAACUCAAUU CAGCAA	echovirus27
3 0 1	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUCUGGUUAUCACGGU ACCUUUGUGCGCCUGUUUAUACUUCUCCCCAACUGCAACUUAAGAAGUAAACAAACCGAUCAAC AGUCAGCGUGGCACACCCAGCCACGUUUUGAUCAAACACUUCUGUUAACCCGGACUGAGUAUCAAUAG ACUGUCUACGCGGUUGAAGGAGAAAACGUUCGUUAUCCGGCCAACUACUUCGAGAAACCUAGUAACG CCAUGGAAGUUGGAGUGUUUCGUCUAGCAGCUUACCCAGUGUAGAUACAGGUUGAGUAGUACCCGCA UUCCCACAGGGUGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCAUUGGGGAAACCCAUUGGACGCU CUUAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUUGCGGCU AAUCCCAACUGCGGAGCAUACUCUCAAGCCAGAGGGUAGUGUGUCGUAUUGGGCAACUCUGCAGC GGAACCGACUACUUUGGGUGUCCGUGUUUUAUUUAUUCUUAUACUGGUGCUUAUUGGUGACAAUUGA AGAGAUGUUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGCAUACAGAGCUAAUUAUUAUUCUUU UGUUGGGUUUAUACCAUAGCUUAGCUUAGAAAGGUUAAACUCUACAUUAUUAUUAUACUGAACCC GCAA	echovirus29
3 1 1	UUAAAAAGCCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUUAUCGUGGUA CCUUUGUGCGCCUGUUUAUACACCCUCCCCAUUUGUAACUUAAGAAGUAAAGCAUAUUGAUCAAUA GCAAGCAUGGCACACCCAGCCGUGUUCAGAUCAAGCACUUCUGUAACCCCGGACUGAGUAUCAAUAGG CUGCCAAAGCGGCUGAAGGAGAAAACGUUCGUUAACCCGGCCAACUACUUCGAGAAGCCAGUAUCAC CAUGAAAGUUGCAGAGCGUUCGUCUAGCAGCUUCCCCCGUGUAGAUUAGGUGCAGUAGUACCCGCAU CCCCACAGGCGACUGUGCGGUGGCGUGCGUUGGCGGCCUACCUACGGGGCAACCCGUAAGGACGCVU UAAUACAGACAUGGUGAGAAGAGCCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUUGCGGAUA AUCCUAAUCUGCGGAGCAUGUGCUGGCAUCCAGCGAGAGGCAUGUCGUAACGGGCAACUCUGCAGCG GAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUGUUGGUGCUUAUGGUGACAAUUGA AAGAUUGUUGCCAUAUAGCUAUUGGGUGCGGCCACUGGUUAUCUACAGAGCAUUAUUUAUUCUUGUU GUUGGUUUUGUACCCUUAUUAUCUAGGUUUCUGAAACCCUUAACUACACAUUGAAACUCAAUUAU UGAA	echovirus31
3 1 2	UGUGGGUUGUCCCCACCCACAGGGCCACUGGGCGCUAGUACACUGGUACCAAGGUACCUUUGUACG CCUGUUUAUACCCUCCCCGAUCCGAAACUUAAGAAGUAAAGCAACCCGAUCAAUAGUAGGUGCG ACACUCCAGUGUAUCUCGAUCAAGCACUUCUGUUUCCCGGACUGAGUAUCAAUAGACUGCCACG CGGUUGAAGGAGAAAACGUCCGUAUCCGGCUAACUACUUCGAGAAACCCAGUAGCACCAGGAAGU UGCAGAGUGUUUCGUCUAGCAGCUUACCCCGUGUAGAUACAGGUGUAGUAGUACCCGCAUUCCCACGG GUGACCGUGGGCGGUGGCGUGGUGGCGGCCUGCCUUGGGGCAACCCAUAGGACGCUCAAUACGGA CAUGGUGUGAAGAGUCUAUUGAGCUAGUUAUUAUCCUCCGGCCCCUGAAUUGCGGCUAAUCCUAAU GCGGAGCACAUACUCCUACUCCAAGGAGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACU ACUUUGGGUGUCCGUGUUUCCUUAUUAACAUUACUAGGUGCUUAUGGUGACAAUUGAAAGAUUGUU ACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGUCAACAGAGCAUUAUGUUAUCCUUCGUGGGUUU AUCCCAUUGACUGUUAAAGAUUAACAACUCUCUACUUGAUAAUACAACUAAACACUGUAAA	echovirus32
3 1 3	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACUCUGGUUAUCACGGU ACCCUUGUGCGCCUGUUUAUAGUCCUCCUACUGUAACUUAAGAAGUAAAGCAACCCGAUCAAUAG AGUCAGCGUGGCACACCCAGCCAGUUUUGAUCAAGCACUUCUGUUAACCCCGGACCGAGUAUCAAACAG ACUGUCUACGCGGUUGAAGGAGAAAGUUCGUUAUCCGGCCAACUACUUCGAGAAACCCAGUAGCACCAGGAAGU ACUGCGUACGCGGUUGAAGGAGAAAGUUCGUUAUCCGGCCAACUACUUCGAGAAACCCAGUAGCACCAGGAAGU UGCAGAGUGUUUCGUCUAGCAGCUUACCCCGUGUAGAUACAGGUGUAGUAGUACCCGCAUUCCCACGG GUGACCGUGGGCGGUGGCGUGGUGGCGGCCUGCCUUGGGGCAACCCAUAGGACGCUCAAUACGGA CAUGGUGUGAAGAGUCUAUUGAGCUAGUUAUUAUCCUCCGGCCCCUGAAUUGCGGCUAAUCCUAAU GCGGAGCACAUACUCCUACUCCAAGGAGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACU ACUUUGGGUGUCCGUGUUUCCUUAUUAACAUUACUAGGUGCUUAUGGUGACAAUUGAAAGAUUGUU ACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGUCAACAGAGCAUUAUGUUAUCCUUCGUGGGUUU AUCCCAUUGACUGUUAAAGAUUAACAACUCUCUACUUGAUAAUACAACUAAACACUGUAAA	echovirus33

	CCAUGGAAGUUGCAGAGUGUUUCCGUCACGACUACCCAGUGUAGAUCAGGUCGAUGAGUCACCGCA UCCCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGCGGCGCUGCCUUAUGGGGGAACCCAUAGGACGCU CUAAUACAGACAUGGUGCGAGAGUCCAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAAUCGCGGAGCAGACACCUUUAAGCCAGGGCAGUGUGUCUAGCCGGCAACUCGCGACG GGAACCGACUACUUGGGGUGCCGUGUUUUAUUUUUAUUAUACUGGCGUCUUAUGGUGACAAUUG AGAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCAGUGACUAGCAGAGCUAUUAUAUACCUCU UGUUGGGUUUAUACCACCUAAUUAUGAAAGAAGUUAACCAUAGAAUUCUAUAUUAUUAUUGAAUACA UUUAA	
3 1 4	UUAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCCACUGGGCGCUGAGCACACUGGUUAUCACGGU 1 ACCUUUGUGGCGCGUUUUUAUACCCCCCACCACCUUUAACUAGAAGCAAAGCAACCCGAUCA 4 AUAGCAGUAGUAGCAAGCCAGUGGCAUCUUGAUCAAGUACUUCUGUCUCCCGGACUCAGCUAAU AGACUGUCUACCGCGGUUGAAGGAGAAAAAGUUCGUUACCCGACCAACUACUCCAGAAACCUAGUAA CACCAGUAAAGUAGCAGGGGUGUUUCGUAUCAGCACAAACCCAGUGUAGUACAGGCGUGAGUGACACCG CAUCCCCACGGGCGACCGUGGCGGUGGCGUGCGUGGCGGCCUGCCUAGGGGCAACCCUAGGGAGC CUCUAAUAUGGACAUUGGUGUGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGG CUAAUCCUAAUCUGCGGAGCAUAAACCCAAAGCCAAGGGGAGUGUGUCUAAACCGGCAACUCUGCA GCGGAACCGACUACUUGGGUGUCGUGUUUCCUUUAUUCUCAAACUGGUGUCUUAUGGUGACAAU CUCAGAGUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGAGCAACAGAGCAAUUGUCUACCAA UUUGUUGGAUUUAUCCACUCUAAUCUAGCGGUUUCAGAACAUUGAAUUAUAUUCUAAACCCUCAACA AGACGA	enterovirusB 69
3 1 5	UUAAAAACAGCCUGUGGGUUGUUCACCCACAGGGCCCCACGGGCGCUGAGCACACUGGUUAUCACGG 1 UACCUUUUGUGCGCCUGUUUUUAUACCCCCCCCCCAAGUAAACCUUAGAAGCAAUACACACCCUGGUCG 5 ACAGUGGUGGUGACAAUCCAGUCACAUCAUGACCAAGCACUUCUGUUAUCCCCGGACCGAGUAUCAU AGGCGUCUACGCGCGCUAAGGAGAAAGCGUUCGUUACCCGCGCAGUACUCCAGAAAGCCUAGUA CACCAGUAAAGUUGCAGAGUGUUUCGUUCAGCACUUCCCCCGUGUAGAUACAGGCGUGAGUGACACCG CAUUCUCCACGGGCGAGGUGGCGGUGGCGUGCGUGGCGGCCUGCCUAGGGGACACCCCAUAGGACG CUCUAAUACCGCAUGGUGGCAAGAGUCUAUUGAGCUAGCUGGUAGUCCUCCGGCCCCUGAAUGGCGG CUAAUCCCAACUGCGGAGCAGCACCCUCUAAACCAGAGGGGAGCGUGUCGUAUUGGGCAACUCUGCA GCGGAACCGACUACUUGGGUGUCGUGUUUCCUUUAUUCUUAUACAUUGGCGUCUUAUGGUGACAAU UGAGAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAAACAAAGCUAUCUAUAUUA UUCAUUGGUUAUCAUACCACUUAACCCUGAAAGAGAUCAAGACCUUGCACUUUAUCUUAUCACUAAACA CUGCAA	enterovirusB 73
3 1 6	UUAAAAACAGCCUGUGGGUUGUUCACCCACAGGGCCCCACUGGGCGCUGAGCACACUGGUUAUCACGGU 1 ACCCUUGUGGCGCGUUUUUAUACCCCCCCCCCAAGUAAACCUUAGAAGCAAUACCCAAUUGGCUAAU 6 AGACAGCCAGUACACCCAGCUGGGUUUCGGCCAAAGUACUUCUGUUAACCCGGACCGAGUAUCAUAG GCUGCUCACGCGCGUGAAGGAGAAACGUUCGUUAUCCGGCCAAUUAUUCUAGAGAAACCUAGUAAACA CCAUGAAAGUUGCGCAGUGUUUUCGUCUCCGCACAACCCAGUGUAGUACAGGUCGAUGAGUCACCGG UUCUCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCUAGGGGCAACCCUAGGACGCU UCAUACUGACAUUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAAACUGCGGAGCAUACCCACUAGCCAGUGGGCGGUGUCGUUACCGGCAACUCCGCGAG GGAACCGACUACUUGGGUGUCGUGUUUCCUUUAUUCUUAUACUGGCGUCUUAUGGUGACAAUUG AAAGAUUGUUGCCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACAAACAGGGCGAUUAUUUAUUGUU UGUUGGUUUUGUUCUUAUAAACUAUAGAGUUUGAAUAUCUAGAAUUAUUAUAAACUACAACACA ACAAA	enterovirusB 74
3 1 7	UUUAAAAAGCUGCUGGGUUGUUAACCCACCCACAGGGCCCCACUGGGCGCCAGCACACUGGUUAUCAC 1 GGUACCUUUGUGCGCGGUUUUAUUAUACCCUUCUCCCUACUAAAGUAAACUAGAAGCAAUUAACAAAG 7 ACCAACAGUAAAGCACAGCGCACCAGCUGUGUUUCGGUACAAGCACUUCGUUAUACCCCGACUGAGUAU CAUAGGCUACUACAGUGGCUAGAAGGAGAAAAAGUUCGUUACCCGACCAUUAUUCUAGAGAGCCUA GUAAACCAUUGGAGGUUGCGCGGCGUUUCGUCUCCACACAACCCAGUGCAGAUACGGUCGAUGAGCC ACCGCAUUCUCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCUAGGGGCAACCCUAGG GACGCUAAUACUAGCAGUGGAGGAGAGUUCUUAUUGAGCUAAUUGGUAGUCCUCCGGCCCCUGAAUG GCGGCUAAUCCCAACUGCGGAGCAGAUACCCACGCGCCAGUGGGCAGUCUGUCGUAUUGGGCAACUC UGCAGCGGAACCGACUACUUGGGUGUCGUGUUUCCUUUAUUCUUAUACUGGCGUCUUAUGGUGAC AAUUGAGAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACGAGCAGAGCUAUUAUUUA CUAUUUGUUGUUUAUACCAUUGAACUAUAAAGUUUUUAAAAACCUUUAACCUUACUUAUAGUACUCA AUACAGCAAA	enterovirusB 75
3 1 8	UUAAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCCCACUGGGCGCUGAGCACACUGGUUAUCCGGU 1 ACCCUUGUGGCGCGUUUUUAUACCCUUCUCCCUACUGCAACUAGAAGUUAACACAGUCACCGUGAU 8 AGGCGGUUCAGUAAACCAACUGGAUCUUGACCAAGCACUUCUGUUAACCCGGACUGAGUAUCAUAA CGUCGUUACUGGCGUGAGGAGAAACGUUCGUUACCCGGCCAAUUAUUCUAGAGAAACCUAGUACCA CCAUGAGGUUGGCGAGCGUUUCGUCUCCACACAACCCAGUGCAGAUACAGGUCGAUGAGCCGCA CUCCUCACGGGCGACCGUGGCGGUGGCGUGCGUGGCGGCCUGCCUAGGGGCAACCCUAGGGACGCU UAUCCCAACUGCGGAGCAGUACACCCACAGUGGGCAGCCUGUCGUAAACGGGCAACUCUGCAGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUCUUAUACUGGCGUCUUAUGGUGACAAUUG AAAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCAGUGACAAACAGAGCAAUUGUCUUAUUAU UGUCGGUUUGUUAACCAUAAAUUCAAAGUCCUAAAAACCCUCAUUAUUAUUGUUAUUAACACA ACAAG	enterovirusB 77

3 1 9	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUUACCCGGU ACCUUUGUGCGCCUGUUUUUAUACCCUCCUCCCAUUAUCGUAACAUAGAAGCAAUGCACUCCCGGUA AUAGCUGACGUGACAAACCAGUUGCGUUUUGACCAAGCACUUCUGUAACCCCGGACCGAGUAUCAU AAGCUGCUGCGCGGCUAAAGGAGAAACCGUCCGUUACCCGGCUAACUACUUCGAGAAACCUAGUAC CACCAUGAACGUGCGGAGUGUUUUCGUCAGCACUUCUCCCGUGUAGAUACGGUCCGAGAGCCACCG CAUUCUCCACGGGCGACCGUGGCGGUGGUGCGUUGGCGGCCUGCCUGUGGGGCAACCCACAGGACG CUCUAAUUAAGACAUGGUGUGAAGAGUCUAUUGAGCUAGUAGUAGUCCUCCGGCCCCUGAAUGCGG CUAAUCCUAACUGCGGAGCACAUACCCCUAAUCCAAGGGGACAGUGUGCGUAACGGGCGACUCUGCA GCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUCUUUAUUAUUGGCGUCUUAUGGUGACAA UUGCAGAGUUGUUAACCAUUAAGCUUUUGGAUUGGCCAUCCGGUGAGCAAUAGAGCUUUUAUUAUCA AUUUGUUGGAUUCGUCGCACUCAACCCACACAGUCUUUAGAACACUCGGAUUUUAUCUUGACCCUAAU ACGAGAAA	enterovirusB 79
3 2 0	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUUACCGGU ACCUUUGUGCGCCUGUUUUUAUACUCCUCCUCCCAACCGUAACUUAAGAAGCAGCACUAACUGAUCAA UAGCCAGCGCGCAUGCCAGCCAGCGUUCGUAUACAAGCACUUCUGUUAACCCGGACUGAGUAUCAUA GACUGCUCACGCGGUGAAGGAGAAAACGUUCGUCUACCCGGCCAAACUACUUCGAGAAACCUAGUAAC ACCAUGGAAGUUGCGGAGUGUUUCGUCAGCACUACCCAGUGUAGAUACAGGUCGAGUAGUACACGCG AUUCCUCCACAGGGGACUGUGGCGGUGGCGUGGUGGCGGCCUGCCUAGGGGAAACCCAUGGGACGC UCUUAUAUAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAAGUCCUCCGGCCCCUGAAUGCGGC UAUCCUAACUGCGGAGCAUACCCCUAAACCCAGAGGGCAGUGUGUGCUAAUUGGGCAACUCUGCAG CGGAACCGACUACUUGGGUGUCCGUGUUUACUUUAUUUCUAAAUUGGCGUCUUAUGGUGACAAU GAGAGAUGUUAACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGACCAAUAGAGCUAUUAUAUAUACUCU UUGUUGGGUUAUACCAUUAAGCUUAGAAAGAGGUUAAAACUCUGCAUUAUUAUUAUUAUUAUUAUUC CGCAA	enterovirusB 80
3 2 1	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUUACCCGGU ACCUUUGUGCGCCUGUUUUUAUACAUCUACCCUCCUCCCAACCGUAACUUAAGAAGCAAGAUACCCGAUACCCAGG AAUUGUGGCAAAACAGUCACAUUCUGAUACAAGCACUUCUGUUAUCCCGGACUGAGUAUCAUAGGCU GUUCACACGCGUGAAGGAGAAAACGUUCGUUACCCGGCCAAACUACUUCGAAAAACCUAGUAUACCCG UGAACGUGCGAGAGUGUUUCGUCAGCACUUCUCCCGUGUAGAUACAGGUCGAGUAGUACCCGCAUUC CCACAGGGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUGCGGGGCAACCCGAGGACGCUCUA AUACAGACACGGUGUGAAGAGUCUAUUGAGCUAGUUGGUAAGUCCUCCGGCCCCUGAAUGCGGCUAAU CCUAACUGCGGAGCGAGCGCUUGUAUCCAACAGGUAGCUCGUCGUAACGGGUAACUCUGCAGCGCA ACCGACUACUUGGGUGACCGUGUUUCCUUUAUUAUUAUACUGGCGUCUUAUGGUGACAAUACAG AGUUGUUAACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGUCAAAUAGAGCAAUUGUGUUUCUUUUGU UGGUUGGUUAACCAUUAAGCUUAGACAGCUACAACACUCUGUCUACAUGUUAAGAUUAGACACAGCA AA	enterovirusB 81
3 2 2	UUAAAAAGCCUGUGGGUUGUACCCACCCAGGGCCACUGGGCGCUAGUACACUGGUUACCCGGU ACCCUUGUACGCCUGUUUUUAUUCUCCUCCUCCCAACCGUAACUUAAGAAGUAAAGCAAACCCGAUC AAUAGCAGGUGCGACACACAGUCGUAUCUUAUCAAGCACUUCUGUAACCCCGGCGGAGUAUCA UAGACUCUCACGCGGUGAGGAGAAAACGUCGUAUCCCGGCUAGUACUUCGAGAAACCCAGUAACA ACACCAUAGAAGUUGCAGAGUGUUUCGUCAGCACUUCUCCCGUGUAGAUACAGGUCGAGUAGUACCC GCAUUCUCCACGGGUGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUAGGGGCAACCCAUGGGAC GCUCUAUAUACAGACAGUGGUGUGAAGAGUCAAUUGAGCUAGUAGUAUUCUCCGGCCCCUGAAUGCG GCUAAUCCUAACUGCGGAGCACGUGCCUCAAUCCAGGGGUGGCGUGUCGUAACGGGUAACUCUGC AGCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUAUUAUUAUUGGCGUCUUAUGGUGACAA UUGAAGGAUUGUUAACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGACUAACAGAGCAAUUAUAUAUUA AUUUGUUGGUUUUAUACCAUUGAAUUAACAAGUUCUAAUACCCUAAAUAUUAUUAUUAUUAUUAU ACGUCAA	enterovirusB 82
3 2 3	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUUACCCGG UACCCUUGUGCGCCUGUUUUUAUACCCUCCUCCCAACCGUAACUUAAGAAGCAAGCAACCUUAUGAUCAA UAGUGGGCGUGGCAUCCAGUCACGUUAUGAUACAAGCACUUCUGUAGACCCCGGACUGAGUAUCAUAG ACUGCUCACGCGGUGAAGGAGAAAACGUUCGUUACCCGGCCAGUUAUUCGAGAAACCCAGUAACA CCAUGAAUGUUGCAGAGUGUUUCGUCAGCACUUCUCCCGUGUAGAUACAGGUCGAGUAGGACACUGCA UUCUCCACGGGCGACCGUGGCGAGUGGCGUGGUGGCGGCCUGCCUAGGGGUCACCAUAGGACGCU CUAAUACUGACAUUGGUGCAAGAGUCUAUUGAGCUAAUUGGUAAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAACUGGAGCACAUUCCUCAAUCCAGGGGCGAGUGUGUCGUAACGGGCAACUCUGCAGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUAUUAUUAUUGGCGUCUUAUGGUGACAAUUA CAGAAUUGUUAACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGACUAACAGAGCAAUUAUUAUUAUUA CGUUGGAUUUGUCCACUCGAUUUCAAAGCAGUCAUACACUCUCAUUAUUGCUAACUCAACACA ACAAA	enterovirusB 83
3 2 4	UUAAAAAGCCUGUGGGUUGUACCCACCCACAGGGCCACCGGGCGCCAGCACACUGGUUACCCGGU ACCCUUGUGCGCCUGUUUUUAUACCCUCCUCCCAACCGUAACUUAAGAAGCACACCAUGUAUACGGUCA AUAGGCGGCGCAGUACACCAACUGGGCCACGACCAAGCACUUCUGUUAACCCCGGACUGAGUAUCAU AAGCUGCUCACGCGGUGAAGGAGAAAACGUUCGUUACCCGGCCAGUUAUUCGAGAAACCCAGUAACA CACCAUGAAGGUGCGCAGCGUUUCGUUCCGCACAACCCAGUGUAGAUACAGGUCGAGUAGUACCG CGUUCUCCACGGGUGACCGUGGCGGUGGCGUGGUGGCGGCCUGCCUAGGGGCAACCCAUGGGACG CUUCAAUACUAGCAUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAAGUCCUCCGGCCCCUGAAUGCGG CUAAUCCUAACUGCGGAGCAGAUACCCACAACACAGUGGCGAGUCUGUCGUAACGGGCAACUCUGCA GCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUAUUAUUAUUGGCGUCUUAUGGUGACAAU	enterovirusB 84

	UGAGAAUUGUUACCAUUAUAGCUAUUGGAUUGGCCAUCCGGUGACAAACAGAGCAGUUGUUUACCUA UUUGUUGGUUUCGUACCGCUGAACCUUAAAGUUCUAAAGACCCUCAUUUUUACUUAAGCACUUAACA CAGCAAA	
3 2 5	UUAAAACAGCCUGUGGGUUGUUCACCCACAGGGCCACCGGGCGCUAGCACACUGGUAUUGCGGU ACCUUUUGUGCGCCUGUUUAUACACCCUCCUUAAUGUAACUCAGAAAGUACAAAUUACCGGUCAAAC AGAUGACUCAGUAAACCAACUGAGCCUCGACCAAGCACUUCUGUUUCCCGGACCGAGUAUCAAUAG GCUGUCACCGCGGCUAGAGGAGAAAGUUCUUUACCCGGCCAAUUAUUCGAGAAAGCCAGUACCA CCAUGGAGGUUGCGCGGCUUUCGCUCCGCACAUCGCCAGUGUAGAUACAGGCUGAUGAGUACCCGCG UUCGCCACGGGUGACCGUGGCGGUGGCGUGCGUGGCGGCCUGCCCAUGGGGUAACCAUGGGACGCU UCAAUACUGACAUGGUGUGAAGAGCCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCUUGAAUGCGGCU AAUCCUAACUGCGGAGCAGAUACCCACACACAGCGGGCAGUCUGUCGUAACGGGCAACUCUGCAGC GGAACCCACUACUUUGGGUGCCGUGUUCUUUUUUUUAUGCUGGCGUCUUAUGGUGACAAUUG AGGAAUUGUUACCAUUAAGCUAUUGGAUUGGCCAUCCAGUGUCCAAACAGAGCGAUCAUUUACCGUU CGUUGGUUUCGUACCAUAAACCACAAGUCACUCAGCACACUAAUUAUUAUUAACUCUCAACACA ACAAG	enterovirusB 85
3 2 6	UUAAAACAGCCUGUGGGUUGUACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUAUUAACGGU ACCCUUGUGCGCCUGUUUAUUAUACCCUCCUCCACGCAACCUAGAAAGCAUGACUUAAGCGGUCAAU AGAUGGCUAGUAUGCCAAACUGAGUUAUGACCAAGCACUUCUGUUACCCGGACCGAGUAUCAAUAG GCUGUCACCGCGGCUAGAGGAGAAAGUUCGUUACCCGGCCAGUUAUUCGAGAAACCUAGUAUCA UCAUGAAGGUUGCGAGGCGUUUCGCUCCACAGACCCCGGUGUAGAUACAGGCUGAUGAGUACCCGCA UUCGCCACGGGCGACCGUGGCGGUGGCGUGCGUGGCGGCCUGCCCAUGGGGCAACCAUGGGACGCU GCAUUAUUGACAUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCUUGAAUGCGGCU AAUCCUAACUGCGGAGCAGACACCCACAUGCCAGUGGGCAGUCUGUCGUAACGGGCAACUCUGCAGC GGAACCCACUACUUUGGGUGCCGUGUUCUUUUUAUUCUUAUUAUGGCGUCUUAUGGUGACAAUUG AAAGAUGGUUCCAUUAAGCUAUUGGAUUGGCCAUCCAGUGACUAACAGAGCAUUAUGUCUACCGUU CGUUGGUUACGACCAUAAUUAUAAAGUUUGAACACACUUAUUUUUAUUAUAGCACUUAACACC UUAAA	enterovirusB 86
3 2 7	UUAAAACAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUAUUAACGGU ACCCUUGUGCGCCUGUUUAUUAUACCCUCCUCCACGCAACCUAGAAAGCAUUAAGCGGUCAAU UAGCAAGCGCAGCACUCCAGCUGUGUUUCGGUCAAGCACUUCUGUUUCCCGGACUGAGUAUCAAUA GGCCACUCACUGGGCCGAAGGAGAAACGUCGCUUACCCGGCCAAUUAUUCGAGAAACCCAGUAAC ACCAUAGAGGUUGCGCAGCGUUCGCUCCGCACUCCCCAGUGUAGAUACAGGUCGAUGAGUACCCG GUUCCCGCACGGCGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGGGUAACCAUUGGACG UUCAAUACUGACAUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCUUGAAUGCGGC UAAUCCUAACUGCGGAGCAGAUACCCACGCUCCAGUGGGCAGUCUGUCGUAACGGGUAACUCUGCAG CGGAACCGACUUAUUGGGUGCCGUGUUCUUUUUAUUCUUAUUAUUGGCGUCUUAUGGUGACAAU GAAAGAUUGUUACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGUCUAAUAGAGCAUUAUUAUUCUGU UUGUUGGUUUUAUACCACUGAAUUAACAGUUUACAAGACCCUCAUUAUACAUCUACGUUUAACAA GACAAA	enterovirusB 87
3 2 8	UUAAAACAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACUGGGCGCUAGCACACUGGUAUUAACGGU ACUUUUUGUGCGCCUGUUUAUUAUACCCUCCUCCUUAUGUAACUUAAGAAUAAACAAACGGUCA AUAGGUAGCUCAGCAAAACAGCUGAGCCUCCGGCCAGCACUUCUGUUUACCCGGACUGAGUAUCAA AAGCUGUUAACACGCGCUGAGGAGAAACGUCGUAUCCGGCCAAUUAUUCGAGAAACCUAGUAUA CACCAUGAAGGUUGCGUAGCGUUCGCUCCACACAACCCAGUGUAGAUACAGGUCGAUGAGUACCCG CUCUCCCGACGGGUGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGAGCACUCCAUUGGGACG CUUCAUACUGACAUUGGUGAAGAGUCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCUUGAAUGCGG CUAAUCCUAACUGCGGAGCAGACACCCACACGCCAGUGGGCAGUCUGUCGUAUUGGGCAACUCUGCA CGGAAACCGACUACUUUGGGUGCCGUGUUCUUUUUAUUCUUAUUAUUGGCGUCUUAUGGUGACAAU CAAGGAAUUGUUACCAUUAAGCUAUUGGAUUGGCCAUCCGGUAUCUAAACAGAGCAUUAUGUUUACCCU UUCACCGGUUUGUACCAUUGAAUUGCAUAGUCUCAACACCCUUGGCUUAUUAUUAUUAUUGAAU CAGCAAA	enterovirusB 88
3 2 9	UUUAAAACAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACUGGGCGCUAGCACaUGGUAUcACGG UACCcUUGUGCGCCUGUUUAUUAUACCCUCCUCCCAAAGCAUUAAGAAAGAACUcAACCGGUCA AUAGACAGCCUAGUACACCAACUAGGCCCUGACCAAGUACUUCUGUUUCCCGGACUGAGUAUCAAU AGGCUGCUUACGCGGCCGAAGGAGAAACGUCGUAUCCGGCCAAUUAUUCGAGAAACCCAGUAC CACCAUGAAAGUUGCGCAGUGUUUCGCUCCGCACAACCCAGUGUAGAUACAGGUCGAUGAGUACCCG CAUCCCCCACAGGUGACUGUGGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGAGCACUCCAUUGGGACG AUCCCCCACAGGUGACUGUGGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGGGCAACCCAUUGGGACG CUUCAUUAACGACAUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCUUGAAUGCGG CUAAUCCUAACUGGAGCAGAUACCCACACACAGUGGGCAGUCUGUCGUAUUGGGCAACUCUGCA CGGAAACCGACUACUUUGGGUGCCGUGUUCUUUUUAUUCUUAUUAUUGGCGUCUUAUGGUGACAAU CAAGGAAUUGUUACCAUUAAGCUAUUGGAUUGGCCAUCCGGUAUCUAAACAGAGCAUUAUGUUUACCCU UUUGUUGGAUUGUACCAUUGAAUUGCAUAGUCUCAACACCCUUGGCUUAUUAUUAUUAUUGAAU CAGCAAA	enterovirusB 93
3 3 0	UUAAAACAGCCUGUGGGUUGUUCACCCACCCACAGGGCCACUGGGCGCCAGCACACUGGUAUUAACGGU ACCUUUUGUGCGCCUGUUUAUUAUACCCUCCUCCCAAUGUAACUUAAGAAUAAACUAGAAAGAUCAAU AGCAAGCAGACAGUAUGCCAAACUGUGUUUAUGAUCAAGCACUUCUGUUUACCCGGACUGAGUAUCAAU GCUGCUAAAGCGGCUAGAGGAGAAACGUCGUAUCCGGCCAAUUAUUCGAGAAACCCAGUACCA CCAUGAAGGUUGCGCAGUGUUUCGCUCCACACAUCCAGUGUAGAUACAGGUCGAUGAGUACCCGCG UUCGCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGGAGUACCCAUUGGGACGCU	enterovirusB 97

	UCAAUACUGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCGUGAAUGCGGCU AAUCCUAAACUGCGGAGCAGAUACCCACACACCAGUGGGCAGUCUGUCGUAACGGGUAACUCUGCAGC GGAACCGACUACUUGGGUGUCCGUGUUUUUUAUCCUCAAUUGGCUUGCUUAUGGUGACAAUUG AUUAGUUGUUAACAUUAGCUAUUGGAUUGGCCAUCCGGUGUCUAAACAGAGCUAUUGUUUUUUUUU UGUUGGUUUCGUGCCACUAAAUUACAAGUUUUUGAAACUCUCAACUUAUUUGACACUCAUAAU GCAA	
3 3 1	UUAAAAACAGCCUGUGGGUUGUCCCCACCCACAGGGCCCAAUGGGGCCAGCAGCUCUGGUUUUACGGU ACCUUUGUGCGCCUGUUUUUAUCCCCUCCCCCAAUUGUAACUUAAGAAGCAACACGACCGACCAAC AGUGGGCGUGGCACACCAGCCACGUCUUGGUCAAGCACAUCUGUUUCCCCGGACUGAGUAUCAAUAG ACUGCUACCGCGGUCGAAGGAGAAAACGUUCGUUAUCCGGCCAAUACUUCGAGAAGCCAGUAACA CCAUGAAUUAUGCGAAGUGUUUCGUCAGCAGCUCUCCAGUGUAGAUACGGUCGAUGAGUCACCGCA UUCCCCACGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUUAUGGGGCAACCCGUAGGACGCU UCAUAUUGGACAUUGGUGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCGUGAAUGCGGCU AAUCCUAAACCGGAGCAGGCACCCACAACCCAGUGGGCAGCUUGUCGUAACGGGUAACUCUGCGGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUCUUAUACUGGCUUGCUUAUGGUGACAAUUG AGAGAUUGUUGCCAUUAGCUAUUGGAUUGGCCAUCCGGUGACAAACAGAGCAAUUGUUUAUCUUAU UGUUGGUUUUAUACCGUUAAGUUUAAGGUCCUAAAAACCCUUAACUUGGUCUACUACUCAACUCA GCAA	enterovirusB 98
3 3 2	UUAAAAACAGUCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGGCCAGCAGCAGCUCUGGUUUUACGGU ACCUUUGUGCGCCUGUUUUUAUCCCCUCCCCAAUUGUAACUUAAGAAGUAGAGCAACACUGAUCA UAGUGGGUGGCAACUGCCAGUUGCAUCACGAUCAAGCAGUUCUGUAUCCCCGGACUGAGUAUCAAU AGGCGUCUACCGCGGUGAAGGAGAAAACGUUCGUUAUCCGGCCAGUUAUUCGAGAAACCUAGUA CAUCAUGGAGGUUGCGCAGUGUUUCGUCUCCGACACCCCCAGUGUAGAUACGGUCGAUGAGUCACCG CGUCCCCACGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUAGGGAUAAACCAUGGGGACG CUUCAUACUGACAUGGUGAAGAGUCUAUUGAGCUAACUGGUAGUCCUCCGGCCCGUGAAUGCGG CUAAUCCUAAACUGUGGAGCAGACACCCACAUGCCAGUGGGCAGUCUGUCGUAACGGGCAACUCCGA GCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUCUUAUUGGCUUGCUUAUGGUGACAAU UGAAAGAUUGUUAACCAUAGCUAUUGGAUUGGCCAUCCGGUGUCAACAGAGCAAUUAUUUUUG UUUGUUGGUUUUAUACCUUGAACUUUAAGUGCUAAAGACUCUCAAUUUUUAUUAUUAUCUCAACA AAUCAGG	enterovirusB 100
3 3 3	UUAAAAACAGCCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGGCCAGCAGCUCUGGUUUUACGGU ACCUUUGUGCGCCUGUUUUUAUCCCCUCCCCACUUGAAACUUAAGAAGUGAAACACACUGAUCAAU AGUUGACAUGGCAAGCCAGCCAGUUCUUGAUCAAGCAGCUCUGUUUCCCCGGACUGAUUAUAAUAG ACUGCUUCGCGGGUUGAAAGAGAAAAGCGUUCGUUAUCCGGCUAACUACUUCGAGAAACUUAUACA CCAUGGAAGUUGCGGAGUGUUUCGUCAGCAGCAUCCCAUGUAGAUACGGCCGAGUAGUCACCGCA UUCGCCACCGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUUGGGGCAACCCACAGGACGCU CUAAUACAACUUGGUGCGAAGAGUCUAUUGAGCUAGUUAUAGUAGUCCUCCGGCCCGUGAAUGCGGCU AAUCCUAAACUGCGGAGCAGCUGCUUCCAUCAGGAGGUGGCGUGUCGUAACGGGUAACUCUGCAGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUCUUGCAUUGGCUUGCUUAUGGUGACAAUUG AGAGAUUGUUAACCAUAGCUAUUGGAUUGGCCAUCCGGUGACAAACAGAGCAAUUAUUAUUCUUA UGUUGGAUACAUUCCAUAUAGAACUACAAGGAGUUAUAAACCCUCAAUUAUUAUUCUGUUCUUGAAUUC GCAAG	enterovirusB 101
3 3 4	UUAAAAACAGCCUGUGGGUUGUCCCCACCCACAGGGCCAAUGGGGCCAGCAGCAGCUCUGGUUUUACGGU ACCUUUGUGCGCCUGUUUUUAUCCCCUCCCCAAUGAAACUUAAGAAGCUUAACUCAACCGGUCGAU AGACAGCCAGUACACCAACUGGGUUUUGACCAAGUACUCCUGUUUCCCCGGACUGAGUAUCAAUAA CGUCGUCGCGGGCGAAGGAGAAAACGUUCGUUAUCCGGCCAAUUAUUCGAGAAACCUAGUACCA CCAUGAAGCUGUGCGAGUGUUUCGUCUCCGCAACACCCAGUGUAGAUACGGUCGAUGAGUCACCGCA AUCCCCACGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUAGGGGCAUCCCAUGGGACGCU UCAUACUGACAUGGUGCGAAGAGUCUAUUGAGCUAAUUGGUAAUCCUCCGGCCCGUGAAUGCGGCU AAUCCUAAACUGGGGCGAGACACCCACAGCCAGUGGCGAGUCUGUUAACGGGUAACUCCGAGC GGAACCGACUACUUGGGUGUCCGUGUUUCCUUUAUUCUUAUUGGCUUGCUUAUGGUGACAAUUG AAAGAUGUUGCCAUUAGCUAUUGGAUUGGCCAUCCGGUGUCUAAUAGAGCAAUUAUUAUUCUUA UGUUGGCUUACUCCGUUAAUAAAGUGAGCUACAGACCCUGGCUUAUUAUCAAUUGAACACA AUAAA	enterovirusB 106
3 5 5	UUUCCCCGAAGGGUCUAACGCUAUGCUGUAUCCACGGCAGAAGUAUGACCUACCGUUAACCGGCCA UGUACUUCGAGAAGCCUAGUACCAUUAUGAAAGUUAUUGAUUGUACGUUCCCCAGCAACCCAGUG UAGUUCAGGUCGAGUAGUCUCGGCAUGCCCCACGGGCGACCGUGGCGUGGCGUGGCGGCGGCG CCUACCAACCUUAUGGUGUAGGACGCUAAGAUACUGACAUGGUGUGAAGAGCCUUAUUGAGCUACGUGG UAGUUCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCCGGAGCGUCGCUAGCAAUCCCGCUGAAG GGCGUCGUAUUGCGCAAGUCUGGAGCGGAACCGACUACUUGGUGUCCGUGUUUCCUGUUUACUUA UUGUUUGGCGUCUUAUGGUGACAACUUA	enterovirusB 110
3 6 6	UUAAAAACAGCCUGUGGGUUGUCCCCACCCACAGGGCCACUGGGGCCAGCAGCAGCUCUGGUUUUACGGU ACCUUUGUGCGCCUGUUUUUAUAAACCCUCCCCAAUGCAACUUAAGAAGCUAACUUAUACCGGUCAAU AGACAGUCCAGUACGCAACUGGGCUUGCAUCAAAGCACUUCUGUUAACCCGGACUAGUAUCAAUAG GCUGCUACGCGGCGAAGGAGAAAACGUUCGUUAUCCGGCCAAUUAUUCGAGAAACCCAGUAACA CUAUGGAGGUUGCGAGUGUUUCGUUCCACACAACCCAGUGUAGAUACGGUCGAUGAGUCACCGCU UUCCCCACGGGCGACCGUGGCGGUGGCGUGGCGGCGCCUAGGGGCAACCCAGGAGGACGCU UCAUACUGACAUGGUGUGAAGAGUCUAUUGAGCUAAUUGGUAAUCCUCCGGCCCGUGAAUGCGGCU AAUCCUAAACUGCGGAGCAGAUACUACAGCCAGUGGGCAGUCUGUUAACGGGUAACUCCGAGC	enterovirusB 111

	GGAAACGACUACUUUGGGUGUCCGUGUUUCCUUUUUUAUUAUUAUUGGUGCUUAUGGGUGACAAUUG AGAGAUUGUUACCAUUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCAAUUGUUUACCUAUU UGUUGGCUUUUGGCCGUUGAACUACAAGGUUUUAAAACUAAUUAACUACAUCUAGUACUUAACACG UUGAA	
3 3 7	UUUCAUCAUAAGCUACUUGGUCUUGCCAACCGGAAGAUAAUUGGUUAUUGUUGGUUUCACAAUUC UGCGAACCUUCACUAAUUUUUAUUAUUGUCCUAAAAUUCAAACAGG	enterovirusB 113
3 3 8	UUAAAAAGCUCUGGGGUUGCACCCGCCCCAGAGGCCACGUGGCGGCUAGUACUCCGGUAUUGCGG UACCCUUGUACGCCUGUUUUUAUACUCCUCCCGUAACUUAAGACGCACAAAACCAAGUUCAAUAGAA GGGGGUACAAACAGUACCAACCAAGCAACAGCACUUCUGUUUCCCGGUGAUUGUUAUAGACUGCU UGCGUGGUUGAAAGCGACGGAUCCGUUAUCCGCUUAUGUACUUCGAGAAGCCAGUACCAACCCUGGA AUCUUCGAGUGCUUUGCGCUCAGCACUACAACCCAGAGUGUAGCUUAGGCUAGUAGUCUGGACAUCC CUCACCCGUGACGGUGGUCUAGGUGCGUUGGCGGCCUACCUAUGGCUAACGCCAUGGGACGCUAGU UGUGAACAAAGUGUGAAGAGCCUUAUAGAGCUACAUAAGAAUCCUCCGGCCCCUGAAUGCGGCUAAUC CCAACCCUGGGCAGGUGGUCACAAACCAAGAGUAGUUGGCGUGUCGUAACGCGCAAGUCCGUGCGGAA CCGACUACUUUGGGUGUCCGUGUUUUUUUUUAUUUAUUGGUGGUGCUUAUGGUGACAAUCACAGAU UGUUAUCAUAAGCGAAUUGGAUUGGCCAUCCGGUGAAAGUGAGAUUCAUUAUCUACUGUUUGCUG GAUUCGCUCAUUGAGUGUGUUACUCUAAGUACAUAUUUCAAGUUAUUUCAAGUUAUUCAUACAGACAAUUGUA UCAUA	poliovirus1- Sabin
3 3 9	UUAAAAAGCUCUGGGGUUGUCCCAACCCAGAGGCCACGUGGCGGCCAGUACACUGGUUAUUGCGG UACCUUUGUACGCCUGUUUUUAUACUCCUCCCGUAACUUAAGAGCACAAUGUCCAAAGUUCAAUA GGAGGGGUACAAACAGUACCAACCAAGCAACAGCACUUCUGUCCCGGUGAGGCUUGUAUAGGCU GUUUCACCGCUAAAAGCGGCUAUCCGUUAUCCGCUAUGUACUUCGAGAAGCCUAGUAUACCCUU GGAAUUCUUGCAGUGCUUGCGCUCACACUACAACCCAGAGUGUAGCUUAGGUGCAUGAGUCUGGACG UUCUCCACCGCGCAGCGUGGUGCAGGCGUGGUGGCGGCCUACCUUGGCGCCAAAGCCACAGGACGC UAGUUGUGAACAAAGGUGUGAAGAGCCUUAUAGAGCUACCUAGAGUCCUCCGGCCCCUGAAUGCGGCU AAUCCUAACCAAGGAGCAGGAGUGGCAUCCAGGACACAGCCUGUCGUAACGCGCAAGUUCGUGGC GGAAACGACUACUUUGGGUGUCCGUGUUUCCUUUUUAUUUAUACAAGGUGCUUAUGGUGACAAUCA UUGAUUGUUAUCAAAAGCAAUUGGAUUGGCCAUCCGGUGAGAAUUGAUUAUUAUUAUACUCUCU UGUUGGAUUGCUCCUUUGAAUUCUUGGACUCACACCUAUGGAAUUAACCUAUGUUAAGAUUA CAUCCACCU	poliovirus2- Lansing
3 4 0	UUAAAAAGCUCUGGGGUUGUACCCACUCCAGAGGCCACGUGGCGGCUAGCACUCCGGUAUUAACGG UACCUUUGUGCGCCUGUUUUUAUACUCCUCCCGCAACUUAAGAGCACGAAACCAAGUUCAAUAGAA GGGGGUACAAACAGUACCAUACGAACAAGCACUUCUGUUUCCCGGUGACAUUGCAUAGACUGCU CACGCGGUUGAAAGUGAUCAUCCGUUAACCGCUUGUGUACUUCGAAAAGCCUAGUAUCCGCUUGGA AUCUUCGACGCGUUGCGCUCAGCACCCGACCCCGGGUGUGGCUUAGGCUAGUAGUCUGGACAUUC CUCACCGGUGACGGUGGUCUAGGUGCGUUGGCGGCCUACCUAUGGCUAACGCCAUGGACGUUAGA UGUGAACAAAGUGUGAAGAGCCUUAUAGAGCUACAUAAGAUCCUCCGGCCCCUGAAUGCGGCUAAUC CUAACCCGGAACAGGCGUGCGGAACAGUAGUAGGCUUGGUGUAACGCGCAAGUCUGGCGGAA CCGACUACUUUGGUGUCCGUGUUUCCUGUUUUUAUUAUAGGUGCUUAUGGUGACAAUCAGAGA UUGUUAUCAUAAGCGAAUUGGAUUGGCCAUCCGGUGAGUGUUGUGUAGGUUAACUACUGUUUGU GGAAACACUGUGUAGCUUUACUUCUUAUUAACCAAUUAUUAUAAAAACAUAACGAGGAUAAAAACA CAAUACUACA	poliovirus2- Sabin
3 4 1	UUAAAAAGCUCUGGGGUUGUUCCCAACCCAGAGGCCACGUGGCGGCUAGUACACUGGUUAUACCG UACCUUUGUACGCCUGUUUUUAUACUCCUCCCGCAACUUAAGAGCAUACAUAUACAGCUCAUAG GAGGGGUGCAAGCCAGCGCCUCCGUGGGCAAGCACUACUGUUUCCCGGUGAGGCGCGAUAGACUG UUCACCGGUGAAAGUGUCCGUAUCCGCUAUGUACUUCGAGAAGCCUAGUAUCCGCUUG GAUUCUUGCAGCGUUGCGCUCAGCACUACAACCCCGAGUGUAGCUUUGGCGGAGUAGUCUGGACAG UCCCAUCUGGCGACAGUGGUCAGGCGUGCGUGGCGGCCACCUUGGCGCCAAAGCCACGGGACGCU AGUUGUGAACAGGGUGUGAAGGCCUUAUAGGCUACAUGAGAGUCCUCCGGCCCCUGAAUGCGGCUA AUUCUAACCAUGGAGCAGGACGUGCAACCCAGCAGCCAGCCUGUUGCUAACGCGCAAGUCCGUGGCG GAACCGACUACUUUGGUGUCCGUGUUUCCUUUAUUCUUGAUGGCGUUAUUGGUGACAAUCAUA GAUUGUUAUCAUAAGCGAGUUGGAUUGGCCAUCCAGUGGAAUACAGAUUAUUAUACUCCUUGUUUG UUGGAUCCACUCCGAAACGUUUUACUCCUUAACUUAUUGAAUUGUUGAAGACAGGAUUCAGUG UCACA	poliovirus3
3 4 2	UUAAAAAGCUCUGGGGUUGUUCCCAACCCAGAGGCCACGUGGCGGCUAGUACUUGGUUAUUGCGG UACCUUUGUACGCCUGUUUUUAUACUCCUCCCGCAAGUACUUAAGAGCAAUUCACAAAGUUCAAUA GAGGGGUACAAACAGUACCAACCAUGAACAAAGCACUUCUGUUUCCCGGUGAAGCUUAUAGACUG UUCACCGGUGCAAAUUGGACAGUCCGUUAUCCGCUACAGUACUUCGAGAAGCCUAGUAUACUUCG GAAUUCUUGCAGCGUUGCGCUCAGCACUACAACCCCGAGUGUAGCUUUGGCGGAGUAGUCUGGACAG GAAUUCUUGCAGCGUUGCGCUCAGCACUACAACCCGAGUGUAGCUUAGGCGGAGUAGUCUGGCGA CCCCACCGGCGACGGUGGCCCCAGGCGUGCGUUGGCGGCCUACCCUAGGCUAAUGUACUUGGACGCUAG UUGUGAACAAAGGUGUGAAGAGCCUUAUUGAGCUACUUGAGAGUCCUCCGGCCCCUGAAUGCGGCUAAU CCUAACCAUGGAGCAAGUGCCCCACAAGCCAGUGGGUGGCUUGUCGUAUUGCGCAAGUCUUAUGGCGGA CCGACUACUUUGGUGUCCGUGUUUCCUUUAUUCUUGAUGGCGUUAUUGGUGACAAUCAUA AUUGUUGUACAUAUAGCUUUUGGAUUGGCCAUCCGGUGAUUUUGAGAUUAUUAUACCUUCUUAUUA CUUCAUAGUACUUAUUUAUUAUACUACUACUGUAGUUA	coxsackievir usA1
3 4 3	UUAAAAAGCUGGGGUUGUUCCCAACCCAGAGGCCACGUGGCGGCUAGUACUCCGGUAUACAGG UACCUUUGUACGCCUGUUUUUAUACUCCUCCCGUAACUUAAGAGCAAUUCACAAAGUUCAAUAG GAGGGGUACAAACAGUACCAACCAUGAACAAAGCACUUCUGUUUCCCGGUGAAGCUUAUAGACUG UUCACCGGUGCAAAUUGGACAGUCCGUUAUCCGCUACAGUACUUCGAGAAGCCUAGUAUACUUCG GAAUUCUUGCAGCGUUGCGCUCAGCACUACAACCCGAGUGUAGCUUAGGCGGAGUAGUCUGGCGA CCCCACCGGCGACGGUGGCCCCAGGCGUGCGUUGGCGGCCUACCCUAGGCUAAUGUACUUGGACGCUAG UUGUGAACAAAGGUGUGAAGAGCCUUAUUGAGCUACUUGAGAGUCCUCCGGCCCCUGAAUGCGGCUAAU CCUAACCAUGGAGCAAGUGCCCCACAAGCCAGUGGGUGGCUUGUCGUAUUGCGCAAGUCUUAUGGCGGA CCGACUACUUUGGUGUCCGUGUUUCCUUUAUUCUUGAUGGCGUUAUUGGUGACAAUCAUA AUUGUUGUACAUAUAGCUUUUGGAUUGGCCAUCCGGUGAUUUUGAGAUUAUUAUACCUUCUUAUUA CUUCAUAGUACUUAUUUAUUAUACUACUACUGUAGUUA	coxsackievir usA11

	UUACCACGGUUGAAAGCGACCGAUCCGUUAUCCGCUAUGUACUUCGAGAAGCCUAGUACUCCUUG GAUUCUUCGAUGCGUUGCGUCAGCACUCAACCCCGAGUGUAGCUUAGGCUGAUGAGUCUGGACAG CCUCUACCGGCGACGGUUGGCCAGGCUUGCGUUGGCGGCCUACUUGUGGCCCAAAGCCACAGGACGCU AGUUGUGAACAGGUGUGAAGAGCCUAUUGAGCUACAUGAGAGUCCUCCGGCCCCUGAAUGCGGCUA AUCCUAAACCACGGAGCAGGCGAUCCGCAAACAGCAAUAGCCUUGCGUAAACGCGAAAGUCUGUGCG GAACCGACUACUUUGGUGACCGUGUUUCUUAUUUUUAUUCAGGUGCUUAUGGUGACAAUACU AGAUGUUUAUCAAAGCGAGUUGGAUUGGCCAUCCGGUAAAAGUCAAGUAUAUACUUAUUGUUU GUAGGAUUUAUCCUUGAACGUCCAUUUACUCAAUUAUCAGAGUCACAUAUACUAGUAAGAUCUC ACUAUCACA	
3 4	UUAAAAAGCAGCUCUGGGUUGUCCCCCCCCAGAGGCCACGUGGCGGCCAGUACUCCGGUAUACGG UACCCUUGUACGCCUGUUUUUAUCCCUUCCCCCGCAACUUAGAAGUUUAAAACCAAGUUCAAUAG AAGGGGUACAGCCAGUACCAACCAAGCAAGCACUUCUGUUUCCCCGGUGAUGUUGUAUAGACUG CUAACCGCGUUGAAACAGCUGAUCCGUUAUCCGCUUGUGUACUUCGAGAAGCCUAGUAUCAUCUUG GAUUCUUCGAUGCGUUGCGUUCAGCGCUCUACCCAGAGCGUAGCUUAGGCUGAUGAGUCUGGGCAU UCCCCACCGGCGACGGUGGCCAGGCUUGCGUUGGCGGCCUACCUAUGGCUAAGCGCAUAGGACGCUA GUUGUGAACAGGUGUGAAGAGCCUAUUGAGCUACACAAGAGUCCUCCGGCCCCUGAAUGCGGCUAA UCCUAAACCAUGGAGCAAGUGAUCACAAUCCAGUGAUAUUCUUGCGUAACGCGCAAGUCCGUGGCGG AACCGACUACUUUGGUGUCCGUGUUUCCUUUUUAUUAUUGGUGCUUAUGGUGACAAUACU GAUUGUUUAUCAUAAAGCGAAUUGGAUUGGCCAUCCGGUGAAAGUGAGACUUAUCAUCUACUUGUU UUGGAUUCACUCCAUGAACACAUUCACUCCAGUAUCAUUAUCAAUAGUAUACAGAAUAGACAGUU GUUACACA	coxsackievir usA13
3 4 5	UUAAAAAGCAGCUCUGGGUUGUCCCCCCCCAGAGGCCACGUGGCGGCCAGUACUCCGGUAUUGCGG UACCCUUGUACGCCUGUUUUACACUCCCUUCCCCGUAACUUAGACGCAUAAACCAAGUUCACUAGG AGGGGUGCAAACAGCACCACCAAGCAAGCACUUCUGUUUCCCCGGUGAUAUUGUAUAGACUGUA UCCACGGUUGAAACGAUUGAUCCGUUAUCCGCUUCUGUACUUCGAGAAGCCUAGUAUCAUCUUGGA AUUCUUCGACGCGUUGCGUCAGCAUUAACCCAGAAUGUAGCUUAGGUGCUAUGAGUCUGGACAUUC CUACCCGGUGACGGUGGCCAGGCGUGCGUUGGCGGCCUACCUUGAGACCAAAGUCACAGGACGCUAG UUGUGAACAGGUGUGAAGAGCCUAUUGAGCUACAAGAGAGUCCUCCGGCCCCUGAAUGCGGCUAAU CCUAAACCAAGGAGCAAGUGCUCACGAACCAAGUGAGUGGCCUUGUGCUAAGCAGCAUUCUGUGGCGGA ACCGACUACUUUGGUGUCCGUGUUUCCUUUUUAUUAUUGGUGCUUAUGGUGACAAUACUUG AUUGUUAUCAUAAAGCGAAUUGGAUUGGCCAUCCGGUGAAAUCAAAGUUGAUCUUAUUGUUGU UGGAUUCACUCCAUAUACUUAUUUCAAUUGACCUAUUAACGUUAUUGUAUUAUAGUUAAGAACAU CAUCACA	coxsackievir usA13; coxsac kievirusA18
3 4 6	UUAAAAAGCAGCUCUGGGUUGUCCCCCCCCAGAGGCCACGUGGCGGCCAGUACUCCGGUAUUGCGG UACCCUUGUACGCCUGUUUUUAUCCCUUCCCCGUAACUUAGAAGCAUAAACCAAGUUCAAUUA GAAGGGGUACAAACAGUACCAACCAAGCAAGCACUUCUGUUUCCCCGGUGAAGUUGUAUAGACU GUACCCACGGUUAAGCAUUGAUCCGUUAUCCGCUUCUGUACUUCGAGAAGCCUAGUACUACCUU GGAAUUCUUGCGUUGCGCUCACACUCUACCCAGAGUGUAGCUUUGGCGUGAUGAGUCUGGACG UUCUCCACCGGCGACGGUGGCCAGGCGUGCGUUGGCGGCCUACCUUGGCGCCAAAGCCACAGGACGC UAGUUGUGAACAGGUGCGAAGAGCCUAUUGAGCUACAAGAGAGUCCUCCGGCCCCUGAAUGCGGCUA AUUCCCAACCAAGGAGCAGUAGUUGCAAGCCAGCAUUCGACUUGCGUAACGCGUAAGUCUGUGGC GGAACCGACUACUUUGGUGUCCGUGUUUCCUUUUUAUUAUUGGUGCUUAUGGUGACAAUCA UUGAUUGUUUAUCAUAAAGCGAAUUGGAUUGGCCAUCCGGUGAGAAUACAAGUUAUUAUUCUUGU CAUUGGAUUUACUCCAUGAGACAACAUACAUACGACCUUAUUAAGAGUCAUUAUACUUAUUGAGACU UCCACCAUA	coxsackievir usA17
3 4 7	UUAAAAAGCAGCUCUGGGUUGUCCCCCCCCAGAGGCCACGUGGCGGCCAGUACUCCGGUAUACGG UACCUUUGUACGCCUGUUUUUAUCCCUUCCCCGUAACUUAGAAGUUUAAACUAAAGUUCAAU AGAGGGGUACAAACAGCACCACCAAGCAAGCACUUCUGUUUCCCCGGUGAAACUGCAUAGACU GUAAACCAAGUUGAAAGCGGUUGAUCCGUUAUCCGCUUGAGUACUUCGAGAAGCCUAGUACACCUU GGAAUUCUCCGAGUGCGUUGCGCUCACACUCUACCCGAGUGUAGCUUAGGCGGAUGAGUCUGGCGU UCCCCACCGGCGACGGUGGCCAGGCGUGCGUUGGCGGCCUACCUAUGGUUAACGCCAUGGACGCUA GUUGUGAACAGGUGUGAAGAGCCUAUUGAGCUACUCAAGAGUCCUCCGGCCCCUGAAUGCGGCUAA UCCCAACCAAGGAGCGAGUGCUCACAAUCCAGUGAGUGGCUUGCGUAACGCGCAAGUCUGUGGCGG AACCGACUACUUUGGUGUCCGUGUUUCCUUUUUAUUAUUGGUGCUUAUGGUGACAAUACUAG GAUUGUUUAUCAUUAAGCUUUUGGAUUGGCCAUCCGGUGACUUCUAAAGUUAUUAUACUUGCU UAAUUUUCAAACACAUUCUAGUUGUUGAAUUAUAAACUUCACA	coxsackievir usA19
3 4 8	UUAAAAAGCAGCUCUGGGUUGUACCCACCCAGAGGCCACGUGGCGGCCAGUACUCCGGUAUACGG UACCCUUGUACGCCUGUUUUUAUCCCUUCCCCGUAACUUAGAAGCACAAAACCAAGUUCAAUAGA AGGGGGUGCAAACAGUACCAACCAAGCAAGCAUUCUGUUUCCCCGGUGAUGUUGUAUAGACUGC CGCGCGGUUGAAAGCAAGGAUCCGUUACCCGCUCAAGUACUUCGAGAAGCCUAGUAUUAACCUUG AAUUCUUGAUGCGUUGCGUUCAGCACUCGACCCUGGAGUGUAGCUUAGGCUGAUGAGUCUGGACGUC CCUACCGGUGACGGUGGCCAGGCGUGCGUUGGCGGCCUACCUAUGGCUAAGCCAUAGGACGCUAG UUGUGAACAGGUGUGAAGAGCCUAUUGAGCUACUAGAGUCCUCCGGCCCCUGAAUGCGGCUAAU CCUAAACCAUGGAGCGGCGUACAGACCAAGUAGCUUUGCGUAUUGCGCAAGUCUUAUGGCGGA ACCGACUACUUUGGUGUCCGUGUUUCCUUUUUAUUAUUGGUGCUUAUGGUGACAAUACUAG AUUGUUAUCAUAAAGCGAAUUGGAUUGGCCAUCCGGUGAAAGCGAGACUUAUUAUACUUAUUGU UGGACUUAACCAACUUAUACAUUUUAUUCUAGGUGCACUUGUAUAGCAUUAAGAAUCAAACAGUUG CAUCAUA	coxsackievir usA20

3	UUA AAAACAGCUCUGGGGUUGU UCCACCC CAGAGGCCACAGUGGCGGCAGUACUCUGGUAUUAACGG	coxsackievir usA21
4	UACCUUUGUACGCCUGUUUUUAUCCCUUCCCGGUAACUUAGAGCUUUAUCAAAGUUCAAUAGC	
9	AGGGGUGUACAAACAGUACCUUCUACGAACAAGCACUUCUGUUUCCCCGGUGUAUUCACAUAGACUGUA	
9	CCACCGGUGUACAAAGUAGUUAUCCGUUAUCCGCUUAGAGUACUUCGAGAAAGCUAGAACCUUGGA	
9	AUCUUCGAGUGCGUUGCGCUC AACACUCUGCCCCGAGUGUAGCUUAGGCGUGAUGAGUCUGGGCACUCC	
3	CCACCGCGCAGCGUGGCCACAGGCUUGCGUUGGCGGCCUACCC AUGGCUUAGUCCGUGGGGACGCUAGUU	coxsackievir usA22
0	GUGAACAAAGGUGUGAAGAGCCUAUUGAGCUACUCAAAGAGUCCUGCGGCCUUGAAUGCGGCUAAUCC	
3	UAACCACGGAGCAACCGCUCACAACCCAGUGAGUAGGUGUUCGUAAUGCGUAAGUCUGUGGCGGAAC	
3	CGACUACUUGGGUGUGCGUGUUUCCCUUUAUAUUCUACUAGUGGCUUGCUUAGUGGUGACAAUUUACAAA	
3	UUGUUACCAUAUAGCUUAUUGGAUUGGCCACCCAGUAUUGGCUAAUAUAUUGGAGUGUUUCUUUCAU	
3	AGCCUUUAUUAACAUACAAUUUUUAUCCAAUAAACAGAGUCA	coxsackievir usA24
3	UUA AAAACAGCUCUGGGGUUGU UCCACCC CAGAGGCCACAGUGGCGGCCAGUACUCCGGUAUUGCGG	
5	UACCUUUGUACGCCUGUUUUUAUCCCUUCCCGGUACCUUUAAGAGUAAACAAACCAAGUUCAAUA	
1	GAAGGGGUGUACAAACAGUACCAACCCAGCAACAAGCACUUCUGUUUCCCGGUGAGCUUGUAUAGACU	
1	GUACCCACGGUUGAAAACACUAGUACCGUUAUCCGCUUUGUACUUCGAGAAGCCAGUAUACCCU	
3	GGAAUCUUCGAGUGCGUGCGCUCAGCACUAACCCAGAGUGAGCUUAGGCUUAGGUGAGUCUGGACA	enterovirus 96
5	AUCCCUACCGCGACCGGUGGUUCCAGGCGUGGCGUUGGCGGCCUACCUGUGGCCCAAAGCCACAGGACGC	
3	UAGUUGUGAACAAAGGUGUGAAGAGCCUAUUGAGCUACGAGAGAGUCCUCCGGGCCUGAUGUGCGCU	
3	AAUCCCAACCAACCGAGCAGGUGAUCACAGCCAGUGUAUAGCCUGUGCUUAGCGCGCAAGUCUGUGGC	
3	GGAAACCGACUACUUGGGUGUGCGUUGUCCCUUAUUUAUUUAUUGGCUUAGUGGACAUUGGUGACACACA	
3	GAUUGUUAUCAUAAAGCGACUUGGAUUGGCCAUCCAGUGAAAUA AAACACAUCAUUUACCUGAUUG	enterovirus 99
5	UUGGAUUUACUCCCGUGACUAUCUACUACACAGGUCUUAUUAAGGUUAUUCUGUGUAUAGGUCACA	
3	AUUUCACA	
3	UUA AAAACAGCUCUGGGGUUGU UCCACCC CAGAGGCCACAGUGGCGGCCAGUACUCCGGUAUACGG	
5	UACCUUUGUACGCCUGUUUUUAUCCCUUCCCGCAACUUAGAAGCUUUAAUUAUCAAAGUCAAUAG	
3	AAGGUGGUAACAAACAGUACCGCCACGAACAAGCACUUCUGUUUCCCGGUGAGUAAUUGCAUAAAGCUG	enterovirus 104
3	UCCACCGGUGAAGGCGUAUUGAUCGCUUACCCGCUUUGUACUUCGAGAGCCUAGGAAGCCUAGUAC	
3	GAUUAUUCGAGUGCGUUGCGCUCAGCACUAACCCAGAGUGUAGCUUAGGCUUAGGAGUCUGGACAU	
3	ACCCACCGCGCAGCGUGGUCUACGGCUGCGUUGGCGGCCUACCUGUGGCCCAAAGCCACAGGACGCU	
3	AGUUAUGAACAAAGGUGUGAAGAGCUUAUUGAGCUACUAAAGAGUCCUCCGGGCCUUGAUGCGGCUA	
3	AUCCUAACCAACCGGAUCAGGCGAUUGCAACCCAGCAAUUAGCCUGUCGUAAUGCGCAAGUCUGUGGC	enterovirus 104
5	GAACCGACUACUUGGGUGUGCGUGUUUCCUUUAUUUAUUCUUAACUUGGCGUCUUAUGGUGACAAUUCU	
3	AGAUGUUAUCAUAAAGCAAGUUGGAUUGGCCAUCCGGUGAGAAUAAAGAUUAUUGUUUAUCUGUUUG	
3	GUUGGAUUUACCCUAAGACCACUCAUUUACCGACUUAUUUAGAAUAAUUAUACUACUUGGAAUUA	
3	ACUAUACACA	
3	UUA AAAACAGCUCUGGGGUUGU UCCACCC CAGAGGCCACAGUGGCGGCCAGUACUCCGGUAUACGG	enterovirus 104
5	UACCUUUGUACGCCUGUUUUUAUCCCUUCCCGGUAACUUAAGAGCUUUUGUACCAACAAGUCAAU	
3	AGAAGGGUGUACAAACAGUACCAACCCAGCAACAAGCACUUCUGUUUCCCGGUGCGAUGACAUAGACU	
3	GUACCCACGGUUAAGGUAACCAUACCGUUAUCCGCUUUGUACUUCGAGAGCCGCUAGUAUUGCCU	
3	GGAAUUCUGAUGCGUGGCGCUC AACACUCGACCCUGAGGAGUGAGCUUAGGCTUGAUGGAGUCUGGACA	
3	UCCCCACCGGUGACGGUGGCCACAGGCGUGCGUUGGCGGCCUACCCAUAGCUA UUGCUAUGGGAUGCU	enterovirus 104
5	AGAGUUGAACAAAGGUGUGAAGAGCCAAUUGAGCUACUAAAGAGUCCUCCGGGCCUUGAUGCGGCUA	
3	AUCCCAACCAACCGGAUCAGUGCGUGCAACCCAGCACGUAGCUUGUGCUUAGGCGCAAGUCUGUGGC	
3	GAACCGACUACUUGGGUGUGCGUGUUUCCUUUAUUUAUUCUUAACUUGGCGUCUUAUGGUGACAAUUCU	
3	GGAUUGUUAUCAUUAAGCGAUUGGAUUGGCCAUCCGGUGAGAAUAAAGAUUAUUGUUUAUCUGUUUG	
3	UGGCUUUUACUCCAUAAACAAAAACACAUUCCCUAAUUAUAAAGUGGUACUAUUGAUCAGAUACGC	enterovirus 104
3	UAGCCUUA	
3	GGUAAACCUUGUACGCCUGUUUUUAUAAUCCCUUCCCGGUAACAGUAGAAGCACAAAGACAAACCGCU	
5	CAUAGAGUAGCGCAUACCUCCAGGUGUGUUAUGAGCAAGCACUUCUGUCCCGGGUUGCGCCCUAAU	
4	CGUGUGCGAAGCCAAAAGAGGCAUAUACGUUAACCCGCUUACGCUACUACGGAACCAACUAGUAUU	
3	CUUCGAUUGAUUUGAGAGUUGCGUUCAGCGCCCAACCCCGGCGUAGCUACGGCUGAUGAGGCUCCG	enterovirus 104
3	AACACCCCAACAGGCAGUGUGUCAGAGCCUGCGUUGGCGGCCUGCCCAUACCCUGCGGUUUGGAGC	
3	CUCACUUCAGUACAUUGGAUAGAGAGUCUACUGUGCUAGUCAAGAUUCCUCCGGGCCUUGAAUGUGG	
3	AUAUUAUACCAACUAGCUUACUUGCGUACUUGCGUUAUUGGAGGCGUUAUUGGCAUUGCGUAGUGG	
3	GCGGAACCAACUACUUGGGUGUGCCUGUUUCCCUUACUUCUACUGAAUGCUUAUGGUGACAAUUG	
3	UUUUGACUAGUUGGCUAAUACUUGCGUUCAGUGUGUAAUUAACAAAACAAGAUCAUCAGUUGUUA	enterovirus 104
3	CUAUCCCUACCGUUGGAGCGCAUUUAUACAAAACAAGAUUUUCUGUCAUA	

3	CCCCUAGGAUCCACUGGAUGUUAGUACACUGGUAUCGCGGUACCUUUGUACGCCUGUUUUAUACCC	enterovirusC 105
5	CUUCCCCGCAACUUUAGAAGUAUCAAAGUACCGCUCAAUAGCCGCCACGCUCCAGUGUGGUUUCG	
5	AGCAAGCACUUCUGUUUCCCCGGUUGCGUCCCAUAGCUGUGCAAACGGCAAAAAGGGACAAUACUG	
	UUACCCGCUUUAUACUACGGGAACCUAGUACCGCCAUUGAUUGUAUUGAGGGUUGCGCUCAUCAC	
	CUUUCCCCGGUUGAGCUCAGGCUAGUAGGGUCAGAAUCCCCACAGGUAACUGUGUCUGAGCCUGCG	
	UUGCGGCCUGCCUCGCUUUAUGGCUGGGACGCUACUACAUGACAUUGGUGCGAAGAGUCUACUG	enterovirusC 113
	UGCUAUACAAGUCCUCCGGCCCCUGAAUUGGUAUAAUCCUAAACCACUGAUCCACGUAACGCAAAAC	
	CAGCGUGUAGUGGGUUAACGCGCAAGUCGUGGCGGAACCGACUACUUUGGGUGACCGUGUUUCC	
	UUUAUUAUCCUUAUGAAUGUUUAUGGUGACAAUUGUUUAUUCAGUUGUUACCAUUGCUACAUAUCAA	
	UAACCUAGUGUCAAACCGAUUGAAUUGCUACA	
3	ACUCUGGUAUUGCGGUACCUUUGUACGCCUGUUUAUUAUCCCUUCCCCAUGUAACUUUAGAAGCAAU	enterovirusC 116
5	UCACAAGGUUCAAUAGAGGGGUACAACACAGUACCAACACGAACAAGCACUUCUGUUUCCCCGGUG	
6	AAGCUGUAUAGACUGUCCCCACGGUCAAAAACGGCAGAUCCGUUAUCCGCUUAUGUACUUCGAGAAG	
	CCUAGUAUACUUCUGGAUUCUUGAUGCGUUGCGUUCAGCACUCUUAUCCGAGUGUAGCUUAGGUCG	
	AUGAGUCUGGGCAUCCCCACCGCGCAGGUGGGCCAGGUCGUGGCGGCCUACCCAUUGGCUAAC	
	GCCAUGGGACGCUAGUUGUGAACAAGGUGUGAAGGCCUUAUUGAGCUACAUGAGAGUCCUCCGGCCC	enterovirusC 117
	CUGAAUUGCGGUAAUCCCAACCGAGAGCAAGUACUACACACCCAGUGGGCAGCUUUGUCUAACGCG	
	UAAGUCUGUGGCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUUUUAUUUUUGAUUUGGUGCUU	
	AUGGUGACAUCAAGAUUGUUUAUCAUAUAGCUUUUGGAUUGGCCAUCCGGUGAUUUUUGAGACUAAA	
	ACGCUUGUUUGCCUUAUCACACAUAUUAGUAUUUUUGUAUUAUUAUACGGUCAUA	
3	UUCCCCUCCAGAGGCCACGUGGCGGCUAGUACUCUGGUAUUGCGGUACCUUUGUACGCCUGUUUU	enterovirusC 118
5	UAACCCUCCUCCCAAGUAACUUUAAGAAGCAAUUCACAAAGUUCAAUAGAGGGGGUACAACACAGUAC	
7	CACCACGAACAAGCACUUCUGUUUCCCCGGUGAAGCUACAUAAGACUGUCCCCACGGUCGAAGGUGGC	
	AGAUCGUUAUCCGCUACAAGUACUUCGAGAAGCCUAGUACCAUUGGAAUUCUUGAUGCGUUGCGC	
	UCAGCACUCUACCCCGAGUGUAGCUUAGGUCGAUGAGUCUGGGCAUGCCCCACCGGUGACGGUGGCC	
	CAGGUCGCGUUGGCGGCCUACCCGUGGCUAACGCCAUGGGACGCUAGUUGUGAACAGGUGUGAAGA	enterovirusC 119
	GCCUUAUUGAGUACUUGAGAGUCCUCCGGCCCCUGAAUUGCGGUAAUCCUAAACCACGGAGCAAGUGC	
	UCACAACACAGUGAGUGGCUUGUCGUAACGCGUAAGUCUGUGGCGGAACCGACUACUUGGGUGUCC	
	UGGUUUCCUUUAUUUAUUAUUGGUGGCUUAGGUGACAAUCUAAAGAUUGUUUAUCAUAUAGCUUUU	
	GGAUUGGCCAUCCGGUGAUUUGAAGAGAUUAUUUACCUAUUGUUCUACCAUAUCUAAUACUCA	
	UCACAGCCACUAUACACA	enterovirusC 120
3	UUUAAAACAGCUUGAGGGUUGUCCCCACCCUCAAGGCCUACUGGGCGCUAGUACACCGGUUUCUGG	
5	UACCCUUGUACGCCUGUUUUUAUCUUCUCCCCCAAAAGUAACUUUAGAAGCAAGUAAACUACACAGCU	
8	CAUAGCGCGUACGACCGCAGUGCGCUAUGAGCAAGCACUCCUGUUUCCCCGGUUGCGUCCUUAU	
	GCUGUGGAAACGGCAGAAGGGGCAAAUUCGUUACCCCGCUUAUCUACUACGGGAAGCCUAGUAACAC	
	CAUUGAACACUUGAGGGUUGCGUUCAGCGCCAAACCCGCGUAGCUCAGGCCGAUGGGGCUUCGGA	
	AUCCCCACGGGUGACCGUGUGGAAGCCCGCGUUGGCGGCCUGCCACACCAUUGGGUGUGGGACGCU	enterovirusD 68
	CAUUAUUGGACAUUGGUGCGAAGAGCCUACUGUGCUAGGUGAGAUCCUCCGGCCCCUGAAUUGGCGU	
	AAUCCCAACCAACCGAACCUGGUGUAGCAUCCAGCAACUAGCGGGCGUAUUGCGCAAGUCGGUGAC	
	GGAAACCAACUACUUGGGUGACCGUGUUUCCUUUAUUCUUAUUGAAUGCUUAUGGUGACAAUUGUUGA	
	GUACAGUUGUUAACCAUACUGCAUUCUGCAAAUAGUGAAUAAACUACGCAACAAUUAACACA	
3	UUAAAACAGCUCUGGGUUGUCCCCACUCAAGGGCCACGUGGCGGCUAGUACUCUGGUAUUCUGGU	enterovirusD 70
6	ACCUUUGUACGCCUGUUUUAAUUCUCCCCCAACGUAACUUAGAAGCUUUUAAACCAAGCUCAAUA	
0	GGUGGAGCGCAACACAGCGCUCUUAUGAGCAAGCACUUCUGUCUCCCCGGUGUGGUUGUAUAGACUG	
	UCCCCACGGUUGAAACAAUUAUCCGUUAACCGCUUAUAGUACUUCGAGAAACCUAGUAUUGCCUUC	
	GGAGUGUUGAUGCGUUGCGUCAGCACUAACCCGUGUGUAGCUUUGGUGCAUGAGUCUGGACGUA	
	CCCCACUGGCGACAGUGGUCAGGCGUGCGUUGGCGGCCUACUCAUGGUGAAAACCAUGAGACGCUAG	enterovirusD 71
	ACAUGAACAAAGGUGUGAAGAGUCUAUUGAGCUGCUAUGAGUCCUCCGGCCCCUGAAUUGCGGCUAAU	
	CCUAACCAUGGAGCAAGUGCUCACAAACAGUGAGUUAUUGUCGUAACGCGCAAGUCCGUGGCGGA	
	ACCGACUACUUGGGUGUCCGUGUUUACUUAUUUAUUGACUGCUAAUGGUGACAAUUAUAUA	
	UUUUUUA	
	AAUUA	
3	UUAAAACAGCUCUGGGUUGUCCCCACCCACAGGCGCCAGUGGCGGCUAGUACUCCGGUACCCCGG	enterovirusD 72
6	UACCCUUGUACGCCUGUUUUUAUACUCCCUUCCCCAAGUAACUUUAGAAGAAUAAACUUAUUGUCAA	
1	CAGGAGGGGUAACAACAGUACCAACACGAACACACACUUCUGUUUCCCCGGUGAAGUUGCAUAGA	
	CUAGUACCAAGGUGUAAAGCGAUGAAUCCGUUACCGCUUAAGUACUUCGAGAAGCCUAGUAUUAUC	
	UUGGAAUUCUUGAUGCGUUGCGAUCAGCACUCUACCCCGAGUGUAGCUUGGGUGCAUGAGUCUGGAC	
	ACCCACACCGGCGACGUGUCCAGGCGUGCGUUGGCGGCCUACCAUGGCUAGCACCAUGGGACGCU	

	AGUUGUGAACAGGUGCGAAGAGCCUAUUGAGCUACCUAGAGAGUCCUCCGGCCCCUGAAUUGCGGCUA AUCCCAACCACGGAGCAAUUGCUCACAAUCCAGUGAGUGUUUGUCGUAUUGCGCAAGUCUGUGGCG GAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUUUUAUUAUUGGUGCUUUGGUGACAAUCUG AGAUGUUAUACUAUAGCUAUUGGAUAGCCAUCCGGUGAUUUCUGAAAUUUGCCAUAAUUUUU CACAAAUCCUACAACAUUACACUACACUUUCUUGAAUAAUUGAGACAACUCAUA	
3 6 2	UUUCCCCGUGAAGUUACAUAGAUGUGCCACGGUGAAAGUGACUGAUCCGUUACCCGUUUAGUA CUUCGAGAAGCCUAGUAAACUUCUGGAUUCUGAUGCGUUGCCGUCAGCACUACCCCCAGAGUGU ACUCCGACCCGAUACGUCUGGACACUCCCCACGGCGACGGUGGCCAGGCGUGCGUUGGCGGCCUAC CUGUGGCCCCAAAGCCACAGGACGCUAGUUGUGAAACAAGGUGUGAAGAGCCUUAUUGAGCUACUAAAGA GUCCUCCGGCCCCUGAAUUGCGGCUAAUCCCAACCACGGAGCAAGUGCUCACAAUCCAGUGAGUGGCU UGUCGUAACGAGCAAUUCUGUGCGGGAACCAACUACUUUGGGUGUUCUGGUGUUCUUUUUAUCAUU GGCUGCUUUGGUGACAAUCUA	enterovirusD 111
3 6 3	UUAAAACAGCCUGGGGUGUACCCACCCCCUGGGGCCACGUGGCGCUAGUACUCUGGUUCGUUAGA ACCUUUGUACGCCUGUUUCCCCUCCUUAACAAUUAAGAUUCUGCCAAUGUGGGGAGUAGUCCG ACUCCGACCCGAUACGUCGACACAGUAGACCGGUUCGCUUAGGACCCUUCUACGGAUUGGUUAGAGU UCCCCACCCCGUAACUUAGAAGUACUAGCAAAACCGACCAAUAGGAGCGUGGCACCCAGCUGCGUUA AGGUCAGCACUUCUGUCUCCCCGGCCAGAAUUGGUGUACCCGCCCUCUACUACAGAGAAGCCU AUUAACCAUUGAAGGCGAUGAGGAGUUGCGCUCACCACAACCCAGUGGUAGCUCUGAGAGAUGGG GCUCCGAGUCACCCUUGUAGUGGCGCAAGAGCCUUAUUGUGCUGGGGAGGUUUUCCCGGAGCCUUGGC CGUUCACUCCAAUCUGUUGUAAGUGGCCAAGAGCCUUAUUGGUCUAGAGAGGUUUUCCCGGAGCGG UGAUGGUGCUAAUCCCAACCCUCCGAGCGUGUGCGCACAAUCCAGUGUUGCUACGUCGUAACGCGCA AGUUGGAGGCGGAACAGACUACUUCUGGUAUCUCCGUGUUUCCUUAUUAUUUAUACAACAUAUUUAG GUGACAUUGACUGAUACUUAUGAGUUCGCCCCGUUGCCAUUGAAUUAUUGCCUUGUAUUAACCUUAUG CAUUUCAAAGCCACAGAUUCACCCUCGAGCUCAUUCACUUGCAGUUGUUGAAUCGCAUACA CAAGACAUUUGAAC	enterovirusE 1;bovineente rovirus1
3 6 4	UUAAAACAGCCUGGGGUGUACCCACCCCCUGGGGCCACGCGGCGCUAGUACUCUGGUUCGUUAG AACCCUUGUACGCCUGUUUCCCCUCCUUAACAAUUAAGAUUCUGCCAAUGUGGGGAGUAGUCC GACUCCGACCCGAUACGUCGACACAGACCGGUUCGCUUAGGACCCUUAUGCGGAUUGGUUAGAG UCCCCCCCCGAAACUUAGAAGCACAUUACAACUUGGCCAAUAGGAGCGUGGUAUCCAACCGCGUUG CGGUCAGCACUUCUGUUCCCCGGUCAGAAUGAUCGUUACCCGCCCUCUCCACUACGAGAAGCCU AUUAACCAUUGAAGUGAUGAUGAGUUGCGCUCAGCCACAACCCCGGUGGUAGCUCUGAGAGAUGGG GCUCGCACCAACCCCGUGGUGACACGGUUGCUUGCCCGGUGUGCUCUGGGUUCAGUUCGCGAGACU UCCUCCUCCUGACUACUUGUAGUGGCGCAAGAGCCUUAUUGUGCUGGGGAGGUUUUCCCGGAGCCU GAAUGGUGCUAAUCCCAACCCUCCGAGCGUGUGCGCACAAUCCAGUGUUGCUACGUCGUAACGCGUAA GUUGGAGGCGGAACAGACUACUUCGGUACUCCGUGUUUCCUUAUUAUUUAUUGUUAUUAUUGGU GACAUUGUCUGACAUUUUGGAAUUCGCUUGCCUGCCACUUAUAUAGCGUUUAUAUCAUCUGAUAGCA UUCCAUAAGACUACAGAACUACUUGGAACUCAUCGACUUGCGUUGUUCUGAAUACUGUUCACA ACCAAAACAAACAAA	enterovirusE 2
3 6 5	UUAAAACAGCCUGGGGUGUACCCACCCCCUGGGGCCACGUGGCGCCAGUACUCUGGUACGCUAGU AACCUUUGUACGCCUGUUUCCCCACCCUUAACAAUUAAGAUUACCAUACUGAGGGGAGUAGUCC GACUCCGCUCCGACACUGCCGCACAGUAUUCGGUACGCUUAGUACCUAAUACGGAGUAGUAGGU AUCCCCACCCCGCAACUUAGAAGCAUGCAACAACCCGACCGAUAGGAGCGCGGCAUCCAGUCGCGU AUAGGUCAAGACUUCUGUUCCCCGGUCCGAAAGGAUCGUUACCCGCCCAACCCACUACGAGAAGC CUAGUAACUGGCCAAACAACCGCAAAGUUGCGCUCAGCCACAACCCAGUGGUAGCUCUGGAAGAUG GGGCUCCGACUCCCGUGGUAACACGGUUGCUUGCCCGCGUGUGCUUCGGGUUCGCGCCUAGUGC CGUUCACUUAUAAAGUUGUUAACAGCAAAGAGCCUUAUUGUGCUGGGAUGGUUUUCCUCCGGGGCCG UGAUGGUGCUAAUCCCAACCCUCCGAGCGUGUGCGCACAAACAGUGUGCUACGUCGUAUAGCCGUA AGUUGGAGGCGGAACAGACUACUUCGUAACCCGUGUUUCCUUAUUGUUAUUAUACAUUUUAGG UGACACUGACUGACAUUUUGGAAUUGCUGCGUUGCCUCUCAAUAUUGCACUAUAUUAACUGAUUGC AUUUCACAACACUAGUAAACUAACCCUUGAACUUAUCGACUUGCUGUUUACCUAAGUCACUUAUUA CACAAGUACAAUA	enterovirusE 3
3 6 6	CGGAGUAGAGGUUAUCCCUUCUCCGUAAACUUAAGAAGUUAACUAAACGACCAUAGCGGUGGUG CAACAGCAUCACAACGGUCAAGCACUUCUGUUUCCCCGGCACGUCAGUGUCGUUACCCGCCACAC UACUACGAGAAGCCUAGUAACUUCCAAUCGACAGCACAGUUGCGUUCGGCCACAACCCUUGUGUA GCUCUGGAAGUUGGGCU	enterovirusE 4
3 6 7	UUUAAAACAGCUCGGGGUGUUCUCCACUCCCGGGGCCACGUGGCGCUAGUACUCUGGUACGUUAG UACCUUUGUACGCCUGUUUCCCCUCCUUAACAAUUAAGAUUACCAUACUGAGGGGAGUAGUCCG ACUCCGCUCCAGCACUGCUGCACCAGUGAAUUGGUUAGAUUAGUACCUUCACGGAGUAGAUUGGUA UUCCUACCCCGCAACUUAAGAAGUUUAUUAACAACCGACCAUAGGUGCACAGUAUCCAGCUGUGUA ACGGUACAAGCACUUCUGUUUCCCCGGUCCACCAGGAUCGUUACCCGCCGACCCACUACGAGAAGCC UAGUUAACUGGCAAGUGACUGCGAAGUUGCGUACCAACCAACCCAGUGGUAGCUCUGGAAGAUGG GGCUCGCAACUCCCGUGGUAACACGGUUGCGUGCCCGCGUGUGCUUCCGGGUUCGCGCCUUGUGCC GUUCACUUCUACGCUAUGUAACUAGCCAAGAGCCUUAUUGUGCUGGGACGGUUUUCCUCCGGGACCGU GAAUGGUGCUAAUCCUUAACUCCGAGCGUGUGCGCACAAUCCAGUGUUGCUACGUCGUAACGCGUAA GUUGGAGGCGGAACAGACUACUUCGGUACCCCGUUGUUCUUCUUUAUUAUUAUUAUUAUUGGUGA CAAUUGCAGAAAUUUGUGAUUAGCGACUUAACCAUUGCACAUAGCAUUAUUAUUGUUGAUUGCUUU CCACAAAACAUUCUGAUUAACCAACUUAUUAUCGAUCUUAUUCUUCUUGCUUGAAGAAUAAA UUAAAGAAGCACG	enterovirusE 5

3 6 8	UUCAAACAGCCUGGGGGUUGUACCCACCCUAGGGGCCACGCGGCUAGUACUCUGGUACGUUAGU ACCUUUGUACGCCUGUUUCCCCUCCUUAUAAUAAUUAAGAUUACCACUACUGAGGGGAGUAGUCC GACUCCGCUCCGACAAUGUCGCACACAGUGCAGUGUACGCUAGUACCUUUUCACGGAGUAGCUGGUA UUCCUCCCGAAACUUAGAAGCACUUAAAUAACCGACCAAUAGGCGCGGUGAGCCAGCCGCGUUA CGGUCAAGCACUCCUGUUUCCCCGGUCCACAGGAUCGUUACCCGCCGACCCACUGCGAGAAACCU AGUAGUUGGCCAAGUGGUAAACGAGGUUGCGCUCAGCCACAACCCAGUGGUAGCUCCGGAAGAUGGG GCUCGCACACCCCGCGGUAAACGGGUGUGCUGGCCGCGUGCUUCCGGGUCGCGUCCGUACCG UUCACCCUAAUACCAACGUAACAGCAAAGAGUCUAUUGCGUGGACGGUUUCCUCCGGGCGCGUG AAUGCUGCUAAUCCCAACUCCGAGCGUGUGCGCACAAUCCAGUGUUGCUACGUCGUAACGCGUAA UUGGAGGCGGACAGACUACUUCGGUACCCGUGUUUCCUUAUCAUCCAUUUCAUUUUGGUGA CAAUUGCUGAGAUUCUGAUUUGGCGACCUUACCACUGAAUUGCCUUAUUAUACUUGGUGCAU CCACAAACCCUCUGACAUACCGUCCUAUUAUCGACCUACUUGUUUCCUCAAUUUAAAGUAUAGA CUACAAUAGCAGA	enterovirusF 2
3 6 9	UUUAAAACAGCCUGGGGGUUGUACCCACCCUAGGGGCCACGUGGCGCCAGUACUCUGGUACAUUAG UACCUUUGUACGCCUGUUUCCCCUCCUUAUAAUAAUUAAGAUUACCACUACUGAGGGGAGUAGUCC CGACUCCGCUCCAGCAAUGCUGCACCAGUGCAGUGUACGCUAGUACCUUUUCACGGAGUAGUAGGU AUCCCCACCCGAAACUUAGAAGCAUGUGAUCAAUAGCACCAAUAGGCGCGCGGUUGCCAGCCGCGU UUCGGUCAAAGCACUUCUGUUUCCCCGGUCCGCAAGGAUCGUUACCCGCCGACCCACUGCGAGAAGC CCAGUAAACUGGCCAAGCGGUAGCGAGGUUGCACUCAGCCACAACCCAGUGGUAGCUCUGGAAGAU GGGCUCCGAGUUCGCCGUGGUAAACAGGAUGCCUGCCGCGUGUGCUUCCGGGUUCCGGUCCUGGUC CGUUAACCUCAAUACCGCUAAACAGUGAAGAGCCUAUUGCGCUGGACGGUUUCCUCCGGGCGCG UGAAUGCUGCUAAUCCUAAACCUUGAGCGUGUGCGCACAAUCCAGUGUUGCUACGUCGUAACGCGUA AGUCAGGGGCGGAACAGACUACUUCGGUACCCGUGUUUCCUUAUUAUUAUUAUUAUUAUUAUUAU GGUGACAAUUGCUGAGAUUUGCAUUAAGCGGUUUUGCCAUUGAACAUUGCUUUGUACUACUUGGU GCAUUUCCACAAACCUCCGACAUACCCAGCUCUACUAGCUGUUGUUUCCUUAUUAUUAAGACA UAAACUAGAGAACACAA	enterovirusF 3
3 6 7 0	UUAAAACAGCCUGGGGGUUGUACCCACCCUAGGGGCCACGUGGCGUAGUACCCUGGUACGCUAGU ACCUUUGUACGCCUGUUUCCCCUCCUUAUAAUAAUUAAGAUUACCACUACUGAGGGGAGUAGACC GACUCCGCUCCAGCACUGCUGCACCAGUGAACUGGUACGCUAGUACCUUUGCACGGAGUAGUAGGUA CACCCACCCCGUAAUUAAGAAGUACUGAACACAACCGACCAAUAGCGGCGUGACAGCCAGUACGCA ACGGUCAAAGCACUUCUGUUUCCCCGGUCCGUAUGGAUCGUUACCCGCCCAAGCUACUACAGAAGCC UAUUAAUACUACUGAAGACUGUUAAGGUUGCGCUAACCAUCCAGUGGUAGCUCGAGUGAGG GGCUCGCGAUUCCCCGUGGUAACACGGUUGCUGUUGCCGCGCGCACUCGGGUCGCGUUCUUGACC GUUACCUUGAGACAAACGUAAGUAGCCAAAGAGCCUAUUGUGCUGGUUUUGGUUAUCCUCCGAGCCGU GAAUGCUGCUAAUCCUAAACCCCAAGCGUGUGCGCACAAACAGUGUUGCUACGUCGUAACGCGUAA GUUGGAGGCGGAACAGACUACUUCGGUACUCCGUGUUUCCUUAUUCUAAUUAUUAUUAUUAUUAU GUGACAAUUGUGUGAUUUGUAAACUUAACUACGUCUUGCAUAAACGCCUUGUUUUAUUCUAGUCG CCUCCAUUAUAAUUAAGCUUUAACCCUUGAGUUAACUACUAGCAAUUUGAGCUCAAUUA CUAGAUUAUAGUA	enterovirusF 4
3 7 1	CUAAAACAGCCUGCGGGUUGUACCCGCCCGAGGGGCCACUGGGCGCUAGUACUCUGGUACGCUAGU ACACUAGCACGUUAGUGCUUUUUGUACGCCAGUACCUUUGUACGCCUGUUUUAUUAUUAUUAUUAU AUGUAACUUAGAAGUUUAGAAACAACAAGACCAAUAGGAGUCCAGCAUCCAGUUGGAUUGCGGU CAAGUACUUCUGUUUCCCCGGACCUAGUAGUGAUAGGCGUACCCACGGCGGAAGAUAAACCGUCC GUUAUCCGGCUACCUACUUCGGGAAGCCUAGUAACAUCUGAAGUCUCUGAGGCGUUUCGUCAGCA GACCCCGGUGUAGUUGGGUGUAGUGGUCUCCGCAUACCCACGGGUGACCGUGGCGGAGGCGCG UUGGCGGCCCGCCUAUGGCGAAAGCCUAAGGACGCCUUAUAGAGACAGGUGUGAAGAGCCUAGCUG AGCUGGGUAGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACACGGAGCGUCCACCGGCAAC CGGUCGGGAGGGCGUGUACCGGUAAUCUCUGGCGGAACCGACUACUUGGGUGUCCGUGUUUCC UUUUGAUCCUUAUUUGGUGCUUAUGGUGACAACGAUAAUGUUAUUAUUAUUAUUAUUAUUAUUAU CACCUGGAAAAAGUUAUCAGUGUUUGAUUAUUGUUGGCUUCACGCGCUCCAAACAAACAGGCCUUA CUUCUUUAUUAUUGCAUUUAUCCAUCCGAGGUUAUUAACAUUCCAUUCUGAAGCUUCUUCACAG UAGCAGUAAA	enterovirusG 2; formerly terovirus10
3 7 2	UUAAAACAGCCUGCGGGUUGUACCCACCCGAGGGGCCACUGGGCGCUAGUGCAGUGGUUAGCUAGU ACCUUGGUUUAACGGUACCAUUGUACGUUAGUACCUUUGCAGCGCGUUUUAUUAUUAUUAUUAUUA GCAACAUUAGAAGUCUAUCUAUUAAGGACCAAUAGGAGUCCAAACAGCCAGUUGGAUUGUGUCAA ACUUCUGUCUCCCGGACUUAAGUAGUAGGCGUUAACCCAGGCGGAAGAUAGAACUGUCCGUUAU CCGGCCAGCUAUCGAGAGCCUAGUAGCACCAAAGAUUCUGUUGGCGUUCGUCACGCGCGUCC CCCCGCGUAGUCCGGCGUAGUGGUCUCCGCAUCCCGACGGUGACCGUGGCGGAGCGCGUGGC GGCCCGCUUAUGGCGAAAGCCUAGGAGCGCAUUAUAGUACAGGGUGUGAAGAGCCUUAUUGAGCUA GUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACAGGAGCGUCCACCGGCAACAGCU GGGAGGGCGUGCUAACCGGUAACUCUGGCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUAU AUCCUUAUUGGCGUGCUUAUGGUGACAACUUAACAGUUGUUAUUAUUAUUAUUAUUAUUAUUAU GGAAAGAUUUAACAGUUUGUUGCUCUUAUUGCUCUUAAGUUUCCAUUAAGGAAAGCAUUAUUAU UCUUAUUUAUUAUUAAGGCUUGACAAGAUUUAAGCGCUCAUUAAGAUUCGCUUAUUAUUAUUAU CAAC	enterovirusG 3
3 7 3	UUAAAACAGCCUGGGGGUUGUACCCACCCACAGGGGCCACGGGCGCUUCGCAACUGGGAGUCCUAA CGUCGGAUUUGCGGAUCCGAUGUUUAUUGCCCGUGCGCGCGUUUAUUAUUAUUAUUAUUAUUAU UAACUUUAAGAAGUUUAUUAUUAACAAAGACCAAUAGGAGUCCAAACUCCAGUUGGAUCCGGUCAGGC ACUUCUGUCUCCCGGACCUAGUAGUAGGCGUUAACCCAGGCGGAAGAUAGAACUGUCCGUUAU	enterovirusG 4

	CCGGCCAGCUACUUCGAGAAGCCUAGUAACAUCAAAGAUUCUGUCUGGCGUUCGCUACAGCGGUUU CCCCCCGCGUAGAUCGGGCGUAGUGGUCUCCGCAUACCCACGGGCGACCGUGGCGGAGGCGCGUG GCGGCCCGCCUUAUGGCGAAAGCCAUAGGACGCCAUUUACAGUACAGGGUGUGAAGAGCCUUAUGAGC UAGUUGGUAGUCCUCCGGCCCCUGAUGCGGCUAAUCCUAACCACGGAGCGUCCACCAGCAAACCAG CUGGAAGGGCGUCGUAACGGGUAACUCUGUGGCGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUU UGAUCCUAUUUGGCGUCUUAUGGUGACAACUACAGUUGUUAUCAUAAAGCUUUUGGAUUGGCCAU CCGGAAGAGUUAACAGUUGUUCGCUUAUUGCUCUCUCAAACUCCAUUAGGAGAAAGCAAUUUAU CUUCUCAUUUCAUUUACAGCUUGAUAAAAUUUCAAGUGCUCUUUAGAGUUACUCAUCCAACUCC AUCAAAU	
3 4 7	UAAACAGUCAGGGGGUUGUACCCACCUCGGGGCCACGUGGCGUCAGUGCUCUGGUGCUUAGUAC CUUUGCAGCGCUGUUUCCUCCUUAACAAUUAAGAUCCAGCGCCGAGGGGAGUAGUCCGA CUCCGCCUCCACUACGGUGCUCACUGGAUUGCUAUUACGGUAGCUCUUCACGGGUGGCGUGGUGCC CCAACCCCGCAACUUAAGAGCCAGCACAGUACGACCAUAGCGGCUCCGACGCCAGUCGGGCAACGG UCAAGCACUUCUGUUUCCCGGUCUUAAGGAUCGUUACCCGCCAGGAUGCUACAGAAACUUAAGU AAACACCCAAGCUUGAGAGGUGGCGUCAGCCACAACCCAGUGGUGAGCUCUGAGUGAUGGGGCU CGUGACUACCCUCUGGUAACAGAGACAUUGCCCGGUGGACUCGGGUUCAUCAGAAUUGGUGU UCACCUCGACCAAGCGUAAGUGUUAAGAGUCUAUUGUGCCAGCCUGGUUUCUCCGGGCGCGUAA UGCUGCUAAUCCUAACCCGAGCGUGUGCGCAAAUCCAGUGUUGCUACGUCGUAACGCGUAAGUU GGAGGCGGAACAGACUACUUCGUAACCCCGUGUUUCCUCAUUAUUAUUCAUUUAUGGUGACAA UUAACAGAGAUUUGGAAUUGGCGAAACUACCACUGAAUUGGCGUUGUAUUGUCUUAUUGCCUCCA CAAACUUCUGAUUAUCCACCAGUUAUAUCGAUCUCCUGGUUAUUCUAAUUAUUAUUAUUAUUG AACCAUAACA	enterovirusG 5
3 7 5	UUAAACAGCCUGUGGGUUGUUCACCCACAGGGCCACUGGCGCUAGUACACUGGUUAGCCUUA CCUCGGUAUCACGGUACCGUUGUAUACAGUACCUUUGUACGCCUGUUUUAAUCCUCCUCCCAAU GUACUUUAAGAUUUAAGCAACAAGACCAUAGGAGUCCAAACACAGUUGGAUUGCGGUCAG CACUUUCUGUCUCCCGGACCUAGUAGUGAUAGGCGUUAACCCAGCGCGGAAGAUAAACCGUCCGUUA CCCGGCCAGCUACUUCGAGAAGCCUAGUAACAUAAGAUCUGUCUUGGCGUUUCGUCACAGCGGUU CCCCCGCGUAGAUCGGGCGUAGUGGUCUCCGCAUACCCACGGGCGACCGUGGCGGAGGCGCGUG CGGGCCCGCCUUAUGGCGAAAGCCAUAGGACGCCAUUUCAGUGACAGGGUGUGAAGGCCUAUUGAGC UAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACCACGGAGCGUCCACCAGCAAACCAG UGGCGAGGGCGUCUAACGGGUAACUCUGUGCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUU UGAUCCUAUUAUUGGUGCUUAUGGUGACAACGAUAGUUGUUAUCAUAAAGCUUUUGGGUUGGCCAC CUGGAAAAAGAUUAUCAGUUAUUGUAUUGUCCGGCUUACCGCUUACCAUUAUAAACAAUCCUUAU UAUUUGAUUUCCUUUAUCCACAGAGGAAUUAUUGAGUUUAUCUCCAGUUGGCUUAUACCUUAU CAGCAAG	enterovirusG 6
3 7 6	UAUUUCCAGGUGGAUUGGGGAUGCAGGCCUGAGUUGUUUUGCCCCAAGGAUAAUUAUUAUUC CCAGAGCCGAGGGGAGUAGUCCGACUCCGCUCCACUACGGUGCUCAGUAGAUUGCUAUUGCGGUAG CUUCUACGGGUGGCGUGUAUUCUACCCUGAAACUUAAGAAGUUUGUACACACCGACCAUAGCG GUCUGGCAGCCAGCCAGGCAACGGUCAAGCACUUCUGUUUCCCGGUCGCAAGGAUCGUUACCCGC CCGAAUUGCUACGAAAGCCUAGUAAGUACCCAGUUGUGCGAGGUGGCUUACGCCACAUAUCCCA GUGGUAGCUCUGAGUGAUGGGGCGUGUAUUAUCCUUCUGGUGACAGAGACAUUGCCCGCGCGUGC ACUUGGGUUCACCAGAAUUGGUGUACCCUAAUCAAAGCGCAAGUACUGAAGAGCCAAUUGUGCCAA UUUGGUCUCCUCCGGGCGGUGAAUGCUGCUAAUCCUAACUCCGAGCGUGUGCGCAUUAUUGAGUG UUGCUACGUCUAACCGUAAGUUGGAGGCGGAACAGACUACUUCGGUACCCCGUGUUUCCUUAU UUAUUAUUAUUAUUAUUGGUGACAAUUGUUGAGAUUUGUGAGCUGACAAAUAUACCAUUGAAUUG GCAUUGUACUGUCUCAUUGCUUUUACAAAACCUACAGAUUCCACCAGUUAUUGAUUUAUUGAG UUCAUUUGAAUUAUACAGUUAACACUGAGACAAGUAAA	enterovirusG 7
3 7 7	UACCUUUGUACGCCUGUUUAUACUCCUUCUCCCUACAUGUAACCUAGAAGUAUAAUAAACAAAG ACCAUAGGAGUCCAAACUCCAGUUGGAUCGCGGUAAGUACUUCUGUUUCCCGGACCUAGUAGUG AUAGGCGUUAACCCACGGCCGAAGAUAGAACCUGUCCUUAUCCGGCCAGCUACUUCGAGAAGCCUAGU AAACUAACAAGAUUCUGUCUUGGCGUUUCGCUACGCGGUUCCCCCGCGUAGAUCCGGCUGAUGGGUC UCCGCAUACCCACGGGCGACCGUGGCGGAGGCCGCGUGGCGGCCGCCUUAUGGCGAAAGCCAUAGG ACGCCAUUUCAGUGACAGGGGUGUGAAGAGCCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCGUAAUG CGGCUAAUCCUAACCCAGGAGCGUUCACCAGCAAACAGCUGGGAGGGCGUCGUAUUGGGCAACUCU GUGGCGGAACCGACUACUUGGGUGUCGUGUUUCCUUUGAUCCUAUUAUUGGUGCUUAUUGGUGAC AACGAUAGUUGUUAUCAUAAAGCUUUUGGUGGCCACCUGGAAAAAGUUAUCAGUUAUUGAUUAU GUUCGGCUUCCACGGCUUACCAUAAUAAACAGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU AGAUUUCUGAAUUAUUAUCCAGUUGCUCUUAACACUACAGCAAA	enterovirusG 8
3 7 8	GCCUGUUUUUAUUCUCCUCCUCCAAACGUAACAUUAGAAGUAUAAACAAACAAAGACCAUAGGAG UCCAACAUCCAGUUGGAUUGCGGUAAGCACUUCUGUCUCCCGGACCUAGUAGUGAUGGCGUAC CCACGGCCGAAGAUAACCCGUCCGUUAUCCGGCCAGCUACUUCGAGAAGCCUAGUAACAUAAGA UCUGUCUUGGCGUUUCGUCACGCGGUUCCCCCGCGUAGAUCCGGGUGAUGGGGUCUCCGCAUACCC CACGGGCGACCGUGGCGGAGGCGCGUGGCGGCCUGCCUAGGCGAAAGCCAUAGGACGCCAUUUA GUGACAGGGUGUGAAGAGCCUAUUGAGCUAGUUGGUAUUGGUGUCCUCCGGCCCCUGAAUGCGGCUAAUCCU AACCACGGAGCAUCCACCAGCAAACAGCUGGAAGGGUGUCGUAACGGGUAACUCUGUGGCGGAACC GACUACUUGGGUGUCCGUGUUUCCUUUGAUCCUAUUAUUGGUGUCUUAUGGUGACAAACGAUAGUU GUUAUCAUAAAGCUUUUGGUGGCGCCACUGGAAAAAGUUAUCAGUGUUUGAUUAUUGUUGGCGUUUC ACGCCUACCAUAAACAAAGCUUUAUACUUGCUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU UUUUACCUAACAUUUUGGUCACCCUACAGUAAA	enterovirusG 9

3 7 9	GGGUUGUCCCCACCCGCGGGGCCACUGGGCGCUAGUACACUGGUACGCCAGUACCUUGGUUUUACG GUACCAUUGUACGUCAGUACCUUUGUACGCCUGUUUUUAUACCUUCCCCCACAUGUAACAUAAGAA GUCAAGUAAACAAAGACCAUAGGAGCCCAACAACCAAGUUGGGUUGCGGUCAGACAUUCUGUCUCC CCGGACCUAGUAGUAGUAGGCGUACCCACGGCCGAAGAUGAACCCGUCGUUACCGGCCAGCUAC UUCGAGAAAGCCUAGUAACAUAAGAUAUCUGUCUUGGCGUUUCGCUACGCGCGUCCCCCGCGUAGA UCGGGCUAGUUGGUCUCCGCAUACCCACGGGCGACCGUGGCGAGGCCGCGUGGCGGCCGCCCUAU GGCGAAAGCCAUAGGACGCCAUUUCAGUACAGGGUGUGAAGAGCCUAUUGAGCUAGUUGGUAGUCC UCCGGCCCCUGAAUGCGGCUAAUCCUACCCAGGAGCAUCCACCAGCAAAACAGCUGGGAGGGUGUC GUAAACGGGUAAUCUCUGUGCGGAACCGACUACUUGGGUGUCCGUGUUUCCUUUGAUCCUAUUUG GGUGCUUUGUGGACAACGAUAGUUGUUAUCAUAAAGCUUUUGGGUUGGCCACCGUAAAAAAUUG UCAGUGUUUGAUUUCGUCGGCUUUCACGCCUACAGUCAAAACAGCUCUUUAUCCUUGUUGGAU UCUACUCAUAAAGGACAUAACUCAGUUUCACAUAAUUCUGUUUGCUUUCAUUGUCAGCAAA	enterovirusG 10
3 8 0	UUAAAACAGCCUGUGGGUUGUCCCCACCCGACGGGCCACUGGGCGCUAGUACACUGGUUUGCCAGU ACAUAUAGUACGCUAGUACUCUUGUAGGCCAGUACCUUUGUACGCCUGUUUUUAUACUCCCCUCCCCAU UUGUAAACAUUAGAAUUAAGUAAUAAGACCAUAGGAGUCCAAACUCCAGUUGGAUUAACGGUUA GCACUUCUGUUUCCCCGGACCUAGUAGUAGGCGUACCCACGGCCGAAGAUGAACCCGUCGGU AUCGGGCCAGUACUUCGAGAAGCCUAGUAAACAUCAAGAUUCUGUUGGCGUUCGCUACGCGCGU UCCCCCGCGUAGUACGGGCGUAGUGGUCUCCGCAUACCCACGGGUGACCGUGGCGGAGGCCGCGU GGCGGCCCGCCUAGGCGAAAGCCAUAGGACGCCAUUUCAGUGACAGGGUGUGAAGAGCCUAUUGAG CUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACACGAGCGUCCACCAACAUCCA GUUGGCGAGGCGGUGUAACGGGUAAUCUUGGGCGGAACCGACUACUUGGGUGUCCGUGUUUCCU UUGAUCCUAUAUUGGCGUCUUAUGGUGACAACGAUAAGUUGUUAUCAUAAAGCUUUUGGGUUGGCCA CCUGGAAAAAGUUGUCAGUGUUUGAUUAUUGUUCGGCUUUCACGCCUACCAUAAAGCAAGCUUUAUA UUUGUAGUUUCAUUUUAUCCAUUAAAGGAUAUACCGAAUUCACUACAGCUUUUGCUACAUAUA UCAGCAAA	enterovirusG 17
3 8 1	UUUAAAACCAACCCGGGGUUGUACCCACCCUGGGGCCACUGGCGGCUAGUACCUUGGUUUUGCUC AGAACCUUUGCAGCGCUGUUUCCCCUCCCCUAAUAAUUAAGCUUGCCGGGACUGAGGGGAGUAC UCCGACUUCGCUCCCUACGGUGCCCCAGUAAACUGGUUACCGGGUACCUUUUACGGAGUACCCGG CAGCCCCCCCCCGCAACUUAAGAAGUUUGAACACAACCGACCAUAGAGGCCCGGCAUCCAGUCGAG CAACGGUCAAGCACUUCUGUUUCCCCGGUCCGCAAGGAUCGUUACCCGCCCGACUACUACGGGAAG CCUAGUAUUCUAGCCAAACUUGGAAAAAGUUGCGUCACGCCACAACCCAGUGGUAGCUCUGAGUGAU GGGGCUCGCGACAACCCCUUGGUAACAAGGACGCUUGCCGCGCGUGCACUCGGGUUACCUUCCG GUGUUCACCCUGACCAAGUGUAACUAGAGAAGCGCUACUGUGCUAGUUGGGUUUUUCCUCCGGGGC GUGAAUUGCUGUAUCCCAACUCCGAGCGUGUGCGACAUAUCCAGUGUUGCUACGCGUAACGCGC AAGUUGGAGGCGGAACAGACUACUUCGGUACCCCGUGUUCUUUAUUGUUUAUAAUUAUUGG GACAAUUGUUUGAGAUUUGUAGUUUAGCGAAUUAACCUUGAAUUGGCAUUAUAUUAUACUUGGUGC AUUUCACAAAACCUCAaAGAUUCCCGCAUCCUAUAUAGAUUCUUCGUCUAUUUGAACUACAGCAC AAAUUAAUCAAUAAA	enterovirusG 20
3 8 2	UUAAAACAGCCUGGGGUGUCCCCACCCACGGGCCACUGGCGCAAGUACUCUGCUUCGUAGA AGCCUUGUAGCCGCUUUUUCCCCUCCUAAUAAUUAAGAUUACCAUCCGAGGGGAGUAGUCCG ACUCCGCUCCGAUACGUCGCUCCAGUGCGCUAGUACGCUUAGUACUUCUACCGGGGAUAGUGGUAC CCUCCCUAGUAAUUAAGACUUGAAAAACCGACCAUAGGCACCUCGCAUCCAGCGGGGUAAG GUCAGGACUUCUGUUUCCCCGGGUCGAGUAGCGAUAGACUGUGCCACGGGUCGAGGUGAAACAA CCGUUAUCCGACUUGUACUUCGGGAAGCCUAGUACCAACAAAGAUUAGCUUGGGGUUUCGCUAG CAGGACCCUGUGUAGUACAGGCUAGGGAUACCGCGUUCUACAGGCGACUUGGCGGUGGUGC CGUGGACGCCUGCCGAGUGGGCAACCAUCCGAGCGCAAGCAUAGACAGGUGUGAAGAGCCUACU GAGCUACAAAGUAUUCUCCGGCCCCUGAAUGCGGCUAAUCCCAACACGGAGCAUUGCUACCAAA CCAGGUAGUGAAUUGCUAACGGGUACUUCUGUGGCGGAACCGACUACUUGGGUGUCCGUGUUUC CUUUUAUUUAUACUUCUGUGUAGGUGACAACUAUAGUGCUAUCUCGAUUUGCAUUAUUAUUGU AGAUUGACACUUAUUAUAGUUGCAUUAUACCUUUGAGUGAGUUUCCUUGAACAGAUUAUUAU UACUUAUCCUGUUUAUUAUACACAGCAGAAUACUUGCAAAG	enterovirusI 2;dromedaryc amelenterovi rus2
3 8 3	AAAAUAGCCUCAGGGUGUCCCCACCCUGAGGGCCACUGGCGCUAGUACUCUGGUUACCGUACC UUUGUAGCGCCUUAUUUAUACCCUCCCCGUAAACUUAAGAGCUUUAACUAUAAGUUAACUAGCAGG GCUACCAACCAAGGUAAGUUAUCGAACAAGCAUUCUGUUACCCCGGUCCUAGUACCUAGAGAGCUGU ACCCACGGCGAAGGGGAAAAAGAUUGUACCCGGCUAUCUACUUCGAGAAGCCUAGUAAAGCUCU GAAACUCUUGACCGUUGCGUUCAGCACAGUACCCCGUGUAGCUUUGGCUAGAGUUCUGGCAU CCCCACUGGCGACGUGGCCAGGCGUGCGUUGGCGGCCAACCGAUCCGGGCAACCGGGUCGAGCUCU CAUUGUGGACAUUGGUGAAGAGCCUAUUGAGCUAGGUAGUUCUCCGGCCCCUGAAUGCGGCU AAUCCCAACUCCGGAGCAAGUGCGCAAAUCCAGUGUGUGGCUUUGCGUAAUGCGUAAGUCUGGAGC GGAACCGACUACUUGGGUGUCCGUGUUCUGUUUCCUUAUUGGCGUCUUAUGGUGACAUAUCU AGUGUUGUUAACUAUAGCGAUUGGAUUGGCCAUCCGGUGAAUAGCUUUGUGACAUAUCCCUACGUU AUUUUUAUAACUCAACUUCACGAGAGUCAUUGAGUUUGACUAAU	enterovirusJ 103
3 8 4	UUAAAAGAGCCUGGGGUGAGUCCAGCCCCAGGACCCACGUGGUGCCAGCACUACUUAUAGUACGAU UAUUUAGUGCGCCUUCUUUAUAUCCCCUCCCCACCCUCCCCUUUCCUUGAAUUCUUUCGUAACAA GUGCCUACGACUAGCGUUUGACGGUGUACUGUAACCAUACGGCAACCGCCUGUAAAGCAGGAAC ACUUCUGUUUCCCCGGUGGUAUUAAGGCGUUCUCCACGGCCGAAAAUGAAGCAUCGUUAUCCCA CUGGUACUGCGGAAAGCUAGUAACACCAUAGAUUAAGAUUAGUAGUUCUUCGCUACAGCAAUUGCCCGA UUGUAGAUCCGGGGAUGGGGCUCCUGUUCUUCACUGGCGACAGUGGAGGAGCCCGCGUGGCCGCC CGCCUACCCCGCAUGGGGUGAGGAGUCCAGAUAAUGACAGGGUGUGAAGACUCAAGUGCGCUACCAAG	enterovirusK 1

	UGCAUCCUCCAGGCCUCUGAAUUGCGGCUAUUCUUAACCCCGGAGCACCCGUCACAGUCCAGUAAGU AGGGUGUCUGUACCGCGUAAGUCUGGGAUGGAACCGACUACUUGAGUGACCUGUGUUUCUUUUUAUUU UUUUUAGGGUUGUUUAUGGGUGACAAGAUUGGCUCUAUCCGUGUAUAGAGUUUACAAGGUGAAAGAGG UUUUUACAGACUUUUUGUUAUCAAUCUUGUCCUAUCCCGCAGGGGAAUUGAUGAUAGAAGUCC UUUUAGAGGUGAUUCAAUUGAUAGCUCAAACAACUGAGAAUUCUUAUUUUUAUUAUACAAGGUCUGAA UUUGAAUAAAGAUCCAGUGAAUUUUUACCCCCACCUAUUAGCGACCAAAUUCACAAGCA	
3 8 5	UUAAAACUGGGUGGGUUUUUUCCACUCUACCCGGAUAGGUGUGUACUUGUUUUUUCGGGUA ACUUUGUACGCCAUUUUCCUCCUCCUCCCAUCCUUUACGUAAUUAACAACUUUAAAACAAGACC AAUAGUAGGCAACUCUCCAGGUUGUCUAAGGUCACAGCACUUCUGUUUCCCGGUGAGCUUGAUUAG CUCUAAACAGGGCAAAAACAACUGAUUUCGUUACCCGCAAGUGCCUACAGAGCUUAGUAGGAUUC UGAAAGAUUUUUGGUGUGUCGUCACGUCUACCCAGCAGUAGACCUGGACAGUAGGCGUGGACAU UCCCCACUGGUUACAGUGUGGCCAGCCUGCGUGGGUGCCUGCACACUCUUAUGAGGUGUGAAGCCAAA GAUUGGACAGGGUGUGAAGAGGCCGCGUGUGCUCACUUUAGUCCUCCGGCCCGGUAAGUGCGGCUAAC CUUAAAACUCGACCCCAUGGCUCUAAUAAACCAUAGAGCUUAUGGUCGUAUAGGACAAUUGCGGGAUGGG ACCGACUACUUUGGGUGUCCGUGUUACUUUUCCUUUAUCAAUUGCUUAUGGUGACAAUUAUUAUAC AUAGAUAUAUUAUGGCAUC	rhinovirusA1 B
3 8 6	UUAAAACUGGAUCCAGGUUGUUCACCUGGAUUUCCACAGGGUGUGUACUCUGUUUUUACGGUA ACUUUGUACGCCAGGUUUUAUUCUCCUUCUCCCAUGUAACUUAAGAUUUUUCAACAAGACCAAUAGC CGGUAAUCAGCCAGAUUACUGAAGGUCACAGCACUUCUGUUUCCCGGUCAAUUGUGAUUAGCUCCAA CAGGGCAAAAACAACUGCGAUUCGUUAAACCGCAAGCGCCUACGCAAGCUUAGUAGCAUUCUUGAAA UCGUUUUGGCGUGGUCAGUCCGCUAAUUCUCCUGGACCCGCGAAGGCGAGUAGGACUAGAAUACCCCAU GGCGACAGUGUUCUAGCCUGCGUGGCGUCCUGCACACCCUAUGGGUGUGAAGCCAAACAUGGACAA GGUGUGAAGAGCCCCGUGUGUCGCUUUGAGUCCUCCGGCCCCUGAAUUGGCGUAAUCCUUAACCCUG CAGCUUAGACGACGUAAACCAUUGAGUUAUGAGCUUUAUGCAAUUGCGGGAUAGGACCAACUACU UUUGGUGUCCGUGUUUACCUUUUCCUUUAUUAUUGCUUAUGGUGACAAUUAUACAUAUUAUUAU UGGCACC	rhinovirusA2
3 8 7	UUAAAACUGGGAGUGGGUUGUCCCAUCACUCCACCCACGUGGUGUGUACUCUGUUUAUCCGGUAA CUUUGUACGCCAGGUUUUUUCCUCCUUCUCCUUAUUUAGUAACUUAAGAACUUAAGAACAAUUGGACCA UAGGCAUUAUACUCCAGAUUAUUGAAGGUCACAGCACUUCUGUUUCCCGGUCAAUAAAGAUUAGCU UUACCCAGGGCAAAAACUUUUGAGAUCGUUACCCGCACAUCAACUACGCAAGGCCUAGUAUACCCU UGAUGUACUAAAGACUGGUCGUGUUCUGCAUCCCCAGCAGUAGACUGGCGAGGCGUGGACGUAC CCACUGGGCAGAGUGGUGCCAGCCUGCGUGGCGUGCCUGCUCGCCCAUUAUGGGCGAGAAGCCUUUAU UUGACAAGGUGUGAAGAGCCGCGUGUGCUUUGGUGUGAGUCCUCCGGCCCCUGAAUUGGUGCUAACCU AACCCUGCAGCCAUUGCUCAUAAUCCAUGAGUUAUGGUGCUAAUGCGCAAGUGCGGGAUGGGACC AACUCUUGGGUGUCCGUGUUUUAUUUUUUAUUAUGCUUUAUGUGUGUUAUUAUGGUGACAAUUUAUUA AUUAAGUUGUCAUC	rhinovirusA7
3 8 8	UUAAAACUGGGUUCAGGUUGUUCUCCAUUCUGAACUUCUCCACGUGGAGUAGUACACUUAUUUACGGUA AUUUUGUACGCCAGGUUUUAUUCUCCUCCUCCCAUUGUAACUUAAGUAAGUACACAAAGCAACACAGG CGGUGGUUAUCCAAACCAUCAAAGGUCACAGCACUUCUGUUUCCCGGUCAAUUGUGAUUAGCUCCAA CAGGGCAAAAACAACUGCGAUUCGUUUAUCCGCAAGAUCCUACGCAAGCUUAGUAAACAUUUGGAAG AUUUUAGGCGUGGUCGUUCCACUUAUACCCCAUAGUAGACCUGGCAGUAGGCGUAGAAUACCCCAU GGUGACAGUCUUCUAGCCUGCGUGGCGUGCCUGCGCACCCUUAUGGCGCAAGCAUUGGACAA GGUGUGAAGAGCCCCGUGUGUCUACUUGAGUCCUCCGGCCCCUGAAUUGGCGUAAACCUUAACCCCG CAGCCAGUGACACAAUCCAGUGUGUUUUUGGUCGUAUAGUUAUUGGUGGAUGGGACCAACUACU UUUGGUGUCCGUGUUUACCUUUUCCUUUAUUAUGCUUAUUGCUAUAUACUUAUUAUUAUUAUGAC ACC	rhinovirusA8
3 9 0	UUAAAACUGGAUCUGGGUUGUUCUCCACUCAGAUUCCUCCACGUGGAGUAGUACACUUAUUUACGGUA AUCUUGUACGCCAGGUUUUAUUAUUCUCCUUAUCCCAAGUAACUUAAGUAACAAAGACCAUAGC GAGAUGAUACUACGAGUACAUAGGUCACAGCACUUCUGUUUCCCGGUCAAAGUUGAUUGCUUACU ACAGGGCAAAAACAAGCUGAGAUUGUUAUCCGCAAAAGCGCCUACGCAAGCCUAGUAACAUUCUGGAA GUUGCUUGGUGUGGUCUCCGCCCAUUAUCCUAGGUAAGCUGGCGAGUAGGCGUAGAAUUCUCCCA UGGUGACAGUGUUCUAGCCUGGUGGCGUGCCUGCCGACCCUUAUGGUGAAGCAAGUAAUGGACA GGGUGUGAAGAGCCCCGUGUGUCGCUUUGAGUCCUCCGGCCCCUGAAUUGGCGUAACCUUAACCCU GCAGCUAGGGCACACAUAUCCAGUGUGUAUCUAGUCGUAAUAGAGCAAUUGCGGGAUGGGACCAACUAC UUUGGUGUGUCCGUGUUUCUUGUUUUUCCUUUAUUAUUGCUUAUUGGUGACAAUUAUAGUAAUUAUUA UUGGCAUC	rhinovirusA9
3 0	UUAAAACUGGAUCCAGAUUGUUCUCCAUUCUGGAUUUCCUCCACGUGGAGUUGUACACUUAUUUACGGUA AUUUUGUACGCCAGGUUUUAUUAUACCCUCCUCCUUAUAGAAUUGUUUAACAAGGACCAUAGGCA UUUUAACCCAGGUUGUUAUAGGUCUACAGCACUUCUGUUUCCCGGUCAAAGUUGAUUAGUCCAA GGGCAAAAACAACUGAGAUUGUUAUCCCGCAAGGUGCCUACGCAAGCCUAGUAACCCUUAAGCAU GUGUGUUGUGUUCUGGCGUACCCACAGUAGACCUGGCAGUAGGCGUAGAUUGUCCUCCACUGG CGCAGUGUUCUAGCCUGGCGUGCCUGCGCACUCUCUGAGUGGCGAAGCCUUAUUAUUAUUAUUA UGUGAAGAGCCCCGUGUGCUCACCUUAGUUCUCCGGCCCCUGAAUUGGCGUAACCUUAACCCUGCA GCUAGUAUAGUACAUAUCCAGCAUAGCAGCUAGUCGUAAACGAGCAAUUGCGGGACGGGACCAACUACU GGUGUCCCGUGUUUACCUUUUCCUUUAUUAUUGCUUAUUGGUGACAAUUAUUAUAGAUUUUAUUAUUG UGACC	rhinovirusA1 0
3 1	UUCAAACUGGAUCCAGGUUGUUCUCCACCUGGAUUCUCCACGUGGAGUUGUACUCUUAUUUUCGGUA AUUUUGUACGCCAGGUUUUAUUCUCCUUCUCCUUAUUAAGAAUGUAGUACUACCCGACCAAAUAG UGGCAUUUAUACGAAGUGCAAGAGCAAGCACCUUCUUGUUUCCCGGUCAAAGUUGAUUAGUCCU GGGCAAAAACAACUGAGAUUGUUAUCCCGCAAGGUGCCUACGCAAGCCUAGUAACCCUUAAGCAU GUGUGUUGUGUUCUGGCGUACCCACAGUAGACCUGGCAGUAGGCGUAGAUUGUCCUCCACUGG CGCAGUGUUCUAGCCUGGCGUGCCUGCGCACUCUCUGAGUGGCGAAGCCUUAUUAUUAUUAUUA UGUGAAGAGCCCCGUGUGCUCACCUUAGUUCUCCGGCCCCUGAAUUGGCGUAACCUUAACCCUGCA GCUAGUAUAGUACAUAUCCAGCAUAGCAGCUAGUCGUAAACGAGCAAUUGCGGGACGGGACCAACUACU GGUGUCCCGUGUUUACCUUUUCCUUUAUUAUUGCUUAUUGGUGACAAUUAUUAUAGAUUUUAUUAUUG UGACC	rhinovirusA1 1

	AAACAGGGGCAAAAACAUUCGCAUUCGUUAUCCGCAAGCAACUACGCAAGCCUAGUAUACAUUUUGA GAGUAUGUGGUUGGUCGCUCCACUAUAAACCCAAUAGUAGACCUGGCAGAUAGGGCUAGAAGUCCCC CACUGGUAAACAGUGUUCUAGCCUGCGUGGCGUGCCUGCGUGCCUCUGGUAGGCACGAAGCCAUUAUU UGACAAGGUGUGAAGAGCCCCGUGUGCUUGCUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUA ACCCUGCAGCUAGUGUGCAUAAUCCAAUGCAUGGCUAGUCGUAAUAGGCAAUUGCGGGAUUGGACCA ACUACUUGGGUGUCCGUGUUCUACUUAUUUCCUUAUUUGCUUAGGUGACAAUAUAUAUAUAC AUUAUAUUGGCAUC	
3 9 2	UUCAAAACUGGGUUGUGGUAUCCACCCACACUGCUCACUGAGCGUAGUGCUCUAUAUCCGGUA UGCUUGCAGCCAGUUUACUCCCCUCCCCUUAUUGUAACUUAAGAAGCAAGCACACCUGAUCAAUA GGAGAUAAUUGCCAGGUUAUCUAAAGAUCAAGCACUUCUGUUUCCCCGGCAAGUGUGAAUAUGCUUU ACCCAAGGGCAAAAGCUGAUUCUUGCGUUAUCCGCAAGGUGCCUACGAGAAGCCUAGUAUUGCUCAG AUGUGGUUUGACUGGUGCGUCCGUGUCCCCACAGUAGACCUGGCAGAUAGGGCUGGACAUUCCCC ACCAGCGAUUGGUGGCCAGCCUGCGUGGCGUGCCUGCACACCCUAGGGUGUGAAGCCAAAUUACUG ACAGGGUGUGAAGAGCCUAUUGUGCUCACCUUGAUUCCUCCGGCCCCUGAAUGUGGCUAAUCCUAAAC CCCGCUCUAUUGCACACAGCAACCAAGUGUGUAAGUAGUCGUAAUAGGCAAUAGUGGGAUGGAAACCA UACUUGGGUGUCCGUGUUCUUGUUUUUUAUUAUUGUCUUAUGGUCACAAUAUAUAUAACGUA UAUUGGAUC	rhinovirusA1 2
3 9 3	UUAAAACUGGGUUCGGGUUGUCCACCCAAACCACCCACGUGGUGUAGUACACUGUUAUCCCGUA ACUUGUACGCCAGUUUUAUUCUCCCCUCCCCUUAUUGUAACUUAAGAAGUUAUGCACAUAUGACCAUA GUGGGCAUAUACAGAUUGUCUAGCGUGCGAGCACUCCUGUUUCCCCGGUUAAGUUGAUUAGCUUCU AACAGGGCAAAAACAAUUGAAUCGUUACCCGCAAGGUGCCUACACAGAACUUAAGUACCAUUCUGGA AAAUUCCUUGGUGGUGCGUCUAGUUAUACUACCCAGUAAUAGACUUGGCAGAUAGGGGCUAGAAGUCCC CACUGGCGACAGUGUUCUAGCCUGCGUGGCGUGCCUGCACACCCUUCUUGGGUGUGAAGCCUAGAAU GGACAGGGUGUGAAGAGUCGCGUGUGCUCAUUCUUGAGUCCUCCGGCCCCUGAAUGCGGCUAACCCUA ACCCUGCAGCCAUUGCCCCACAACCCAGUGGGUUAUGGUCGUAAUAGGCAAUUGCGGGAUUGGACCG ACUACUUGGGUGUCCGUGUUCACUUAUUUAUUAUUAUUGCUUAUGGUGACAAUAUAUAAGUAACA UAUAUUGACACC	rhinovirusA1 3
3 9 4	UUAAAACUGGGUUCAGGUUGAUCCACCCUGAACCUCUCCAAAGCGGAGUAGUACUCUGUUAUUCGGUA ACUUGUACGCCAGUUUUAUUCUCCCCUCCCCUUAUUGUAACUUAAGAAGUAGUACACAGACCAAUAGC CGGCAUAUACCCAGAUUGUCAUUGGUCACAGCACUUCUGUUUCCCCGGUACACGUUGAUUAGCUCCAA CAGGGCAAAAACAACUGAGAUUCGUUAUCCGCAAGUGCCUACGCAAGCCUAGUUAUGCUUUCGAG UCAUUGGUUGGUGCGUCCGCGUGUAACCCACAGUAGACCUGGCAGAUAGGGCUAGAAACUCCCCACU GGCGACAGUGUUCUAGCCUGCGUGGCGUGCCUGCACACCCUUAUGGGUGUGAAGCCAAAUUACGGACAA GGUGUGAAGAGCCCCGUGUGCUCACUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACCCUG CAGCUAGAGUGCACAAUCCAGUGUAUUUCUAGUCGUAAUAGGCAAUUGCGGGAUUGGACCAACUACU UUGGGUGUCCGUGUUCACUUCUUAUUUAUUCUUAUUAUUGCUUAUGGUGACAAUAUAUUGAAUUA UAUAUUGACACC	rhinovirusA1 5
3 5	UUAAAACUGGAUCUGGGUUGUCCACCCAGAUACCUACAUUGGUGUUGUACACUAUUAUACGGUA AUCUUGUACGCCAGUUUUAUACUCCUUCUCCCAAUUGUAACUUAAGAAGCAAGCAAAACUGAUCAAUA GGGGACAUAUACCCAGAUUGUCAAAAGAUCAAGCACUUCUGUUUCCCCGGUCAAAGUUGAUUAGCUCC ACAAGGGCAAAAACAACUGAGAUUCGUUAUCCGCAAGAUCCUACGCAAGCCUAGUUAUAUAUUGAA AGAUACUUGGUUGGUCGUCACAGCUGUUAACCCAAAGUAGACCUGGUAGAUAGGGCUAGAGAUUCCC CUCGGGACAGGUUUCUAGCCUGCGUGGCGUGCCUGCACACCCUUAUGGGUGUGAAGCCAAAGUAUUGG ACAAGGUGUGAAGAGCCGCGUGUGCUCUUCUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAA CCUGCAGCCAGUGCACACAAUCCAGUGUGUAGCUGGUGUAUAGGCAAUUGCGGGAUUGGACCAAC UACUUGGGUGUCCGUGUUCACUUAUUUUUCCUUAUUAUUGCUUAUGGUGACAAUAUAUUAUUGUAUA AUAGCAUAUAUAUUGUUAUC	rhinovirusA1 6
3 6	UUAAAACUGAAUCCAGGUUGUCCACCCUGGAUUUCCCAUGUGGAGUAGUACACUGUUAUUCGGUA ACUUGUACGCCAGUUUUAUAAACCCUUCUCCUUAUUGUAACUUAAGAAGUAAGCACAAAGACCAAUA GGUGGUGGCUAUCCAGUCACUAAAGGUCAAGCACUUCUGUUUCCCCGGUACACGUUGAUUAGCUCC ACCAGGGCAAAAACAACCGAGAUUCGUUAUCCGCAAGUGCCUACGCAAGCCUAGUAGCACUUGAA UAGCAUUGGUUGGUGCGUCCACUAUCCCCAUAGUAGACCUGGCAGAUAGGGCUAGAAUCCCCCA CUGGCGACAGUGUUCUAGCCUGCGUGGCGUGCCUGCGUGUCCUACCGGACACGAAGCCAUUUGUUGA CAAGGUGUGAAGAGCCCCGUGUGCUCUUAUUGAGUCCUCCGGCCCCGAAUUGGCUAACCUUAAAC CUGCAGCUGGGGCAUACAAUCCAGUGUGAUCUAGUCGUAAUAGGCAAUUGCGGGAUUGGACCAACU ACUUGGGUGUCCGUGUUCACUUAUUUUUAUUAUUGCUUAUGGUGACAAUAUAUAUUAUUGUAUA UUGGCACC	rhinovirusA1 8
3 9 7	UUAAAACUGGAUCUGGGUUGUCCACCCAGAUACUCCAGUGAUGUUGUGCUCUAUUAUUCGGUA AUUUUGCAGCCAGUUUUUAUUCUCCCCUCCCCUUAUUGUAACUUAAGAAGCUUAAACAAUAGACCAAUA GGAAGCAAAAACACAGUUGCGAUAGGUCAAGCACUUCUGUUUCCCCGGUCAAUGUUGAUUAGCUCC ACGGGGCAAAAACAGCAGCGAUUCGUUAUCCGCAAGGUGCCUACGCAAGCCUAGUAGCGCUCUUGAA AUUGUUUGGUUGGUGCGUCAAGUACAACCCAGUACUAGACCUGGCAGAUAGGGCUAGAAACCCCC ACUGGCGACAGUGUUCUAGCCUGCGUGGCGUGCCUGCACACCCAUUUGGGUGUGAAGCCAAACACGG ACAGGUGUGAAGAGCCCCGUGUGCUCUUAUUGAGUCCUCCGGCCCCGAAUUGGCUAACCUUAAAC CCUGCAGCCAUUGCUCACAAUCCAGUGAGUUGGUGUGCUAAUAGGCAAUUGCGGGAUUGGACCGAC UACUUGGGUGUCCGUGUUCACUUAUUUACCUGUUUAUUGCUUAUGGUGACAAUAUAUAUAGUAGUAU AUUAUGACACC	rhinovirusA1 9

3	UUA AAAACUGGAUCCAGGUUGUCUCCACCUGGAUCUCCUACAUGGAGUUGUACUCUAUUAUUCGGUA	rhinovirusA2 0
9	AUUUUGUACGCCAGUUUUUAUCUCCCCUACCCCUUACGUAAACUUAGAAGUUUUAACACACACCGACC	
8	AAUAGGUGGUUUUACCCAAAUAACCAAAGGUCAGCACUUCUGUUACCCCGGUCAAAUAUGAUUG	
	CUCCACCAGGGCAAAAACAAUUGCGAUCGUUAUCCGCAAGCAACUACGCAAGCCUAGUACCAUUU	
	UUGAUGAUGUGUGGUUGGUCGUUCCACUAUGAACCCGAUAGUAGACCUGGCAGAUAGGGCUAGAAAU	
	UCCCCACUGGCAACAGUGUUCUAGCCUGCGUGGCUGCCUGCGUACCCCUUGUGGGUACGAAGCCUAU	rhinovirusA2 1
	CAUUUGACAAGGUGUGAAGAGCCCCGUGUGCUUGCUUUAGAGUCCUCCGGCCCCUGAAUGUGGCUAAC	
	CUUAACCCUGCAGCUAGUGUGUACAACUCCAGUACAUAAGCUAGUCGUAAUGAGCAAUUGCGGGAUGGG	
	ACCAACUACUUUGGUGUCCGUGUUUCUGUUUUUCUUAUUAUUGCUUAUGGUGACAAUAUAUAU	
	AACACAUUAUUGGCACC	
3	UUA AAAACUGGUGCCAGGUUGUCCACCUGGAUCUCCUUAUUGGAGUUGUACUCUAUUAUUCGGUA	rhinovirusA2 2
9	AUUUUGUACGCCAGUUUUUAUCUCCCCUACCCCUUAGCAACUUAGAAGUUUUAACAAUAGACCAUAU	
9	GGCGGUAGUUAGCCAAACUACAGAGGUCAGCACUUCUGUUUCCCGGUCAAAGUUGAUUAGCUCC	
	AACAGGGCAAAAACAAUCGAGAUUGUUUACCGCAAGUGCCUACGCAAGCCUAGUAAACACCUUUGA	
	AGAUAUUGGUGUGGUCGUUCCGCUAUUUUCCCAUAGUAGACCUGGCAGUAGGCUAGAAAUUCGCCA	
	CUGGCGACAGUGUUCUAGCCUGCGUGGCUGCCUGCGCACCCCUUGGGUGCGAAGCCAUACAUGGAC	rhinovirusA2 3
	AAGGUGUGAAGAGCCCCGUGUGCUCAUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAGCC	
	UGCAGCUAGUGCAUGUAAUCCAACAUGUUGCUAGUCGUAAUGAGCAAUUGCGGGACGGGACCAACUA	
	CUUUGGGUGUCCGUGUUUACUUUUUCCUUUAUAUUGCUUAUGGUGACAAUAUAUAUAAACAUAU	
	AUAUUGACACC	
4	UUA AAAACUGGAUCUGGGUUGUCCACCACCCAGAUACCCAAUGGUGUAGUGCUUGUUUUUUCGGUA	rhinovirusA2 4
0	ACCUUGUACGCCAGUUUUUAUCUCCCCUACCCCUUAGCAACUUAGAAGCUUUUAAACACAGACCAUAUG	
0	GGAGCGACAAACAGUCGUGAAAGGUCAGCACUUCUGUUUCCCGGUCAAUGUUGAUUAGCUCCUC	
	CAGGGCAAAAACAAACAGCAUCGUUAACCCGCAAGAUUGCCUACGCAAGCCUAGUAAAUUUCGAG	
	AUACUUGGUUGGUCGCUCAUGUGCAACCCAGCAUAGACCUGGCAGUAGGCUAGAAAUUCGCCA	
	CUGGCGACAGUGUUCUAGCCUGCGUGGCUGCCUGCACACCCUAAACGGGUGUGAAGCCAGUAUUGGA	rhinovirusA2 5
	CAAGGUGUGAAGAGCCCCGUGUGCUCAUCUUGAGUCCUCCGGCCCCUGAAUGCGGCUAACCUUAAcC	
	CUGCAGCCAUUGCUCAAAUCCAGUGAGUCUAUGGUCGUAAUGAGCAAUUGCGGGAUGGGACCGACU	
	ACUUUGGGUGUCCGUGUUUCUUCUUUUUCCUUUAUAUUGGUGACAAUAUAUAUAGUGAUUAU	
	AUAUUGGCAUC	
4	CCAAAGUAGUUGGUCUCCGUCUCCGCAUGCAACUAGAAGCUUUGCACAAGAGCCAAUAGCCAGUAAUC	rhinovirusA2 6
0	AUCCAGAUUAUCUUAAGGUCAGCACUUCUGUUUCCCGGUCAAAGUUGAUUAGCUCCACCAGGGCAA	
1	AAACAACUGCGAUUCGUUAUCCGCAAGGCGCCUACACAAAGCUUAGUAGCAUUCUGAAAGUUGUCUGG	
	UUGGUCGUUCCACCAUUCUCCCGUGUAGACCUGGCAGAUAGGCUAGAAACUCCCCACUGGCGACAG	
	UGUUCUAGCCUGCGUGGCUGCCUGCACACCCUUGGGUGUGAAGCCAUACGAUGGACAGGGUGUGAA	
	GAGCCCCGUGUGCUGCGUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAACCCUGCAGCUAGA	rhinovirusA2 7
	GCACACAUAUCCAGUGUGUUUCUAGUCGUAAUGAGCAAUUGCGGGACGGGACCAACUACUUUGGGUGU	
	CCGUGUUUCCUGUUCUUCUUGAAUAUCUGCUUAUGGUGACAAUAUAUACGUUAUAUUGACACC	
4	UUA AAAACUGGAUCUAGGUUGUCCACCACUAAAUUCCUACAUGGAGUUGUACUCUAUUAUUCGGUA	
0	AUUUUGUACGCCAGUUUUUAUCUCCCCUACCCCUUAAAGUAACUUAAGAAGUUGGACUAUACGACCAAU	
2	AGGUAACAUUAUCCAAAUGUUAAGGUCAGCACUUCUGUUUCCCGGUCAAAGUUGAUUGUGUUC	rhinovirusA2 8
	UAACAGGGCAAAAACAGCUGCGAUUCGUUAUCCGCAAGCAACUACGCAAGCCUAGUAGCAUUUUU	
	GAGAGUAUGUGGUUGGUCGCUCCACUAUAACCCAAUAGUAGACCUGGCAGUAGGCUAGAAAUUC	
	CCCACUGGUAAACAGUGUUCUAGCCUGCGUGGCUGCCUGCGUGCUCCUUGUGAGCACGAAGCCAUUAU	
	UUUGACAAGGUGUGAAGAGCCCCGUGUGCUUGCUUUGAGUCCUCCGGCCCCUGAAUUGGCUAACCU	
	UAACCCUGCAGCUAGCGUGUACAUAUCCAGUACAAGCUAGUCGUAAUGAGCAAUUGCGGGAUGGGAC	rhinovirusA2 9
	CAACUACUUUGGGUGUCCGUGUUUCCUGUUUUUUAUAUUGCUUAUGGUGACAAUAUAUAUAUAUAU	
	AACAUAUAUAUUGGCAUC	
4	UUA AAAACUGGAUCCAGGUUGUCCACCUGGAUCUCCUACAUGGAGUUGUACUCUAUUAUUCGGUA	
0	AUCUUGUACGCCAGUUUUUAUACCCCUUACCCCUUAGAAAGUUAAGCACAUAGACCAAUAG	
3	GCGGUGGCCACCCAGGUCGCUAACGGUCAAGCACUUCUGUUUCCCGGUUAAGAUUGAUACGCUCCA	
	ACAGGGCGAAAACAGUCUGAAUCCGUUAACCGCAAGUGACUACGCAAGGCUUAGUAUUGCCUUGAAG	
	AAUUUAUGACUGGUGUCCGCUUAUACCCCUAGUAGACCUGGCAGUAGAGCUAGAAAUACCCAC	rhinovirusA2 10
	UGGUAACAGUGUUCUAGCCUGCGUGGCUGCCUGCACACCCUUCGCGGUGUGAAGCCAUUAUUGGAC	
	AAGGUGUGAAGAGCCCCGUGUGCUCAUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACCC	
	UGCAGCUAGUGUAUGUAAUCCAACAUAUGGCUAGUCGUAAUGAGCAAUUGCGGGAUGGGACCAACUA	
	CUUUGGGUGUCCGUGUUUACUUUUUCCUUUAUUAUUGCUUAUGGUGACAAUAUAUAUAUAUAUAU	
	AUUGGCAUC	
4	UUA AAAACUGGGUCUAGGUUGUCCCGCCUAGAUUACCCUAGUGGUGUAGUACUCUAUUAUACGGUA	rhinovirusA2 11
0	AUUUUGUACGCCAGUUUUUAUACCCUCCUCCUUGUUAACUUAAGAAGUUUUAACUUGGAGACC	
4	AAUAGGCGGUAUUAUCCAGAUCAUCAACGGUCAAGCACUUCUGUUUCCCGGUUAAGAUUGAUACG	
	CUCCAACAGGGCGAAAACACUCUAUUCGUUAUCCGCAAGAUCCUACGCAAGCCUAGUACCAACCU	
	UCGAUUCUAUUUGGUUGGUGCUUCAGCUACAAAACCCAGUAGUAGACCUGGCAGUAGAGGCUAGAAA	
	UAACCCUACUGGCGACAGUGUUCUAGCCUGCGUGGCUGCCUGCACACCCUUCGCGGUGUGAAGCCAUUAUUGGAC	rhinovirusA2 12
	AGUUGACAAGGUGUGAAGAGCCCCGUGUGCUCAUCUUGAUUCCUCCGGCCCCUGAAUGUGGCUAACCC	
	UUAAACUUGCAGCCAAUUGUACACAUAUCCAGUGUAUUAUUGGUGCUAAUGAGUAAUUGCGGGAUGGGA	
	CCAACUACUUUGGGUGUCCGUGUUUCUUCUUUUUACCUUUAUAUUGCUUAUGGUGACAGUAUAUGG	
	UAUAUAUAUAUUGUCACU	

4 0 5	UUA AAAACUGGAUCUGGGUUGUCCCAACUGGAUCUCCUACAUGGAGUUGUACUCUUAUUAUCCGGUA AUCUUGUACGCCAGUUUUAUACCCCUUCCCUUUGUAACUUAGAAGUUAAACACACAGACCAAUAG GUGAUGGCUAUCCAAGUCAUUAACGGUCAAGCACUUCUGUUUCCCGGUUUAUUAUAGUACGCUCCA CCAGGGCGAAAAACAAUUUAAAUCGUUAUCCGCAAAUAGACUACGCAAGCCUAGUAACAUCUUGAAG AACUUUAUGGUUUGGUCGUUCCACUAAACCCCAUAGUAGACCUGGCAGAUAGGCUAGAAAUACCCAC UGGAGACAGUUGUCUAGCCUGCGUGGCCUGCCUGCACACCCUAUGGGUGUGAAGCCAUUUUUGACA AGGUGUGAAGAGCCCCGUGUGCUCAUUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACCCU GCAGCUAGUGCAUGCAAUCCAGCAUGUUGCUAGUCGUAAUGAGCAAUUGCGGGAGUGGACCAACUAC UUUGGGUGUCGUGUUUACUUUUCCUUUUUUAUUGCUUAUGGUGACAAUUAUUAUAGUACAUAUUAU AUUGGCAUC	rhinovirusA2 9
4 0 6	UUA AAAACUGGAUCCAGGUUGUCCCAACUGGAUCUCCACAGGGAGUGGUACUCUGUUAUUAACGGUA ACUUUGUAGGCCAGUUUUAUCCUCCCAACCCUUAUGUAACUUAGAAGUUUACAUAAAGACCAUAG GUAUAAUACAUCAGAUUACUUAAGGUCAAGCACUUCUGUUUCCCGGUUGAUGUUGAUUAGCUCCA ACAGGGCAAAAAACAAGAAAUCGUUAUCCCGCAAAGUGCCUACGCAAGCCUAGUAGCACUUAGAA GUUGUCUGGCUGGUCGCGUCCGCCAUUUCCCGUGUAGACCUGGCAGAUAGGCUAGAAAUACCCAC UGGCGACAGUUGUCUAGCCUGCGUGGCCUGCCUGCACACCCUAUGGGUGUGAAGCCAUACAUGGACA AGGUGUGAAGAGCCCCGUGUGCUACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACCCU GCAGCUAGGGCACACAUCAGUGUGUAUCUAGUCGUAAUGAGCAAUUGCGGGACGGGACCAACUAC UUUGGGUGUCCGUGUUUCUUAUCCUUAUGUUUGCUUAUGGUGACAAUUAUGAAUUAUUAUUG ACACC	rhinovirusA3 0
4 0 7	UUA AAAACUGGGUGCGAGUUGUCCCAUCUGGCACUCCUACUGGAGUAGUACACUUAUUAUACGGUA AUUUUGUAGGCCAGUUUUAUAAACCCUUCUCCCUUUUGUAACUUAGAAGUUUAGACUUGAGACCAUAG GUGGUGAUCAACCAGAUCCGCAAGGUCAAGCACUUCUGUUUCCCGGUCAACAUAUAGUAUAGCUCCA CCAGGGCAAAAAACAAUUGCGAUCCUUAACCCGCAAAGUGCCUACGCAAGCCUAGUAAUUAUCUUGAAC GAUUAUUGGUUUGGUCGCUCCACUAAUCCCAUAGUAGACCUGGCAGAUAGGCUUGGAAACUCCCCAC UGGCGACAGUUGUCCAGCCUGCGUGGCCUGCCUGCGCACUCUACAGAGUGCGAAGCCAUUAUUAUGAC AAGGUGUGAAGAGCCCCGUGUGCUACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACCC UGCAGCUAGUGCAUACAACAGUAUGUGGCUAGUCGUAAUGAGCAAUUGCGGGACGGGACCAACUAC UUUGGGUGUCCGUGUUUCUUAUCCUUAUGUUUGCUUAUGGUGACAAUUAUGAAUUAUUAUUG ACACC	rhinovirusA3 1
4 0 8	UUA AAAACUGGAUCCGGGUUGUCCCAACCCGGAUCUCCACGUGGAGUUGUACACUUAUUAUACGGUA AUUUUGUACGCCAGUUUUAUUAUCCCUUACCCCAAGUAACUUAGAAGUAAACACACAGACCAUAGG AGGUGGCUAUCCAGAUCAUUAAGGGUCAAGCACUUCUGUUUCCCGGUCAAGGUUGAUUAUGCUCCAA CAGGGCAAAAAACAGCUGAGAUCAUUAUCCGCAAGUGCCUACGCAAGCCUAGUAAUUAUCUUGAAA GUACUUGACUGGUCGCUCCGCCAUACCCCAUGGUAGACCUGGCAGAUAGGCUAGAAAUCCCCACU GGUGACAGUGUUCUAGCCUGCGUGGCCUGCCUGCACACCCUUAUGGGUGUGAAGCCAAUUAUGGACA GGUGUGAAGAGUGCGAGUGGUCUACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCCUAAACCCUG CAGCUAGGGGCACACAUCAGUGUGUAUCUAGUCGUAAUGAGCAAUUGCGGGAGUGGACCAACUACU UUUGGGUGUCCGUGUUUCUUGUUUUCAUUAUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAUUAU UGGCGUC	rhinovirusA3 2
4 0 9	UUA AAAACUGGAUCUAGGUUGUCCCAACCUAGAUCUCCUACAGGAGUUGUACUCUGUUAUUCGGUA ACUUUGUACGCCAGUUUUAUUAUCCCUUCCCUUUGUAACUUAGAAGCUAACAUCAGACCAAUAGG GGCGGCAUUCACACAGAUCAUUAAGGGUCAAGCACUUCUGUUUCCCGGUCAAGGUUGAUUAUGCUCCAA AACAGGGCAAAAAACAGCCUUAUCGUUAUCCGCAAGACAACUACGCAAGCCUAGUUAUCAUCUUUGA AAGUAUUGGUUUGGUCGCUCCGCUAUAUAAUCUAAUAGCAGACCUUGGCAGAUAGGCUAGAAACUCCC CACUGGCGACAGUUGUUCUAGCCUGCGCGGCGUCCUGCGUGCCUUAUACGGGUACGAAGCCAUUAUUA UGACAAGGUGUGAAGAGCCCCGUGUGCUUUGUUGUAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUA ACCCUGCAGCUGGUAGCAUUAUCCAGUACUAGGCCAGUCGUAAUGAGCAAUUGCGGGAGUGGACCA ACUACUUUGGGUGUCCGUGUUUCCUUAUUUAUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAUUA CAUUAUUAUUGGCAUC	rhinovirusA3 3
4 1 0	UUA AAAACUGGGUCCAGGUUGUCCCAACUGGACUCCCAACUGGAGUUGUGCACUGUUAUUCGGUA ACUUUACACGCCAGUUUUAUUCUCCCUCCCAUUAUGUAACUUAGAAGUUAACACAGCGACCAUAGG UGGCAACUAGCCAAAGUUGCUUAAGGUCAAGCACUUCUGUCCCGGUCAAGUUGUUAUACGCUUCC CAGGGCGAAAAACAACUGAGAUUGUUAUCCGCAAGGCGCCUACACAAAACCUAGUAAUUAUUAUG GUAUGUGGCUUGGUCGUUCCGCUAAACCCCAUAGUAGACCUGGCAGAUAGGCUAGAAAUUCCCAU GGCGACAGUUGUUCUAGCCUGCGUGGCCUGCGUGUCCUUAACGGACACCAAGCCAUUAUUAUUAU AGGAGAGUAGUCCGUGUGCUUGUUGUAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACCCU GCAGCCAGUGCGCACAAACAGUGUGUAGCUGGUCGUAAUGAGUAAUUGCGGGAUGGACCAACUAC UUUGGGUGUCCGUGUUUACUUUUUUACUUUUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAUUA UUGACACC	rhinovirusA3 4
4 1 1	UUA AAAACUGGGAGUGGGUUGUCCCAUCACUCCACCAUUGCGGUGUUGUACACUGUUAUUAACGGUA ACCUUGUACGCCAGUUUUGCCACCCUCCCAUUAUGUAACUUAGAAGUUUUAACAACAAGACCAU AGGUGAUAAUUAUCCAGAUUAUCAAAGGUCAAGCACUUCUGUUUCCCGGUCAAGUUGUUAUACGCUUCC UACCCAGGCCAAAAACCUUAAAGAUUGUUAUCCCAACACUGCCUACACAGAGCCUAGUACCAUUAU GAUGUAAUUAAGGUUGGUCGCUCCCGCAAACCCAGCAGUAGACCUGGCAGAUAGGCUUGGACACUCC CCACUGGUGACAGUGGUCAGCCUGCGUGGCCUGCGUCACCCUUAUUGGGUGAGAGGCCUACUUAU UGACAAGAUUGUAGAGCCGCGUGUGCUAGUGUGCUUCCUCCGGCCCCUGAAUGUGGCUAACCUUA ACCCUGCAGCGCGUCCCAUUAUCCAAUUGGGUUUACGGUCGCAACGCGCAAGUGCGGGAUGGACCA ACCCUGCAGCGCGUCCCAUUAUCCAAUUGGGUUUACGGUCGCAACGCGCAAGUGCGGGAUGGACCA	rhinovirusA3 6

	ACUACUUUGGGUGUCCGUGUUUCCUAAUUUUUUUUUUUUUGCAUUUUUUGGUGACAAUUUAUAGUAU AUAGAUGUCGUC	
4 1 2	UUAAAACUGGAUCUGGGUUGUCCACUCAGAUUCUCCAUUGCGGAGUUGUACUCUAUUUUCGGUA AUCUUGUACGCCAGUUUUUACACCCUCCCCCUUGUAACUUAGAAGAAAGUACAUAACGACCAUAG GUGGCAAUUAAACCAAUUGUCAACGGUCAAGCACUUCUGUUUCCCCGGUCAAGAUUGAUACGCUCCA ACAGGGCGAAAACAAUUGAGAUCCGUUAUCCGCAAGUGCCCACGCAAGCCUAGUAAACAUUUGAAA GGUCUAUGGUUGGUCGUUACGCUAAACCCCAUAGUAGACUUGGCGAGUAGGCUAGAAAUCCCCAC UGGCGACAGUUGUCCAGCCUGCGUGGCGUGCCUGCACACCCUCACGGGUGUGAGCCAUUAACUCGAC AAGGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAACC CGCAGCUAGUGCAUAACAACAGUAUGUGGCUAGUCGUAUAGGCAAUUGUGGGAUGGGACCAACUA CUUUGGGUGUCCGUGUUUACUUUUUCCUUUUGUUGCUUUGGUGACAAUAUAUAGUAUAUAU UGACACC	rhinovirusA3 8
4 1 3	UUAAAACUGGAUCUGGGUUGUCCACCUAGAUACCUACAGGGUGUAGUACACUGUUUUUACGGUA ACUUUGUACGCCAGUUUUUAUCCCCCCCCCAACGUAAACUUAGAAGCAAGCACAAACGUCCAUAG GCGGUGAUCAUCCAAACUACUACGAGCAAGCACUUCUAGUCUCCCCGGCAACUGCUGAUUAGCUC AACAGGGCAAAAACAGCGGGAGUCGUUAUCCGCAAAUUGCCUACGCAAGCCUAGUAUACCCUAGA AGGUGUGUGGUGGUGGUGCUCAGCCGUGAACCCGACGGUAGACCUUGGCGAGUAGGCUAGAAUACCC CACUGGCGACAGUUGUACCCUGCGUGGUGGUGCCUGCACACCCUACAGGGUGUGAAGCCAUUUAUUG GACAGGGUGAGAAGAGCCAGUGUGCUCAUUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAA ACCUGCAGCCAAUGCACACAUAUCCAGUGUGUAUUUGGUGUAACGAGCAAUUGCGGGAUGGGACCAA CUACUUUGGGUGUCCGUGUUUUCUUUUUAUUCUUAUACAUUUUGCCUAUGGUGACAAUAUAUACUGUUU AUUUGUUAUC	rhinovirusA3 9
4 1 4	UUAAAACUGGAUCCAGGUUGUCCACCUUGAUUUUUAUAGGAGUUGUACUCUGUUUUUCCGGUA ACCUUGUACGCCAGUUUUUAUCCCCCCCCCAUUGUAACUUAGAAGUUUAGCAUACAGACCAUAGGU GGUGAUCCUCCAGAUACUAACGGUCAAGCACUUCUGUUUCCCCGGUUUGAAUUGAUACGCUCCAAC AGGGCGAAAACAGUUUAAUUCGUUAUCCGCAAGUGCCUACGCAAGCCUAGUAGCACUUUAGAAA UCUAUGGUGUGUCGUUACGCUUAUACCCCUAGUGGACCUUGGCGAGUAGGCUAGAAUACCCACUG GCGACAGUUGUUCUAGCCUGCGUGGUGCCUGCACACCCUUAUGGGUGUGAAGCCAUUAUUGACAA GGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAACCUG CAGCUAGUGCAUGUAUCCAACAUGUUGCUAGUCGUAUAGGCAAUUGCGGGAUGGGACCAACUACU UUGGGUGUCCGUGUUUACUUUUUCCUUUUAUUGCUUUAUGGUGACAAUAUAUAUAAACAUAUAU AUUGGCAUC	rhinovirusA4 0
4 1 5	UUAAAACUGGGUUGGGUUGAUCCACCCAAACCACCCAGUGGUGUAGUACACUGUUUUUCCGGUA ACUUUGUACGCCAGUUUUUAUCCCCCCCCCUUGUAACUUAGAAGUACUAGCACAUAAGACCAU AGCAGGCAAUCAUCCAGAUUGUCAUCGGUCAAGUACUUCUGUUUCCCCGGUUCAGAUUGAUUGCUC CAACAGGGCAAAAACAACCUGGAUCGUUAUCCGCAAGAUGCCUACGCAAGCCUAGUAUACACACUGG AAGAUUCUGGGUUGGUGCUCAGUUAACAACCCAGUAACAGACCUUGGCGAGUAGGCUAGAAUUC CCACUGGCGACAGUUGUUCUAGCCUGCGUGGUGCCUGCACACCCGUUUGGGUGUGAAGCCUAGAAU GGACAGGGUGCGAAGAGUCGAGUGGUCUACUUUGAGUCCUCCGGCCCCUGAAUGCGGCUAACCUU AACCUGCAGCCAUUGGCCCAAAACCAAGUGGGUCUAUGGUUGUAUAGGCAAUUGCGGGAUGGGACCG ACUACUUUGGGUGUCCGUGUUUACUUUUUACUUUAUUAUUGCUUUAUGGUGACAAUAUAUAGAUACA UAUAUUGACACC	rhinovirusA4 1
4 1 6	UUAAAACUGGGUCCAGGUUGUCCACCUUGAUACUCAGUGUUGUUAUCUGUUUUUACGGUA ACUUUAUACGCCAGUUUUUAUCCCCCUUACCCCAUUGUAACUUAGAAGUUUAGCAUACCGACCAU AGUGAGCAAUCAUCCAGAUUGUUUAAGGUCAGACUUCUGUUUACCCGGUCAGCGUUGAUUGCUC CAACAGGGCAAAAACAACAAUGAUCCGUUAUCCACAAGAUCCUACACAGAGCCUAGUAACAUCUAG AAAAAUUCUGGUUGGUGCUCUAAUUGCAAAACCCAGCAUAGACCUUGGCGAGUAGGCUAGGACAUCC CCACUGGCGACAGUGGUCAGCCUGCGUGGUGCCUGCACACUCCUUUGGAGUGUGAAGCCAUAAU UGGACAGGGUGCGAAGAGCCGUGUGGUCUACUUUGAGUCCUCCGGCCCCUGAAUGCGGCUAACCUU AAACUGGACGCCAUUGUUUGCAUCCAGCAGAUUAUUGGUGUAAUAGGCAAUUGCGGGAUGGGACC GACUACUUUGGGUGUCCGUGUUUCCUUUAUUAUCCAUUUUAUUGCUUUAUGGUGACAAUAUAUAGUAUA UAUUGACACC	rhinovirusA4 3
4 1 7	UUAAAACUGGGUGUGGUGUUGUCCACACGACUAUCUACGGAUGUAGUGCUCUUGUAUUCGGUAC GCUUGCACGCCAGUUUUUGUACCCCCCUUACAUUGUAACUUAGAAGAUAAACACAAGACCAUAG GCGGCAUAGAACCACAUAUGUCAACGGUCAAGCACUUCUGUUUCCCCGGUCAUUCUAGAUUAGCUUUA CCAAAGGCAAAAACUGGAGCGAUCGUUAUCCGCAAGGUGCCUACGCGAAACCCAGUAGCAUUCUUA UUUGAUUCUGGUUGGUGCUCAGCCGUUAAACCCAAACGGUAAACCUUGGCGAGUAGGCGUGAAGAU CCCACACGCGAUGGUGUUCAGCCUGCGUGGUGCCUGCACACCCGAAAGGGUGUGAAGCCUUUUA AAGACAGGGUGCGAAGAGUCUACUGUGCUCACCUUGAUUCCUCCGGCCCCUGAAUGUGGCUAAUCCU AACCCUGAGCUGUGGUGGCAUUCAGCACAUUCGAGUCGUAUUGGUAACUGCGGGAUGGAACC AACUACUUUGGGUGUCCGUGUUUCCUUUAUUAUUGCUUUAUGGUGACAAUUAUACCGAU AUAGUUGCAUC	rhinovirusA4 5
4 1 8	UUAAAACUGGAUCUGGGUUGUCCACCCAGAUUUCCAUUGGAGUAGUACUCUAUUUUCGGUA AUUUUGUACGCCAGUUUUUAUUAUCCUACCCCGAGUAACUUAGAAGUAAGCACACAAGGCCAAUAG UGGGUGAUUAGCCAAAUCAUCAAAGGUCAGCACUUCUGUUUCCCCGGUCAAGGUUGAUUAGCUCU ACCAGGGCAAAAACAACUAGAUCCGUUAUCCGCAAAUUGCCUACACAAGCCUAGUAUACUUGGA AGGACAGAGUUGGUGCUCUACUUAUCCCCAUAGUAGACCUUGGCGAGUAGGCUAGAAAAUCCCCA UGGCGACAGUUGUUCUAGCCUGGUGGUGCCUGGUGCCUUGCGAGGACGAAGCCUUGUUGA CAAGGUGUGAAGAGCCCCGUGUGCUCAUUUUGAUUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAACC	rhinovirusA4 6

[illegible]

	CAUGUAGCUAGUCGUAUAUAGCAAUUGCGGGAUGGGACCAACUACUUUGGGUGUCCGUGUUUACUU UUUACCUUAUUAUUGCUUAUGGUGACAAUUAUUAUUGCAUAUAUUAUUGGCAUC	
4 2 6	UUAAAACUGGgGUUCAGGUUGUUCACCUGGACUUCUUAUGUGGAGUUAACACUGUUAUCCCGUA ACCUUGUACGCCAGUUUUUAUACACCUUCCUGUAACUUAAGUUAUAAGCACGAGGACCAUAGG UAGUGACUUUCCAGGUCACUAAAGGUCAGCACUUCUGUCUCCCGGUCAAAGUUGAUUAGCUCUCC CCAGGGCAAAAACAAUUGAGAUCGUUAUCCGCAAUUGGCCUACGCCAAAGCCUAGUAAUUAUUAUAGU GAUUAUUGGUUGGUGCUCUCCGCCAUUUCCUAGUGAGACUUGGAGAGGCCUAGAAACUCCCCAC UGGCGACAGUGUUCUAGCCUGCGUGGCCUGCCGCGCACUCUCUGAGUGCGAAGCCAUUAUUGGAC AGGUGUGAAgAGCCCCGUGUGCUCACAUUGAAUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAACCU GCAGCUAGUGCACAAACCAGUGUGUAGCUAGUCGUAACGAGCAAUUGCGGGAUGGGACCAACUAC UUUGGGUGUCGUGUUUACUUUUUCCUUUUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAUACAU UAUUGACACC	rhinovirusA5 6
4 2 7	UUAAAACUGGAUCCAGGUUGUUCACCUGGAUCUCCUACGGGAGUUGUACUCUGUUAUUCGGUA CCUUGUACGCCAGUUUUUACUCCUCCUCCCAUUGUUAACUUAAGAAUUAACACAGGACCAUAG GCAGUGAUUAACGAGUACUUAAGGUCAAGCACUUCUGUUUCCCGGUCUGAGUUGAUUAGUCUUA ACAGGGCAAAAACAAAGUCAGAUCCGUUAUCCGCAAGUGCCUACGCCAAAGCCUAGUAUCUUCUUGAA AAUGUAGUGGUUGGUAUUCGCGUGAACCCACAGUAGACUUGGAGAGGCCUAGAAACCCCCAC UGGCGACAGUGUUCUAGCCUGCGUGGCCUGCCUGCACACCCUUUAGGGUGUGAAGCCAUUAUUGAC AAGGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAACCC UGCAGCUAGUGCAUACAAGCCAGUAUGUUGCUAGUCGUAACGAGCAAUUGCGGGAUGGGACCAACUA CUUUGGGUGUCCGUGUUUACUUUUUACCUUAUACUGCUUAUGGUGACAAUUAUUAUUAUAGCAUUA UAUUGACACC	rhinovirusA5 7
4 2 8	UUAAAACUGGGAGUAGGUUGUUCACCACUACUACCCACGUGGUGUUGUACACUGUUAUACCGUA ACUUUGUAGGCCAGUUUUUGCCACCCUUCUCCGUAGCGUAACUUAAGAUUUUACACUUAAGACCAU AGGCGAUUAUUAUCCAGAUUAUCAAAGGUCAAGCACUUCUGUUUCCCGGUCAAUGAGGAUAUGCUU UUCGCAAGGCAAAAACCUUAGAGAUUCGUUAUCCCAUUAUGCCUACACAGAGCCUAGUAUACCUUU GAGUUAUUUAGACUGGUCUCCUGCAAACCCAGCAGUAGACUUGGAGAGGCCUAGAGGCCUGGACGUCC CCACUGGCGACAGUGGUCAGCCUGCGUGGCCUGCCUGCCACCCAUUUGGGUGGGAAGCCUAACUUAU UGACAUGGUGUGAAGAGCCGCGUGUGCUCAAUGUGCCUCCUCCGGCCCCUGAAUGUGGCUAACCUUA ACCCCGCAGCCGUGCCCAUGAUCCAGUGGGUAUACGGUCGUAACGCGCAAGUGUGGGAUGGGACCA ACUACUUUGGGUGUCCGUGUUUCCUGUUUUUUUUUAUGCAUUUAUGGUGACAAUUAUUAUUAUGGU AUUAGAUUGUCAUC	rhinovirusA5 8
4 2 9	UUAAUACUGGAUCCAGGUUGUUCACCUGGAUCUCCCAAGCGGAGUGGUACACUUAUUAUCCCGUA AUUUUGUACGCCAGUUUUUAUACCCUCCUCCUUUGUAACUUAAGAUUAACACAGACCAUAGG UGGUAACUUAACCAAGUUAUAAAGGUCAAGCACUUCUGUUUCCCGGUGAUUAUUGAUUAUGCUCAC CAGGGCAAAAACAAUAGAGUCCGUUAUCCGCAAGUGCCUACGCAAAAGCCUAGUAAACUUAUUAU GUUUUAGGUUGGUGCUCAGCUAUUCCUCCUAGUAGACUUGGAGAGGCCUAGAAUUAUCCCCACU GGUAAACAGUGUUCUAGCCUGCGUGGCCUGCCUGCGCACCCUCCGGUGCGAAGCCAUACACUAGACA AGGUGCGAAGAGCCCCGUGUGCUCACUUUGAAUCCUCCGGCUCUGAAUGUGGCUAACCUUAAACCU GCAGCCAGUGCACAUAAAGCCAAUGUGUAUCUGGUCGUAUAGAGUAAUUGCGGGAUGGGACCAACUAC UUUGGAUGUCCGUGUUUCCUGUUUUUCCUUAUUUUGCUUAUGGUGACAAUUAUUAUAGUAAUUAUA UUGGCAUC	rhinovirusA5 9
4 3 0	UUAAAACUGGAUCCAGGUUGUUCACCUGGAUCUCCUACGGGAGUAGUACACUUAUUAUACCGUA AUUUUGUACGCCAGUUUUUAUAGAUCCUCCUCCUUUGUAACUUAAGAUUAACAAUUAAGACCAU AGGCGGCAAUUCCAGAUUGUCAAGGUCAAGCACUUCUGUUUCCCGGUCAAAGCUGAUUAGCUC CAACAGGGCAAAAACAGCUGAGAUCCGUUAUCCGCAAGUGCCUACGCAAGCCUUAUUAACACUUUG AAGUUAACGUGGUUGGUGCUCAGCCAAACCCUAGGUAGACUUGGAGAGGCCUAGAGGCCUAGAAUACCC ACUGGCGACAGUGUUCUAGCCUGCGUGGCCUGCCUGCACACCCUUGGGUGUGAAGCCAUUAUUGGA CAAGGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACC CUGCAGCUAGUGCAUGCAAUCCAGUAUGUAUCUAGUCGUAUAGCAAUUUGCGGGAUGGGACCAACU ACUUUGGGUGUCCGUGUUUCCUUUUUAUCCUUAUUAUGUUAUUGGUGACAAUUAUUAUUAUUA UAUUGACaCC	rhinovirusA6 0
4 3 1	UUAAAACUGGGUCUGGGUUGCUCUCCACCCAGACCACCAUGUGGUGUUGUACACUGUUAUUCGGUA ACUUUGUACGCCAGUUUUUAUACUCCUCCUCCUUUGUAACUUAAGAGCUAAACACAUUGACCAAU AGCAGGCAAUUACCCAGAUUGCUUAUGGUCAGUACUUCUGUUUCCCGGUCUCCUUGAUUAGCUC UACCAGGGCAAAAACAGAGUAGAUCCGUUAUCCGCAAGAUCCUACGCAACGCCUAGUAAACACUUUGA UGUGUUUUUGGUGUGCUCAGUUAACAACCCAGUAAUAGACUUGGAGAGGCCUAGAGGCCUAGAUUCC CCACUGGCGACAGUAGUACUAGCCUGCGUGGCCUGCCUGCACACCCUUUAUGGGUGUGAAGCCAAACAU AUGACAAGGUGCGAAGAGCCACGUGUGCUCACUUCUAGUCCUCCGGCCCCUGAAUGCGGCUAACCU AACCCUGCAGCCAUUGCUCACAAGCCAGUGGGUAUGUGGUCGUAUAGAGAAUUGCGGACGGGACC GACUACUUUGGGUGUCCGUGUUUACUUUUUACCUUAUUUUGCUUAUGGUGACAAUUAUUAUUAUACAG AUUAUAUUGACACC	rhinovirusA6 1
4 3 2	UUAAAACUGGAUCCAGGUUGUUCACCUGGAUcUCCUACGUGGAGUUGUACUCUUAUUAUUCGGUA ACCUUGUACGCCAGUUUUUAUUAUCCUACCCUCCUUUGUAACUUAAGAGCUAAACACAUUGACCAUA GGCGGUGACCACCAAGUCACUAACGGUCAAGCACUUCUUGUUUCCCGGUUAAGAUUGAUACGCUCC AACAGGGCGAAAACAAUUAUAAUUGCUUAACCGCAAGUGACUACGCAAGCUUAGUAAUACCUUGAA GAAUUUAUGGUGUGCUGUCCGCUUAUAAACCCUAGUAGACUUGGAGAGGCCUAGAGGCCUAGAAUACCCCA UUGGUAACAGUGUUCUAGCCUGCGUGGCCUGCCUGCACACCCUUUUGGGUGUGAAGCCAUUAUUAUUG ACAAGGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAAC	rhinovirusA6 2

	CCUGCAGCUAGUGUAUGUAAUCCAAACUAUGGCUAGUCGUAUAGGCAAUUGCGGGAUGGGACCAAC UACUUUGGGUGUCCGUGUUUACUUUUUCCUUUUUUUUUGCUUUAUGGUGACAAUUAUUAUUAUUAU AUUUGGCAUC	
4 3 3	UUAAAACUGGGUCCAGGUUGUCCACCUGGACCUCUAAGUGGAGUGGUACACUAUUUUCGGUA AUUUUGUACGCCAGUUUUUACUCCCCCCCCUUUUUUAAGAAUUAUGAACACAGACCAAU AGGUGGCAACUAGCCAAGUUGUAAAGGUCACAGCUCUUGUCUCCCCGGCUGACGUUGAUUGCUC CAACAGGGCAAAAACAAUAGAAUGUCUUUACCGCAAAGUGCCUACGCAAGCCUAGUAACACUUUGA AAUGUCUAUGGUGGUGGUCGUCAGCUAUUCCCCAUAGUAGACCUGGCAGAUAGGGCUAGAAUACCCC ACUGGUAACAGUGUUCUAGCCUGGUGGUGCCUGCGCACCCUCCCCGGGUGCGAAGCCAUUAUUAU GACAAGGUGCGAAGAGCCCCGUGUGCUCACUUUGAAUCCUCCGGCUCUGAAUGUGGCUAACCUUA CCCCGCGACCCAGUGCACAUAAGCCAUGUGUGGUGGUGCUAAUGAGCAAUUGUGGGAUGGGACCAA CUACUUUGGAUGUCCGUGUUUCCUGUUUUUUCUUUUUUUUUUGCUUUAUGGUGACAAUUAUUAUGUUU AUUAUUAUUGGCAUC	rhinovirusA6 3
4 4 4	UUCAAACUGGAUCUGGGUUGUCCCCACCCAGAUACCCCAAGUGGUGUAGUACUCUGUCAUUCGGUG ACUUUGUACGCCAGUUUUUACACCCCCCUUUGUAACUUAGAAGAAUUAAGUUCAGACCAAUAG GAAAACGACAAACCAGUCGUUUAAGGUCAAGCACUUCUGUUUCCCCGGUCAAUGCUGAUUAGCUUA CCAAGGCAAAAACAGCAGCGAUCCGUUACCCGCAAGACGCCUACGCAAGCCUAGUAUUGCCUUUGAA GAUUAUUUGUCUGGUCACUCAGGUACAAUCCAGUACUAGACCUGGCAGAUAGGGCUGGAAAAUCCCC ACUGGCGACAGUGUCCAGCCUGCGUGGUGCCUGCACACCCAUACGGGUGUGAAGCCAAAUUUGG ACAAGGUGUGAAGAGCCCCGUGUGCUCGUCUUGAGUCCUCCGGCCCCUGAAUUGCGGCUAACCUUAAC CCCACAGCCAUUGCUUGCAUUCAGCAAGUCUAUGGUGCUAAUGAGCAAUUGUGGGAUGGGACCGAC UACUUUGGGUGUCCGUGUUUACUUUUUACCUUUUAUUAUGCUUAUGGUGACAAUUAUUAAGCGUUUA AUUAUUGGCAUC	rhinovirusA6 4
4 5 3	UUAAAACUGGGUGUGGGUUGCACCCACCCACACUGUCUACGGACGUAUGUCUCUCAUUAUACGGUAU GCUUGCAGCCAGUUUGUUUUCCCCUUUUUUUUUGUAACUAUAGAAGAUUGCACUUAUUCUGGACC AAUAGGCGAGUAGUACACACUACUACGGUCAAGCACUUCUGUUACCCCGGUAUUGCGGAGAAUAG GCUUCGAAAAAGGGUGAAGCUUCAAUUAUCGUAUCCGGCAAUUACUACGCAAGCUUAGUAACUU CCCAAGAAAUUCUUUGGUUGGCCUUCGCGUGGUACCCACCCAGUAGGCUUGGCAGAUAGGGCUGAG UAAUCCCCACUUGGCGACAGUGACUCAGCCUGCGUGGUGCCUACACACCUUGUAACAGGUGUGAAGC CAUAGAAGGACAGGGUGUGAAGAGCCCCGUGUGCUAAUUGUGAUUCCUCCGGCCCCUGAAUGCGGC UAUUCUAACCCACAUCCAUUGCGCAUAAGCCAUGCGUAUUGGUGCUAAUUGGGCAACUGUGGGAC GGAACCGACUACUUUGGGUGUCCGUGUUUUCUUUUUAUUCUUUAUUAUUGUUAUGGUGACAAUUAU AAGUGUUUAUUAUUGCAUC	rhinovirusA6 5
4 3 6	UUAAAACUGGAUCUAGGUUGUCCCCACCUAGAUUCUCCCAUGUGGAGUAGUACUCUAUUAUACGGUA AUUUUGUACGCCAGUUUUUAUACCCCUUCCCCUUUGUAACUUAAGAAGUCUGAACACAGACCAUA GUGGGGUAUCACCCAGAUCCGUUAAGGUCAAGCACUUCUGUUUCCCCGGUACAGUUGAUUAGCUUC ACCAGGGCAAAAACACGGAGAUCCGUUAUCCGCAAGGUGCCUACGCAAGCCUAGUAACCUUUGGA GGAUUAUUGGUGGUGGUCAGCUAUUACCCAUAGUAGACCUGGCAGAUAGGCUAGAAACUCCCCA CUGGGCAGAGUGUUCUAGCCUGCGUGGUGCCUGCGCACUCUUAUGAGUGCGAAGCCAUUAUUAUUGA CAAGGUGUGAAGAGCCCCGUGUGCUCACCUUGAGUCCUCCGGCCCCUGAAUUGGGCUAACCCUAACC CUGCAGUAGUGCAUGAAUCCAUCAUGUUGCUAGUCGUAUUGAGCAAUUGCGGGGAUGGGACCAACU ACUUUGGGUGUCCGUGUUUACUUUUUCCUUUUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAACAU AUUAUUGGCAUC	rhinovirusA6 6
4 3 7	UUAAAACUGGAUCUAGGUUGUCCCCACCUAGAUUCUCCCAUGUGGAGUGGUACACUAUUUACCGUA AUUUUGUACGCCGGUUUUUACUCCCCUACCCCAAGUAACUUAAGAAGUUAUUAUUAAGACCAACAG GCAGUGAUUAUCCAGAUACUAAAGGUCAAGCACUUCUGUCUCCCCGGUCAAAGUUGAUUAGCUCCA ACAGGGCAAAAACACAGAGUAGCUUUAUCCGCAAAGUGCCUACGCAAGCCUAGUAACUACCUUGAA AUUACUUGGUUGGUGUCCACCAAACCCCAUGGUAGACCUGGCAGAUAGAGCCUAGAAAUUCCCCAC UGGUAACAGUGUUCUAGCCUGCGUGGUGCCUGCACACCCUUUGGGUGUGAAGCCAAAGUUAUUGGACA GGGUGUGAAGAGCCGAGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUUGGCUAACCUUAACCCU GCAGCUGAGAGCGCGCAAACAGCGUGUAUCUAGUCGUAUUGAGCAAUUGCGGGACGGGACCAACUAC UUUGGGUGUCCGUGUUUUCUUUAUUCUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAUUAUUAU UUGGCACC	rhinovirusA6 7
4 3 8	UUAAAACUGGGUCCAGGUUGUCCCCACCUAGAUUCUACUGGAGUUGUACUCUAUUAUUCGGUA AUUUUGUACGCCAGUUUUUACUCCCCUACCCCUUUAUGUAACUUAAGAAGUUUAUGCAGCGGCCAA UAGGUGACAUUAUCCAAAAGUGCAAAGGUCAAGCACUUCUGUUAACCCCGUUAAGAAUUGAUUUGCU CCAACAGGGCAAAAACAAUUAUUGAUUCGUUAUCCGCAAGGCAACUACGCAAGCCUAGUACCAUUUU GAUGAUGUGUGGUGGUGGUCUCCACUGUGAACCAGCAGUAGACCUGGCAGAUAGGCUAGAAAUUC CCCACUGGUAACAGUGUUCUAGCCUGCGUGGUGCCUGCGUACCUUUUGCGGGUACGAAGCCAUAC AUUUGACAAGGUGUGAAGAGCCCCGUGUGCUUGCCUUGAGUCCUCCGGCCCCUGAAUUGGCUAACC UUAAACCCUGCAGCUAGUGGCACAAUCCAGUGCAUGGCUAGUCGUAACGAGCAAUUGCGGGGAUGGGA CCAACUACUUUGGGUGUCCGUGUUUUCUUUUUAUUAUUGCUUAUGGUGACAAUUAUUAUUAUUAU UAAUUAUUAUUGGCACC	rhinovirusA6 8
4 3 9	UUAAAACUGAGUGUGGGUUGUCCCCACCCACACUGUCUACGGACGUAUGUCUCUCAUUAUUCGGUAU GUUUGCAGCCAGUUUGCCUCUCCUUUUUCCACGUAACGUUAAGAAGCUUAUGCACUAAUUGGACCAC UAGGUGAGUGUAAACCAUAUACUUCUUUGGUCAAGCACUUCUGUUUCCCCGGUUAUGCAGGGAUAG GCCCCCCCCAGGGCUGAAGCCUUAUUAUUCGUUAUCCGCAAUUACUACGCAAGCCUUAUAGCGU UCUGGAAGUUCACUGGCGUGGUGUCCGCUAGACUAAACCGUUAUAGUAGACCUGGCAGAUAGGGCU GAUAAACCCACUGGCGACAGUAUCCAGCCUGCGUGGUGCCUGCACACCUUGGAACAGGUGUGAA	rhinovirusA7 1

	GCCCAUGAAUUGGACAGGGUGUGAAGAGCCCCGUGUGCUAAGUGUGAGUCCUCCGGCCCCUGAAUUGCG GCUAAUUCUAAACCCACAGCUAAUACACACAAACAGUGUGGAUUUAGUCGUAUUGGUAACUGUGGG ACGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUUAUUAUUGCUUAGUGGACAGGU AUUAUGAUUAUUAUUGUUAAC	
4 4 0	UUAAAAACUGGGUUUGGGUUGAUCCCAACCCAAACCCUCCAAAGCGGAGUGGUACUCUGUUAUUUCGGUA ACUUUGUACGCCAGUUUUUAUUCUCCUCCUCCUUUAAACUGUAAGCUAAGCACACAGACCAUAG CUGACAAUACUCCAGAUUGUUAUUGGUAAGCAGUUCUGUUUCCCGGUAACGUUGAUUAGCUCUCC ACAGGGCAAAAAACACUGCGAUCGUUAUCCGCAAGUGCCUACGCAAGCCUAGUAGCACCUUUGAA GUCAUUUGGUUGGUCGUCUCCGUGUUUCCCAACAGUAGACCUGGCAGAUAGAGGUAGAAUACCCAC UGGUGACAGUUGUUCUAGCCUGCGUGGUGCCUGCACACCCUUUGGGUGUGAAGCCAAUACUGGACA AGGUGUGAAGAGCCCGUGUGUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGUUACCCUUAACCCU CGACUAGAGUGACAAUCCAGUGCAUUUCUAGUCGUAUUGAGCAAUUGCGGGAUGGGACCAACUAC UUUGGGUGUCCGUGUUUACUUUUUACCUAUUAUUUGCUUAGUGGACAAUUAUUAUUUAUUAU UGGCACC	rhinovirusA7 4
4 4 1	UUAAAAUGGAUCCAGGUUGUCCCAACCCUGGAUACCCAAUGGGUGUAGUAUACUGUUAUUACGGUA ACUUUAUACGACGGUUUAUUCUCCCCCCCCUUUUGUAACUUAAGAAGUUUUGCACACCGACCAAU AGUGGACAAUACAGCCAGAUUGUUUAGGUCAGCACUACUGUUUCCCGGUAACGUUGAUUAGCU CCGACAGGGCAAAAACACAGCGAUCGUUAUCCGCAAGGUGCCUACACAGAGCUUAGUACCGCCUUA GAAGGAUUUGGUUGGUCGUCAGUUGCAACCCAGCAUAGACCUGGCAGAUAGAGCUGGACAUUC CCACUUGGUGACAGUGGUCAGCCUGCGUGGUGCCUGGCACACCUCUAGAGGUGUGAAGCCAGAAU UUCGACAAGGUGCGAAGAGCCGCGUGUGUCUACCUUGAAUCCUCCGGCCCCUGAAUGCGGCUAACCU UAAACCCGCAACCAUGGUCACAAUCCAGUGGAUGUAUGGUGCUAAUGAGUAAUUGUGGGAGGGAC CGACUACUUUGGGUGUCCGUGUUUCCUGUUUUUCCUUUAUACUGCUUAGUGGACAAUUAUUAUGA UAUUAUUGGCACC	rhinovirusA7 5
4 4 2	UUAAAAUGGAUCCAGGUUGUCCCAACCCUGGAUCCUACAAGGAGUUGUACUCUUAUUUUGCGGUA AUUUUGUACGCCAGUUUUUAUUCUCCCCCCCCGUAUUGUAACUUAAGAAGCUAACAAGACCAUA GGUGACAUUUUAUCCAAAUGUUAAGGUCAGCACUUCUGUUUCCCGGCCAAAGUUGAUUAGCUC CAACAGGGCAAAAACAGCUGCGGUGCUUAUCCGCAAAACAACUACGCAAGCCUAGUACCCUUCU AGAGCAUGUGGUGGUGCGUCCACUAUAAACCCAAUAGUAGACCUGGCAGAUAGAGCUAGGAGCUCC UCACUGGUAACAGUUGUCCUAGCCUGCGUGGUGCCUGCGCACCCUGACAGGUGCGAAGCCAAUGUU UGACAAGGUGUGAAGAGCCCCGUGUGCUUUGUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUA ACCCUGCAGCUAGUGUGUACAUAUCCAGUGCAUAGCUAGUCGUAUUGAGCAAUUGCGGGAUGGGACCA ACUACUUUGGGUGUCCGUGUUUCUUAUUUUUCCUUAUUAUUUGCUUAGUGGACAAUUAUUAUUA CGUAUAUUAUUGGCAUC	rhinovirusA7 6
4 4 3	UUAAAAUGGAUCCAGGUUGUCCCAACCCUGGAUCCUACAAGGAGGUGUACUCUUAUUUACGGUA AUUUUGUACGCCAGUUUUUAUUCUCCCCUACCCUUUGUAACUUAAGAAGUUUAACACAGACCAUA GGCGACGAUCAUCCAGAUCCGUUAAGGUCAGCACUUCUGUUUCCCGGUAAGCUGAUUAGCUCU ACCAGGGCAAAAACAGUAGAGUCCGUUAUCCGCAAGGUGCCUACACAAGCUUAGUAACAUUUUGAA GAGCAUUUGGUUGGUCGUCAGCUUAUACCCCAUAGUAGACCUGGCAGAUAGAGCUAGAGUUCUCC ACUGGCGACAGUGUUCUUGCCUGCGUGGUGCCUGCGCACUCUCAUGAGUGCGAAGCCAUAGUUUG ACAAGGUGUGAAGAGCCCCGUGUGCUACCUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAAC CCUGCAGCUAGUGCAUGAAUCCAUCAUGUAACUAGUCGUAUUGAGCAAUUGCGGGAUGGGACCAAC UACUUUGGGUGUCCGUGUUUCUUAUUUUUCCUUAUUAUUUGCUUAGUGGACAAUUAUUAUAGCAUA UAUUGACACC	rhinovirusA7 7
4 4 4	UUCAAACUGAAUCAUGGUUGUCCCAACCAUGAUUACCUACGGGUGUAGUACACUUGUAUCCCGGUAC AUUUUGUGCGCCAGUUUUUAUUCUCCUACCCAAUUGUAACUUAAGAAGUUAAGCAACACAAGAUCAA AUGGUAACUAACCAAGUUUAUCAAAGGUCAAGCACUUCUGUUUCCCGGUAAGCUGAUUAGCUCU UAACCAAGGCAGAAGCUAAUCCUACUGUUAUCCGCAAGGUGCCUACGGGAAGCCUUAUUAACAUUG AAAGAAUUUGGUGGUGCGUCAGUUGCGCAAUCCACGCAUAGACCUGGCAGAUUGGGGCUAGACA AACCACACAGCGAUGGUGGUCUAGCCCGGUGGUGCCUGGCACACCCUUUUUGGGUGUGAAGCCA GAAAUUGGACAGGGUGCGAAGAGCCUACUGUGCUCACCUUGAGUCCUCCGGCCCCUGAAUGUGGCUA AUCCUAAACCCUGCAGCUGUUGCAAGAUCCAUUGUGUAUACAGUCGUAUUGAGCAAUUGCGGGAUG GAACCAACUACUUUGGGUGUCCGUGUUUCCUUGUUUUUUAUUAUUGCUUAGUGGACAAUUAUUAU UAUAGAUUAUUGUCAUC	rhinovirusA7 8
4 5	UUAAAAUGAAUCCAGGUUGUCCCAACCCUGGAUUUCCACAGGAGCAGUACUCUACUUAUCCCGUA GUUUUGUACACCAGUUUUUAUUCUCCUCCCGUGUAACUUAAGAAGCAAGCACUUCAGACCAUAGG AUGGUACCUAACCAAGUUUAUCAAAGGUCAAGCACUUCUGUUUCCCGGUAACAAUUGAUUAGCUC CCAGGGCAAAAACAGUUUUGAUCCGUUAUCCGCAAGGUGCCUACGCAAGCUUAGUAACACUUGAAG AAUGUGUGGUUGGUGUCCGCUAUUCCCAUAGUAGACCUGGCAGAUAGAGGUAGAAAUCCUCCAC UGGUGACAGUUGUUCUAGCCUGCGUGGUGCCUGCGCACCCUUUGAGGUGCGAAGCCAUUCCUUGA CAAGGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACC CUGCAGCUAGUGCAUGCAUCCAGCAUGUCUCUAGUCGUAUUGAGUAAUUGCGGGAUGGGACCAACU ACUUUGGGUGUCCGUGUUUCUUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU UAUUAUUGACACC	rhinovirusA8 0
4 4 6	UUAAAAUGGAUCCAGGUUGUCCCAACCCUGGAUCCUACAUGGUGUUGUACACUUAUUUACGGUA AUUUUGUACGCCAGUUUUUAUACCCUCCUCCAAUUGUAACUUAAGAAGCAUACAACUGAUCAUA GGAAGCAUAUCCAGAUUGCUAAGCAUACAGCACUUCUGUCCCGGUAAGGUUGAUUAGCUCC ACAGGGCAAAAACAGCUAAGAUCCGUUAUCCGCAAGGUGCCUACGCAAGCCUAGUAACACUUGAAG AAUGUGUGGUUGGUGUCCGCUAUUCCCAUAGUAGACCUGGCAGAUAGAGGUAGAAAUCCUCCAC UGGUGACAGUUGUUCUAGCCUGCGUGGUGCCUGCGCACCCUUUGAGGUGCGAAGCCAUUCCUUGA CAAGGUGUGAAGAGCCCCGUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCUUAACC CUGCAGCUAGUGCAUGCAUCCAGCAUGUCUCUAGUCGUAUUGAGUAAUUGCGGGAUGGGACCAACU ACUUUGGGUGUCCGUGUUUCUUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU UAUUAUUGACACC	rhinovirusA8 1

	CUCUGGUGACAGAGUUCUAGCCUGCGUGGCGGCCUGCACACCCUAUUGGGUGUGAAGCCAAUUAUUG GACAAGGUGUGAAGAGCCCCGUGUGCUAUCUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACC UUA ACCUGCAGCCAGUGCACACAAACCAGUGUGUAGCUGGUCGUAACGAGCAAUUGCAGGAUGGGACCAA CUACUUUGGGUGUCGUGUUUACCUUUUAUUAUUGCUUAUGGUGACAAUAUAUAUGUAUA UAUAUAUUGUUAUC	
4 4 7	UUAAAACUGGGAGUGAGUUGUCCCCACUCACUCUACCCACGCGGUGUUGUACACUGUUAUUAACGGUA ACUUUGUAUGCCAGUUUUUUGUCCCCUAUCCCCUUUAAAGUAACUUAGAAGCAUGAACUUACAGAC CAAAAGGUGAUGAUCACCCAGAUCAUUAAGGUCAGCACUUCUGUUUCCCCGGUCAUUGUUAUAU GCUUCAAACAGGCAAAACACAGCGAUCGUUAUCCGCAAGAUCCUACACAAAGCUUAGUACCUCC UUUGAUGGUUAUUGUCUGGUCGUCAGGUGUUUCCCCAACACUAGACCUGGCAUGAGGCUAGAAA CUCCCCACUGGCGACAGUUGUAGCCUGGCGUGGCGGCCUGCACACCCUUUCGGGUGUGAAGCCAAA UACUGGACAAGGUGCGAAGAGCCCCGUGGUCUACUUCUAGUCCUCCGGCCCCUGAAUGCGGCUAAC CUUAAACCCCGAGCCAUUACUCACAUCAGUGAGCAGAUGGUCGUAUAGAGCAAUUGUGGGAUGGG ACCGACUACUUUGGGUGUCCGUGUUUACUGUUUUAUUGACUUGCUUAUGGUGACAAUAUAUAG AGUAUAUAUUGGCAUC	rhinovirusA8 2
4 4 8	UUAAAACUGGAUCCAGGUUGUCCCCACCUGGAUUUCCUACAUGGAGUUGUACUCUGUUAUUCGGUA ACCUUGUACGCCAGUUUUUAUACCCUCCCCUUGCAACUUAAGAAGUUAAGCACGCGACCAUAGG UGGUGAUUAUCCAAUACUAAACGGUCAAGCACUUCUGUUUCCCCGGUUUAAACUGAUACGCUCCAA CAGGGCGAAAACAGUGCAAUUCGUUAUCCGCAAGGUGCCUACACAAAGCUUAGUAGCGCUUUGGAGA AUGUGUGGUUGGUCGUCAGCUUACCCCCUAGUAGACCUGGCGAUGAGGCUAGAGCACCCCCACU GGCGACAGUGUUCUAGCCUGCGUGGCGUGCCUGCACACCCUUUUGGUGUGAAGCCAUACAUAUUGACA AGGUGUGAAGAGCCCCGUGGUCACUUUGAGUCCUCCGGCCCCUGAAUGGCGUAACCUUAACCU GCAGCUGGUGCAUGUAUCCAACAUUGUUGCUAGUCGUAUAGAGCAAUUGCGGGAUGGGACCAACUAC UUUGGGUGUCCGUGUUCACUUUUUACCUUUUAUUGCUUAUGGUGACAAUAUAUACACAUAU AUAUUGGCAUC	rhinovirusA8 5
4 4 9	UUAAAACUGGGAGUGAGUUGUCCCCACUCACUCUACCCACGCGGUGUUGUACACUGUUAUUAACGGUA ACUUUGUAUGCCAGUUUUUUGUCCCCUAUCCCCUUUAAAGUAACUUAGAAGCAUGAACUUACAGAC CAAAAGGUGAUGAUCACCCAGAUCAUUAAGGUCAGCACUUCUGUUUCCCCGGUCAUAGAGGAUAU GCUUUACCCAAAGGCAAAACCUUAAAGAUCCGUUAUCCGCAACUUAACGCAAGCCUAGUAUAGCG CUUUGAUGUAUUAAGACUGGUCGUUCCUGCAUACCCAGCAGUAGACCUGGCGAUGAAGCUGGACG UUCCCCCACUGGCGACAGUGGUCAGCCUGCGUGGCGUGCCUGCUACCCAUUUGGGUGAGAAGCCUUA UUUUGACAAGGUGUGAAGAGCCCCGUGGUCUUAUUGUGAAUCCUCCGGCCCCUGAAUGGCGUAACCU CUUAAACCCCGAGCCGUUACUCACAUCAGUAAUAGCGGUCGUAACGCGAAAGUGUGGGAUGGG ACCAACUACUUUGGGUGUCCGUGUUUCCUUAUUUUUCCUUAUUGUGUUAUAUGGUGACAAUUUAU AGCGUAUAAGUUGCAUC	rhinovirusA8 8
4 5 0	UUAAAACUGGGAGUGGUGUUGUCCCCACUCACUCCACCCAUGCGGUGUUGUACUCUGUUAUUAACGGUA ACUUUGUACGCCAGUUUUUCCCAACCUUCCCCAUAAUGUAACUUAGAAGUUUGUACAAUAUGACCAA UAGGUGACAUAUCACAGACUGUCAAAGGUCAGCACUUCUGUUUCCCCGGUCAUAGAGGAUAUGCU UUACCCAAGGCAAAACCUUAGAGAUCCGUUAUCCCCACACUGCCUACACAGAGCCAGUACCAUUUU UGAUUAUUAUUGGGUUGGUCGUCUCCUGCAAACCCAGCAGUAGACCUGGCGAUGAGGCUAGACAUUC CCACUUGGCGACAGUGGUCAGCCUGCGUGGCGUGCCUGCUACCCUUCUUGGGUGAGAAGCCUAAUU AUUGACAAGGUGUGAAGAGCCCGUGGUCUAGUGGUCUCCUCCGGCCCCUGAAUGGCGUAACCU UAACCCUCCGAGCCGUUGCCAUAAUCCAUGGGUUUGCGGUCGUAUAGCGUAAGUGCGGGAUGGGAC CAACUACUUUGGGUGUCCGUGUUUCCUGUUUUUUCUUUGAUUGCAUUUAUUGGUGACAAUUUAUAGU GUUAUAGAUUGCAUC	rhinovirusA8 9
4 5 1	UUAAAACUGGGUUCAGGUUGUCCCCACCUGAAUCUCCUACAUGGAGUUGUACUCUUAUUAUCCCGUA AUUUUGUACGCCAGUUUUUAUCCUCCCCUCCCAUGUAACUUAGAAGCUGUAACAACGACCAACAGG UAACAUUACUCCAGAGUGUUAAGGUCAGCACUUCUGUUUCCCCGGUCAAAAGUUGAUUAGCUCCAA CAGGGCAAAACAAUCGCGAUCGUUAUCCGCAAGCAACUACGCAAGCUUAGUACCAUUAUUAUG GUGUAUGGCGUGGUCUCCACUAUAAACCCAUAUAGUAGACCUGGCGAUGAGGCUAGAAACUCCCCA CUGGUAAACAGUGUUCUAGCCUGCGUGGCGUGCCUGCGUCCUUUUGGGCACGAAGCCUUACAUAUUGA CAAGGUGCGAAGAGCCCGUGUGCUUGCUUUGAGUCCUCCGGCCCCUGAAUGGCGUAACCUUAACC CUGCAGCUAGUGUGGCAUUCAGCACAUUGGCUAGUCGUAUAGAGCAAUUGCGGGGAUGGGACCAACU ACUUUGGGUGUCCGUGUUUCCUUUUUUUUAUUAUUGCUUAUGGUGACAAUAUAUAUAGAAUAU AUAUAUUGGCAUC	rhinovirusA9 0
4 5 2	UUCAAACUGGAUCUGGGUUGUCCCCACCAGAUACCCAAAGAGGUGUAGUGCUCUGUCAUUCGGUG ACCUUGCACGCCAGUUUUUAUACCCCUUCCCUUGUAACUUAGAAGAAUUAACUCAGACCAAUAG GAAGCGACAAGCCAGUCGUGACAGGUCAAGCACUUCUGUUUCCCCGGUCAAAAGUUGAUUAGCUUCAC CAAGGCAAAACAAACAGCGAUCGUUAUCCGCAAGAUCCUACGCAAGCCUAGUAGUCCUUUGAAA AUUAUUUGUCUGGUGCUCAGUAACAACCCAGUACUAACCCUGGCGAUGAGGCUAGGAGCUUCCCCA CUGGUAAACAGUGUUCUAGCCUGCGUGGCGUGCCUGCGUCCUUUUGGGCACGAAGCCUUACAUAUUGA CAAGGUGCGAAGAGCCCGUGUGCUUGCUUUGAGUCCUCCGGCCCCUGAAUGGCGUAACCUUAACC CUGCAGCUAGUGUGGCAUUCAGCACAUUGGCUAGUCGUAUAGAGCAAUUGCGGGGAUGGGACCAACU ACUUUGGGUGUCCGUGUUUCCUUUUUUUUAUUAUUGCUUAUGGUGACAAUAUAUAUAGAAUAU AUAUAUUGGCAUC	rhinovirusA9 4
4 5 3	UUAAAACUGGGUCUGGGUUGUCCCCACCUAGAUCACCCAAAGUGGUGUAGUACACUGUUAUUCGGUA ACCUUGUACGCCAGUUUUUAUACCCCUCCCCAAUUGUAACUUAGAAGCUGUACACAACGACCAUAU GAGGGUAUUAUUAUCCAGAUUACUACGGUCAAGCACUUCUGUCCUCCCCGGUUUACUUUAUUGCUCC AACAGGGCAAAACAAAGUAUAGUUAUCCGCAAGAUCCUACGCAAGCCUAGUAAACUCCUGAA	rhinovirusA9 6

214

[illegible]

	ACCCUAAACCCCGAGCCACGCGGCAUAAUCCAAGUGUCGUAAGGUGCUAAUAGUAAUUCUGGGGAGG GGACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUCCAAAUAUUCUUAUGGUCACAGCAUAAUA AUAGAUAUAUACUGUGAUC	
4 6 8	UUUAAACAGCGGAUGGGUAUCCACCAUCCGACCCACUGGGUGUAGUACUCUGGUAAUUUGUACCUU UGUACGCCUGUUUCCUAAUGUACCCACCCUAAAACUUCUACCCAAAGUAACGUUAGAAGUUUCAU CAACAAGUACAAUAGGAAGCAUACAUCCAGUGGUGUUUUGUACAAGCACUUCUGUUUCCCGGAGC GAGGUUAAGGUGUACCCACUGCCGAAAGCCUUUAAACCGUUUACCGCCAAACACUACGUAAAAGCU AGUAUCAUGUUAUAAAAGGCGUUCGAUCAGGUGGAUCCCCCUCCACUAGUUUGGUGCAUGA GGCUAGGAACUCCCCACGGGUGACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCUUCUGUGGGA CGCCUUUUUAUGGACAUGGUGUGAAGACUCGCAUGUGCUUGGUUGUGACUCCUCCGGCCCCUGAAUG CGGCUAACCUUAAACCCCGAGCCUGUGUUGCAAUCCAGUAACAUUAGGGUCGUAAUGAGCAAUUC GGGACGGGACCGACUACUUUGGGUGUCCGUGUUUCUAAUUUUUUAUUUAUUGUCUUAUGGUCACAG CAUUAUAUAACGUUAUACUGUGAUC	rhinovirusB3 7
4 6 9	UUAAAACAGCGGGUGGGUAUCCACACCCGACCCACUGGGUGUGGUACACUGGUACUUUGUACCUU UGUAUGCCUGUUUGCAUCCCUUCCUUUACCUAACAUCUCCUCCCAAGUUUAUUUUUGCGGUAAAG UUAGAAGAAGUGAACaUAGUGCAACAGGAAGUAUACACCACAGUGAUUAACGCACAAGCAAUUCUU GUUCCCCGAGCUGGAUAUAGACUGCUAACGUGGUUGAAAGUCCUUAACCGUUUACCGCCAGUUAAC UGCGAAAAGACUAGUAACAUCAUGCACAUUUCUAGGUGUUCGAUCAGGUGAAACCCCAUUCACUAG UUUGGUGCAUGAGGUGAGAGUACCCACAGGCGACUGUGUCUACGCGUGCGUGGCGGCCAACCCGA CUUAUGUCGGGACGCUUAUUAUAGACAUGGUGUGAAGUCCUUAUUGCGCUUAGCUGUGAGUCCUCC GGCCCCUGAAUGCGGCUAACCUUAAACCCCGAGUCUUGUGAUGUAAGCCAAAUUUGCAAGGUCGUA AUGAGUAUUUCUGGGAUGGAGCCGACUACUUUGGGUGUCCGUGUUUCUAAUUUUUUAUUGUCUUAUGGUG UUUCAUGGUUACAAGCAUUGUUUGUAAUC	rhinovirusB4 2
4 7 0	UUAAAACAGCGGAUGGGUAUCCACCAUCCGACCCACUGGGUGUAGUACUCUGGUACUCUGUACCUU UGUACGCCUGUUUCUCCUACACCCUCCCCUACCCGUUAUAUUCUCAGUAACUUAAGAAGUAUUAU ACCGCACAAUAGGGACCGUUGUAACACGCAACGGUCAGUGCAAGCACUUCUGUUUCCCGAGCGGA GUUAUAGACGGCCACUGCUGUCAAAGCUCUUAACCGUUAACCGCCAGUUAACUACACAACAGCUAGU AACAUUCUGUUGUUCUUAAGGCGUUCGAUCAGGUGAGUUUACCCCCUACUAGUCUGGUGCAUGAGG UGAGAGAAUCCCAACGGGUGACCGUGUCUACGCCUGCGUGGCGGCCAGCCAGCUUAUGCUGGGACG CCUUAAGUAUGACAUGGUGUGAAGACUCGCGUGUGCUUAGCUGUGAGUCCUCCGGCCCCUGAAUGCG GCUAACCUAAACCCCGGAGCCUUGAGGCACAAUCCAGUGCCCGGAGGUCGUAACGAGAAUUCUGG GACGGGACCGACUACUUUGGGUGUCCGUGUUCCUUUAUUCUUUUGAUGUUCUUAUGGUUACAAUA UAUAAAGUAUAUUGUAAUC	rhinovirusB4 8
4 7 1	UUAAAACAGCGGAUGGGUAUCCACCAUCCGACCCACUGGGUGUAGUACUCUGGUAAUUAUGUACCUU UGUACGCCUGUUUCUCCCCCCCCAUUACCCUCCCCCAUUAUCCAGUAACUUAAGAAGAAUUAUUAU UUGCAGCAAAUAGGAGCUUAUACAACCCAGUAUAUAGUUAAGUUAAGUUAUUCUUGUCCCGGAGCGGAG UAUAGAUGGCCACCGCUGUCAAAAGCUCUUAACCGUUAUCCGCCAGUUAACUACGUAAACGGCUAGUA ACAUCUUUGUUGUCUUAAGGUGUUCGAUCAGGUGAAGAGACACCCCCUACUAGUCUGGUGCAUGA GGCUGAGAAAUCCCCACGGGCGACCGUGUCUACGCCUGCGUGGCGGCCAGCCAGCUUAUGCUGGGA CGCCUUAAGUAUGACAUGGUGUGAAGACCCGCAUGUGCUUAACUGUGAGUCCUCCGGCCCCUGAAUG CGGCUAACCUUAACCCUGGAGCCUUGGGGUAACAUCCAGUGCUAACAGGUCGUAACGAGCAAUUC GGGACGGGACCGACUACUUUGGGUGUCCGUGUUCCUUUAUUUUAUUAUUGUUCUUAUGGUUACAA UAUAUAGAGUAUAUUGUAAUC	rhinovirusB5 2
4 7 2	UUAAAACAGCGGAUGGGUAUCCACCAUCCGACCCACUGGGUGUAGUACUCUGGUAAUUAUGUACCUU UGUACGCCUGUUUCUCCCCUCCUCCUCCAAACCAUUAACUUAUAGAAAGUUUUGAUUAUUC GCACAAUAGGGACUUAUUGCAACACAGCAUAGUACAGUACUUCUGUUUCCCGGAGCGGAGUG UAGAUGGCCACCGCUGUCAAAAAGCUCUUAACCGUUAUCCGCCAGUUAACUACGCAACGGCUAGUA GCCUUGUGCUUUUAUGGCGUUCGAUCAGGCGAGAUUAACCCUCCGCUAGUCUGGUGAUGAGGCU GAGAAUUCGCCACGGGCGACCGUGUCUACGCCUGCGUGGCGGCCAGCCAGCCUUAUGCUGGGAGCC UUGAUUUGCGACAUGGUGUGAAGACCCGCGUGUGCUUAACUGUGAGUCCUCCGGCCCCUGAAUGCGGC UAACCUAAACCCUGGAGCCUUGAAACACAACCCAGUGUUCGCAAGGUCGUAACGAGCAAUUCGGGA CGGACCGACUACUUAUGGGUGUCCGUGUUCCUUUAUUCUUUCAAUUGUUCUUAUGGUCACAAUA AGUGUUAUAUUGUGAUC	rhinovirusB6 9
4 7 3	UUAAAACAGCGGAUGGGUUCUCCACCAUCCGACCCAAUGGGUGUAGUACUCUGGUAAUUCUGUACCUU UGUACGCCUGUUUCUCCCCACCCUCCUCCUCCUAAUUAUCCAGUAACUUAAGAAGUUUUGAAUUAUC GUACAAUAGGUGCUGUCAUACCAUGGAGUAGUACAGCACUUCUGUUUCCCGGAGCGGAGUA UAGAUGGCCACCGCUGUCAAAAAGCUCUUAACCGUUAUCCGCCAAUUAACUACGCAACGGUUAAC AUCUUGUCAUUUUAGAGCGUUCGAUCAGGUGAGUACACCCUACUAGUCUGGUGCAUGAGGCUA GAAUUCUCCACGGGCGACCGUGUCUACGCCUGCGUGGCGGCCAGCCAGCUUAUGCUGGGAGCCU AAUUGUGACAUUGGCGAAGACCCAGUGUGCUAAUUGUGAGUCCUCCGGCCCCGAAUGCGGCUA ACCUAAACCCCGGAGCCUUGAGACACAUCCAGUGUUAAGCAAGGUCGUAACGAGCAAUUCGGGA CGGACCGACUACUUAUGGGUGUCCGUGUUCCUUUAUUCUUUCAAUUGUUCUUAUGGUCACAAUA AGUGUUAUAUUGUGAUC	rhinovirusB7 0
4 7 4	UUAAAACAGCGGAUGGGUUCUCCACCAUCCGACCCACUGGGUGUAGUACUCUGGUAAUGAUGUACCU UUGUACGCCUGUUUAUUCUCCAAACCCUCCCGUUAUUCGCCCCAACCCUGUAACGUUAAGAAGUUUG GAACUUAUUGUACAAUAGGGAGCACCAUUAUCCAAUGGUGUAACGUAAAGCACUUCUGUUUCCCGG GAGCGAGGUUAAGGUGGUACCCACUGCCGAAAGCCUUUAAACCGUACUCCGCCAAUUAACUACGCA AGCUUAGGUGUACUUCUUGUUGUAGUAGGACGUUCGAUCAGGUGUCCUACAGGUGUCCUCCGGCCCC GGACCGACUACUUAUGGGUGUCCGUGUUCCUUAUUCUCAAUUGUUCUUAUGGUCACAAUAUAG UAAUUAUAUAUUGUGAUC	rhinovirusB7 2

	GGACGCCUUUCAAUGACAUGGUGAAGACUCGCAUGUGCUUGAUUGUGAGUCCGCCGCCCCUGA AUGCGGCUAACCCUAACCCUGGAGCCUUGCACCACAAUCCAGUGGUGUCUGGUCGUAAUGAGUAAU UCUGGGUAGGACCGACUACUUUGGGUGUCCGUGUUUCUAAUUUUUUAAUUGUCUUUAGGUGCA CAGCAUUAUGAUUAUUAUACUGUGAUC	
4 7 5	UUAAAAACAGCGGACGGGUAUCCCAUCCGGCUCACUGAGCGUAGUACUCUGGUAUUGUGUACCUU UGUACGCCUGUUUUCCCCAAAUCCCUACCCUUUAUUCUCCUACCCCGUAACGUAAAGAUUUUCA AACAGUACAAUAGGAAGCACCACAUCCAGUGGUGUUAUGUACAAGCACUUCUGUUUUCCCCGGAGCGA GGUUAUAGGUGGUCACCACCGCCAAAAGCCUUUAACCGUUAACCGCCAAACUACGUAAUGGCUAG UAACAUCAUUGAUUUAGAUUGGCGUUCGAUCAGGUGGUAUCCCCCACUAGUUUGGUCGAUGAGGCU AGGAAUACCCACGGGUGACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCAUACCGUGGGACGCC AGUUUACUGACAUGGUGCGAAGACUUGCAUGUGCUUGGUUAGAGUCCUCCGGCCCCUGAAUGCGGC UAACCUUAACCCCGAGCCUCGCAGCAUAAUCCAAUGUUUGAGAGGUCGUAAACGAGUAAUUCUGGGA UGGACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUUUUAAAAUGUCUUUAGGUCACAGCAUA UAUACUGUAUAUACUGUGAUC	rhinovirusB7 9
4 7 6	UUAAAAACAGCGGAGUGGUAUCCCAUCCGACCCACUGGGUGUAGUACUCUGGUAUUUUGUACCUU UGUACGCCUGUUUUUCCCUUUGUACCCGCCCUUAAACUCCACCCACGUAAACGUUAGAAGUUUA UACAAGCAGUACAAUAGGUUGCACCACUAAACAGUGGUGUUUCGUACAAGCACUUCUGUUUCCCGGA GCGGGGUAUAGACGGUACCCACUGUCAAAAAGCCCUUAAACGUUAUCCGCCAAUACUACGUAAUGG CUAGUAACACCAUGAAGGUAGUAGCUUCGAUCAGGUGGUAACCCACUAGUUUGGUCGAUGAG GCUAGGAUUUCCCCACGGUAACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCAUUUGUGGAC GCCAAUUUACUGACAUGGUGUGAAGACUUGCAUGUGCUUGAUUGUAGUCCUCCGGCCCCUGAAUGC GCUAAACCUUAACCCCGAGCCUUGCAGCAUAAUCCAAUGUUUGGAGGUCGUAAUGAGCAAUUCUG GGUUGGACCGACUACUUUGGGUGUCCGUGUUUCCUGUUUUUUUAAAGUUGUCUUUAGGUCACAGCA UAUUAUACUUAUUAUACUGUGAUC	rhinovirusB8 3
4 7 7	UUAAAAACAGCGGAUGGGUAUCCCAUCCGCCACUGGGCGUAGUACACUGGUAUUUUGUACCUU UGUACGCCUGUUUUUGCGCCCCUCCCCUUUUUAUACACAUUUUGUGGUCGUGCAACGUUAGAAGAU UUUAAAGUAGUGCAAUAGGAAGCAUCACAACAGUGGUGUUUAGCACAAACACUUCUGUUUCCCGG AGCGAGGUUAGACGGUCCCCACUGUCAAAAAGCCUUUAACCGUUAUCCGCCAACCAACUACGAGACA CUAGUAUAGUCCCUUGUUCUCCUGCAAGGCGUUCGAUCAGACAGAUUUCCCUUGUUAUGUCUGGUCG AUGAGGCUAGAAGUACCCACAGGUAAUCUGUUCUAGCCUGCGUGGCGGCCAGCCAGCUUCUGCU GGGACGCCUUUCAAAGAGACAUGGUGUGAAGACUCAAUUGUGCUUGGUUGUGAGUCCUCCGGCCCCUG AAUGCGGCUAACCUUAAACCCGGAUCCAUUGCUAUGCAAACAGCAUAGUUUAGGUCGUAAUGGGCAA CUUGGGAUGGGACCGACUACUUUGGGUGUCCGUGUUUCCAUUUUUAUUAUUGGUCAUUUUAGGU UACAAUUAAAGUUGUAACC	rhinovirusB8 4
4 7 8	UUAAAAACAGCGGAUGGGUAUCCCAUCCGACCCACUGGGUGUAGUACACUGGUAUUUUGUACCUU UGUACGCCUGUUUUUCCCUUUAUUGUACCUUCCCUCAUCAAUCCUAAACCCUGUAACGUUAGAAGUG CAACUCUGUGCAAUAGGAGAUACUGCAACAGUGGUAUCAAGCACAGAACUUCUGUUUCCUCGGAGC GGAGUAUAGAUUGGUCCCCACCGUCAAAGGCUUUAACCGUUAUCCGCCAAUACUACGUGACGACU GGUAGCAUCAUUUAGACAUUUAGCGUUCGAUCAGGUAGGCUUACCCACCUACUAGUUUGGCCGAU GAGGCUGGGAGUCCCUACAGGUAAUCUGUGUCCAGCCUGCGUGGCGGCCAACCCAGCAAUUGGUGG GACGCCAAUUUUGUAGCAUGGUGUGAAGAUUUAAUGUGCUUGGUUGUGAGUCCUCCGGCCCCUGAA UGCGGCUAACCUUAAACCCCGAGCCUUGUGUCACAAGCCAGUGACAUUAAGGUCGUAAUGAGUAAU UCGGGAUGGGACCGACUACUUUGGGUGUCCGUGUUUCCAUUUUUAUGAUGUUGCUUAUGGUA CAGCAUUAUGUAUUUAUUAUACUGUGAUC	rhinovirusB8 6
4 9 7	UUAAAAACAGCGGAUGGGUUUCCCAUCCGACCCACUGGGUGUAGUACUCUGGUAUUUUGUACCUU UGUACGCCUGUUUUUCCCUCCCUCCCAACCUAUAACAACCCGGCAACUUAAGAAGAUUUGAACUAAU GUACAAUAGGCGUGUCAUCCAGUGGCGGCUAGUACAAGCACUUCUGUUUUCCCCGGAGCGGAGUA UAGAUGGCCACCGCUGUCAAAAGCUCUUAACCGUUAUCCGCCAAUUAACUACGCAACGGCUAGUAAU AUCAUGUAACUUUAGGGUGUUCGAUCAGGUGAGUGUACCCCUACUAGUCUGGUCGAUGAGGCUGA GAUUUCCCCACGGGCGACCGUGUCUACGCCUGCGUGGCGGCCAGCCAGCUUAUGCUGGGACGCCUU AAUUGUGACAUGGUGUGAAGACCCACGUGUGCUUAAUUGUGAGUCCUCCGGCCCCUGAAUGCGGCUA ACCUAAACCCUGGAGCCUUGAAACACAUAUCCAGUGUUAGCAAGGUCGUAAUAGAGUAAUUCGGGACG GGACCGACUACUUUGGGUGUCCGUGUUUCCUUUUAUUUUUAAAGUUGUCUUUAGGUCACAAUUAAG UAAUUAUUAUUGUGAUC	rhinovirusB9 1
4 8 0	UUAAAAACAGUGGAUGGGUUUCCCAUCCGACCCACUGGGUGUAGUGCUCUGGUAUUUUGUACCUU UGCACGCCUGUUUUUCCCAUUGUACCCUCCUUAUUAUCCUCCCAUGUAACGUUAGAAGUUUAAGAA CAUAAUAGUACAAUAGGAAGCAUCCAGUGGUGUAUGGUACAAGCACUUCUGUUUAUCCCGGAG CGAGGUUAAGCGGUACCCACUGCUGAAAGCCUUUAACCGUUAACCGCCAAUACUACGUAAACGGC UAGUAUUUACCUUGUUUGCAAUCUGGCGUUCGAUCAGGUGGUUCCCCCACUAGUUUGGUCGAUGAGG CUAGGAGUUCUCCACGGGCGACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCAUUUUGCGGGACG CCUUUUUUGUAGACAUGGUGUGAAGACCUAGCAUGUGCUUGAUUGUGAGUCCUCCGGCCCCUGAAUGCG GCUAACCUUAAACCCCGAGCCUUGUGACAUAUCCAAUGUCAUUAAGGUCGUAAUGAGCAAUUCUGG GACGGGACCGACUACUUUGGGUGUCCGUGUUUCCUUUUAUUUUUAAAGUUGUCUUUAGGUCACAAUUAAG UAUUAUUAUUAUUGUGAUC	rhinovirusB9 3
4 8 1	UUAAAAACAGCGGAUGGGUUUCCCAUCCGACCCACUGGGUGUAGUGCUCUGGUAUUUUGUACCUU UGCACGCCUGUUUUUCCCUUUGUACCCAUCCGAAUUAUCCUCCCUUGCAACGUUAGAAGUUUGUGAA AUUAAAGUACAAUAGGAAGCAUCACAUCCAGUGGUGUUCAGUACAAGCACUCCUGUUUUCCCCGGAG CGAGGUUAAGCGGUACCCACUGCUGAAAGCCUUUAACCGUUAACCGCCAAUACUACGUAAACGGC UAGUAUUUACCUUGUUUGCAAUCUGGCGUUCGAUCAGGUGGUUCCCCCACUAGUUUGGUCGAUGAGG CUAGGAGUUCUCCACGGGCGACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCAUUUUGCGGGACG CCUUUUUUGUAGACAUGGUGUGAAGACCUAGCAUGUGCUUGAUUGUGAGUCCUCCGGCCCCUGAAUGCG GCUAACCUUAAACCCCGAGCCUUGUGACAUAUCCAAUGUCAUUAAGGUCGUAAUGAGCAAUUCUGG GACGGGACCGACUACUUUGGGUGUCCGUGUUUCCUUUUAUUUUUAAAGUUGUCUUUAGGUCACAAUUAAG UAUUAUUAUUAUUGUGAUC	rhinovirusB9 7

1	CUAGGAUUUCCCCACGGGCGCAGCGUGUCCUAGCCUGCGUGGCGGCCAGCCACAGCUUAUUGCUGGGACG CCUUUUUAAAGACAUUGGUGUGAAGACCUGCAUGUGCUUGAUUUGAGUCCUCCGGCCCCUGAAUGCG GCUAACCUUAACCCUGGAGGCCCGACAGCAUAUCCAAUGUUGUUUGGGUGCUAAUGAGGCAAUCCCG GAUGGGACCGACUACUUUGGUGGCCGUGUUUUUUUAUUCUUAUUGUCUUAUGGUCACAGCAU AUUAGUAUAUAUACUGUGAUC	
4	UUAAACAGCGGGUGGGUUAUCCACACCCCGACCCACUGGGUGUAGUACACUGGUUAUUUGUACCUU 8 UUAUAGCCUGUUUGGCAUACCCUCCUCCUCCUUAACCUUACCUUUAUUUUGGCUAAUUUUAGAG 2 UUUUUAUAUACAGCGCAAUAGGGUAUACCCAUCCAGUGGUUAUGAAGCACAAGCAAUUCUUGUCCCC CGGAGCUGGUAUAUAAACUGCUAACGUGGUUUAAGAGUCCUUAACCGGUUAUCCGGCCAGUUAACUGCGGAG AAGACUAGUAAACCAUGUUUUUUUUUAAGCGUCCGAUCAGGUGGAUACCCCAUCCAUAGUUUGGU CGAUAGGCGUGGGAAUCCCCACAGGUAAACUGUGUCCAGCCUGCGGUGCGGCCAACCCGACCUUUUG CUGGGAGCGCUUACUGACAGACAUGGUGUGAAGAUCUUAUUGCGCUUAACUGUGAGUCCUCCGGCCCC UGAAUGCGGCUAACCCUCAACCCCGGAGCCUUGUGGUGUAACCAACACUCGCAAGGUCGUAAACGAGU AAUUUCUGGGACGGACCAUACUUUGGUGUCCGUGUUUCCUUAUUUUGUUAUUGAUAUGUCCCUA UGGUUACAAAUUACUGUAAUC	rhinovirusB9 9
4	UCCCCUCCACUAGUUUGGUGCAUGAGGCUAGGAAUCCCCACGGGUGACCCGUGUCCUAGCCUGCGUG 8 CGGGCCAACCCAGCAUACGUGGGAGCGCCUUUUCAGAGACUUGGUGUAGAGACUGCAUGUGCUUGA 3 UUGUGAUUCCUCCGGCCCCUUAUUGCGUUUUAACUACCCCGGAGCCUUGUAUUGCAAUCCAGCAA UAUUAAAGGUCGUAAUGAGCAAUUCUGGGACGGGACCGACUACUUGGGUGUGCCGUGUUUCCUGUUUU UCAUAAAUUUUUUCUUAUGGUCACAGCAUAUAGUACAUAUAUACUGUGAUC	rhinovirusB1 01
4	AAAAAGCGGAUGGGUUAUCCCAUCCACCCGACCCACUGGGUGUAGUACUCCUGUACUUUGUACCUUG 8 UACGCCUGUUCUCCCCAUUGUACCCUCCCUUAACUCCAAACCAAGCAACAUAAGAAGUUAUCAAUU 4 UUCAGUACAUAAGAAGACUACAUCCAGUGAUGUUGUACAAGCACUUCUUCUCCCGGAGCGCA GGUAUAAGCAGUAGCCACUGCUAAAGCCUUUAACCGUUAACCGCCAUAACUACGUAAGCUGAG UAUCAUAGUUUUUUGACUGGACGUUCGAUCAGGUGGUUCCCCCCACUAGUUUGGUGCAUGAGGCU AGGAUAACCCACGGGUUACCGUGUCCUAGCCUGCGUGGGCGGCCAACCCAGCAUUUGCUGGGACGCC CUUUUAAAGACAUGGUGUGAAGACUCGCAUGUGCUUGAUUGUGAUUCCUCCGGCCCCUGAAUGCGGC UAACCUUUAACCCCGGAGCCUUGUGUCAAUCCAGCAAAAGCAAGGUCGUAUAGCAAUUUCUGGGA CGGACCGCAUACUUUGGUGUCCGUGUUUCUCAAUUUAACAUAUACUUUUCUUAUGGUCACAGCAUAU AUAAUAAAUUACUGUGAUC	rhinovirusB1 02
4	UCUGCGUGUGCAUCAGGUGGAUUUCCCCUCCACUAGUUUGGUGCAUGAGGCUAGGAAUCCCUACG 8 GCGACCGUGUCCUAGCCUGCGUGGCGGCCAACCCAGCAUAUGCUGGGACGCCCCUUUAUAGACAUGG 5 UGUGAAGACUCGCAUGUGCUUGGUUGUGAUUCCUCCGGCCCCUGAAUGCGGCUAACCUUAAACCCUGG AGCCUUGUAUUAACAACCAUGAAUAUUAAGGUCGUAAUGAGCAAUCCCGGACGGGACCGACUACUU UGGGUGUCCGUGUUUUAUUUAUUAUUAUUGUCUUAUGGUCACAGCAUAUUGUUAUUAUACU GUGAUC	rhinovirusB1 03
4	UUA AAAACAGCGGAUGGGUUAUCCACCAUCCGAGCCACUGGGUGUAGUACUCUGGUUAUUAUGUACCUU 8 UUAUGCCUGUUUUUCCUCCUCCUACUACCCCGCCUACAGUACCUUGAUAUUAAGAAUUGUAAAA 6 CCGCACAAACAGGAGCUAUUACAACAGUAUAUAGUUAUGGCAAGCACUUCUGUUUCCCGGAGCGGAG UAUAGACGGCCACCGCUGUCAAAAGCUCUUAACCGUUAUCCGCCAUUAACUACGUAACGGCUAGUA ACAUUCUUGUGUCUUAAGGGUGUCCGAUCAGGUGAAGAAAAACCCCUACUACUGUUGGUCGUA GGCUGAGAAGUCCCGCCAGGCGACCGUGUCUACGCCUGCGGCGGCCAGCCAGCAUAUUGCUGGGA CGCCUUAAGUAUGCAUGGUGUGAAGACCCGCAUGUGCUUAACUGUGAGUCCUCCGGCCCCUGAAUG CGGCUAACCUUAAACCCUGGAGCCUUGGAGCACAAGCCAGUGUCUUAACAGGUCGUAUAGGACCAAUCC GGGACGGGACCGACUACUUUGGGUGUCCGUGUUUCCUUAUUAUGUUAUUGUUAUUGUUAUAGCA UAUAUAGAGUAUAUUGUAAUC	rhinovirusB1 04
4	GUGACAAUAGCCAGAUUGUUAACGGUCAAGCACUUCUGUUUCCCGGUACCCUUGUAUACGCUUA 8 CCCGAGGCGAAAAGUGAGGUUAUGUUUACCGCAAAGUGCCUACGAGAAGCCUAGUAGCACUUAU 7 AGCCUUAUGGUGUGGUCGUUAACUGUUUACCCAGGACCCUGGAGUAGGCGUAGAGUGUUCUCC ACCAAGCGAUGGUGAUCUAGCCUGCGUGGCGUGCCUGCACACUCUUAUUGAGUGUGAAGCCAGAAUGGG ACAAGGUGUGAAGAGCCUAUUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAAUCCUAA CCCGUAGCUGUGCAUGUAUAAUCCAAACAGUUCUGCAUGCUGUAUUGGCAUUGGGAUUGGAACCA UACUUUGGGUGUCCGUGUUUCUGUUUUUUCUUAUGCUGUCUUAUGGUGACAACUGUAGUUAUUA UUUGUUAACC	rhinovirusC1
4	ACCGCAGCCAGUUGGCAACCCCUAACCAUCCGUAACUUAAGAAGCUUAACUFCGACCAAUAGGCGG 8 UGGCAACCAUACCAUUAACGGUCAAGCACUCCUGUUUCCCGGUUAGCGAGGAAUAGACUCCUACA GGGUGAGAGCCUAGUAUUGCAUCCGCAUUGGUUACGCAAGCUUAUGAGUGCCUUGAAAGUC CCUUGUUGGUGCGUCCUGCUAGUUUCCUCCUAGUAGACUUGGCAGAUAGGCGAGGACAUUCCCA CGCAAGUGGUGUCCUGCGUGGUGGCGUGCGCACCCUAGGGGUGUGCAAGCAAGUAGGACGACAA GGUGUGAAGAGCCCCGUGUGCUACUUAUAGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCAAACCC AGCUAUUGCACACAGCCAGUGUGUAUAUAGUGCUAAUAGAGCAAUUGUGGGACGGAACCGACUACUU UGGGUGUCCGUGUUUCCUUUAUUCUUAUCAUUCUGCUUAUGGUGACAUAUACUGUGAAUAGUGUUG UUACC	rhinovirusC2
4	UUAAGCUGGAUCAUGGUUGUCCCAACCAUGAUUACCCACGCGGUGACAGGUGUCUUGUAUUACGGUA 8 CAUUUCCUACUACCAUUAUACACCCACCCGAAACUUAUAGAAGUUUUGUACCAAUAGACCAAUAG 9 GUUUGGCCUAACAGGUGCUUAAGUGUCAAGCAUUCUGUUUCCCGGACCCUUGUAAGCUUUA CCCGAGGCGAAAAAUAGGUGUGUCGUUAUCCGCAAAGUGCCUACGAAAAGCCUAGUAACACUUUGAA AACCCAUGGUUGGUGCGCUACAGCUUUUCCCAACAGUAGACUCCGCGAGUAGGCGUAGACUAUCC ACCAAGGAGGUGGUGUCCUAGCCUGGUGGCGUGGCGUCCAGCCUCCGCGGUGUAGGCGAGGAAUGG GUG	rhinovirusC3

	ACAAAGGUGUGAAGAGCCUAUUGUCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAACCCUAAAC CCCGUAGCUGUUGCAUGUAACCCAAACAUUAUGCAGUCGUAUUGGGCAACUAUGGGAUGGGACCAAC UACUUUGGGUGGCCGUGUUUCCUGUUUUACUUUUUACUUUUUACUUUGGUGACAAUUGUAUCUGAUACA CUUGUUUACC	
4 9 0	UUAAAACUGGGCUCGGGUUGUACCCACCCGAGCUGCCCAUGUGGCUGGUGUCUUUGUAUCCCGGUA CACUUGCAGCCAGUUUGCCUCUCCUUACCCCGUAACAUUCUAGAAGAAUGUGCAUACGGACCAA UAGGUGUGGGCAACCCAUAGCCACUGACGGUCAAGUACUUCUGUUUCCCCGGUGUGAUGUGAAUAGA CUCUCUCAGGGUUGAAGCAUAAGCAUCGUUAUCCGCUACGUGCCUACGCAAAACCUAGUAGCAUUU UGAAGCUUACUUGGUUGGCCGUCAGCCACUAACCCGUGGUAGGCCUGGCAGAUAGGCGUGGGCGCA CCCCACUGGCACAGUGGCCCAGCCUGCGUGGCUGCCUGCCCCUUGAGGGGAAGCCAUUUUAAGU GACAAGGUGUGAAGAGUCUCGUGUGCUCACGUAGAUAUCCUCCGGCCCCUGAAUGUGGCUAAUCCU AACCCCGCAGCUAUUGCACACAUAUCCAGUGUGUUGAUAGUCGUAUUGGGCAACUGUGGGAUGGAACC AAUACUUUGGGUGCCGUGUUUCCUGUUUUUUAAAACUUGCUUAUGGUGACAAUAGUGUAUA CUGUUUGACACC	rhinovirusC4
4 9 1	UUAAAACUGGGUACAGGUUGUCCACCCUGUACCACCCAAAGUGGUGUAGUGUCUUUGUAUCCCGGUA CAUUUGCAGCCAGUUUGCCACCCCUACCCUCUAUGUAACUUAGAAGCUUACACAAGACCAUAGG CGGUGGUAACCCAUACCCAUACCGGUCAAACACUUCUGUUUCCCCGGCAUUGAGGAAUAGGCUCC AAAAGGGCUGAAGCCACUAGUGUCGUUAUCCGCAUUGGUACUACGCAAGCCUAGUAUCCGGCCUGAA AAUUUCUUGGUUGGUGGUUGGCCAGUUUACCCACUGGUAGACCUGGCAGAUAGGCGAGGACUUACC CCACUGGCGACAGUGGUCUGCCUGCGUGGCUGCCUGCACACCCUUUAGGGGUGUGAAGCCUAGGAA CGGACAGGGUGUGAAGAGCCCCGUGUGCUACCAUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCU ACCCCGCAGCUGUUGCACACAUAUCCAGUGUGUAUGCAGUCGUAUAGGCAAUUGUGGGAUGGAACCG ACUACUUUGGGUGCCGUGUUUCCUGUUUUUUAAAACUUGCUUAUGGUGACAAUUGUAUUAU CAGUGUUGUCAUC	rhinovirusC5
4 9 2	UUGAAACUGAAUCAUGGUUGUCCACCAUGAUUACCCACGUGGUGCAGUGGUCUUUGUAUACGGUA CAUUUCCAUACCCAGUUUUUAUACACCCUACCCCGAAACAUUAAGAAGCUUUGUACACAUGACCAUAG GUGGUGACCAUCCAGGUUACCGAGCGUCAAGCACUUCUGUUUCCCCGGUACCCUUGUAUACGCUUCA CCAGAGGCGCAAAACGAGGUUAUCGUUAUCCGCAAGUGCCUACGAGAAGCCUAGUAGCACUUUCGA AACCUAUGGCGUGGUCGUCAGCUGCUCACCCAGCAGUAGACCUGGCAGAUAGAGGCUAGACAUCCCCC ACCAGUGAUGGUGGUCUAGCCUGCGUGGCUGCCUGCACACCCUAUCCGGGUGUGAAGCCAGAAUGUG ACAAGGUGUGAAGAGCCUAUUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAAUCCUAA CCCGUAGCUGUUGCAUGCAACCCAGCAGUGUAUGCAGUCGUAUUGGGCAACUAUGGGAUGGAACCA UACUUUGGGUGCCGUGUUUCCUGUUUUUACUUUUAUUGCUUAUGGUGACAAUUGUAUUAUAC UUGUUACC	rhinovirusC6
4 9 3	UUAAAACUGGGUUAUAGUUGCUCUCCACUAUAACCCACCGCGUGUUGUGGACUUGUAUUCGGUA CACUUCUACUGCCAGUUUUUAUCUUCUCCCCUGCAACACUUAGAAGAAUGUACUUUAGGACCAUAGG CGGCGGCCACCCAGGUCUGUAGCGUCAAGCACUUCUGUCUCCCCGGUGAUACUGGAUACGCUUUAC CCAAGGCGAAACCCGGGCUCAUCGUUAUCCGCAAGUGCCUACGAGAAGUCUAGUAGCACUCUGAAA ACCUAUGACUGGUCGUCUACUGCAUCCAGCAGUAGACCUGGCAGAUUGGGGCUAGACAUAACCCCA CCAGCGAUGGUGUCUAGCCCGCGUGGCUGCCUGCACACCCUAUUGGGUGUGAAGCCAGAAUGUGAC AGGGUGUGAAGAACCUAUUGUGCUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAAUCCUAA UGCAACCCGUGCGCACAAACCCAGUGCGUAUGCGGUCGUAUUGAGCAAUUGCGGGAUGGAACCA CUUUGGGUGGCCGUGUUUCCUGUUUAUUCUUUAUUGCUUAUGGUGACACUGUAUAGCUUAUACA UCAUC	rhinovirusC7
4 9 4	UUAAAACUGAGUAUGGGUUGUCCACCCAUACUGCCUACGGGCGUAGUACUUCUUGUGCUCGGUAC AUUUGUACGCCAGUUUUUACCCCCACCCCUAGGUAACUUCUAGAAGCACUUGCACACAGACGAAU AGGCAGCCAAACACUCCAGUUGGUAACGUAAGCACUUCUGUUUCCCCGGUAUUGGGGUAAGGCU UUUCAAGAGGCGUAAGCCUGAAGUAUCGUUAUCCCGCGCAAUUACUAAGCAGAGCCUAGUAUCACUUC GACCCUCCACUGGCGUGGUCGUCAUAGAGAAAACCCACUACUAGACCUGGCAGAUUGGGGUGGAAA UACCCACUGGCGACAGUUGUCCAGCCCGGUGGCUGCCUGUCCCAUUAUUGGGGAAAGCCAAUG UAAGGACAAGGUGUGAAGAGCCGCGUUUGCUAAUUGAGUCCUCCGGCCCCUGAAUGCGGCUAAU UUACCCACAGCUAUUGCAUGGAAGCCACCAUGUGGAUAGUCGUAACGGGCAACUGUGGACGGAAC CGACUACUUUGGGUGCCGUGUUUCCUUUAUUCUUAUUAUUCUACUUGGUGACAAACGUAUACAGA UUAUAUGUUGUUUACC	rhinovirusC8
4 5	UUAAAACAGCCGUGUGGUUGUACCCACCCAGGGCAUUAUGUGCGUGGUACACUGGAAUUAACGGUUC CUUUGUACGCCUGUUUUUCCUCCACUUCUUCUCCUACUUCUUAUUGUAACAUAAUUA AGCAUGACAACAGUCCAAUAGGAGGCGGUAAUCCAUAAACCGUCAUAGACAAGCACUUCUGUUUCC CCGGAUGUGUAAUAGGCUUCCACCCAGGCGUAGAGCUGUAGCAUUCGUUAUCCGCACACGUACUACG CGAAACUAGUAACCCUUAUUGGUGUGUGGUGUUUACUCCGGCGUACCCACGCGUAGAUACAGG GCAUUGGGGCUACGGAACCCACCCGCGACGGUGUCGUAGCCCGCGUGGUGCCUCCCGAGAUCCU UGCGAUCUGGGACACCAUAUUAUAGAGGGUGUGAAGAUAUAGUGUGCUACGUGUGAGUCCUCCG GCCCGAAUUGUGGCUAACCUUAACCCCGCAGCUAUGGUGUGCAAAACAGCAUUCGUAGUCGUA UGGGCAACUGUGGGAUGGACCAACUACUUGGGUGUCCGUGUUUCCUUUAUUCUUAUAAUUAUUC UAUAGGUUACAGCUUAUGGUUAUAGGUUUGUAACC	rhinovirusC9
4 9 6	CAUGGUUACCCACGAGGUGCAGUGGUCUUGUAUACGGUACAUUCCAUACAGUUUUUAUAAACCCU ACCCUGAAACUUAUAGAAGCUUUGUACACACCCACCGACAGGUGGCAUUAUCCAGAUUGUGCAAGGU CAAGCACUUCUGUUUCCCGGUUAUCCUUGUAUACGUUACCCCGAGGCGAAACCGAGGUUAUCGUU ACCCGCAAGUGGCCUACGAGAAGCCUAGUAUACCUUUGGAAACCUAUGGUUGGUGGUCUCAGCUGUU UCCCAACAGUAGACCUGGCAGAUAGGCGUGGACGCUCCCAACAGCAGUUGGUGGCCAGCCUGCGUG	rhinovirusC10

	GCUGCCUGCACACCCUGUGGGUGUGAAGCCAGAAAGUGGACAAGGUGUGAAGAGCCUAUUGUGCUC ACUUUGAGUCUCCCGGCCCCUGAAUGUGGCUAAUCCUAACCCCGUAGCUGUUGCAUGUAAUCCAACA UGUAUGCAGUCGUAUUGGGCAACUAUGGGAUGGAACCAACUACUUUGGGUGUCCGUGUUUCUGUUU UUCUUUAUAUACUGCUUAUGGUGACAAUAGUAAUUAUACUCUUGUJACC	
4 9 7	UUAAAACUGGAUACAGGUUGUCCCAACCGUAUCACCCAAGUGGUGUGGUGCUCUUGUAUUUCGGUA CGUUUGCAGCCAGUUUGCUACCCCUCCUUUACGUAACUUAAGAAGUUUACAAAGACCAAUAG GCGGUGGUAUAAUCCAUACCAUACCGGUCAAUACUUCUGUUUCCCCGGCAUGCGAGGAAUAGGCUC CAAAAGGGCUGAAGCCACUAGUGUCGUUAUCCGCAUUGGUACUACGCAAGCCUAGUAUUACCUUGA AAAUUUCUUGGUGUGGUGCUCACCAGAUACCCACUGGUAGACCUGGCAGAUAGGCAGGACUUA CCACUGGCGACAGUGGUCUGCCUGGUGGUGGUGCUGCACACCCUACGGGGUGUGAAGCCGAGA AACAGACAAGGUGUGAAGAGCCCGUGUGCUACUAGUGAGUCCUCCGGCCCCUGAAUGCGGCUAAU UUACCCACAGCUGUUGCAGCAAAACAGCGUGUAUGCAGUCGUAUAGAGCAAUUGGGGAUGGAAC CGACUAGUUUGGUGUCCGUGUUUUUAUCCUAUUUAUUUGCUUAUGGUGACAAUUAUGUAUU AUCAGUGUGUCAUC	rhinovirusC1 1
4 9 8	UUAAAACUGGUAUAGGUUGUCCCAACCGUAUACCCACAGUGGUGUAGUGCUCUUGUAUUUCGGUA CACUUGCAGCCAGUUUGCCUCCUCCUCCUGUAACGUAAGAAGCAAUGUAAACGGACCAAAGGGGAC GGCAAAACACGUCGUUCAAGGUAAGCACAUCUGUUUCCCCGGCAUGCGAGGAAUAGACCUUCUCAA GGGUUGAAGCCACAAGUGUCGUUACCCGCAUUGGUACUACACAGAGCCUAGUACUGCUCUAGAAACU UAUUGGUUGGUGCUAGCCAUUUCCCGUGUAGACCUGGCAGAUAGGCAAGAAGUUCGCCACUGG UAACAGUGUUUUGCCUGGUGGUGGUGCUGCACACCCUAACGGGUGUGAAGCCAAUAGUGGACAGG GUGUGAAGAGCCGUGUGCUACCAUAGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCAACCCACA GCUAUGGCAUGCAAGCCAGCAUGUGCGUAGUCGUAUAGUAUUUGGGGACGGAACCGACUACUU GGGUGUCCGUGUUUCCUUAUUCUCUUAUGUUUUGCCUAUGGUGACAAUUAUAGUAUUAAUUGC UACC	rhinovirusC1 5
4 9 9	UGCACUUCUGUUUCCCCGGUACUAGGGUGUAAACUUCCCAGAGGUGUAAGCCCGAUGUAUCGUUA CCCGCAUGGUUACUAAAGCAGAGCCUAGUAGCAUCUUGCUUUGCAUAGAGCUGGUCGUUCAGUCAGU AAACCCCGUGUAGACCUGGCAGAUAGGGCUGGAAUACCCACUGGUAACAGUGUCCAGCCUGCG UGGUGCCUGCCCCUUCUUUGAGGGGAAGCCCAUAGUAGACAAGGUGUGAAGAGCCGCGUUUGCUA ACUAUGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCAACCCACAGCUCUACUGCGCAGAAACCAUG GUAGGUGUCGUAUAGAGCAAUUGGGGACGGAACCGACUACUUGGGUGUCCGUGUUUCCUGUUUU ACUUUAUUGCUCUAGGACACACAUAUAUAAAUCGUAUAUGUUGUCC	rhinovirusC1 6
5 0 0	CACUUCUGUUUCCCCGGAUGUGUGAAUAGGCUUACCCGAGGCGUAAACUGUAGCAUUCGUUAUCCG CAUUGGCUACUACGCGACUGCUAGUAGCAUUCUGUAUGACAUUGGGGUGUUCGCUACAGCAGAGACCA GUCUGUAAGUAGGGUAAUAGGGCUACACAUAUCCCAUUGGCGACAGUGGCUAGCCUGCGUGGUGC CCUGCCCCGAGGUUCGUCCUGGGACCCUCCAUUGUAGACAGGGUGUGAAGACACUAGUGUGCUAGUC ACUAUGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCAACCCACAGCUCUACUGCGCAGAAACCAUG AGGUGGUCGUAUAGAGCAAUUGCGGGAUGGAACCGACUACUUGGGUGUCCGUGUUUCCUUAUAC AUUAUUGUGUCUUAUUGGUUAUUAUUGGUGUAUAC	rhinovirusC1 9
5 0 1	UUAAAACUGGGUGUGGUGUUGUCCCAACCAUAUACCCUACGUGGUGUUGUAGUCUUGCAUUUCGGUG CACUUGCAUGCCAGUUUAUUCUCCCAAGCCCCGUAACUUAAGAAGAAAGAACACAGACCAAUAGGA GGUGGUAACCCAGACACAUUACGGUCAGCACAUCUGUUUCCCCGGUUGAUUAGAUUAGCUUUACC CGAGGCAAAAACUAAAUCAUUGCUUAUCCGCAAGUGCCUACGGGAAGCCUAGUAGCACUCUGAAAG CAUUAUGGUGGUGGUGCUCUUAUUGCAAAACCCAGCAAUAGACCUGGCAGAUAGGCGUAGACGUACCCAC CAGCGAUGGUGGUGUAGCCUGCGUGGUGCCUGCACACCCUACUCCGGUGUGAAGCCAGAAAGUGGA CAGGGUGUGAAGAGCCUACUGUGCUCACUUGAGUCCUCCGGCCCCUGAAUUGGGCUAAUCCUAACC CUGCAGCUGUUGCACACAAGCCAGUGUGUAUGCAGUCGUAACGAGCAAUUGCGGGAUGGAACCAACU ACUUAUUGGGUGCCGUGUUUCUUAUUCUUAUGUUGACCUAUGGUGACAAUGUGCAUUAUACAGU UUGUCACU	rhinovirusC2 2
5 0 2	UUAAAACUGGGCAGGAGUUGUCCCAUCUUGCCACCCACAGGGUGUGGUGCUCUUGUAUUACGGUA CACUUGCAUGCCAGUUUAUUCUCCCAAGCCCCGUAACUUAAGAAGCAAUACACAUAGACCAAU AGGAAGUGGUAAGCCAUACCAUUAAGGUCAAGCACAUCUGUUUCCCCGGUAUGCGUGGAAUAGGCC UCUUAAGGGGUGAAGCCACAUAUACGUUAACCCGCAUGUGAACUACGCGAAACCUAGUACUACCUU GAAGUACAUUGUCUGGUGCUCAGCCAAUCCCAUGGUAGACCUGGCAGAUAGGCAAGAGGUACC CCACUGGUAACAGUGUUCUUGCCUGCGUGGUGCCUGCACACCCUACCGGGUGUGAAGCCAAUGUA UAGACAGGGUGAGAAAGAGUCGAGUGUGCUUACAUAGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCA ACCCCGCAGCUGUAGUAUGCAAGCCAGCAUAGUGCAGUCGUAUUGGUAACUUGCGGACGGAACCG ACUACUUGGGUGUCCGUGUUUCUUAUUCUUAUUGUUGACCUAUGGUGACAAUUAUAGGAUU AAUAUUGUCACU	rhinovirusC2 3
5 0 3	AGCUCUGCAGAUUUCUCCCGGACGUGUGAAUAGGCCCUACACGGGUGAAGCUGUAGCGUUCGUUA UCCGCAUGCCUACUGCGCAAGGUGUAGUAACUUAAGUAGUAGUUGGUGUUUCGUCACGACAGUAC CCCACUUGAGUACAGGGUAAUAGGGUACGCAUCCCCACUGGCGACAGUGGCGUAGCCUGCGUGGU GCCUGGCCAGGUUAUACUGGACACCAUUAUUGGACAGGGUGUGAAGGCCAGCGUGCGCUAGG CAUGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACCCUGCAGCCAUUGCAUGCAAACAGCAUG UUGGUGUGCUAAUAGCAAAUUGCGGGAUGGAACCGACUACUUGGGUGUCCGUGUUUCUUGUUUU CCAUUAUUGUCUUAUGGUUACAAAGUUUGUGCAACAAUUGUGACC	rhinovirusC3 5
5 4 0	UUAAAACUGAAUCAUGGUUGUCCCAACCAUGAUUGUCUACUGGACGUUGUACUCUUGUAUUUCGGUA CACUUGUACGCCAGUUUAUAAACCCUCCCUUUGCAACUUAAGAAGAAAGUAAACACUGGCCAAU GGAAAUAGCAAACACGCUAAUUAACGGUCAAAGCACAUCUGUUUCCCCGGUGACGGGAUAGCUUC ACCCGAGGCAGAAACUCAGAUAGUCGUUAUCCGCAAGUGCCUACGGGAAGCCUAGUAGCACUUGG	rhinovirusC3 9

	AAAGUUGUUGGGUGGUGGUCUCAAUUGCUGACCCCGAGCAAUAGACUUGGCAGUAGAGGUGGACGUG CCCCACCAGUGAUGGUGGUGCCAGCCUGCGUGGCGUACUCACCCUAAACAGGGUGAGAAGCCUUAUC AAUAGACAAGGUGUGAAGAGGCCUACUGUGGUCACUUUGAUUCCUCCGGCCCCUGAAUGUGGCUAAUUC CUAACCCCGUAUCUUAUGCGAUGCACACCAGCAUGUAGGUAGUCGUAUUGGGCAACUACGGGACGGAA CCAACUACUUUGGGUGGUGCCGUGUUUCCUUGUUCUUUUAAUACUUGCGUAUGGUGACAAUUGUAUGAA AUACGAUUGACACC	
5 0 5	GACCCGACUAGUACUACGCAAAGCCUAGUAAUUCUUGAAGCUCUUCUGGUUGGUGUCCACUAG UCACCCCUUAGUAGACCUGGCAGUAGAGGCAGGACACUCCCCACUGGUGACAGUGGUCUGCCUGCG UGGUGCCUGGACACCCUCCGGGGUGUGAAGCCAAAAGAAGGACAAGGUGUGAAGAGCCCCGUGUGC UACUAGUGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCUUACCCACAGCUAUGGACACAAUCCAGU GUGUAUGUAGUCGUAUAGUAAUUGUGGGACGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUU AUUCUACUGUUUUUGCUUAGGUGACAAUACUGACAAUAGUGUUUUAACC	rhinovirusC4 0
5 0 6	UUAAAACUGGGUAGGGGUUGUCCCCACCCUACCACCCACUGGUGUAGUGCUCUGUAUUACGGUA CACUUGCAGCCAGUUUGCCUCCUACCCCUUCCCAUUCACACCCACCCUGCGUAAUUAAGAAGC UAGUCACAAGACCAAUAGGUGGUAACAUAUCCUUGUUAACUAAAGGUAAGUACUUCUGUUUCCCCGGC AUGCGUGGAAUAAACAGUGCAAACGGUUGAGGCCACAAUUGUCUUACCCGCGCAUGAACUACGCAG AGCCUAGUAUCCCCUGAAAGUCCAUUGGUUGGUGCGUCAGCUAAUCCCAUAGUAGACCUUGGCAGA UGAGGCAAGACGCACCCACUGGCGACAGUGGUCUUGCCUGCGUGGUGCCUGCACACCCUUCGGGG UGUGAAGCCAAUUGAUAGACAGGGUGUGAAGAGCCGAGUGUGCUUACUGUGAGUCCUCCGGCCCCUG AAUGCGGCUAAUCCAAACCCGCGAGCUGUGGCGACAAUCCAGUGUGUAUGCAGUCGUAACGGGCAC UGCGGACGGAACCGACUACUUUGGGUGUCCGUGUUUCCUUUUUUCUUUAAUCAAAGUACUUAUGGU GACAAUAAUAAUUAUUGUUAACC	rhinovirusC4 1
5 0 7	UUAAAACUGAGUGGGGUUGUCCCCACCCACACUGUCUACGGGACGUUGUACUCUUGUAUUCGGU ACACUUGUACGCCAGUUUGCCACCCUAAACACCCCGUAACAAUAGAAGUCUACGCACACUAGACC AAUAGGGGGUGUUAUCCAGUUCACCCAAAGGUCACAGCACUUCUGUUUCCCCGGCAUAGGGGUGUAG GCUUCCGCCCGAGGCCAAAGCUCUGAUUGUCGUUAUCCGCGACAAUACUGGACAAAGCCUAGUAGCAC CCUACGAGUGUAGCAUGGUUGGUUGUCCGCCAGUAUCCCGUGUAAACUUGGCAGUUGGGGCUAGAA GUUCCUACUGGCGACAGUGUUCUAGCCCGCGUGGUGCCUGCUCUCCUUUAGGGGAGAAGCCAUUCUA UGACAAAGGUGUGAAGAGCCCCGUACGCUAGUUGUAUCCUCCGGCCCCUGAAUGCGGCUAAUCCCA ACCCUGCAGCCAUUACUUGGAAGCCACCAUGUGGAUGGUCGUAUAGUAUUAUUGCGGACGGAACCG ACUACUUGGGUGUCCGUGUUUCUGUUUUUUCUUUACUGACGUUUAUGGUGACAAUUAUUAUUA CGUAUAUUAUUGUUAUC	rhinovirusC4 2
5 0 8	GGUUGGUGCGUCCGUGGCAACCCAGCAACAGACCUGGCAGUAGAGGUGAGCAUCCCCACCAGCG AUGGUGGUCUACCCUGCGUGGUGCCUGCGCACCCUUCGGGGUGCGAAGCCAGAAUGGACAAAGGU GUGAAGAGCCUAAUUGGUCACUUUGAGUCCUCCGGCCCCUGAAUGUGGCUAAUCCUAAACCCCGCAG CUACUACACAAUAGCCAAUGUGCAGGUAGUCGUAACGGGUAACUGUGGGAUGGAACCAACUACUUUG GGUGUCCGUGUUUCCUUGUUAUCCUUUUUAAAUGCCUUAUGGUGACAAUUGUACAGUAUACAUUGU UACC	rhinovirusC4 3
5 0 9	GUGGUAAAUCCAUACCACUAAUGGUCAAAACACUUCUGUCCCCCGGCAUGUGAGGAAUAGGCUCCAA AAGGGCUAAAGCCACUAGUGUCGUUAUCCGCAUUGCUACUACGCAAAAGCCUAGUAUCCUUGGAAG UCUCUCGGUUGGUGCUCUCCACAGGUACCCACUGGUGAGACCUGGCAGUAGAGGCAGGACUACCCCC ACUGGCGACAGUGGUCUCCUGCGUGGCGUGCCUGCACACCCCUACGGGGUGUGAAGCCUAGAGAU GACAGGGUGUGAAGAGCCCCGUGUGCUACCAUAGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAC CCACAGCUGUUGCAGCGAAUCCAGCGUGUAUGCAGUCGUAUAGGACAAUUGUGGGAUGGAACCGAC UACUUUGGGUGUCCGUGUUUCCUUUAUUCUUAUUAUUGCUUAUGGUGACAAUUAUGAUAGCAUCA GUGUUGUCAUC	rhinovirusC4 5
5 1 0	AACUGGGCUUGGGUUGUGCCACCCAAAGCUGCCACGUGGCGUUGUGUCUUGUAUUCGGUACACU UGCAGCGCCAGUUUGCCUCCCUACCCUGCAACGAAUUAGAAGAAUGGACACAAUAGACCAUAGG UGGUGACAAUCCAUGUACUAAUUGGUCACAGCACUUCUGUUUCCCCGGUGUGAUGUGGAAUAAUCC CCAGGGGUUGAAGCCAAUAGCAUCGUUAACCCGUAACGUGCCUACGCAAAACCUAGUAGCAUUCUGGA GUCUACUUGGUUGGCCGUUCAGCCACUAAACCCGUGGUAAGGCCUGGCAGUAGAGGUGGGCGUACCCC ACUGGUAAACAGUGGCCACGCCUGCGUGGCGUGCCUGCCCCUUAUGAGGGGAAGCCAAUAGGUGACA GGGUGUGAAGAGUCCCGUGUGCUCACGUAGAGUCCUCCGGCCCCUGAAUGUGGCUAAUCCUAAACCCC GCAGCUAUUGCACACAAUCCAGUGUGUUAUGUAGUCGUAAUGGGCAACUGUGGGAUGGAACCAACUAC UUUGGGUGUCCGUGUUUCCUUUUUUAUUCUUAUUAUUGCUUAUGGUGACAAUUAUGAUAGCAUCA GACACC	rhinovirusC4 9
5 1 1	UUCCCCGGACGUGUGAAUAGGCUCCACCCAGGGGUGAAGCUGUAGCGUUCGUUACCCGCACUGCUAC UACGCAAGGGUUGUAUUAACUGAAACAAUUGGAGUUAUCGUCACGAGAACCCACUGUAGAUUC GGGGUAAUGGGGCUACACAACCCCCACCGGCGACGGUGGUGUAGCCCGCGUGGUGCCCCGCCUGAGG CUUGUGCCUCAGGACGCCACAUGUUGACAGGGUGAGAAGGCCUACGUGUGCUAGCAGUGAGUCCUC CGGCCUCUGAAUGCGGCUAAUCCUAAACCCUAGCCUACACACAAUCCAGUGUGCAGGUGGUGCGU AACGGGUAACUAGGGAUGGAACCGACUACUUUGAGUGUCCGUGUUUCCUUAUUUACCUUUACAUGU CUUAUGGUCACAAUUAUGUGUAUACAUAUUGUGAUC	rhinovirusC5 0
5 1 2	UUAAAACAGCCGUGUGGUUGUACCCACCCAGGACACACUGUGUGUUGGACUGGAUUCUGGUUC CUUUGCAGUCCGUGUUUUAAACCCCCUCCCCACUCUUUCUAUUAUUCUUUCCUUAUUGUAAC UUGAGACGUUAUACACAGUCCAAUAGGACGUGGCCAAUCCAGGCCAUGACGACAAAGCACUUCUGUC UCCCCGGGCGUGUGAAUAGGCCUCAACAAGGCGUAGGUGUAGCGUCUUAUCCGCAUUUGUACUA CGCGAUGGCUAGUAUCCUUCUUGGUGUUGGCGUUUCGCUACGCGUGGUAACCCCCACGUGAG AUCAGGGCAUAGAGGCGUCAUAGCCCCACCGGCGACGGUGAUGCAGCCUGCGUGGUGCCCCGCCUG	rhinovirusC5 1

	GGUUAACCCCCAGGACGCCACAAUACAAGACAGGGUGAGAAGACCUGAGUGUCUACAGAUAGUCCU CCGGCCCCUGAAUGCGGCUAAUCCUAAACCCUGCAGCUAUCGUAUGCAAACCAGCAUUAUGGUAGUCG UAACGAGCAAUUGCGGGACGGAACCGACUACUUGGGUGUCCGUGUUUCUCUUUUUAUUUGAAUUGU UGCUUAUGGUUACAAGUUUUGGUAAUAAACUUGUGACC	
5 1 3	UUAAAAACAGCUGUGUGGUUUUCCACCACAGGCACACUGUGCGUUUACACUGGGAUUCCGGUCA CUUUUGUACGCGUUGUUUGCUAUCUUUUUACUUUAUUAUUAGAAUGUACACAAACGCCCAUUAUAG GAUGCGGCCAAACAGGUCGCUUAGGGCAAGCAGUUCUUGUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU ACAGGGUUGAAGCUGUAGCAGUCGUUUUCCGCGCAACUACUACGCGAAUGUUAGUAGCAUCCUGUGU UGCAUUUGGGAUUUCGCUCCGACAGAAAACCCCAUCUGUAGAUUAGGGCAUAGAGGCUACACAUACCC CACUGGCGACAGUGUGUAGCCUGCGUGGUGCCUACCCAGGCCUUCUUGGCCUGGGAUUCCACUUA CAAGACAGGGUGUGAAGGCACUAGUGUGCUAGUUGUGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCU UAACCCCGUAGCCCCCGCAAGUAAACCAACUUGUAGGUGGUCGUAUAGAGUAAUUACGGGAUGGAAC CGACUACUUUGGGUGUCCGUGUUUCCUUUAUUCUUUAUUUUGCAUCCUAGGUUACAACAUAAGU AAUC	rhinovirusC5 4
5 1 4	GCACUUCUGUUUCCCGGACGUGUGAAUAGGCCUCACCAAGGCUGAAGCUGUAGCGUUCGUUAUCCG CUCACCUACUGCGCAAGACUAGUAACUAGAUACUUAUGUGGUUUUCCGUCAGCAGUACCCCA CUGUAGUACGGGGCAUAGAGGCUACGCAUCCCCACUGGCGACAGUGGCGUAGCCUGCGUGGGUGCC UGCCAGGCUUCGGUCUGGGACACCACUAGUUGACAGGGUGUGAAGACCGGCGUGCGUAGGGUGUG ACUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACCCCGAGCCACCGCAUGCAACCCAGCAUGUAGA UGGUCGUAACGAGUAAUUGGGGAUGGAACCGACUACUUGGGUGUCCGUGUUUCCUGUUUUUUUUUU AUUGUUGUCCUAGGUUACAACUUUGUGUAUCAGAUUUGUGACC	rhinovirusC5 5
5 1 5	UUAAAAACUGAAGCAUUGGUUUUCCACCAUGGUUACCCACGUGGUGCAGUGGACUUGUAUACGGUA CAUUUCCAUACCCAGUUUUUAACCCCUACCCUGAACUUUUAGAAGCAUGUACACAUACAGACCAAC AGGUGGCAAUACAGCCAGAUUGCUAACGGUCAAGCAGUUCUGUUUCCCGGUAUUCUUGCAUACGCUU CACCCGAGGCGAAAAGUGAGAUUAUCGUUAUCCGCAAGUGCCUACGAGAAGCCAGUAGCACAUUG AAAGCCUAUGGUUGGUGCGUCAGCUGUUUACCCAAACAGUAGACCUUGGCGAGUAGGCUAGAUUUCC CCACAGCGAUGGUGAUCUAGCCUGCGUGGCGUGCCUGCACACCCUAUCGGGUGUGAAGCCAGAAAGU GGACAAGGUGUGAAGAGCCUAUUGUGUCACUUGAGUCCUCCGGCCCCUGAAUGGGCUAAUCCUA ACCCCGUAGCUGUUGCAUGUAACCCAAACAGUUCUGCAGUCGUAUUGGGCAACUACGGGAUGGAACCA ACUACUUGGGUGUCCGUGUUUCCUGUUUUUUUUUAUUAUUGCUUAGGUGACAAUUGUAGAUACUA CACUUGUUACC	rhinovirusC5 7
5 1 6	ACCCUACUUAGGCAACCCCAACCCUUCUCUACGGGAGCUAGGGCCAAGGAUUCUUAUGUAGUGGUG UUGUAUCCCUUACCUUAUACCCAUUCCCAACCAACCGGACUGUAUUAACGGUUUAGACUGAAUAAU GGUUUAUGAAUUAACCCUGUUAUCUACUACAGGGGAAAAUAGAAGACAUUAUUCUAGGAUUGGACCU GUUAUGACCAAGAGUUGGCGUCUUGCGCGUAGCCGUUAGGACAGUUUAUUGGUUACUUCUUGGAUU AUGAUUUCAGGGUUGGCCGCCAGGAUAAAAGGCAAUUGUGGCAAGUGAUGUUAGUCAUUGGCAUU GAAACCCUGCCUAGAACAUCUAGGUCUACGCUUGUGCGGGCCGAAGUAAGUCUAGGAACAACAGGGAG UAUGCUUUUUUUGUUUACCCAAACAGGACCGUACACGAAAGAGCUAGAGGCAUCUGGAGGCAAG GGAAAGAAUUUCUAGCCGAAUGUUAUAGAGUCCUUGACGAAUGCGUCCCUUCUGUCCCGAGG GUGAGGCGUAUGGCGCAUGCUCUAGGCAUUAUCCAAAGGUGUAUCUAGAGGGGGGGCCCCCUUAC UUAUGUCUAGUGCUACCUAGCAGGGCCGCGGCGUGGUCGUUUGUGUAUGGUAUAACAGUAGUAUCC CCAUGGAUUGCUUAAAAACCCCUCCCCCUCAACAAGACAUCUCUAAAG	simiansapelo virus3
5 1 7	CCGACCAUGUAGUGUCCGAAUUGCUAGAGUGGACUUCUCCCUUCCCUUGGUGUAUCUUGCCAGUU UCCUCCCCCAACCCUGACGUGCCACAGGCGUGUGAAAACACCCCGGAAAGCUGCCAAAAGUGGGAAU UGUGGGUCCCCCUUUGGCAAGGAGUCGAGUUCUUCUCCUUAAGGCUAGCCCAUCAGUGAAUUC UGUUGGGCAACUAGUGACGCCACUGCAUGCCUCCGACCUCCGCGCGGAGUGCGCCCCCAAGUCG UGCCCCUGACCAAGUUGUGCUGUCUGGCAACAUUGUCUGUGAGAAUGUCCGUGUGGCGUGCCA AGCCUGGUAACAGGCGUCCCAUGUGCGUAGUUCUACUCCAGACUUCGGUCUGGCAACUUGCUGUU AAGACACGGCGUAAGGGGCGUGGCCAACGCCUGGAAACGAGUGUCCACUUAUUAACCCCGAGGAAU GCUACGUAGGUACCCUGGUUCCGAGGGAUCUGAGCGUAGGCUAAUUGUCUAAGGUAUUUUCUUAU UCCACUCUCUCUUCUUGUUCUAU	aichivirusA2 ;caninekobuv irus
5 1 8	GGAAUUAACGGGCACCCACUACCCCAUCCCCCUUGUAACUUAAGUGUGUCUGGACCUUG ACUCCUGCGGCUUACCCGACCCGUCAGUGAACAACUAGCUAGGUGACUCCUUCUCCCCCUUGGC CUGGAUGGCCACAUCCAGGUCUUCUUCUCCCGAGCCGACAUUGUAUGGGCUGUUUGGAAUCCCGC GAAAGCUGCUAACGUGGCAAUUGUGGGUCCCCUUCACAAAGACACCGAGUUCUUCUCCCUUAAGG CUAGCCCGUGUGGUGAAUUCACCCGGCAACUAGUGGUGUCAAUGCAUGCCUCCGACCUCCGCGC GGAGUGCUGCCCCCAAGUUGCGAAACUGACCAAGACUGUGCGUCCUGGCAAGCAUAGUCUGAGGAG AUGUUCGCGUGUGGUUGCCAAACUGGUAAACAGGUGCCCCAGUGGCGUAGUCUUCUCCGGUUCGU CCGGCUGUCCUUGGUAAAGACACAAGUUAAGGGGCGUGUGCCAACGCCUUGGAAAGAGUGAACUUC UACUGCCCUAGGAUCCGACGCGAGUACCCUGGUUCCACAGGGAUCUGAGCGUGGGCUAAUUG UCUAACGGGUAUUAUUAUUCUCCUUCUUCUUGUUGACUAC	aichivirusA4 ;felinekobuv irus
5 1 9	UUCGGUACCGUUGGAGGAAUUAACGGGCACCCUACCCAAAGGCUUAUGGGACUCCAAUCUCC UUUCCCCGUAAUCUUAAGUAAUGUGUCGUAAUUCUUGACCCUCCGGGUGUGUGCCCCCGUGAGUG AACAAAUAGUAGACACGCUCCUCCUCCUUGGUGGGCACACGCCACUACACAUCCACCCCCCA GCACAACAUUGUGUGUGGUUGAUGCACCUGGAAGCGCGAAAGCUUCGGUUAGUAGUGGCAAC CGGUUGACCUACACAACCAAUGCUCAGAGUCUUCUCCCUUAAGGCUAGACCCCUUGGUAAUUC ACGGUGGCAACUAGUUGAGCAACAUGUCCCAAGUGCGGCGCGGAGUGUUCUCCCAAGCCU GUGUGUCCUUGGUAAGUUGUGCGGCAAGCAUUCUAGAGAAGAUUCCGCGUGCGUGGCGCAA CCUGGCAACAGGUGACCCAGUGUGCGUAGUCUUCUCCGUCUACGACGACUUCUUAAGGAGGCC	aichivirusA7

	CUGGUGUAAAGGUACAUGUGCCAAACGCCCAUGCGUGGUGUUCUUAAGGCCCUAGGAAUGACCCGC AGGUACCCUGUCUGCGCAGGGAUCUGAGCGGUGGCUAAUUGUCUACGGGCCUUUUAUUCUCCAUU CUUCUUGACUGUCAUCAUGGCCCAUCACACGGAUCUCGUUUGGUGUCCAAUGCUUCUCAGCUUCUU UCCUCCUAUGAUAGACUCUACAGAACUGUCAGGGACUGGCGUGACACUGAUAACUGCGCCAGACCC GAAUGCCCUACCUUCUUCUGUUGCCCCUGGUGAGUGCAUUCUUGCUACCUUGAGCGGACCGAACG GGACCCCAUGUCCAUCACGACGAUGGGUACCCUGGACGGUUUGCCAGACUGCAAGACCCUCCC AAACGCAACCCCCCCCCCAAGAGUUUCUGUAUGACUCUGACGUCAGUCUGGUUUGAAGA CCAGACUCCUGUGACUGACAGAGAACAACAUAUUUUGACCAAAUAGCC	
5 2 0	GCCAGGUCGCAUCGCCUGGUUGGUGACCAACUGCUAGAUGCUCCUCCUUUCCUUGGGGUGGCG CAGCGCAUCCUGCAUCCUCCCCAGCCUGACAUAGCUACAGGCUUGGCAAGAACCCTCGGAAAGC UGCCACCGUGGCAUUGUGGGUCCCUUCUUGUCAAGGCACAGAGUCUUUCUCCUUUAGGCUAGAC CCUUCUGUGAAUUCAGAGCGGCAACUAGUUGGCCACUACAUGCCUAGACUCCGUGCGCGAGUG CUGUCCCCAAACCCUGACACUGCCGCGUGAGUAUGCUGCUUGGCAAGCAUUGCCUGGAUGGUCGUU CCGCGUGGGGUGCCAGGUCUGGUAACAGACCCGUCUCCAGUGGCUAGACCAUCCAGCUUCGGCUG CAGGUCGCGCUUUUCCGAGACUAGCUCGGAUUCAAAAGUUCUUGGCACCCUUUACCCUUCAGGC UGCCUAGGAAUGCUACGCAAGUACCCUGCUCCUUGGGAGCCAGGGAUCUGAGCGUAGGCUAAUU GUCUGCGGGUAGUUUAUUCACCCUUCUAGACUGUUCACC	aichivirusA8 ;murinekobuv irus
5 2 1	UUUGAAAAGGGGUGGGGGGGGCUUCGGCCCCUACCCUUCUUUCCGGCGGCGCAUGUUCGCGUCGU CGUAAAGUCUGGAUCCCAAGGCCCAUCGGUUCAACUUCGGUUUCCGGACAAUACAAGAACCUCA GUCCUCUUGGUACUUUCUGCCUGAGUCUACAAGCGAGAAACUUGCCCUUAACGCCAGACGAGC GGCAUAAACUCGAACUUCUGGCACGUUCCACCACCCUUCUCCUAUCCCAACCCCAUUGCGCUCU CAAGGUCGCGCUUUUCCGAGACUAGCUCGGAUUCAAAAGUUCUUGGCACCCUUUACCCUUCAGGC CCUUAAGGUAGGAACUGACCUUGUGCUGUGAUCUCGGUGCGGAAGUCUACUGCGUAGUGAUUGUA GAUCCUUUUGUGGUUCUGCCUGGCAAGGCUACAGAGUGCUGUGAUCCGCGUGCGGAUGCCAUCCUG UAACAGGACCCCAUGUGCGCAACAGUAUGUUCACGGUCUUCUGGUAACCAUUCGGAACACU GCUUUCUGGAAACAGUGUGUCCAAUCCUUGUGAUCAGUAUACACACACCCUAGGAAUGCUAG GAAGGUACCCCGUUCGCGGGUUCUGAUCCUAGGCUAAUUGUCUACGGUGGUGCUUUUUUAUUU UCCACUUCACUCUGGUUACAACUGCUUGAUUUUCUGUUUGCUGUUUUCUUGCUCUCACUGC CAUUCUACAAGUGUUCACACUGUCCAAGGUCUUUGGUUGUUCGCUUCCACUUGCCACUGCAACUCU UGUC	aichivirusB1 ;bovinekobuv irus
5 2 2	CCAUGGUGUACAUCGGUUCUCCGGGCAAAUACAAGAACCAGUUUUCUGGCAAUUCCUUGACUGGG UUAAAGAAUUAAGGAACCAAGCCCCUUAACGCCAGAGCGGCUUUAACUUGAACUUCUGGCACA CUUUUCUCCACUACCCUUUCCCAACUCGGACUAGCUCACUGGAUUUUGACCCACUCCCCUUUGU CUGACCAUCUCUUUGGAUCUCCUCCAAUAGCCGUGGUGCCUGGCGGAAAAGCAACAAGUUGUGU GAGCAACACCCCAAGGUCACCGGAACGGUCUGUGCCGCGAGAUUAGUACGUGCGGAUCCUUGG UUGCUUCCACUGCUCGCGUGUGAGUAAGUAGGUGCUGGUGUGAUGCGUUCUUGGACGCCUGCGGC AUAGCAUUGCUCCCCGAGAAUACAAGCAUCCAGAUUUUCUGGCCUCUGAAGUCACGCCUGGACGC GUAG	aichivirusB2 ;ferretkobuv irus
5 2 3	UUUGAAAAGGGGUGGGGGGGGCUUCGGCCCCUACCCUUCUUUCCGGUGGCCACCCGCGGGGCCA CCGUUACUCCACUCCACUCCUUCGGGACUGGUUUGGAGGAACAUAACAGGGCUUCCCAUCCUGUUU ACCCUUAUCUCACUACCCUUCUCCUUGACCAACCUUAUCCACACCCCAUCUGACUGACUCCUUUGGA UCUUGACCUCGGAAUGCCUACUUGACUCCCAUUGCCUUCUCCUUUCCGGAUUGCGGUGGUGCCU GCGGGAAGAAAGCACAAGUGUGUUGGCUACCAACUCCUACCCGACAAAGGUGCGUGUGCGCGUG CUGAGUAUUGGGAUAGGAGAUCCAAUAAACAGGCUCGCCCAUGAGUAGAGCAUUGGACUGCGGUGCAU GUGACUUCGGUACACAGGGGCAUAGCAUUGCUACCCUUGAUAUACAAGUACUGAGAUUUUCUGACC UUGCUAAGUGCUCUGUGGUUGCGUGGUGGGAUCCACGCUUGACCAUGUACUGCUUGAUAGAGUCCG GGCUGGCCGACUCAUGGGUUAAGUACUUGACAAGACAC	aichivirusC1 ;porcinekobu virus
5 2 4	CAUCCCAAAUUGUUCGGUGAGGAUUGCUACACGUAACUAUUGAUAACCAUUGACUGUUCUGCC CUGGCAAGGCUAUAAGAUACAGUUCUCCGUGUGGAUGCCGUCUGGUAACAGGAUCCCAUGUGUGU GUAGCUGUAUGUUAACUUCGGUUGUUCACAGUCCUUGGUCUAGCUUUUAGUGAAUUGGUGCGCC AUUACCUUACAUUUGGGUGCCACCUCUACACCCUAGGAAUGCUAGGAAGGUACCCACCCUCCGGU GGGAUCUGUCCUAGGCUAAUUGUCUAGGGGUGUUGUUCUUUUCACCAUUAACUGGCU CACAGAUUGUUGUUGUUGUCCUGCUCUCUAGCUCCCAACCGGCACACAGUCUCAGCCACACC GCCUUUGCAACACAUACAGAUUUGUUCUUGACUCCUUGAUUUCACAUUUCACUAGCAACAG AC	aichivirusD1 ;kagovirus1
5 2 5	GAAAGGGGCGCGCGGGCAACACCCGUGCGUCCUUCUUCGUCUCCAGCCGCGGACGUUAAACUG UGGACGGUUGCGCUAGGUGCAACCCUGCUACUAUUGGCGGUACCCUGCCUUAUGUAUGCAUGU UGUAUGAAUGGAUUCUGUAGUGAGUCCUUGGUGGUCUUAACACACUACAGUGAUGGCUAG AGCUCCUACUUCGGCUACCUUAGUAGUAGCAUUGGAAGUAUAUUGUAACAGGGUUCACGCCC GAUAAAGCGUGUAUUCGCGUACUGGAGUGGACCUAGGUAACUAGGGCAUGAGUGUGUGAAUC AGCACCCGCAAGGGAUAGCAUAUUGGAAGGUGCUCUGAAGGCAAGGUGGAGUUAUAGCAGCU CUCAGUCUACAAGGUCAGCUUCUAGACAUUGGCCGAGUACCCGUAAGGGAUCUGAGCGGUAGCUA AUUAGUCCUGGAGGUGGUGACUUCUUCUUAUUCUUCUUUACUUGCGAAUUGUGUGGAUUGGU AUUUUAGAGCUUGGGAGUUGUCUCAUAUUGAAUUAUAUUAUUUACUUAUACUACACUCAAACA CGGCUCUCUGGUCAGUUAAGAAAUAUUAUAGAAUUAUAAA	oscivirusA2; turdivirus3
5 2 6	UUAAAACAGCGGAUGGGUACCCACCAUCCGACCCACUGGGUGUAGUACUUGGUACUUCGUACCUU UGUACGCCUGUUCUCCAUUGUACCCUUCUGAACUUCACCAAGUAACGUAGAAGCUAACA UUUAGUACAACAGGAAGCACAACCAUGGUGUUAUAGUACAAGCAUUCUGUUUCCCGGAGCGA GGUAUAGGCGUGUACCCACUGCCAAAACCUUAACCGUUAUCCGCCAACCAACUACGUAAAAGCUAG	HRV-B3

	UAGUAUUUUGUUUUUUAACUAGGGGUUCGAUCAGGUGGAUUUCCCUCCACUAGUUUGGUCGAUGAGG CUAGGAAUUCACCGGUGACCGUGUCUAGCCUGCGUGGGGCCAACCCAGCUUAUGCUGGGACG CCUUUUUAUAGACAUGGUGUGAAGACUCGCAUGUGCUUGGUUGUGAUUCCUCGCGCCCGAAUUGCG GCUAACCUAUACCCUGGAGCCUUGUGUCACAAACAGUGAUGAUAGGUCGUAAUGAGCAAUUCGCG GACGGGACCGACUACUUUGGGUGUCCGUGUUUCUUAUUUUUUAUUUUGUCUUUUGGUCACAGCA UAUAUAUAACAUUACUGUGAUC	
5 2 7	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACAAUAGAUGACUUAACAUAUUCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUCAAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAGCCA AUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAUCGUGUAUUAACAGCCUGUGGGUUGAUCC ACCCACAGGCCAUUGGGCGCUAGCACUCUGGUAUCACGGUACCUUUGUGCGCCUGUUUAUACCC CUCCCCAACUGUAACUUAAGAAGUAAACACACCCGAUCAACAGUCAGCGUGGCACACCAGCCACGUU UUGAUCAAAGCACUUCUGUUAACCCCGACUGAGUAUCAUAGACUGUCACCGGUUGAAGGAGAAAG CGUUCGUUAUACCGGCCAACUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUGUUUCGU CAGCACUACCCAGUGUAGAUACAGGUCGAUGAGUCACCGCAUUCACCGGGCGACCGUGGCGGUGG CUGCGUUUGCGCGCCUGCCCAUGGGGAACCCAUUGGACGCUUUAUACAGACAUGGUGCGAAGAGUC UAUUGAGCUAGUUGGUGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACUGCGGAGCACACCCCU CAGCCACAGAGGCGAGUGUGCUAACGGGCAACUCUGCAGCGGAACCGACUACUUUGGGUGUCCGUG UUUCAUUUAUUCUUAUACUGGCGUCUUAUGGUGACAAUUGAGAGAUUGUACCAUAUAGCUAUUGG AUUGGCCAUUCGUGACUAAUAGAGCUAAUUAUUAUCCCUUUGUUGGGUUUAUACCAUUAAGCUUA AAGAGGUUAAACAUAUACAAUUAUUGUUAAGUUAUAGAAUACAGCAAAUAGUUAUAGAAUGCAGCUGCUG UCUUGCAUCGCGCCUGAGCCUGGCGCCUGGUCACAAACAGCGCUCUUAACAGCUCUAGCACCAAGAAA CCAGCUGCAGCUCGAGCACCCUGCUGGACCGUGCAGAUAGUACUGAACGGCAUUAACAACUACAAA GAACCCAAAGCUGACCAGAAUGCUGACAUUAAGUUCUACAUGCCUAGAAGGCCACAGAGCUGAAG CACCUGCAUUGUCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUAUUCUGGCUAGAGCAAGA ACUUCACCCAGAGACCUCCGGACCGUAGCAGCGACAUAACGUGAUCGUGUGGAACUGAAAGGCGA CGAGACAACCUUAUGUGCGAGUACCGCGAUGAGACCGCCACCAUCUGGAGUUCUGAAUCGGUGG AUCACCUUUCGCGAGUCCAUCAUCCACCCUGACCUGAUAAUAGAGACGCUACGGACUUAUUAU UGAGCCUUAAGAAGAAAUUCUUUAAGUGGAUGCUCUAAACUCAGGGAAACCUAAAUUAGUUAUA GACAAGGCAUUCUGAGCCAAGCCGAAGUAGUAUUAUAGUAAAGACAGUGGACAAUCCAGCGAAUACA GCAUAUCUAG	AnaX-CVB3- IL2 precursor
5 2 8	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACAAUAGAUGACUUAACAUAUUCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUCAAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAGCCA AUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAGAAGUAAACACACACCGAUCAACAGUCAGCGUGGCAC ACCAGCCACGUUUUGAUCUAGCACUUCUGUUAACCCCGGACUGAGUAUCAUAGACUGCUCACGCGGU UGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCA GAGUUUUCCGUCAGCACUACCCAGUCUAGAUACAGGUCGAUGAGUACCGCAUUCGCCACGCGGCGA CCGUGGCGGUGGCGUGCGUUGGCGGCCUGCCCAUGGGGAACCCAUUGGACGCUUUAUACAGACAUG GUGCGAAGAGUCUUAUGAGCUAGUUGGUGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCUGCGG AGCACACACCCUCUAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUU UGGGUGUCCGUGUUCAUUUUUAUCCUUAUACUGGCGUCUUAUGGUGACAAUUGAGAGAUUGUACCA UAUAGCUUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUUAUCCCUUUGUUGGGUUUAUAC CACUUAAGCUUAAAGAGGUUAAACAUAUACAUAUUGUUAAGUUAUACAGCAAAGCCACCAUG UAUAGAAUGCAGCUGCUCUUGCAUCCGCGCCUGAGCCUGGCGUAGGUCACAAACAGCGCUCUUAACA GCUCUAGCACCAAGAAAACCCAGCUGCAGCUCGAGCACCGCUGCUGGACCGCAGAUAGUCCUGAA CGGCAUUAACAACUACAAGAACCCAAAGCUGACAGAAUGCUGACAUUAUAGUUCUACAUGCCUAAAG AAGGCCACAGACCUCAAGCCAGAGCACCUUGCAUUGCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUA AUCUGGCGCAGAGCAAGAACUUCACCCUGAGACCUCCGGACCGUAGCAGCGACAUAACGUGAUCGU GCUGGAACUGAAAGCAGCGAGACAACCUUAUGUGCGAGUACGCCAUGAGACCGCCACCAUCGUG GAGUUCUGAAUCCGUGGUAUACCUUCGCCAGUCCAUCAUCCACCCUGACCUGAUAAUUAUAAAC AGCCUGUGGGUUGAUUCCACCCACAGGCCCAUUGGGCGCUAGCACUUCGUGUAUACGGUACCUUUGU GCGCCUGUUUAUACCCCUCCCCCAACUGUAACUUAUUAUAGAGCCUUAAGAAGAAAUUCUUU AAGUGGAUGCUCUAAACUCAGGGAAACCUAAAUUAGUUAUAGACAAGGCAAUCCUGAGCCAAGCC GAAGUAGUAUUUAGUAAAGCCAGUGGACAAUCCAGCGAAUACAGCAUAUCUAG	AnaX-CLEAN- PIE-IRES- rawCVB3- CVB3-IL2 precursor
5 2 9	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACAAUAGAUGACUUAACAUAUUCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUCAAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAGCCA AUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAUAUACAACACACCGAUCAACAGUCAGCGUGGCAC ACCAGCCACGUUUUGAUCUAGCACUUCUGUUAACCCCGGACUGAGUAUCAUAGACUGCUCACGCGGU UGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCA GAGUGUUUCGUCAGCACUACCCAGUGUAGAUACAGGUCGAUGAGUACCGCAUUCGCCACGCGGCGA CCGUGGCGGUGGCGUUGGCGGCCUGCCCAUGGGGAACCCAUUGGACGCUUUAUACAGACAUG GUGCGAAGAGUCUUAUGAGCUAGUUGGUGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCUGCGG AGCACACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUU UGGGUGUCCGUGUUCAUUUUUAUCCUUAUACUGGCGUCUUAUGGUGACAAUUGAGAGAUUGUACCA UAUAGCUUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUUAUCCCUUUGUUGGGUUUAUAC CACUUAAGCUUAAAGAGGUUAAACAUAUACAUAUUGUUAAGUUAUACAGCAAAGCCACCAUG UAUAGCUUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUUAUCCCUUUGUUGGGUUUAUAC CACUUAAGCUUAAAGAGGUUAAACAUAUACAUAUUGUUAAGUUAUACAGCAAAGCCACCAUG UAUAGAAUGCAGCUGCUCUUGCAUCCGCCUGAGCCUGGCGGCGGUCACAAACAGCGCUCUUAACA GCUCUAGCACCAAGAAACCAAGCCAGCUGACAGAAUGCUGACAUUAUAGUUCUACAUGCCUAAAG CGGCAUUAACAACUACAAGAACCCAAAGCUGACAGAAUGCUGACAUUAUAGUUCUACAUGCCUAAAG AAGGCCACAGACCUCAAGCCAGAGCACCUUGCAUUGCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUA AAGGCCACAGAGCUGAAGCACCUGCAUUGCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUA	AnaX-CLEAN- PIE-IRES- rawMutation2 -CVB3-IL2 precursor

	AUCUGGCUAGAGCAAGAACUCCACCGUGAGACCUCGGGACCUGAUCAGCGACAUAACGUGAUCGU GCUGGAACUGAAAGCGAGCGAGACAACCUUACUGUGCGAGUACGCCGAGAGACC GCCACCAUCUGUG GAGUUCUGAAUACGGUGGUAUACCUUCUGGCCAGUCCAUUCCACCCUGACCUGAUAUUAUAAAC AGCCUGUGGGUUAUCCACCCACAGGCCCAUUGGGCGCUAGCACUCUGGUAUACGGUACCUUUUGU GCGCCUGUUUUAUACCCCCUCCCCAACUGUAACUUAUUAUUGAGCCUUAAGAGAGAAUUCUUU AAGUGGAUGCUCUAAACUCAGGGAACCUAAUUCUAGUUAUAGACAAGGCCAUCCUGAGCCAAGCC GAAGUAGUAUUAAGAACCCAGUGGACAUCGACGGUAACAGCAUAUCUAG	
5 3 0	GGGAGACCCUCGACCGUGCAUUGUCCACUGGUAACAAUAGAUGACUUAACUAUCCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUAAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCA AUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAACCUCGUCAGAGCAAGAACUCCACCGUGAGACCU CGGGACUGAUCAGCGACAUAACGUGAUCGUGCGGAAACUGAAAGGCAGCGAGACAACCUUACUGU GCGAGUACGCCGAGAGACCGCCACCAUCGUGGAGUUCUGAAUCGGUGGAUACCUUCUGCCAGUC CAUCAUCUCCACCCUGACUGAUAUUAUAAACAGCCUGUGGGUUGAUCACCCACAGGCCCAUUGG GCGCUAGCACUCUGGUAUCACGGUACCUUUGUGCGCCUGUUUAUACCCCCUCCCCAACUGUAACU UAGAAGUAACACACACCGAUCAACGAGUCAGCGUGGCACACCGCCACGUUUUGAUCAGCAAGCACUUCG UUACCCCGGACUGAGUAUCAAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCA ACUACUUCGAAAAACCUAGUAACACCGUGGAGUUGCAGAGUGUUCGUCUACGACUACCCACAGUGU AGAUCAGGUCGAGUAGUCACCGCAUUCUCCACGCGGCGACCGUGGCGGUGGUGCGUUGGCGGCCUGC CCAUGGGGAAACCCAUUGGACGCUUAAUAACAGACAUGGUGCGAAGAGUUAUUGAGCUAGUUGGUA GUCCUCCGGCCCCUGAAUUGCGGUUAUCCUAAACUGCGGAGCACACCCUCAAAGCCAGAGGGCAGUG UGUGCUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUAUUAUUAUUCUUAU AAGUGGAGUCUCUAAACUCAGGGAACCUAAUUCUAGUUAUAGACAAGGCCAAUCCUGAGCCAAGCC GAAGUAGUAUUAAGAACCCAGUGGACAUCGACGGUAACAGCAUAUCUAG	AnaX-CLEAN- PIE-Cargo- PIE1-CVB3- IL2 precursor
5 3 1	GGGAGACCCUCGACCGUGCAUUGUCCACUGGUAACAAUAGAUGACUUAACUAUCCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUAAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCA AUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAUCUCUGACAUAACGUGAUCGUGCGAAGUGGUA GGCAGCGAGACAACCUUACUGUGCGAGUACGCCGAGAGACCGCCACCAUCGUGGAGUUCUGGUAUC GGUGGAUCACCUUCUGCCAGUCCAUCAUCUCCACCCUGACCUGAUAUUAUAAACAGCCUGUGGGUUG AUCCCAACCCACAGGCCCAUUGGGCGUAGCACUCUGGUAUCACGGUACCUUUGUGCGCCUGUUUAU ACCCCCUCCCCAACUGUAACUUAAGUAACACACACCGAUAACAGUCAGCGUGGCACACCCAGCC ACGUUUUGAUAAGCACUUCUGUUAACCCGGACUGAGUAUCAAUAGACUGCUCACCGCGUUGAAGGA GAAGCGUUCUUAUCCCGGCAACUACUUCGAAAACCUAGUAACACCGUGGAAUGUGCAGAGUGU UCGCUACGACUACCCACAGUGAUAUCAGGUCGAGAGUCACCGCAUUCUCCACGGGCGACCGUGGC GGUGGUGUGUGUGCGGCGUGCCAUUGGGGAACCCAUUGGAGCGUCUUAUACAGACAUGUGGCGAA GAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACUGCGGAGCACAC ACCCUAAAGCCAGAGGGCAGUGUGUACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGU CCGUGUUUAUUAUUAUCCUAUACUGGCGUUAUUGGUGACAUAUAGAGAUCCGUACCAUAUAGCU AUUGGAUUGGCCAUCCGUGGACUAAUAGAGCUAUAUUAUUAUCCUUGUUGGGUUUAUACACUUAU CUUGAAAGAGUUAAACAUUAACAAUUAUUAAGUUAAGUUAACAGCAAGCCACCAUGUAUAGAA UGCAGCUGCUGUUAUUGCAUCGCCUGAGCCUGGGCCUGGUCACAACAGCGCUCUACAAGCUCUAG CACCAGAAACCCAGCUGCAGCUCGAGCACCUGCUGCGGACCUGCAGAUUAUCCUGAACGGCAU AACAAUACAAGAACCCAAAGCUGACCCAGAAUGCUGACAUAUUAAGUUAUUAUCCUUAAGAGGCCA CAGAGCUGAAGCACCUGCAUUGUCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUAUUCUGGC UCAGAGCAAGAACUCCACUGACACCUGAGACUUAUUAUUAUAGGCCUUAAGAGAAAUUCUUU AAGUGGAGUCUCUAAACUCAGGGAACCUAAUUCUAGUUAUAGACAAGGCCAAUCCUGAGCCAAGCC GAAGUAGUAUUAAGAACCCAGUGGACAUCGACGGUAACAGCAUAUCUAG	AnaX-CLEAN- PIE-Cargo- PIE2-CVB3- IL2 precursor
5 3 2	GGGAGACCCUCGACCGUGCAUUGUCCACUGGUCUCCGUGCAUUGUCCACUGGUAACAAUAGAUGACU UACAACUAUCCGGAAGGUGCAGAGACUCCAGCGGAGCUACCCUAACGUAAGACGAGGGUAAAGAGA GAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAACCCCCCAUUAUAA ACAGCCUGUGGUGUAUCCACCCACAGGCCCAUUGGGCGCUAGCACUCUGGUAUACACGGUACCUUU GUGCGCGUGUUUAUACCCCCUCCCCAACUGUAACUUAAGAAUACACACACCGAUAACAGUACAG CGUGGCACACCCAGCCAGUUUUAUUAAGCAUUCUGUUAACCCCGGACUGAGUAUCAAUAGACUGCU CACCGGUGUGAAGGAGAAAGCGUUCGUUAUCCGGCCAACUACUUCGAAAAACCUAGUAACACCGUGG AAGUUGCAGAGUGUUCGUCAGCACUACCCAGUGUAGAUACGGUCGAGUACCGCAUUCUCC ACGGGCGACCGUGGCGGUGGCGUGCGUUGGGCGCCUGCCAUUGGGGAACCCAUUGGAGCGCUUAUA CAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCU AACUGCGGAGACACACACCCUAAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUUCGAGCGGAAC GACUACUUGGGUGUCCGUGUUUAUUAUUAUCCUUAUACUGGCGUUAUUGGUGACAUAUAGAGAU CGUUAACCAUAUAGCUAUAUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUAUAUUAUCCUUGUUGG GUUAUAACCAUUAAGCUUAGAAAGAGGUUAACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA CCACCAUGGCGAGCAGUUAAGAAAGAGCUCUUCUGGAGGGCAGUGGUGGCGGAGUUCUGGCCAC GACCCUGUUUGUCUUAUAGCAUCGUGUUCUCCCGUGGCUUCAAUACCCGUGGGGAAACACACAG	AnaX-Gcv3- 2xhomo-CVB3- AQF1 precursor

	ACGGCGGUCCAGGACAACGUAAGGUGUCGUGGCCUUCGGGUGAGCAUCGCCACGUGGCGCAGA GUGUGGGCCACAUCAGCGGGGCCACCUCACACCCGGCUGUCACACUGGGGUGCUGCUCAGCUGCCA GAUCAGCAUCUUCGUGGCCUCUAGUACAUCAGGCCAGUGGUGGGGGCCAUUCGUGGCCACCGCC AUCCUCUCAGGCAUACCCUCCUGACUGGGAUCUGCUUGGGCCGAAUAGCCUGGUGAUGGUG UGAACUCGGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGGACCCUCCAGCUGGUGCUAUGCGUGCUGGC UACUACCGACCGGAGGCGCGUGACCUUGGUGGCUAGCCCCCUUGGCAUCGCGCCUCUGUAGCC CUUGGACACCCUUGGCUAUGACUACACUGGCUUGGGGAUUAACCCUGCUCGGUCCUUUGGCUCCG CGGUGAUCACACACAACUUCAGCAACCACUGGAUUUCUGGGUGGGGCCAUUCAUCGGGGAGCCCU GGCUGUACUACUACGACUUAUCUUGGCCCCACGCAGCAGUGACCUACAGACCCGCGUGAAGGUG UGGACCACGGGCCAGGUGGAGGAGUAUGACCUGGAUGCCGACGACAUCAACUCCAGGGUGGAGUA AGCCCAAUGAUAAAGACGGGGGGGCUAAAUAUUGAGCCUUAAGAAGAAAUCUUUAAGUGGA UGCUCUCAAACUCAGGGAACCUAAAUUAGUUAUAGACAAGGCAUCCUGAGCCAGCCGAAGUAG UAAUAGUAAGACCAGUGGACAAUCGACGGGACCAGUGGACAAUCGACGGAUAACAGCAUUCUAG	
5 3 3	GGGAGACCCUCGACCGUGCAUUGUCCACUGGUCGUGCAUUGUCCACUGGUAACAUAAGAUGACU UACAACCUAAUUGGGAAGGUGCAGAGACUCGACGGGAGCUACCCUACGUAAGACGAGGUAAGAGA GAGUCCAUAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAACUGGAAACUCGCUUG GGCGCAUAGACCGUGCUGAUGGUGUAGAACUCGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGGACCCU CCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCGGAGCGCGCGUGACCUUGGUGGCUAGCCCCC CUUGCCAUUCGGCCUCUCUGUAGCCCUUGGACACCUCCUGGCUAUGACUACACUGGCUUGGGAUUA ACCCUGCUGGUGCCUUUGGCUCCGCGGUGAUCACACACAACUUCAGCAACCACUGGAUUUCUGGGU GGGGCAUUCUACGGGGGAGCCUGGCGUACUACUACGACUUAUCCUGGGCCACGCAGCAGU GUUUGCGCCUGCCUGAAGGUGGAGCCAGCGGCCAGUGGAGGAGUAUGACUUGGAUGCCGACG ACAUAACUCCAGGGUGGAGAUAGAAGCCCAAUGAUAAUAAAACAGCCUGUGGGUUGAUCCACCC ACAGGCCCAUUGGGCGCUAGCACUCUGGUAUACCGGUACCUUUGUGCGCCUUGUUUAUACCCCUCC CCCAACUGUAACUUAAGAUAACACACACCGAUACAACAGUCAGGUGGACACACAGCCACGCUUUGA UGGACGACUUCUGUUAUCCCGGACUGAUUCAAUAGACUGCUACGCGGUUGAAGGAGAAAGCGUU CGUUAUCCGGCCAACUACUUCGAAAACCUAGUAACACCGUGGAAGUUGCAGAGUGUUUCGUCAGC ACUACCCAGUGUAGAUACAGGUGCAUGAGUACCCGCAUUCGCCACGGGCGACCGUGGCGGUGGUGC GUUUGCGCCUGCCUGCCAUUGGGGAACCCAUUGGGAGCGUCUAAUAACAGACAUGGUGCGAAGAGUCU GAGCUAGUUGGUGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCUGCGGAGCACACCCUCAAG CCAGAGGGCAGUGUGCUAAGCGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUC AUUUUAUCCUUAUACUGGCGUUAUUGGUGACAUAUGAGAGAUUGUACCAUAUAGCUAUGGUAUG GCCAUCCGGUGACUAAUAGGCUAUAUAUAUCCCUUUGUUGGGUUUAUACCAUAGCUUGAAAGA GGUAAAACAUUACAAUUAUUGUUAAGUUGAAUAACGCAAGGCCACCAUGGCCAGCGAGUUAAGA AGAAGCUCUUCUGGAGGGCAGUGGUGGCCGAGUUCUUGGCCACGACCUCUUGUCUUAUCAGCAU CGGUUCUGCCUGGGCUUCAAUAUACCCGUGGGGAACAACAGAGCGCGUCCAGGACAACGUGAAG GUGUGCGUGGCCUUCGGGCUAGCAUCGCCACGCGUGGCGCAGAGUGUGGGCCACAUCAGCGCGCGCC ACCUAACCAGCGGUGUACACUGGGGCGUGGUGCUCAGCUGCCAGAUACGACUUCUUGCGGCCUACU GUACAUCAUCAGCCCAUGGCGUGGGGGCCAUUGGCGCCACCGCCAUCCUUCAGGGCAUACCUCCAGC UUAAUAUUGAGCCUUAAGAAGAAAUUCUUUAAGUGGAGUCUCUCAAACUCAGGGAAACCUAAAU CUGAUUAUAGACAAGGCAUUCUGAGCCAGCCGAGUAGUAUAUAGUAAGACCAGUGGACAUCGA CGGGACACGUGGACAUCGACCGGAUAACAGCAUUCUAG	AnaX-GCv3- 2×homo- CLEAN-PIE- Cargo-PIE1- CVB3-AQP1 precursor
5 3 4	GGGAGACCCUCGACCGUGCAUUGUCCACUGGUAUAUAGAGAAAGUGCAACGCAUUAUCCGAUAGGA AGUAGGGUCAAGUGACUCGAAAUGGGGAUUAACCUUUAAGGUAUGAGUAUAGUCUGAACAUUAUUG GAACAUUAUAGAGGAUAGGAGUAACGAACCUAUCGUAACAUAAUUGAACCAGCUCUUGAGUGGG AACUCGCUUGGGCCGCAUAGACCGUGGUGGUAAGGUAACUGGGCCAGGGCCUGGGCAUCGAGAUCA UCGGGACCCUCCAGCUGGUGCUAUGCGUGGCUACUACCGACCGGAGGCGCGGUGACCUUGGUGG CUCAGCCCCCUUGCCAUCCGGCCUCUCUGUAGCCCUUGGACACCUCCUGGCUAUGGACUACACUGGC UGUGGGAUUAACCCUGCUGCGUCCUUGGCUCCGCGGUGAUCACACACAACUUCAGCAACCAUGGA UUUUUCUGGGUGGGGCAUUAUCGGGGGAGCCUGGCGUACUACUACGACUUAUCCUGGCCCC ACGCAGCAGUAGCCUACAGACCCGCGUAAGGUGUGGACAGCGGCCAGGUGGAGGAGUAUAGCCUG GAUGCCGACGCAUCAACUCCAGGGUGGAGAUAGAAGCCCAAUGAUAAUUAACAGCCUGUGGGUU GAUCCCAACCCACAGGCCAUUGGGCGUAGCACUCUGGUUAUCGGAUCCUUGUGCGCCUGUUUA UACCCCUCCCCCAACUGUAACUUAAGAUAACACACACCGAUACAACAGUCAGCGUGGCACACCGC CAGCUUUGUAUCAAGCACUUCUGUUAACCCCGACUGAGUAUCAAUAGACUGCUCACGCGUUGAAGG AGAAAGCGUUCGUUAUCCGGCCAAUCUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUGU UUCGCUACGACUACCCAGUGUAUAGAGUACAGGUCGUAUGAGUACCGCAUUCGCCAGGGCGACCGUGG CGGUGGCGUGGUGGCGCCUGCCAUUGGGGAACCCAUUGGGACGCUUAUAACAGACAUGGUGGGA AGAGUCUAUUGAGCUAGUUGGUAUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCGCGGAGCACA CACCUCUAGGCGAGGGCAGUGUGGUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUG UCCGUGUUAUUAUUAUCCUUAUACUGGCGUUAUUGGUGACAAUUGAGAGAUUGUACCAUAUAGC UAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUAUAUAUCCCUUUGUUGGGUUUAUACCAUUA GCUUGAAGAGGUAAAACAUUAACAUUAGUUAAGUUAAGUAUACGCAAGGCCACCAUGGCCAGC GAGUUAAGAAAGAGCUCUUCUGGAGGGCAGUGGUGGCCGAGUUCUUGGCCACGACCCUCUUGUCU UCAUCAGCAUCGGUUCUGCCUGGGCUUCAAUACCCGGUGGGGAACACACAGCGCGGUGCCAGGA CAACGUGAAGGUGUGCUGGCGUCCGGGCGUGAGCAUCGCCACGCGUGGCGCAGAGUGUGGGCCACAU AGCGGCGCCACCUCAACCCGGCUGUACACUGGGGCGUGGUGCUCAGCUGCCAGAUACGACUUCUCC GUGCCCUAUGUACAUCAUCGCCAGUGCGUGGGGGCCAUUGGCGCCACCGCCAUCCUUCUGGCAU AAUUAUUGGCUUUAUACAGUAUUGUAUUGGAAAAUCCUUAUUCAGGGAACACCUAAACAA	Twort-PIE-6- CLEAN-PIE- Cargo-PIE1- CVB3-AQP1 precursor

	ACUAAAGUAGGCAAUCCUGAGCUAAGCUCUAGUGACCAGUGGACAAUCGACGGUAACAGCAUAUCUAG	
5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUGCCGUCGAUUGUCCACUGGUCAACAAUAGAUGACU	AnaX-GCv3- 2×homo- CLEAN-PIE- IRES- rawCVB3-AQP1 precursor
3	UACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAGACGAGGGUAAAGAGA	
5	GAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAGAAGUAAACACACACCG	
	AUCAACAGUCAGCGUGGCACACACAGCCACGUUUUGAUAAGCACUUCUGUUACCCCGACUGAGUAU	
	CAAUAGACUGUCACCGGUGAAGGAGAAAGCGUUCGUUAUCCGCGCCAAUCUACUUCGAAAAACCUA	
	GUAAACCCGUGGAAGUUGCAGAGUGUUUCGUCAGCACUACCCAGUGUAGAUACAGGUCGAUGAGUC	
	ACCGCAUUCGCCACGGGACCGUGGCGGUGGUCGUUGGCGGCCUGCCCAUGGGGAAACCCAUUGG	
	GACGCUUAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGCCCCUGAAU	
	GCGGCUAAUCCUACUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUC	
	UGCAGCGGAACCCGACUACUUGGGUGUCCGUGUUUCAUUUAUUCUUAUACUGGCGUCUUAUGGUGA	
	CAAUUGAGAGUUCGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUUAUAUA	
	UCCCUUUGUUGGUUUUAUACCACUAGCUUGAAAGAGGUUAAAACAAUACAUAUUGUUAAGUUG	
	AAUACAGCAAAGCCACCAUGGCCAGCGAGUUAAGAAGAAGCUCUUCUGGAGGGCAGUGGUGGCCGA	
	GUUCCUGGCCACGACCCUUCUUGUCUUAUCAGCAUCGGUUCUGCCUGGGCUCAAUAUACCCGGUG	
	GGGAACACCCAGACGGCGGUCAGGACAACGUGAAGGUGUGCGUGGCCUUCGGGCGUAGCAUCGCCA	
	CGUGGGCAGAGUGUGGGCCACAUCAGCGGGCGCCACCUCACCCGGCUGUCACACUGGGGUGCU	
	GCUCAGUGCCAGAUACAGAUUUCGUGCCUCAUGUACAUCUACGCCAGUGCGUGGGGGCCAUUC	
	GUCCGCCACCGCAUCCUCUACAGCAUACCCUCCUUGACUGGGAACUCGCUUGGCCGCAUAGACC	
	UGGUGAUGGUGAUCUCCGGCCAGGGCCUGGGCAUCGAGAUCAUCGGGACCCUCCAGCUGGUGCU	
	AUGCGUGCGUACUACCGACCGGAGCGCGUGACCUUGGUGGUCAGCCCCCUUGCCAUCGGC	
	CUCUCUGUAGCCCUUGGACACCUCCUGGCUAUUGACUACACUGGCGUGGGGAUUAACCCUGCUCGGU	
	CCUUUGGCUCCCGGUGAUCACACACAAUUCAGCAACCCUUGGAUUUCUGGGUGGGGCCAUUCAU	
	CGGGGAGCCUUGGCUUACUACUACGACUUAUCCUGGCCACGACGAGUGACCUACAGAC	
	CGCGUGAAGGUGUGGACCGCGCCAGUGGAGGAGUAUAGCCUGGAUGCCGACGACAUAACUCCA	
	GGGUGGAGUAAGCCCAAUAGUAUUAUAAACAGCCUGUGGUUGAUCCACCCACAGGCCCAUUG	
	GGCGCUAGCACUCUGGUUAUCAGGUACCUUUGUGCGCCUGUUUAUACCCCUCCCCCAUCUGUAAC	
	UUAAAUAUUGAGCCUUAAGAAGAAAUUCUUUAAGUGGAGUCUCUAAACUCAGGGAAACCUAAAU	
	CUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAAGUAGUAUUAUAGUAGACCAGUGGACAAUCGA	
	CGGGACCAUGGGACAUCGACGGUAACAGCAUAUCUAG	
5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUGCCGUCGAUUGUCCACUGGUCAACAAUAGAUGACU	AnaX-GCv3- 2×homo- CLEAN-PIE- IRES- mutationCVB3 -AQP1 precursor
3	UACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAGACGAGGGUAAAGAGA	
6	GAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAGAAUUAACAACACACCG	
	AUCAACAGUCAGCGUGGCACACACAGCCACGUUUUGAUAAGCACUUCUGUUACCCCGGACUGAGUAU	
	CAAUAGACUUCGACCGCGUUGAAGGAGAAAGCGUUCGUUAUCCGCGCCAAUCUACUUCGAAAAACCUA	
	GUAAACCCGUGGAAGUUGCAGAGUGUUUCGUCAGCACUACCCAGUGUAGAUACAGGUCGAUGAGUC	
	ACCGCAUUCGCCACGGGACCGUGGCGGUGGUCGUUGGCGGCCUGCCCAUGGGGAAACCCAUUGG	
	GACGCUUAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCUUGGAU	
	GCGGCUAAUCCUAAUCGCGGAGCACACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUC	
	UGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUCAUUUAUUCUUAUACUGGCGUCUUAUGGUGA	
	CAAUUGAGAGUCCGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUUAUAUA	
	UCCCUUUGUUGGGUUUAUACCACUAGCUUGAAAGAGGUUAAAACAAUACAUAUUGUUAAGUUG	
	AAUACAGCAAAGCCACCAUGGCCAGCGAGUUAAGAAGAAGCUCUUCUGGAGGGCAGUGGUGGCCGA	
	GUUCCUGGCCACGACCCUUCUUGUCUUAUCAGCAUCGGUUCUGCCUGGGCUCAAUAUACCCGGUG	
	GGGAACAACCAAGACCGCGGUGCCAGGACAACGUGAAGGUGUGCGUGGCCUUCGGGCGUAGCAUCGCCA	
	CGCUGGCGCAGAGUGUGGGCCACAUCAGCGGGCGCCACCUCACCCGGCUGUCACACUGGGGUGCU	
	GCUCAGUGCCAGAUACAGAUUUCGUGCCUCAUGUACAUCUACGCCAGUGCGUGGGGGCCAUUC	
	GUCGCCACCGCCAUCCUCUACAGCAUACCUCCUCCUGACUGGGAACUCGCUUGGCCGCAUAGACC	
	UGGUGAUGGUGUAAUCUGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGGACCCUCCAGCUGGUGCU	
	AUGCGUGCGUGCUACUACCGACCGGAGGCGCGUGACCUUGGUGGUCAGCCCCCUUGCCAUCGGC	
	CUCUCUGUAGCCCUUGGACACCUCCUGGCUAUUGACUACACUGGCGUGGGGAUUAACCCUGCUCGGU	
	CCUUUGGCUCCCGGUGAUCACACACAACUUCAGCAACCCUUGGAUUUCUGGGUGGGGCCAUUCAU	
	CGGGGGAGCCUUGGCGUACUACUACGACUUAUCCUGGCCACGACGAGUGACCUACAGAC	
	CGCGUGAAGGUGUGGACCGCGCCAGUGGAGGAGUAUAGCCUGGAUGCCGACGACAUAACUCCA	
	GGGUGGAGUAAGCCCAAUAGUAUUAUAAACAGCCUGUGGUUGAUCCACCCACAGGCCCAUUG	
	GGCGCUAGCACUCUGGUUAACCGUACCUUUGUGCGCCUGUUUAUACCCCUCCCCCAUCUGUAAC	
	UUAAAUAUUGAGCCUUAAGAAGAAAUUCUUUAAGUGGAGUCUCUAAACUCAGGGAAACCUAAAU	
	CUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAAGUAGUAUUAUAGUAGACCAGUGGACAAUCGA	
	CGGGACCAUGGGACAUCGACGGUAACAGCAUAUCUAG	
5	GGGAGACCCUCGAGGACACCCCUAGACCCGGUCUCAGAAAGCCUGAAAGCAGAAUAGUCAUUAAGAA	AnaX-MEHA1- CVB3-IL2 precursor
3	GGGUGGUUGGCUUGGUCGGCAUAGACUUGAGCAGAAAGAGGUUGAAAAUGUUGAGCCUGAUUUCUC	
7	UUAGGCCCUUCUGCAGUGUCUGUUGGAGGGCCAGAUACGUAACUGCUUCCGCUUUUUUUGGUCUCA	
	UUCAGGUGGACAAAUCCCUUACAGUUCUACCAAGACAAGCAGCUGAUGAGGUUCCAGCUUUGC	
	AAUAGUAGUCUACCAUGAGGACUCUCUUCUGAGAUUUCUACAGGAACAAUAGAUGACUUAACUUA	
	AUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAGACGAGGGUAAAGAGAGUCCAA	
	UUCUCAAGCCAAUAGGCAGUAGCGAAGCUGCAAGAGAAUGAAAACCCCAUUAUAAACAGCCUG	
	UGGGUUGAUCCACCCACAGGCCAUUGGGCGCUAGCACUCUGGUUAUCAGGUACCUUUGUGCGCCU	
	GUUUUAUACCCCUCCCAACUGUAACUAGAAGUAACACACCCGAUACAACAGUCAGCGUGGCAC	

	ACCAGCCACGUUUUGAUCACAGCACUUCUGUUUACCCCGGACUGAGUAUCAUAGACUGCUCACGCGGU UGAGGAGAAAGCGUUCGUUAUCCCGGCCAACUACUUCGAAAAACCUAGUAACACCUGGAGAGUUGCA GAGUGUUUGCGUCAGCACUACCCGAGUGUAGAUACAGGUCGAUGAGGACCCGCAUUCCCCACGGGCGA CCGUGGCGGUGCGUGCGUUGGCGGCGUGCCCAUGGGGAAACCCUGGAGCGUCUUAAGACAGACAUG GUGCGAAGAGUUCUAUUGAGCUGAGUUGGUAGUCCUCCGCCCCUGAAUUGCGGCUAAUCCUAACUGCGG AGCACACACCUCUAAAGCCAGGGGACAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCCGACUACU UGGGUGCCGUGUUAUUAUUUAUUAUACUACUGGCGCUUAUUGGAGACAUAUGGUAUCGUUUAACCA UUAUAGCUAUUGGAUUGGCCAUCCCGGUGACUAUAGAGCUAUUAUAUUAUCCCUUUGUUGGGUUUAUAC CACUUAAGCUUGAAAGAGGUUAAAACAUUACAUAUUGUUAAGUUAUACAGCAAGGCCACCAUG UUAUAGAUAGCAGCUGCUCUUGCAUCGCCUGAGCCUGGCCUGGCUUGGUCACAAACAGCGCCUACAA GCUCUAGCACCAAGAAAACCCAGCUGCGAGCCAGCCUGCUGGAGCUGGAGUACUAGCUGAUCCUGAA CGGCAUUAACAACUACAAGAACCCAAAGCUGACCAGAUAUGCUGACAUUAUAGUUCUACAUGCCUAAG AAGGCCACAGAGCUGAAGCACCCUGCAAUGUCUGGAAGGAAACUGAAGCGCCUGGAAGAGGUGCUGUA AUCUGGCUACAGACGACAACUCCACCUGAGACCCUGGGACCCUGAUGACGACUACACAGGUAUCGU GCUGGAACUGAAAGGCAGCGAGACAACCUUAUGUGCGAGUACGCCGAGUAGAGACCGCCACCAUCGUG GAGUUCUGUAUUGCGUGGACUCCUUCUGCCAGUCCAUAUCUCCACCUGACCGUUAUAGAGCGGG GGGGCUUUAUUAUUGAGCUCUUAAGAAGAAUUAUUAAGUUCUUAACUACUACAGGGAAC CUAAAUCUAGUUAUAGACAAGGCAUUCUGAGCCAAAGCCAGUAGUAUUAUACCUUGAUGAAAU CUCAGGAAGACAGUCCUCAUGGAUGACUACAUAUUGCAAGCUGGAACCUCAUCAGCUGCAUUGUGUC GUGAGAAACAUAGAAGGGGAUUUGCUCACCUGAAUGAGACCAAAAAAGCGGGAAGCAGUUAACGUUAC UGGCCUCCACACAGACAGCAGACUGGAGGGGCCUAGAGAGAAUACAGGCAACAUUUCAACCUUUCU UGCUCAAAGUCUAUGCCGACCAGGCCAACCCUUCUUAUAGACUAAUUCUGCUUUCAGGCUUUCUG AGACCGGGUCAGGGGGUGUCCAUAACAGCAUAUUCAG	
5 3 8	GGGAGAGCCUCGAGGACACCCUGACCCGCGUCAGAAAGCCUGAAAGCAGAAUUGACUAUAGAA GGGUGGUUGCCUGUGGCGGCAUAGACUUAUGAGCGAAGAGGUGUAAACAAUAGAGUACUUAACAUC AAUCGGAGAGGUCAGAGACUCGACGGGAGCUACCCUAAACGUCACAGACGAGGGUAAAGAGAGAGUCCA AUUCUCAAGGCCAAUAGGCGAGGAGCAAAGCUGCAAGAGAAUAAACACCCCCUUAUAAACAGCCU UUGGGUUGAACCCACCCACAGCCAGCCAUUGGGCGUAGCACUCUGUAUCAGCGUACCUUUGGCGCC UGUUUUAUACCCCUCCCCCAACUGUAACUUAAGAAGUAACACACCCGAUCAACAGUCAGCGUGGCA CACCGAGCCAGGUUUUGAUCACAGCACUUCUGUUAACCCCGGACUGAGUAUCAUAGACUGCUCACGCG UUGAAGGAGAAAGCGUGUUAUUGCCGGCAACUACUUCGAGAAACCUAGUAACACCGUGGAAGUGCG AGAGUGUUCGCUCAGCACUACCCAGUGUAGAUACGGUCGAUGAGUCACCGCAUUCCCACGCGGCG ACCGUGGCGGUGGCGUGCGUUGCGCGCCUGCCAUUGGGGAAACCCAUUGGAGCGCUCUAAUACAGACA GGUGCGAAGAGUUCUAUUGAGCUGAGUUGGUAGUCCUCCGGCCCGUAUUGCGGCUAAUCCUAAACUGCG GAGCACACACCCUACAAGCCAGGGGACAGUGUGUCGUAACGGGCAACUCUGAGCGGAAACCGACUACU UUGGGUGUCCGUGUUAUUAUUAUCCUAUACUGGCGUCUUAUGGUGACAAUUGAGACAUCCGUUACC AUUAUGCUAUUUGGAUUGGCGAUCCCGGUGACUAAUAGAGCUUAUUAUUAUCCCUUUGUUGGUGUUAU CCACUUAAGCUUUGAAGAGGCUAAACAAUUAACAUAUUGUUAUUGAUAAGUACAGCAAGGCCACCAU GUUAAGAUGCAGCUGCUGUCUUGCAUCGCCCCUGAGCCUGGCGGUGGUCACAAACAGCGCUCUACA AGCUCUAGCACCAAGAAAACCCAGCUGCAGCUCGAGCACCCUGCUGUGGACCCUGCAGAUAGUCCUGA ACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAAGUGCAGUACAUAUAGUUCUACAGGCCUAA GAAGGCCACAGAGCUGAAGCACCUAGGCAUUGUCGGAAGAGGAACUUGAAGCGGAGAGGUGCUU AAUCUGGCUCAGAGCAAGAAUUCACCUGAGACCUCGGGACCUGAUCAGCGACAUCAACGUGAUCCG UGCUGGAACUGAAAGGCAGCGAGACAACCUUAUGUGCGGAGUACGCCGAGUAGAGACCGCCACCAUCUG GGAUGCCUGAUAUCCGUGGAUCACCUUUCGCCAGUCCAUAUCUACCCUGACCCUGCAUUAAGAGCAAGCGG GGGGGGCUAAUUAUUGAGCCUUAAGAAGAAUUCUUAUAGUGGAUGCUCUCAAACUCAGGGAAA CCUAAAUUCGUUAUAGACAAGGCAUUCAGGACCAAGCCGCAAGUAGUAUUAUGUAUACACCUUCU UCUGCUCAAAGUCUAUGCCGACCAAGCCACCCUUCUUAUAGACUAAUUCUGCUUUCAGGCUUUC UGAGACCGGUCAGGGGUGUCCAUAACAGCAUAUUCAG	AnaX-MEHA2- CVB3-IL2 precursor
5 3 9	GGGAGACCCUCGAAAAGAGACAGUCUCGCUUAUCGCCAGCGUGGAGUGUAGAAUGCAGUGCUGU GAUUCUUGGCUUCAGCGCAACCUUACUCCCGGAGCUACAAGCGAUUCUUCGCUUCAGCCUCCAGUA GCUAGGAUUAACAGGCGCGCCACACACCAAGCUUAUUUUUAUUAUAGAGAGUAGUGGGUAGU ACCAUGUUGACCAGGCGUGGUCUCAAACUCCUGAGCUUAAGUGACGUGCCCGCCUCAGCCUCCCAAAG UGCUUGGAUUAUAGGUGUGAGGCCACUGCAACAACAAUAGAGUACUUAACUUAUUCGGAAGGUGCAGA GACUCGACGGGAGCUACCUAACGUCAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUA GGCAUGAGCGAAAGCGUACAGGAUAUAAACACCCCCUUAUAAACAGCCUUGGGUUGAUCCACCC CACAGGCCCAUUGGGCGCUAGCACUCUGGUAUCACGGUACCUUUGUGCGCCUGUUUAUACCCCUUC CCCCACUGUAACUUAAGAAGUAACACACCCGAUCAACAGCAGCGUGGCGACACAGCCACCGUUAUUG UGAAGACAUUCUGUUAUACCCGAGUAGUAUCAUUAAGACUGCAGCAGCGGCUACAGGAGAAAGCGU UCGUUAUCCGGCCAAUACUUCGAAAAACCUAGUAACACCCGUGGAAGUUGCAGAGUGUUCGCGUCAG CACUACCCCAUGUGUAUAGUGCUGAUGACUACCGCAUUCGCCACGGGCGACCGUGGCGUGGCGU CGUUGCGGCGUGCCUAGGGGAAACCAUUGGAGCGUCUUAUACAGACAGUUGGUGCAAGAGUCUUAU UGAGCUAGUUGGUAGUUCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACUGGAGCACACACCCUCAA GCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUAUGGGUGUCCGUGU CAUUUAUUCUUAUACUGGCGUCUUAUGGUGACAAUUGAGAGAUUUACCAUUAAGCUUAUUGGAU GGCCUACCCGUGACUAAUAGAGCUUAUUAUUAUCCCUUUGUUGGUGUUAUACCUAGCUUUAAGAG AGGUUAAACAUUAACAUAUUAUUGUUAAGUUAUAGCAAGCAAGGCCACCAUGUAUAGAAUGCAGCUG CUGUCUUGCAUCGCCUGAGGCGUGGCCUGGUCACAAACAGCGUCCUACAAGCUCUAGACCAAGAA AAACCCAGCUGCGACGUCGACACUACUGCUGGAGCUGGAGUAGUACUUAAGAGGCAACGCAACUUA AAACCCAGCUGCGACGUCGACACUACUGCUGGAGCUGGAGUAGUACUUAAGAGGCAACGCAACUUA	AnaX-MEHA3- CVB3-IL2 precursor

	CAAGAACCCAAAGCUGACCAGAAUGCUGACAUUUAAGUUCUACAUGCCUUAAGAAGGCCACAGAGCUG AAGCACCUGCAAUGUCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUAAUCUGGCUCAGAGCA AGAACUUCACCUGAGACCUCGGGACCUGAUCAGCGACAUCACGUGAUCGUGCUGGAACUGAAAAGG CAGCGAGACAACCUCAUGUGCGAGUACGCCGAUGAGACCGCCACCAUCUGGAGUCCUGAAUCCG UGGAUCCUUCUGCCAGUCCAUCUCCACCUCUGACCUGAUAAAGACGGGGGGGCUUAAUAAU UGAGCCUUAAGAAGAAUUCUUAAGUGGAUGCUCUAAACUCAGGGAAACCUAAUUCUAGUUAUA GACAAGGCAAUCCUGAGCCAAAGCCGAGUAGUAAUAGUAAACUUCUACUGAAAAUACAAAAAU UAGCCGGGCGUAGAGGCAUGAGCCUGUAGUCCAGCUACUUGGAGGUGAGGCGAGGAGAAUGGCGU GAACUCGGGAGGCGAGCUGCAGUGAGCCGAGAUACACCACUGCACUCCAGCCUGGGCGACAGAG CGAGACUCCGUCUUAACAGCAUUCUAG	
5 4 0	GGGAGACCCUCGAAGAGAGUCUCGCUUCUGCCACGGCUGGAGUGAGUGGUGAUCUGGCUCACUG CAACUCUCCUCCGUCAGCAUUCUCUGCCUCAGCCUCCAGUAGCUGGAUACAGGCCCCACACCCG CUAUUUUUGUAUUUUCAGUAGAGAUACAUAUGAGUACUUAACUAAUCGGAAGGUGCAGAGACU CGACGGGAGCUACCUAACGUCUAGAGCAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCA GUAGCCGAAGCUGCAGAGAAUAGAAACCCCCCAUUAAGAACAGCCUGUGGGUUGAUCCCAACCA GGCCCAUUGGGCGCUGACACUCUGGUUACACGGUACCUUUGUGCGCCUGUUUUAUACCCCUCCCCC AACUGUAACUUAAGAGUAAACACACACCGAUCAACAGUCAGCGUGGCACACAGCCAGCGUUUGAUCA AGCACUUCUGUACCCCGACUGAGUUAUAGUAGUCUACGCGGUGAAGGAGAAAGCGUUCGU UAUCCGGCCAAACUACUUCGAAACACCUAGUAAACACCGUGGAAGUUGCAGAGUUGUUCGUCAGCACU ACCCAGUGUAGAUACAGGUGCAGUAGUACCCGCAUUCGCCACGGGCGACCGUGGCGGUGGCGUGCU GGCGGCCUGCCAUUGGGGAAACCAUGGGACGCUUAUACAGACAUGGUGCGAAGAGUCUUAUGAG UAAGUUUGUAUUCUCCGGCCCGUAAUGCGGCUAAUCCUAAACUGCGGAGCACACCCCUAAGCCA GAGGGCAGUGUGCUAAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUAU UUUUAUCCUUAUACUGGCUGCUUAUGGUGACAAUUGAGAGAUCCUUAUAGCUUAUUGGAUUGGCC AUCCGGUGGUAUUAAGAGCUUAUUAUUAUCCUUAUUGGGUUUAUACCAUUAAGCUUUAAGAGAGU UAAACCAUUAACAAUUAUUGUUAAGUUAAGUUAACAGCAAAGCCACCAUUAUAGAAUGCAGCUGCU CUUGCAUCGCCUGAGCCUGGCCUGGUCACAAACAGCGUCCUACAAAGCUCUAGCACCAGAAAGAAC CCAGCUGCAGCUCGAGCACCUGCUGCUGGACCUGCAGAUAGUCCUGAACGGCAUUAACAAUACAAAG AACCCAAAGCUGACAGAAUGCUGACAUUAAGUUCUACUAGCCUAAAGAGGCCACAGAGCUGAAGC ACCUGCAUUGUCUGGAAGAGGAACUGAAGCCCGUGGAAGAGGUGCUUAAUUCUGGCUCAGAGCAAGAA CUUCCACCCUGAGACCUCGGGACCGUAGCAGCGACAUCACGUGAUCUGGUGGAAACUGAAGGGCAGC GAGACAACCUUAUGUGCGAGUAGCGCGAGAGACCGCCACCAUCGUGGAGUUCUGAAUCCGUGGA UCACCUUCUGCCAGUCCAUCUCCACCCUGACCUGAUAAAGACGGGGGGGCUUAAUUAUUGAG CCUUAAGAAAGAAUUCUUAAGUGGAGUCUCUAAACUCAGGGAAACCUAAUUCUAGUUAUAGACA AGGCAUCCUGAGCCAAAGCCGAGUAGUAAUAGUAAACUUCUACUGAAAAUACAAAAUAGCGGG UGUGGGCGCCUGUAUCCAGCUACUGGGAGGCGAGGCGAGAGAAUGCUGACGGGAGGAGAGUUGCAGU GAGCCAGAUACACCACUCACUCCAGCCUGGGCGAAGAGCGAGACUCUUAUACAGCAUUAUCUAG	AnaX-MEHA4- CVB3-IL2 precursor
5 4 1	GGGAGACCCUCGAGGCCCCUCUGCAGUUCUCUUGUGGAGGCCAGAUACGUAACUGCUUCCGCAACA AUAGAGACUUAACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCUUAACGUCAGAGCAG GGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCGAGUAGCGAAAGCUGCAAGAGAAUGAAAACCC CCCCAUUAAACAGCCUGUGGGUUGAUCCACCCACAGGCCCAUUGGGCGCUGAGCUCUGGUUAUCA CGGUACCUUUGUGCGCCUGUUUAUACCCCUCCCCCAACUGUAACUUAAGAAUACACACACCGAU CAGACAGUACGGGUGGCACACACGCCACGUUUUGAUCAAGCACUUCUGUUAACCCGGACUGAGUAUCA AUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCUUAUCCGGCCAAUACUUCGAAAAACCUAGU AACACCGUGGAGUUGCAGAGUGUUCGUCUAGCACAUCACCCAGUGUAGAUACGGUCGAGUAGUCAC CGCAUUCGCCACGGGCGACCGUGGCGUGGCGUUGGCGGCCUGCCUAGGGGAAACCCAUUGGGA CGCUCUAAUACAGACAUUGGUGCGAAGAGUCUUAUAGGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGC GGCUAAUCCUAAUCUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUG CAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUAUCCUUAUACUGGCUGCUUAUGGUGACA AUUGAGAGAUCCUUAACUUAUAGCUUAUUGGAUUGGCAUCCGGUGACUAAUAGAGCUAAUUAUUAUC CCUUGUGUGGUUUAUACCAUUAAGCUUGAAGAGGUUAAACAUUAACAAUUCUUAAGUUAAGUUA UACAGCAAGGCCACCAUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCCUGAGCCUGGCCUGGUGCA CAAACAGCGCUCCUACAAGCUCUAGCACCAAGAAACCCAGCUGCAGCUGCAGCACCUGCUGCGGA CCUGCAGAUAGUCCUGAACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAAGUAGCAUUAU AAGUUCUACAUGCCUAAAGAGGCCACAGAGCUGAAGCACCUGCAUUGUCUGGAAGAGGAACUGAAGC CCUGGAAGAGGUGCUUAAUUCUGGCUCAGAGCAAGAACUCCACUGAGACCUCCGGGACCUAGUACG CGACAUCAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAAACUUAUGUGCGAGUACGCCGAG GAGACCGCCACCAUCGUGGAGUUCUGAAUCGGUGGAUCACCUUCUGCCAGUCCAUCAUCUCCACCC UGACCUGAUAAAGACGGGGGGGCUUAAUUAUUGAGCCUUAAGAAAGAAUUCUUAAGUGGAUGC UCUCAACUCAGGGAAACCUAAUUCUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAGAGUAA UUAGUAAGCGGAAGCAGUUAACUUAUCUGGCCUCCACAACAGACACUGCAGAGGGGCCAUAAACAGCAU AUCUAG	AnaX-MEHA5- CVB3-IL2 precursor
5 4 2	GGGAGACCCUCGAUAGACUUAUGAGCAGAAAGAGGUUAGAAUUGUAGCCUGAUUUCUUAAGACA AUAGAUAGACUUAACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCUUAACGUCAGAGCAG GGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCGAGUAGCGAAAGCUGCAAGAGAAUGAAAACCC CCCCAUUAAACAGCCUGUGGGUUGAUCCACCCACAGGCCCAUUGGGCGCUGAGCUCUGGUUAUCA CGGUACCUUUGUGCGCCUGUUUAUACCCCUCCCCCAACUGUAACUUAAGAAUACACACACCGAU CAGACAGUACGGGUGGCACACACGCCACGUUUUGAUCAAGCACUUCUGUUAACCCGGACUGAGUAUCA AUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCUUAUCCGGCCAAUACUUCGAAAAACCUAGU AACACCGUGGAGUUGCAGAGUGUUCGUCUAGCACAUCACCCAGUGUAGAUACGGUCGAGUAGUCAC CGCAUUCGCCACGGGCGACCGUGGCGUGGCGUUGGCGGCCUGCCUAGGGGAAACCCAUUGGGA CGCUCUAAUACAGACAUUGGUGCGAAGAGUCUUAUAGGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGC GGCUAAUCCUAAUCUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUG CAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUAUCCUUAUACUGGCUGCUUAUGGUGACA AUUGAGAGAUCCUUAACUUAUAGCUUAUUGGAUUGGCAUCCGGUGACUAAUAGAGCUAAUUAUUAUC CCUUGUGUGGUUUAUACCAUUAAGCUUGAAGAGGUUAAACAUUAACAAUUCUUAAGUUAAGUUA UACAGCAAGGCCACCAUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCCUGAGCCUGGCCUGGUGCA CAAACAGCGCUCCUACAAGCUCUAGCACCAAGAAACCCAGCUGCAGCUGCAGCACCUGCUGCGGA CCUGCAGAUAGUCCUGAACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAAGUAGCAUUAU AAGUUCUACAUGCCUAAAGAGGCCACAGAGCUGAAGCACCUGCAUUGUCUGGAAGAGGAACUGAAGC CCUGGAAGAGGUGCUUAAUUCUGGCUCAGAGCAAGAACUCCACUGAGACCUCCGGGACCUAGUACG CGACAUCAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAAACUUAUGUGCGAGUACGCCGAG GAGACCGCCACCAUCGUGGAGUUCUGAAUCGGUGGAUCACCUUCUGCCAGUCCAUCAUCUCCACCC UGACCUGAUAAAGACGGGGGGGCUUAAUUAUUGAGCCUUAAGAAAGAAUUCUUAAGUGGAUGC UCUCAACUCAGGGAAACCUAAUUCUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAGAGUAA UUAGUAAGCGGAAGCAGUUAACUUAUCUGGCCUCCACAACAGACACUGCAGAGGGGCCAUAAACAGCAU AUCUAG	AnaX-MEHA6- CVB3-IL2 precursor

	AACACCGUGGAGUUGCAGAGUGUUUCGUCAGCACUACCCAGUGUAGAUACGGUCGAGUAGUCAC CGCAUCCCCACGGGCGACCGUGGGCGUGGCGUGUUGGGCGCCUGCCCAUGGGGAAACCAUGGGA CGCUCUAAUACAGACAUUGUGCGAAGAGUCUUAUAGAGCUAGUUGGUAGUCCUCCGGCCCGUGAAUGC GGCUAAUCCUAAUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUG CAGCGGAACCGACUACUUUGGGUGUCCGUGUUUCAUUUUUUAUCCUUAUACUGGCGUCUUUGGUGACA AUUGAGAGAUUGUUACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUAUUC CCUUUGUUGGGUUUAUACCAUUAAGCUUUAAGAGGUUAAAACAUUACAUAUUAUUAAGUUGAA UACAGCAAAGCCACCAUGUAUAGAAUGCAGCUGCUGUCUUGCAUCCGCGUGAGCCUGGCCUGGCUA CAAACAGCGCUCUACAAGCUCUAGCACCAAGAAAACCCAGCUGCAGCUCGAGCACCUGCUGGGA CCUGCAGAUAGUCCUGAACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAGAAUGCUGACAUUU AAGUUCUACAUGCCUUAAGAGGCCACAGAGCUGAAGCACCUGCAAUGUCUGGAAGAGGAACUGAAGC CCUGGGAAGAGGUGCUAAUUCUGGCUACAGCAAGAACUCCACCUGAGACCUCGGGACCUGAUCAG CGACAUCAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAACCUUCAUGUGCGAGUACGCCGAU GAGACCGCCACCAUCGUGGAGUCCUGAAUCCGUGGUAUCCUUCGCCAGUCCAUCAUCCACCC UGACCUGAAUAGACGGGGGGGCUUAAUUAUAGCCUUAAGAAAGAAUUCUUUAAGUGGAUCC UCUCAAACUCAGGGAAACCUAAUUCUAGUUAUAGACAAGGCAAUCCUGAGCCAGCCGAAGUAGUAA UUAUUAACUAAAGAGAAUACGGCUACAUAUUUACCCUUCUUGCUCUAAAGUCUAAUAAACAGCAU AUCUAG	
5 4 3	AAAAUCCGUUGAUUAAAACAGCCUGUGGGUUGAUCCCAACCCACAGGCCCAUUGGGCGUAGCACUCU GGUAUCACGGUACCUUUGUGCGCUGUUUUUAUACCCUCCCAACUGUAAUUAAGAAUAAACACA CACCAGUACAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGUUAACCCGGACUG AGUAUCAUAGAGCUGCUCACGGGUUAGAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAA ACCUAGUAACACCGUGGAAGUUGCAGAGUGUUCGUCAGCACUACCCAGUGUAGAUACGGUCGAU GAGUACCCGCAUUCGCCACGGGCGACCGUGGCGUGGCGUGGCGUGGCGGCGUGCCAUUGGGGAAACC CAUGGGAGCGCUAAUUAACUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGCUAAGCGGC UGAAUCCGCUAAUCCUAAUACUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGCUAAGCGGC AACUCUGCAGCGGAACCGACUACUUUGGGUGUCCGUGUUCAUUAUUAUCCUUAUACUGGCGUCUUAU GGUGACAAUUGAGAGAUUGUACCAUUAAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUU AUUAUCCCUUUGUUGGGUUUAUACCAUUAAGCUUAGCUUAGAAAGAGGUUAAAACAUUACAUAUUAU AGUUGAAUACAGCAAAUUGUAUAGAAUGCAGCUGCUGUCUUGCAUCCGCGUGAGCCUGGCCUGGUC ACAAACAGCGCUCCUACAAGCUCUAGCACCAGAAACCCAGCUGCAGCUCGAGCACCUGCUGCUGG ACCUGCAGAUAGUCCUGAACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAGAAUGCUGACAUU UAAGUUCUACAUGCCUAAAGAGGCCACAGAGCUGAAGCACCUGCAAUGUCUGGAAGAGGAACUGAAG CCCCUGGAAGAGGUGCUAAUUCUGGCUACAGCAAGAACUCCACCUGAGACCUCGGGACCUGAUC GCGACAUCAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAACCUUCAUGUGCGAGUACGCCGA UGAGACCGCCACCAUCGUGGAGUCCUGAAUCCGUGGAUCCUUGGCCAUCCUUAUCCUUAUCCACC CUGACCUGAUAAUAGAGACGCUACGGACUU	AnaX-CVB3- IL2 circRNA
5 4 4	AGAAGUAAACACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGU UACCCCGGACUGAGUAUCAUAGACUGCUCACGGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA CUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUGUUCGUCACGACUACCCAGUGUA GAUCAGUUGCAGUAGUACCGCAUUCGCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGGCGCCUGCC CAUGGGGAAACCCAUUGGACGCUUAAUACAGACAUGGUGCGAAGAGUCUUAUUGAGCUAGUUGGUAG UCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGU GUCCUAAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUUUAUCCUUAU UGGCGUCUUAUUGGAGACAUAUAGAGAUUGCUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACU AAUAGAGCUAUUAUUAUCCUUGUUGGGUUUAUACCAUUAAGCUUAGAAAGAGGUUAAAACAUUAC AAUUCAUUGUUAAGUUGAAUACAGCAAAGCCACCAUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCC CUGAGCCUGGCCUGGUCACAAACAGCGCUCCUACAAGCUCUAGCACCAGAAACCCAGCUGCAGC UCGAGCACCUGCUGGACCUGCAGAUAGUCCUGAACCGCAUUAACAACUACAAGAACCCAAAGCU GACCAGAAUGCUGACAUAUUAAGUUCUACAUGCCUAGAAAGGCCACAGAGCUGAAGCACCUGCAUUGU CUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUAAUUGGCUACAGCAAGAACUUCACCUGA GACCUCGGGACCUGAUCAGCGACAUAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAACCUU CAUGUGCGAGUACGCCGAUGAGACCGCCACCAUCGUGGAGUCCUGAAUCCGUGGGAUCCUUCUGC CAGUCCAUCAUUCACCUGACCUGAUAUUAACACAGCCUGUGGGUUGAUCCACCCACAGGCC AUUGGGCGCUGACACUCUGGUUAUCAGGUACCUUUGUGGCCUGUUUAUACCCUCCCCCAACUG UAACUU	AnaX-CLEAN- PIE-IRES- rawCVB3- CVB3-IL2 circRNA
5 4 5	AAAAUUAACAACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGU UACCCCGGACUGAGUAUCAUAGACUGCUCACGGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA CUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUGUUCGUCACGACUACCCAGUGUA GAUCAGUUGCAGUAGUACCGCAUUCGCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGGCGCCUGCC CAUGGGGAAACCCAUUGGACGCUUAAUACAGACAUGGUGCGAAGAGUCUUAUUGAGCUAGUUGGUAG UCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGU GUCCUAAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUUUAUCCUUAU UGGCGUCUUAUUGGAGACAUAUAGAGAUUGCUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACU AAUAGAGCUAUUAUUAUCCUUGUUGGGUUUAUACCAUUAAGCUUAGAAAGAGGUUAAAACAUUAC AAUUCAUUGUUAAGUUGAAUACAGCAAAGCCACCAUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCC CUGAGCCUGGCCUGGUCACAAACAGCGCUCCUACAAGCUCUAGCACCAGAAACCCAGCUGCAGC UCGAGCACCUGCUGGACCUGCAGAUAGUCCUGAACCGCAUUAACAACUACAAGAACCCAAAGCU CUGGCGUCUUAUUGGAGACAUAUAGUCCUAGAAAGGCCACAGAGCUGAAGCACCUGCAUUGU CUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUAAUUGGCUACAGCAAGAACUUCACCUGA GACCUCGGGACCUGAUCAGCGACAUAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAACCUU CAUGUGCGAGUACGCCGAUGAGACCGCCACCAUCGUGGAGUCCUGAAUCCGUGGGAUCCUUCUGC CAGUCCAUCAUUCACCUGACCUGAUAUUAACACAGCCUGUGGGUUGAUCCACCCACAGGCC AUUGGGCGCUGACACUCUGGUUAUCAGGUACCUUUGUGGCCUGUUUAUACCCUCCCCCAACUG UAACUU	AnaX-CLEAN- PIE-IRES- rawMutation2 -CVB3-IL2 circRNA

	CUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUGCUUAAUCUGGCUAGAGCAAGAACUCCACCUGA GACCUCGGGACCUGAUCAGCGACAUAACGUGAUCGUGGUGAACUGAAAGGCAGCGAGACAACCUU CAUGUGCGAGUACGCCGAUGAGACCGCCACCAUCGUGGAGUUCUGAAUCGGUGGUAUCACCUUCUGC CAGUCCAUAUCUCCACCCUGACCUGAUAAUAAACAGCCUGUGGGUUGAUCCACCCACAGGCC AUUGGGCGCUAGCACUCUGGUAUCACGGUACCUUUGUGCGCCUGUUUUAUACCCCUCCCCAACUG UAACUU	
5 4 6	AAACCUCGCUAGAGCAAGAACUCCACCUGAGACCUCGGGACCUGAUCAGCGACAUAACGUGAUC GUGCUGGAACUGAAAGGCAGCGAGACAACCUUACUGUGCGAGUACGCCGAUGAGACCGCCACCAUCG UGGAGUUCUGAAUCGGUGGAUACCUUCUGCCAGUCCAUAUCUCCACCCUGACCUGAUAAUAAA ACAGCCUGUGGUUGAUCCACCCACAGGCCAAUUGGGCGCUAGCACUCUGGUAUCACGGUACCUUU GUGCGCCUGUUUUAUACCCCUCCCCAACUGUAACUUAAGAAGUAACACACACCCGAUACAACAGUCAG CGUGGCACACCAGCCACGUUUUGAUCAGCACUUCUGUUACCCCGGACUGAGUAUCAUAGACUGCU CAGCGCGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUGGAAAAACCUAGUAACACCGUGG AAGUUGCAGAGUGUUCGCUAGCACUACCCAGUGUAGAUACGGUCGAGUGACUACCCGAUUCCCC ACGGGCCACCUGGGCGGUGGCGUGGCGGCCUGCCCAUGGGGAAACCAUGGGACGCUUAUA CAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAUCCUCCGGCCCCUGAAUGCGGCUAAUCCU AACUGCGGAGACACACCCUCAAGCCAGAGGGCAGUGUGGUAACGGGCAACUCUGCAGCGGAACC GACUACUUGGGUGUCCGUGUUCAUUUUUAUCCUUAUACUGGCGUCUUAUGGUGACAUAUAGAGAU CGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUUAUAGAGCUAUUAUAUACCCUUUGUUGG GUUUUAUACCAUUAAGCUUGAAAGAGGUUAAACAUUACAUAUUGUUAAGUUGAAUACAGCAAG CCACCAUGUAUAGAUGCAGCUGCUGUCUUGCAUCGCCUGAGCCUGGGCCUGGUCACAAACAGCGC UCCUACAAGCUCUAGCACCAAGAAACCCAGCUGCAGCUCGAGCACCUGCUGGACCCUGCAGAU AUCCUGAACGGCAUUAACAACUACAAGAACCACAAAGCUGACCAGAAUGCUGACAUUAAGUUCUACA UGCCUAGAAGGCCACAGAGCUGAAGCACCUGCAUUGUCUGGAAGAGGAACUGAAGCCCCUGGAAGA GGUCUU	AnaX-CLEAN- PIE-Cargo- PIE1-CVB3- IL2 circRNA
5 4 7	AAUCUCUGACAUAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAACCUUACUGUGCGAGUAC GCCGAUGAGACCGCCACCAUCGUGGAGUUCUGAAUCGGUGGUAUCACCUUCUGCCAGUCCAUAUCU CCACCUGACCUGAUAAUAAAACAGCCUGUGGGUUGAUCCACCCACAGGCCAAUUGGGCGCUAGC ACUCUGGUAUCACGGUACCUUUGUGCGCCUGUUUAUACCCCUCCCCAACUGUAACUUAAGAUA ACACACACCGUAACAGUCAGCGUGGCACACCGCCACGUUUUGAUCAGCACUUCUGUUAACCCG GACUGAGUAUCAUAGACUGCUCACGCGGUUGAAGGAGAAAGCUUCGUUAUCCGGCCAAUACUUC GAAAAACCUAGUAACACCGUGGAGUUGCAGAGUUGUUCGUCAGCACUACCCAGUGUAGAUACAGG UCGAGAGUACACCGCAUUCCCCACGGGCGACCGUGGCGGUGGCGUGCGUUGGGCGCCUGCCAUUGGG AAACCCAUUGGACGCUCAUUAACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAUCCCG GCCCCUGAAUGCGGCUAAUCCUAAUCUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGCGUAA CGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUAUCCUUAUACUGGCU CUUAUGGUGACAUAUAGAGAUCCGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAG CUUAUUAUUAUCCCUUGUUGGGUUUAUACCACUUAAGCUUGAAAGAGGUUAAACAUUACAUAUACU UGUUAAGUUGAAUACAGCAAGGCCACCAUGUAUAGAAGCAGCUGCUGUCUUGCAUCGCCUGAGCC UGGCCUGGUCACAAACAGCGCUCCUACAAGCUCUAGCACCAAGAAACCCAGCUGCAGCUCGAGCA CCUGCUGCUGACCUGCAGAUUCCUGAACGGCAUUAACAACUACAAGAACCACAAAGCUGACCAGA AUGCUGACAUUAAGUUCUACAUGCCUAGAAGGCCACAGAGCUGAAGCACCUGCAAUGUCUGGAAG AGGAACUGAAGCCCGUGGAAGAGGUGCUUAUUCUGGCUAGAGCAAGAACUCCACCUGAGACCUAG AGACUU	AnaX-CLEAN- PIE-Cargo- PIE2-CVB3- IL2 circRNA
5 4 8	AAAACCCCCCAUUAACAGCCUGUGGGUUGAUCCACCCACAGGCCCAUUGGGCGCUAGCACUCU GGUAUCACGGUACCUUUGGCGCCUGUUUAUACCCCUCCCCAACUGUAACUUAAGAAGUAACACA CACCGAUACAAGUCAGCGUGGCACACCGCCACGUUUUGAUAAGCACUUCUGUUAACCCGGACUG AGUAUCAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAA ACCUAGUAACACCGUGGAAGUUGCAGAGUUGUUCGUCAGCACUACCCAGUGUAGAUACAGGUCGAU GAGUACCCGAUUCCCCACGGGCGACCCUGGCGGUGGCGUGCGUUGCCGGCCUGCCCAUGGGGAAACC CAUGGGACGCUCAUUAACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAUCCUCCGGCCCC UGAAUGCGGCUAAUCCUAAUCUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGC AACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUAUCCUUAUACUGGCGUCUUA GGUGACAUAUAGAGAUCCGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAU AUUAUCCCUUGUUGGGUUUAUACCAUAGCUUAGAAAGAGGUUAAACAUAUACAUAUUGUUA AGUUGAAUACAGCAAGCCACCAUGGCCAGCGAGUUCAGAAGAAGCUCUUCUGGAGGGCAGUGGUG GCCGAGUUCUGGCCACGACCCUCUUGUCUUAUCAGCAUCGGUUCUGCCUGGGCUUCAAUACCC CGGUGGGGAACAACAGAGCGCGUCCAGGACAACGUGAAGGUGUCGUGGCCUUCGGGCUAGCAU CGCCACGCGUGGCGCAGAGUGGGGCCAUACAGCGGGCCCAACCUCAACCCGGCGUCACACUGGGG CUGCUGCUCAGCUGCCAGAUACGAUCUUCGUGGCCUUAUGUACAUAUCGCCAGUGCGUGGGGG CCAUCGUCGCCACCGCCAUCCUCUAGGCAUACCUCCUCCUGACUGGGAACUCGCUUGGCCGCAA UGACCUGGCUAGUGGUGAUCUGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGGACCCUCCAGCUG GUGCUAUGCGUGGUGGCUACUACCGACCGGAGGCGCCGUGACCUUGGUGGCUACGCCCCCUUGCCA UCGGCCUCUCUGUAGCCCUUGGACACCUCCUGGCUAUUGACUACACUGGCGUGGGGAUUAACCCUGC UCGGUCCUUUGGCUCCGCGUGAUCACACAAUUCAGCAACCACUGGAUUUUCUGGGUGGGGCCA UUCAUCCGGGGAGCCUGGCUUACUACUACGACUUAUCCUGGCCCCACGCAGCAGUGACCUC CAGACCCGUGAAGGUGGAGCAGCGGCCAGGUGGAGGAGUAUGACUGGACCGGACGACAUCAA CUCCAGGGUGGAGAUAGAGCCAAUAGUAUAGACGGGGGGGCUU	AnaX-GCv3- 2×homo-CVB3- AQPl circRNA

5 4 9	AACUGGAAACUCGCUUGGCCGCAUAGACCUGGCUAUGGUGUGAACUCGGGCCAGGGCCUGGGCAUC GAGAUCAUCGGGACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCCGAGGCGCCGUGACC UUGGUGGCUACGCCCCCUUGCCAUCCGGCCUCUCUGUAGCCCUUGGACACCUCCUGGCUAUUGACUA CACUGGCUUGGGGAUUAACCCUGCUCGGUCCUUGGCUCCGCGGUGAUCACACACAACUUCAGCAAC CACUGGAUUUUCUGGGUGGGCCAUUUAUCGGGGAGCCUUGGCUUACUACUACGACUUCUACUCC UGGCCCCACGCGAGCUGUACUACAGACCCGCGUGAAGGUGUGACCAGCGGCCAGGUGGAGGAGUA UGACCUGGAUGCCGACGACAUCAACUCCAGGGUGGAGUAAGGCCAAAUGAUAAUUAACAGCCU GUGGGUUGAUCCCAACGACAGGCCAUUGGGCGCUAGCACUCUGGUAUACGUAACCUUUGGCGCC UGUUUAUACCCCCUCCCCAACUGUAACUUAAGUAACACACACCGAUCAACAGUCAGCGUGGCA CACCAGCCACGUUUGAUCAAGCACUUCUGUACCCCGACUGAGUAUCAAUAGACUGUCACGCGG UUGAAGGAGAAAGCGUUCGUUAUCCGGCCACUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGC AAGUGUUCGCUACGACUACCCAGUGUAGAUACGGUCGAGUAGUACCCGAUUCGCCACGGCGG ACCGUGGCGGUGGCUUGGUGGCGGCCUUGCCAUUGGGGAAACCAUGGGACGCUUAAUACAGACAU GGUGCGAAGAGUCUAUUGAGCUAGUUGGUGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCUGC GAGCACACACCCUACAGCCAGAGGGCAGUGUCGUAACGGGCAACUUGCAGCGGAACCGACUACU UUGGUGUCCGUGUUAUUAUCCUUAUACUGGCGCUUUGGUGACAAUUGAGAGUACGUUACC AUUAGUUAUUGGAUUGGCCAUCCGUGGACUAAUAGAGCUAAUUAUUAUCCUUGGUGUUAUA CCACUUAAGCUUAGAGGUUAAACAUUAACAAUUAUUAAGUUAAGUUAAGUUAAGCAAGCCACCAU GGCCAGCGAGUUAAGAAGAAGCUCUUCUGGAGGGCAGUGGUGGCCAGUUCUGGCCACGACCCUC UUUGUCUUAUCAGCAUCGUGUCCUGGCCUUCAAUACCCGUGGGGAAACACACGACGCGCGG UCCAGGACACGUGAAGGUGUCGUGGCCUUCGGGUGAGCAUCGCCACGUGGCGCAGAGUGUGGG CCACAUACGCGGGCGCCACCUCACACCCGGCUGUCACACUGGGGCGUGGCUACGUGCCAGAUACG AUCUUCGUGCCCUCAUGUACAUAUCGCCCGAGUGCGUGGGGCCAUCGUGGCCACCGCCAUCCUCU CAGGCAUCACUCCAGCUU	AnaX-GCv3- 2×homo- CLEAN-PIE- Cargo-PIE1- CVB3-AQP1 circRNA
5 5 0	AACAGCUCUCCUGACUGGGAAACUCGCUUGGCCGCAUAGACCUGGCUAUGGUGUGAACUCGGGCCAG GGCCUGGGCAUCGAGAUCAUCGGGACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCCGA GGCGCCGUGACCUUGGUGGCUACGCCCCCUUGCCAUCCGGCCUCUCUGUAGCCCUUGGACACCUCCU GGCUAUUGACUACACUGGCUUGGGAUUAACCCUGCUCGGUCCUUGGCUCCGCGGUGAUCACACAC AACUUCACAGCAACCAUGGAUUUUCUGGGUGGGCCAUUUAUCGGGGAGCCUUGGCUUACUACU ACGACUUAUCCUGGCCCCACGCGAGCUGACCUCACAGACCCGCGUGAAGGUGUGGACGACGCGCCA GUGGGAGGAGUAGACCUGGAUGCCGACGACAUCAACUCCAGGGUGGAGUAAGGCCAAUUGAUAA UUAACACAGCCUGUGGGUUGAUCCCAACCCAGGCCAUUGGGCGCUAGCACUCUGGUAUACGAGUA CCUUGUGGCGCCUGUUUAUACCCCCUCCCCAACUGUAACUUAAGAUAACACACACCGAUCAACA GUCAGCGUGGACACCCAGCCACGUUUGAUCAAGCACUUCUGUACCCCGGACUGAGUAUCAAUAGA CUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAAACCUAGUAACAC CGUGGAAGUUCGAGAGUUCUUCGCUACGACUACCCAGUGUAGAUACGGUCGAGAGUACCCGCAU UCCCCACGGGCGACCGUGGCGGUGGCUUGGCGGCCUGGCCAUGGGGAAACCCAUGGGACGCU UAAUACAGACAUUGGUGCGAAGAGUUAUUGAGCUAGUUGGUAUCCCGGCCCGUGAAUUGCGGCUA AUCCGAACUGCGGAGCAGACACCCUACAGCCAGAGGGCAGUGUGUACGUAACGGGCAACUCUGCAGC GAACCGACUACUUGGGUGUCCGUGUUUAUUAUCCUUAUACUGGCUUUAUGGUGACAAUUGA GAGAUUGUAACAUUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUUAUCCU GUUGGGUUUAUACCAUUAAGCUUAAGAGGUUAAACAUUAACAAUUAUUAAGUUAAGUUAUACAG CAAAGCCACCAUGGCCAGCGAGUUAAGAAGAAGCUCUUCUGGAGGGCAGUGGUGGCCAGUUCUG GCCACGACCCUUCUUGUCUUAUCAGCAUCGUGUCCUGGCCUUCAAUACCCGUGGGGAAACA ACCAGACGGCGGUGCCAGGACAACCGUAAGGUGUGCGUGGCCUUCGGGUGAGCAUCGCCACGCGUGGC GCAAGUGUGGGCCCAUACAGCGCGGCCACCUCACCCGGCUGUCACACUGGGGCGUGGCUACG UGCCAGAUACGACUUCUCCGUGCCCUCAUGUACAUAUCGCCCGAGUGCGUGGGGCCAUCGUGGCCA CCGCCAUCCUCUUGGCAU	Twort-PIE-6- CLEAN-PIE- Cargo-PIE1- CVB3-AQP1 circRNA
5 5 1	AGAAGUAAACACACCCGAUCAACAGUCAGCGUGGCACACCGCCAGCUUUUGAUCAAGCAGUUCUGU UACCCCGGACUGAGUAUCAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA CUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUUGUCCUCAGCACUACCCAGUGUA GAUCAGGUCGAGUACACCGCAUUCUCCACGGGCGACCGUGGCGGUGGCGUGGUGGCGGCCUGCC CAUGGGGAAACCCAUGGGACGCUUAUUAACAGACAUUGGUGGGAAGAGUUAUUGAGCUAGUUGGUAG UCCUCCGGCCCCUGAAUGCGGCUAAUCCUUAACUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGU GUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUAUUAUCCUUAUA CUGGCGCUUAUUGGUGACAAUUGAGAGUACGUUAACCAUUAUGCUUAUUGGAUUGGCCAUCCGGUACU AAUAGAGCUUAUUAUUAUCCUUGUUGGUGUUAUACCAUUAAGCUUAGAAAGAGGUUAAACAUUAC AAUUAUUGUUAAGUUGAAUACAGCAAAGCCACCAUGGCCAGCGAGUUAAGAAGAGCUCUUCUGG AGGGCAGUGGUGGCCAGUUCUGGCCACGACCCUUCUUGUCUUAUCAGCAUCGGUUCUGCCUGG GCUUCAAUAACCCGUGGGGAACAACAGACGGCGGUCCAGGACAACGUAAGGGUGUGCGUGGCCU CGGGCUGAGCAUCGCCACGUGGCGCAGAGUGUGGCCACAUCAGCGGCGCCACCUCACACCCGGCU GUCACAUUGGGGCGUGCUGUCUACGUGCCAGAUCAUCUUGGUGCCUUAUGUACAUCUACGCCCC AGUGCGUGGGGCCAUCGUGCGCCACCGCAUCCUCUAGGCAUACCUCCUCCUGACUGGGAACUC GCUUGGCCGCAUAGACCUUGGCUAUGGUGUGAACUCGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGG ACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCGGAGGCGCGGUGACCUUGGUGGCGUAG CCCCCUUGCCAUUGGCCUUCUGUAGCCCUUGGACACCUCCUGGCUUAUUGACUACACUGGCGUGG GUAUAAACCCUGCUGGCUUUGGCGCGGUGAUCACACACAAUUAAGCAACCAUGGAUUAUUAU UGGGUGGGGCCAUUAUCGGGGGAGCCUGGCUUACUUAUCGACUUAUCCUGGCCCCACGCA GCAGUGACCUACAGACCGCGUGAAGGUGUGGACGCGGCCAGGUGGAGGAGUAUAGACUGGAUGC	AnaX-GCv3- 2×homo- CLEAN-PIE- IRES- rawCVB3-AQP1 circRNA

	CGACGACAUCACUCCAGGUGGAGAUAGAGCCAAUUGAUAAUUAACAGCCUGUGGGUUGAUCC CACCCACAGGCCCCAUUGGGCGCUAGCACUCUGGUUAUCACGGUACCUUUGUGCGCCUGUUUAUACCC CCUCCCCCAACUGUAACUU	
5 5 2	AAAAUUAACAACACCCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGU UACCCCGGACUGAGUAUCAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA CUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUUGUUCGUCAGCACUACCCAGUGUA GAUCAGUUGCAGUAGUCACCGCAUUCACCGGCGACCGUGGCGGUGGCGUGGCGUGGCGGCCUGCC CAUGGGGAAACCCAUUGGACGCUCAAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAG UCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACUGCGGAGCACACCCCUCAAGCCAGAGGGCAGUGU GUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUCAUUUAUCCUUA CUGGCGUCUUAUGGUGACAUAUAGAGAUUGUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACU AAUAGAGCUAUUAUUAUCCUUGUUGGGUUUAUACCAUAGCUUGAAAGAGGUUAAACAUAUAC AAUUAUUGUUAAGUUGAAUACAGCAAGGCCACCAUGGCCAGCGAGUUAAGAAGAGCUUUCUGG AGGGCAGUGGUGGCCGAGUUCUGGCCACGACCCUUCUUGUCUUAUCAGCAUCGGUUCUGCCUGG GCUUCAAUAUACCGGUGGGGAACAACAGACGCGGUGCCAGGACAACGUGAAGGUGUCGUGGCCU CGGGCUGAGCAUCGCCACGCGUGGCGCAGAGUGUGGGCCACAUCAGCGCGGCCACCUCACCCGGCU GUCACACUGGGGCGUGCUGCUCAGCUGCCAGAUACAGCAUUCUGGCGCCUACUGUACAUACUGGCC AGUGCGUGGGGGCCAUUGCUGGCCACCGCCAUCCUCUCAGGCAUACCCUCCUCCUGACUGGGAACUC GCUUGGCCCGCAUGACCCUGGCUAGUGGUGAUCUCGGGCCAGGGCCUGGGCAUCGAGAUCAUCGGG ACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCGGAGGCGCGGUGACCUUGGUGGCGCAG CCCCCUUGCCAUUGGCGCUCUGUAGCCUUGGACACCCUUGGCUAUUGACUACACUGGCGUGG GCUUAACCCUGCUCGUGCCUUGGCGUGGUGAUCACACACACUUCAGCAACCAUGGAUUAUUC UGGGUGGGGCAUUAUCUGGGGAGCCUGGCGUACUUAUCUACGACUUAUCCUGGCCCCACGCA GCAGUGACCUACAGACCGCGUGAAGGUGUGGACGCGGCCAGGUGGAGGAGUAUAGCUGGAUGC CGACGACAUCAACUCCAGGUGGAGUAAGGCCAAUUGAUAUUAACAGCCUGUGGGUUGAUCC CACCCACAGGCCCAUUGGGCGCUAGCACUCUGGUUAUCACGGUACCUUUGUGCGCCUGUUUAUACCC CCUCCCCCAACUGUAACUU	AnaX-GCv3- 2×homo- CLEAN-PIE- IRES- mutationCVB3 -AQP1 circRNA
5 5 3	AAAACCCCCCAUAAAAAGCCUGUGGGUUGAUCCACCCACAGGCCCAUUGGGCGCUAGCACUCU GGUAUCACGGUACCUUUGUGCGCCUGUUUAUACCCCUCCCCAACUGUAACUUAAGAAGUAACACA CACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGUUAACCCGGACUG AGUAUCAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAA ACCUAGUAACACCGUGGAAGUUGCAGAGUUGUUCGUCAGCACUACCCAGUGUAGAUACAGGUCGUA GAGUACACCGAUUCCCAACGGGCGACCGUGGCGGUGGCGUUGGCGGCCUGCCCAUGGGGAAAC CAUGGGACGCUUAUACAGACAUUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUGUCCCGGCCCC UGAAUGCGGCUAAUCCUAACUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGUGCUAACGGGC AACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUAUUAUCCUUAUACUGGCGUCUUA GGUGACAUAUAGAGAUUGUUAUUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAUU AUUAUACCCUUGUUGGGUUUAUACCAUAGCUUAGAAAGAGGUUAAACAUAUACAAUUAUUGUUA AGUUGAAUACAGCAAAGCCACCAUGUAUAGAAUGCAGCUGCUGUUGCAUCGCCUGAGCCUGGCC CUGGUCACAAACAGCGCUCUUAACAGCUCUAGCACCAAGAAAACCCAGCUGCAGCUCGAGCACCUGC UGCUGGACCUAGCAGAUUCCUGAACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAAGAUUCU GACAUUAUAGUUCUACAUUGCCUUAAGAAGGCCACAGAGCUGAAGCACCUGCAAUGUCUGGAAGAGGAA CUGAAGCCCCUGGAAGAGGUGCUUAUUCUGGCUAGAGCAAGAACUCCACCUGAGACCUCCGGGACC UGAUCAGCGACAUCAACGUGAUCGUGCUGGAACUGAAAGGCAGCGAGACAACCUUACUGUGCGAGUA CGCCGAUGAGACCGCCACCAUCUGGAGUUCUGGAUUCGUGGUAUACCUUCUGCCAGUCCAUCAUC UCCACCUGACCCUGAUAAAGACGGGGGGGCUU	AnaX-MEH1- CVB3-IL2 circRNA
5 5 4	AGAAGUAACACACACCCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGU UACCCCGGACUGAGUAUCAUAGACUGCUCACGCGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA CUACUUCGAAAAACCUAGUAACACCGUGGAAGUUGCAGAGUUGUUCGUCAGCACUACCCAGUGUA GAUCAGGUGCAGUAGUCACCGCAUUCACCGGCGACCGUGGCGGUGGCGUGGCGUGGCGGCCUGCC CAUGGGGAAACCCAUUGGACGCUCAAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUGAG UCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGU GUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUCAUUUAUCCUUA CUGGCGUCUUAUGGUGACAUAUAGAGAUUGUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACU AAUAGAGCUAUUAUUAUCCUUGUUGGGUUUAUACCAUAGCUUAGAAAGAGGUUAAACAUAUAC AAUUAUUGUUAAGUUGAAUACAGCAAA	AnaX-CLEAN- PIE-IRES- rawCVB3- 3', AnaX- GCv3-2×homo- CLEAN-PIE- IRES- rawCVB3-3'
5 5 5	UUAAAAAGCCUGUGGGUUGAUCCACCCACAGGCCCAUUGGGCGCUAGCACUCUGGUUAUCAGGUA CCUUGUGCGCCUGUUUAUACCCCUCCCCAACUGUAACUU	AnaX-CLEAN- PIE-IRES- rawCVB3- 5', AnaX- CLEAN-PIE- IRES- rawMutation2 -CVB3- 5', AnaX- GCv3-2×homo- CLEAN-PIE- IRES-

		rawCVB3-5', AnaX-GCv3-2×homo-CLEAN-PIE-IRES-mutationCVB3-5'
5 5 6	AaAAUUaCaACACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUGAUCAGCACUUCUGU UACCCCGGACUGAGUAUCAAUAGACUGCUCACGCGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA CUACUUCGAAAACCUAGUAACACCGUGGAAGUUGCAGAGUGUUCGCUCAGCACUACCCAGUGUA GAUCAGGUCGAGUAGUACCCGAUUCACCGGGGACCGUGGCGGUGGCGUGCGUUGGCGGCCUGCC CAUGGGGAAACCCAUUGGACGCUCUAAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAG UCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUGCGGAGCACACCCUCAAGCCAGAGGGCAGUGU GUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUCAUUUUUAUCCUAUA CUGGCGUCUUAUGGUGACAAUUGAGAGAUUGUUAACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGACU AAUAGAGCUAUUAUAUACCCUUGUUGGGUUAUACCACUAGCUUGAAGAGGUUAAAACAUUAC AAUUCAUUGUUAAGUUGAAUACAGCAA	AnaX-CLEAN-PIE-IRES-rawMutation2-CVB3-3', AnaX-GCv3-2×homo-CLEAN-PIE-IRES-mutationCVB3-3'
5 5 7	aAAcCUcGCUCAGAGCAAGAACUCCACCUGAGACCUGGGACCUGAUCAGCGACAUCAACGUGAUC GUGCUGGAACUGAAAGGCAGCGAGACACCUUUAUGUGCGAGUACGCCGAGAGACCGCCACCAUCG UGGAGUUCUGAAUCGGUGGAUACCCUUCUGCCAGUCCAUAUCUCCACCCUGACC	AnaX-CLEAN-PIE-Cargo-PIE1-3'
5 5 8	AUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCCUGAGCCUGGCCUGGUCACAAACAGCGCUCCUA CAAGCUCUAGACCAAGAAACCCAGCUGCAGCUCGAGCACCUGCUGUGGACCUGCAGAUAGUCCU GAACGGCAUUAACAACUACAAGAACCCAAAGCUGACCAGAAUGCUGACAUUAAGUUCUACAUGCCU AAGAAGGCCACAGAGCUGAAGCACCUGCAUUGUCUGGAAGAGGAACUGAAGCCCCUGGAAGAGGUcU U	AnaX-CLEAN-PIE-Cargo-PIE1-5'
5 5 9	aAUCUcUGACAUCAACGUGAUCGUGCGGAACUGAAAGGCAGCGAGACAACCUUCAUGUGCGAGUAC GCCGAGUGAGACCGCCACCAUCGUGGAGUUCUGAAUCGGUGGAUACCUUCUGCCAGUCCAUAUCU CCACCCUGACC	AnaX-CLEAN-PIE-Cargo-PIE2-3'
5 6 0	AUGUAUAGAAUGCAGCUGCUGUCUUGCAUCGCCUGAGCCUGGCCUGGUCACAAACAGC	AnaX-CLEAN-PIE-Cargo-PIE2-5'
5 6 1	aACUGGaAACUCGCUUGGCCGCAUAGACCUGGCGUAUGGUGUGAACUCGGGCCAGGGCCUGGGCAUC GAGAUCAUCGGGACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCGGAGGCGCCGUGACC UUGGUGGCUACGCCCCUUGCCAUCCGCCUCUCUGUAGCCUUGGACACUCCUGGCUAUUGACUA CACUGGCUUGGGGAUUAACCCUGCUGGUCUUGGCUCCGCGGUGAUCACACAAACUACGCAAC CACUGGAUUUUCUGGUGGGGCAUUAUCGGGGAGCCUGGCGUACUACUACGACUUAUCC UGCCCCACGACGACUGACCUCACAGACCGCGUGAAGGUGUGGACCAGCGGCCAGGUGGAGGAGUA UGACCUGGAUGCCGACGACAUCAACUCCAGGUGGAGAUAGAAGCCCAAUGAUAA	AnaX-GCv3-2×homo-CLEAN-PIE-Cargo-PIE1-3'
5 6 2	GCCACCAUGGCCAGCGAGUUAAGAAGAAGCUCUUCUGGAGGGCAGUGGUGGCCAGUUCUGGCCA CGACCCUCUUGUCUUAUCAGCAUCGGUUCUGCCUGGGCUUCAAUACCCGGUGGGGAACAACCA GACGGCGGUCAGGACAACGUGAAGGUGUCGUGGCCUUCGGGCGAGCAUCGCCACGCGUGGCGCAG AGUGUGGGCCACAUACGCGGCGCCACCUCAACCCGGCUGUACACUGGGGCGUGCUGCUCAGCUGCC AGAUCAGCAUCUCCUGGCCUUAUGUACAUCAGCCCGAGUGCGUGGGGCCAUCGUGGCCACCGC CAUCCUUCAGGCAUACCUCCagCUU	AnaX-GCv3-2×homo-CLEAN-PIE-Cargo-PIE1-5'
5 6 3	aACCagUCCCCUGACUGGGAACUCGCUUGGCCGCAUAGACCUGGCGUAUGGUGUGAACUCGGGCCAG GGCCUGGGCAUCGAGAUCAUCGGACCCUCCAGCUGGUGCUAUGCGUGCUGGCUACUACCGACCGGA GGCGCCGUGACCUUGGUGGUCAGCCCCCUUGCCAUCGGCCUCUCUGUAGCCCUUGGACACCUCCU GGCUAUUGACUACACUGGCUUGGGAUUAACCCUGCUGGUCUUGGCUCCGCGGUGAUCACACAC AACUUCAGCAACCAUGGAUUUUCUGGGUGGGGCAUUAUCGGGGAGCCUGGCGUACUUAUCU ACGACUUAUCCUGCCCCACGACAGUGACCUCACAGACCGCGUGAAGGUGUGGACCAAGCGGCA GGUGGAGGAGUAUGACCUGGAUGCCGACGACAUCAACUCCAGGGUGGAGUAAGCCCAAUGAUAA	Twort-PIE-6-CLEAN-PIE-Cargo-PIE1-3'
5 6 4	GCCACCAUGGCCAGCGAGUUAAGAAGAAGCUCUUCUGGAGGGCAGUGGUGGCCAGUUCUGGCCA CGACCCUCUUGUCUUAUCAGCAUCGGUUCUGCCUGGGCUUCAAUACCCGGUGGGGAACAACCA GACGGCGGUCAGGACAACGUGAAGGUGUCGUGGCCUUCGGGCGAGCAUCGCCACGCGUGGCGCAG AGUGUGGGCCACAUACGCGGCGCCACCUCAACCCGGCUGUACACUGGGGCGUGCUGCUCAGCUGCC AGAUCAGCAUCUCCUGGCCUUAUGUACAUCAGCCCGAGUGCGUGGGGGCCAUUCGUGCCACCGC CAUCCUUCUGGCAU	Twort-PIE-6-CLEAN-PIE-Cargo-PIE1-5'
5 6 5	GGGAGACCCUCGACCGUGCAUUGUCCACUGGUCACAAUAGAUGACUUAACAUAUCCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUAAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCA AUAGGCAUGAGCGAAAGCUGCAAGAGAAUGAAAUCGUGUAGCCGGCAGCCAAUCUCGACAAAAUG UCCAUGAUGGGGCAUCGGGUCGCAUUCUCCCAUGGUCUACCUUUCGUUGAAUUCUAGUGACAA CUCCGUACAAUGCAGACUUAUGACGGGGAUGAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCG	Anabaena pre-tRNA-Leu gene GROUP I INTRON

	AGCAGAGAUCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCACCCCCAGAGCAAUCGGCCUGUC AUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGAGCUCUUCUGGAGC GGGUGGAACCGCACAUUGUGAUGGGGACAUUGUUAUCUUCAAAGACGCUACGGACUAAAUAUUG AGCCUUAAGAAGAAUUCUUUAAGUGGAGUCUCUCAAACUCAGGGAAACCUAAUUCUAGUUUAGA CAAGGCAUCCUGAGCCAAGCCGAAGUAGUAAUAGUAAGACCAGUGGACAAUCGACGGAUAACAGC AUAUCUAG	POLR2A precursor
5 6 6	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACCCAGCAAGGGAUGGUGUCAAUUCGGCGAAA CCUAAGCGCCCCCGGGCGUAUGGCAACGCGGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGU AGAGACUAGACGGCACCCACCUAAGGCAAAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAG UCACACAAACCGGaaUccgUUggUgcUgCCGGCAGCCAACUCUGCACAAAUGUCCAUGAUGGGGCA UCGGGUCCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUUCUAGUGUGACAACUCCGUACAUGCA GACUUUAGACGGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCAGG AGCUGGGCAUGGUUCCUCGAUGAUUGUCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCA GGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCAC AUUGUGAUGGGGACAUGUUAUUCUCAAagagcagcggacUcaUAUUUCGAUGUGCCUUGCGCCGGG AGACCAGUGGACAAUCGACGGAUAACAGCAUAUCUAG	Azoarcus olearius BH72 GROUP I INTRON-PIE-2 POLR2A precursor
5 6 7	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACCGCGAGCCAAGCUUCGGCGCCUGCGCCGAUGA AGGUGUAGAGACUAGACGGCACCCACCUAAGGCAAAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGG CGAAAGUCACACAAACCGGaaUccgUUggUgcUgCCGGCAGCCAACUCUGCACAAAUGUCCAUGAU GGGGCAUCGGGUGCCGAUUCUCCCAUGGUCUACCUUUCGCUUGAAUUCUAGUGUGACAACUCCGUAC AAUGCAGACUUGACGGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGA UCCAGGAGCUGGGCAUGGUUCCUCGCAUGAUUGUCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAU UGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAA CGGCACAUUGUGAUGGGGACAUGUUAUUCUCAAagagcagcggacUcaUAUUUCGAUGUGCCUUGC GCCGGGAAACACGCAAGGGAUGGUGUCAAUUCGGCGAAACCUAAGCGCCCGCCCGGGCGUAUGGC AGACCAGUGGACAAUCGACGGAUAACAGCAUAUCUAG	Azoarcus olearius BH72 GROUP I INTRON-PIE-4-5 POLR2A precursor
5 6 8	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCGCGGGCGUAUGGCAACGCCGAGCCAAGCUUCGGC GCCUGCGCCGAUGAAGGUGUAGAGACUAGACGGCACCCACCUAAGGCAAAACGCUAUGGUGAAGGCAU AGUCCAGGGAGUGGGCAAGUCACACAAACCGGaaUccgUUggUgcUgCCGGCAGCCAACUCUGCAC AAAUGUCCAUGAUGGGGCAUCGGGUCCGAUUCUCCCAUGGUCUACCUUUCGCUUGAAUUCUAGUG UGACAACUCCGUACAAUGCAGACUUGACGGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGA GACGCGAGCAGAGAUCCAGGAGCUGGGCAUGGUUCCUCGCAUGAUUGUCACCCCCAGAGCAAUCGG CCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUC UGGAGCGGGUGGAACGGGCAUUGUGUGAUGGGGACAUGUUAUUCUCAAagagcagcggacUcaUAUU UCGAUGUGCCUUGCGCCGGGAAACACGCAAGGGAUGGUGUCAAAUUCGGCGAAACCUAAGCGCCCG CGACCAGUGGACAAUCGACGGAUAACAGCAUAUCUAG	Azoarcus olearius BH72 GROUP I INTRON-PIE-5 POLR2A precursor
5 6 9	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCUGCGCGGAUGAAGGUGUAGAGACUAGACGGCACCC CACCUAAGGCAAAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACACAAACCGGaaUc cgUUggUgcUgCCGGCAGCCAACUCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGAUUCUCC CAUGGUCUACCUUUCGCUUGAAUUCUAGUGGACACUCCGUAACAAUGCAGACUUGACGGGGAUGA GUGAUAUCUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCU CGCAUGAUUGUCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAG UGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCAUUGUGUGAUGGGGACA UGUUAUUCUCAAagagcagcggacUcaUAUUUCGAUGUGCCUUGCGCCGGGAAACACGCAAGGGAUG GUGUCAAUUCGGCGAAACCUAAGCGCCCGCCCGGGCGUAUGGCAACGCCGAGCACAAGCUUCGGCGC CGACCAGUGGACAAUCGACGGAUAACAGCAUAUCUAG	Azoarcus olearius BH72 GROUP I INTRON-PIE-6 POLR2A precursor
5 7 0	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCG AAAGUCACACAAACCGGaaUccgUUggUgcUgCCGGCAGCCAACUCUGCACAAAUGUCCAUGAUGG GGCAUCGGGUCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUUCUAGUGUGACAACUCCGUACAA UGCAGACUUGACGGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUC CAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUG UGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACG GCACAUUGUGAUGGGGACAUGUUAUUCUCAAagagcagcggacUcaUAUUUCGAUGUGCCUUGCGC CGGGAAACACGCAAGGGAUGGUGUCAAAUUCGGCGAAACCUAAGCGCCCGCCCGGGCGUAUGGCA CGCCGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGACUAGACGGCACCCACCUAAGGCA AGACCAGUGGACAAUCGACGGAUAACAGCAUAUCUAG	Azoarcus olearius BH72 GROUP I INTRON-PIE-8 POLR2A precursor
5 7 1	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAAACCGGaaUccgUUggUgcUgCCGGCAGCCAAC UCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGAUUCUCCCAUGGUCUACCUUUCGCUUGAAU CUUAGUGUGACAACUCCGUACAAGCAGACUUGACGGGGAUGAGAUAACUUGCACCUGCCACAGU CUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCACCCCCAGAG CAUUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGAC GUCUUCUGGAGCGGGUGGAACGGGCAUUGUGUGAUGGGGACAUGUUAUUCUCAAagagcagcggac UcaUAUUUCGAUGUGCCUUGCGCCGGGAAACACGCAAGGGAUGGUGUCAAAUUCGGCGAAACCUA GGCCCGCCCGGGCGUAUGGCAACGCGGAGCCAAGCUUCGGCGCCUGCGCCGAUGAAGGUGUAGAGAC CUAGACGGCACCCACCUAAGGCAAAACGCUAUGGUGAAGGCAUAGUCCAGGGAGUGGCGAAAGUCACA CGACCAGUGGACAAUCGACGGAUAACAGCAUAUCUAG	Azoarcus olearius BH72 GROUP I INTRON-PIE-9 POLR2A precursor
5 7	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAAUUAUCGAAAAUCCUCUAAUUCAGGGAAAC ACCUAAACAAACUAGAGUAAGGCAUCCUGAGCUAAGCUCUUAAGUAUAAGAGAAAGUGCAACGAC UAUUCGAUAGGAAGUAGGUGCAAGUGACUCGAAUUGGGGAUUAACCUUCUAGGGUAGUGUAUAGU	Staphylococ us phage Twort GROUP

	CUGAACAUUAUGGAAACAUUAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAUUGaacUUU UagUUaUCCGGCAGCCAACUCUGCACAAAUGUCCAUAGAUGGGGCAUCGGGUCCGCAUUCUCCCAUG GUUACCUUUUCGUUGAAUCUUAUGUGGACAAUCUCGUACAUAUGCAGACUUUGACGGGGAUAGAGAUG AACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCA UGAUUGUACACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUUGCAGGACACACUCACAGCAGUGCG CAAAUUCACCAAGAGAGACUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGACAUAUGUU AUCUUCAAacUacUgaaagcaUAAAUAUUGUGCCUUUAUACAGUGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	I INTRON- PIE-2 POLR2A precursor
5 7 3	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAUCCUGAGCUAAGCUCUUAGUAUAAGAGAAAGU GCAACGACUAUUCGAUAGGAAGUAGGGUCAAGUGACUCGAAUUGGGGAUUACCCUUCUAGGGUAGU GAUAUAGUCUGAACAUUAUUGGAAACAUUAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAU UGaacUUUUagUUaUCCGGCAGCCAACUCUGCACAAAUGUCCAUAGAUGGGGCAUCGGGUCCGCAUU CUCCCAUGGUUACCUUUUCGUUGAAUCUUAUGUGGACAAUCUCGUACAUAUGCAGACUUUGACGGGG AUGAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGU UCCUCGCAUGAUUGUCCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUUGCAGGACACACUCACA GCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGG ACAUUGUUUAUCUCAAacUacUgaaagcaUAAAUAUUGUGCCUUUAUACAGUAAUGUAUAUCGAAA AAUCCUCUAUUCAGGGAACCCUAAACAAACUAAGAUGUAGGCAGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	Staphylococ- us phage Twort GROUP I INTRON- PIE-4-5 POLR2A precursor
5 4	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAGAUGUAGGCAUCCUGAGCUAAGCUCUUAGUA AUUAGAGAAAGUGCAACGACUAUUCGAUAGGAAGUAGGGUCAAGUGACUCGAAUUGGGGAUUACCC UUCUAGGGUAGUGUAUAGUCUGAACAUUAUUGGAAACAUUAGAAGGAUAGGAGUAACGAACCUAU UCGUAAACAUAAUUGaacUUUUagUUaUCCGGCAGCCAACUCUGCACAAAUGUCCAUAGAUGGGGCAU CGGGUCCGCAUUCUCCCAUGGUUACCUUUUCGUUGAAUCUUAUGUGGACAAUCUCCGUACAAGCAG ACUUUGACGGGGAUGAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGA GCUUGGCCAUGGUUCCUGCGCAUGAUUGUCCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUUGUGCA GACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGGCACA UGUGUGAUGGGGACAUGUUUAUCUCAAacUacUgaaagcaUAAAUAUUGUGCCUUUAUACAGUAA UGUAUUCGAAAAAUCCUCUAUUCAGGGAACCCUAAACAAACUGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	Staphylococ- us phage Twort GROUP I INTRON- PIE-5 POLR2A precursor
5 5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAAUAGAGAAAGUGCAACGACUAUUCGAUAGGA AGUAGGGUCAAGUGACUCGAAUUGGGGAUUACCCUUCUAGGGUAGUGUAUAGUCUGAACAUUAUUG GAAACAUUAUAGAAGGAUAGGAGUAACGAACCUAUUCGUAAACAUAAUUGaacUUUUagUUaUCCGGCA GCCAACUCUGCACAAAUGUCCAUAGAUGGGGCAUCGGGUCCGCAUUCUCCCAUGGUUACCUUUUCGC UUGAAUCUUAUGUGGACAAUCUCGUACAUAUGCAGACUUUGACGGGGAUGAGAUGAACUUGCACCUGC CAGAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCCACCCC CCAGAGCAAUCGGCCUGUCAUGGGUAUUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAG AGAGACGUCUUCUGGAGCGGGUGGAACGGGCACAUCUGUGAUGGGGACAUGUUUAUCUCAAacUac UgaaagcaUAAAUAUUGUGCCUUUAUACAGUAAUGUAUUAUCGAAAAUCCUCUAUUCAGGGAACA CCUAAACAAACUAAGAUGUAGGCAUCCUGAGCUAAGCUCUUAUGUAGCCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	Staphylococ- us phage Twort GROUP I INTRON- PIE-6 POLR2A precursor
5 7 6	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCUAGGAAGUAGGGUCAAGUGACUCGAAUUGGGGAU UACCCUUCUAGGGUAGUGUAUAGUCUGAACAUUAUUGGAAACAUUAGAAGGAUAGGAGUAACGAA CCUAUUCGUAACAUAAUUGaacUUUUagUUaUCCGGCAGCCAACUCUGCACAAAUGUCCAUAGAUGG GGCAUCGGGUCCGCAUUCUCCCAUGGUUACCUUUUCGUUGAAUCUUAUGUGGACAAUCUCCGUACA UGCAGACUUUGACGGGGAUGAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAU CAGGAGCUGGCAUGGUUCCUCGCAUGAUUGUCCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUUG UGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACG GCACAUUGUGAUGGGGACAUGUUUAUCUCAAacUacUgaaagcaUAAAUAUUGUGCCUUUAUAC AGUAUUGUAUUCGAAAAAUCCUCUAUUCAGGGAACCCUAAACAAACUAAGAUGUAGGCAUCCU GAGCUAAGCUCUUAUGUAUAGAGAAAGUGCAACGACUAUUCGAGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	Staphylococ- us phage Twort GROUP I INTRON- PIE-7 POLR2A precursor
5 7 7	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCCUAGGGUAGUGUAUAGUCUGAACAUUAUUGGAA ACAUUAUAGAAGGAUAGGAGUAACGAACCUAUUCGUAACAUAAUUGaacUUUUagUUaUCCGGCAGCC AACUCUGCACAAAUGUCCAUAGAUGGGGCAUCGGGUCCGCAUUCUCCCAUGGUUACCUUUUCGUUG AAUCUUAGUGUGACAACUCCGUACAUAUGCAGACUUUGACGGGGAUGAGAUGAACUUGCACCUGCCAC AUUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCCACCCCCA GAGCAAUCGGCCUGUCAUGGGUAUUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGA GACGUCUUCUGGAGCGGGUGGAACGGGCACAUGUGUGAUGGGGACAUGUUUAUCUCAAacUacUga aagcaUAAAUAUUGUGCCUUUAUACAGUAUUGUAUUCGAAAAUCCUCUAUUCAGGGAACCCU AAACAAACUAAGAUGUAGGCAUCCUGAGCUAAGCUCUUAUGUAUUAAGAGAAAGUGCAACGACUAU CCGAUAGGAAGUAGGGUCAAGUGACUCGAAUUGGGGAUUACCCUUGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	Staphylococ- us phage Twort GROUP I INTRON- PIE-8 POLR2A precursor
5 7 8	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACAAUUAAGAAGGAUAGGAGUAACGAACCUAUUC GUAACAUAAUUGaacUUUUagUUaUCCGGCAGCCAACUCUGCACAAAUGUCCAUAGAUGGGGCAUCG GGUCCGCAUUCUCCCAUGGUUACCUUUUCGUUGAAUCUUAUGUGGACAAUCUCCGUACAUAUGCAGAC UUUGACGGGGAUGAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGC UGGCCAUGGUUCCUCGCAUGAUUGUCCACCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUUGUGCAGGA CACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGGCACAUG	Staphylococ- us phage Twort GROUP I INTRON- PIE-9 POLR2A precursor

	UGUGAUGGGGACAUGUUAUUCUCAAacUacUgaaagcaUAAAUAAUUGUGCCUUUAACAGUAAUG UAUAUCGAAAAAUCCUCUAAUUCAGGGAACACCUAAACAAACUAAAGAUAGGCAAUCCUGAGCUAA GCUCUUAAGUAAUAGAGAAAGUGCAACGACUAUUCGGAUAGGAAGUAGGGUCAAGUGACUCCGAAUAG GGGAUUAUCCUUCUAGGGUAGUGAUUAUAGUCUGAACAUUAUAGGAGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	
5 7 9	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCUAACGAACCUAUCGUAACAUAAUUGaacUUUUA gUUAUCCGGGACGCCAACUCUGCACAAGUUGCCAUAGUGGGGCAUCGGGUCCGCAUUCUCCCAUGGU CUACCUUUCGCUUGAAUCUUAUGUGACAACUCCGUACAAGCAGACUUUGACGGGGAUGAGAUGAA CUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUG AUUGUCACCCCCACAGCAUUCGCGCUCUAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCA AAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGACAUGUUUAU CUUCAAAcUacUgaaagcaUAAAUAAUUGGCCUUUAUACAGUAAUGUAUUCGAAAAUCCUCUA UUCAGGGAACACCUAAACAAACUAAAGUAGGCAUCCUGAGCUAAGCUCUUAAGUAAUAAAGAGAAA GUGCAACGACUAUUCGAUAGGAAGUAGGGUCAAGUGACUCCGAAUUGGGGAUUAUCCUUCUAGGGUA GUGAUUAGUCUGAACAUUAUAGGAACAUAUAGAAGGAUAGGAGGACCAGUGGACAAUCGACGGAU AACAGCAUAUCUAG	Staphylococ- cus phage Twort GROUP I INTRON- PIE-9.1 POLR2A precursor
5 8 0	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCGCGCCCGAGUACCAGGACCGUCGGGCCCCGAGA AGGGGGGGCCUAGGAUUGCGGCAAGACGACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCG AGCGGAGGUGCCGUAGCGGGCACCCGUCGUACGCGCGGUAGGGCGUCGGUCCCCUCCACGGCG GAGGGGAGGCUUAAAGUACGUGCUAAACCCCCAGCGAUUGGGCCUGUAGGAAAGCCUUGUACGGC GAAGCCUACGGGGACCGCCGAUGGGCGGUGCGAUGCCGGCGGGCCACGGAGCCCCCGGCGUCGAUUG GUUUUACCGGAGCCAAACUCUGCACAATAUGUCCAUAGUAGGGGCAUCGGGUCGCAUUCUCCAUUGG UCUACCUUUCGCUUGAAUCUUAUGUGGACAACUCCGUACAAGCAGACUUAUGACGGGGAUGAGAUGA ACUUGCACCUUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAU GAUUGUACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGC AAAUUACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGACAUGUUUA UCUCAAUAACAGGUUGACGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUCUGGUGUUUUUU UCUUCGCUAGUCGAUCAUCCGUGACGGCCGGGCAAGACCAGUGGACAAUCGACGGAUAAACAGCAUA UCUAG	Scytalidium dimidiatum GROUP I INTRON-PIE-2 POLR2A precursor
5 8 1	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCGCGUGGCCGAUCCCGAGGCGAGGUGCCGUAGCG GGCACCCGUCGUAAACGCGCGGUAGGGCGUCGUCUCCCCUCCACGCGGAGGGGAGGCUUAAAGGUA CGUGCUAAACCCCGAGCAUGGGGCCUGUAGGAAAGCCUGGUACGGCGAAGCCUACGGGGACCGC CCGAUGGCGGUGCGAUGCCGGCGGGCCACGGAGCCCGCGGCGUGCAUUGcUgUUAACGGCAGCCAAC UCUGCACAATAUGUCCAUGAUGGGGCAUCGGGUCGCAUUCUCCAUUGGUCUACCUUUCGCUUGAAU CUUAGUGUGACAACUCCGUACAAGCAGACUUGAGCGGGGAUGAGAUAACUUGCACCUGCCACAGU CUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUUGUACCCCCCAGAG CAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGAC GUCUUCUGGAGCGGUGGAACGCGCAUGUCUGAUGGGGACAUGUUUAUCUCAAUaacaggUUGA CGACAACAGGCCUGUGACAGCGGGGCUUCGAACUUCGUGUUUUUUUUUCUUCGCUAGUCGAUCA UCCGUGACGGCCGGCAAGCGCCCGAGUACAGGACCGUCGGGCGGAGAGAGGGGGGGCCUAGGA UGCGGCAAGACGACCCGCUUCGGGGAACGCCAGCGUGACCAUGGACAAUCGACGGAUAAACAGCAUA UCUAG	Scytalidium dimidiatum GROUP I INTRON-PIE-5 POLR2A precursor
5 8 2	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCGCGGGCACCCGUCGUAAACGCGCGUAGGGCGUCG GUCCCCCUCUCCACGGCGGAGGGGAGGCUUAAAGUACGUGCUAAACCCCGAGCAUGGGGCCUGUAG GAAAGAGCCUUGUACGGCGAAGCCUACGGGGACCGCCGUAUGGGCGUGCGAUGCGCGGGCCACGG AGCCCCGGCGUGCAUUGcUgUUAACGGCAGCCAAUCUCGACAAAAUGUCCAUGAUGGGGCAUCGG GUCCGCAUUCUCCAUUGGUCUACCUUUCGCUUGAAUCUUAUGUGGACAAUCUCCGUACAAUGCAGACU UUGACGGGGAUGAGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAUCCAGGAGCU GGCAUGGUUCCUCGCAUGAUUGUACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGAC ACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUGU GUGAUGGGGACAUGUUUAUCUCAAUaacaggUUGACGACAACAGGCCUGUGACAGCGGGGCUUCGA ACUUCUCCUGUUUUUUUUUCUUCGCUAGUCGAUCAUCCUGACGCGCCGGGCAAGCGCCGAGUACC AGGACCGUGCGGCCCCCGAGAAAGGGGGGCCUAGGAUGCGGCAAGACGACCCGGUUCGGGGAAACGG AGCGUGCGUGGCCGAUCCCGAGGCGAGGUGCCGUAAGACCAGUGGACAAUCGACGGAUAAACAGCAUA UCUAG	Scytalidium dimidiatum GROUP I INTRON-PIE-6 POLR2A precursor
5 8 3	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCGCGGAGGGGAGGCUUAAAGUACGUGCUAAACCC CCAGCGAUGGGGCCUGUAGGAAAGCCUUGGUACGGCGAAGCCUACGGGACCGCCGUAUGGGCGGUC GGAUGCGCGGGCCACGGAGCCCCCGGCGUCGAUUGcUgUUAACGGCAGCCAAUCUCGACAAAAU GUCCAUGAUGGGGCAUCGGGUCCGCAUUCUCCAUUGGUUACCUUUCGCUUGAAUCUUAUGUGGACA ACUCCGUACAAGCAGACUUAUGACGGGGAUGAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAUCCAGGAGCU GAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUACCCCCCAGAGCAAUCGGCCUGU CAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAG CGGGUGGAACGGCACAUGUGUGAUGGGGACAUGUUUAUCUCAAUaacaggUUGACGACAACAGGCC UGUGACAGCGGGGCUUCGAACUUCGUGUUUUUUUCUUCGCUAGUCGAUCAUCCUGACGCGCCGGGCAAGCGCCGAGUACC GGGCAAGCGCCGAGUACAGGACCGUCGGGCCCCGAGAAAGGGGGGCCUAGGAUGCGGCAAGACGACCCGGUUCGGGGAAACGG ACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGCGAGGUGCCGUAAGCGGGCACCCGU CGUAACGCGCGUAGGGCGUCGUGCCCCCUCCACGACCAGUGGACAAUCGACGGAUAAACAGCAUA UCUAG	Scytalidium dimidiatum GROUP I INTRON-PIE-8 POLR2A precursor

5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCGGAGCCCCGGCGUCGAUUGcUgUUAACGGCAGC	Scytalidium dimidiatum GROUP I INTRON-PIE-9 POLR2A precursor
8	CAACUCUGCACA AAAUGUCCAUGAUGGGGCAUCGGGUCCGAUUCUCCCAUGGUCUACCUUUCGCUU	
4	GAAUCUUAGUGUGACAACUCCGUACA AUGCAGACUUGACGGGGAUGAGAUAACUUGCACCUGCCA	
	CAGUCUCUGGAGACGCGAGCAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCACCCCC AGAGCAAUCCGCCUUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAG AGACGUUCUCCUGGAGCGGGUGGAACGGCACAUGUGUAUGGGGACAUGUUAUCUUAUaacagg UUGACGACAACAGGCCUGUGACAGCGGGGCUUCCGAACUUCUGGUGUUUUUUUUUCUUCGCUAGUCG AUCAUCCGUGACGGCCGGCAAGCGCCCGAGUACCAGGACCGUCGGGCCCCGAGAGAGGGGGGCGCU AGGAUGCGGCAAGACGACCCGGUUCGGGGAACGCCAGCGUGCGCUGGCCGAUCCCGAGGCGAGGUGC CCUAGCGGGCACCUCGUAAACGCGCGUAGGGCGUCGGUCCCCCUCCACGGCGGAGGGGAGGCU UAAGGUACGUGCUAAACCCAGCGAUGGGGCCUGUAGGAAAAGCCUGGUAACGGCGAAGCCUACGG GGACCGCCCGUAGGCGGUCGGAUGCCGGCGGCCACGACCAGUGGACAUAUCGACGGAUAAACAGCAUA UCUAG	
5	gggagacccUcgaCCGUCGAUUGUCCACUGGUCAAUGGAGAAGGUGUAGAGACUGGAAGGCAGGUAC	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-6 POLR2A precursor
8	CCUAAACAGUAAAGCUGAGGUAAAGGGACAGUCCAGACCAAAACUGGUUAAUCCAGGCAGUGAAAA	
5	CUGUAGAUGGUAAAGcaUaacccCCGGCAGCCAACUCUGCACA AAAUGUCCAUGAUGGGGCAUCGGGU	
	CCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGAGACAACUCCGUACA AUGCAGACUUU GACGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUAUCAGGAGCUGG CCAUGGUUCCUCGCAUGAUUGUCACCCCCCAGAGCAAUCCGGCCUGUCAUGGGUAUUGUGCAGGACAC ACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUGUGU GAUGGGGACAUGUUAUCUUAAGggcUCAACGAGUAAGAUAUACUAAACGCUUAAAGUUAUUCU ACUAUGGGCGGUACGUAAAGAAACUACGUUUGUUAUCCUGUCAACUCGGGGAAGCCAAUAGCGCG GUAAUCCCGAACCAGCUCAGAGACCAGUGGACAUAUCGACGGaUaacagcaUaUcUag	
5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCACUUAAGUAUGUUUACCGUCACAAUCUGGGGAA	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-2 POLR2A precursor
8	GCCAAUAGCGCGUAAUCCCGAACCAAGCUCAGAAAUGGAGAAGGUGUAGAGACUGGAAGGCAGGU	
6	ACCCUAAACAGUAAAGCUGAGGGUAAAGGGACAGUCCAGACCAAAACUGGUUAAUCCAGGCAGUGAA	
	AAUCUGUAGAUGGUAAAGCAUAACCCCGCAGCCAACUCUGCACA AAAUGUCCAUGAUGGGGCAUCGG GUCCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGAGACAACUCCGUACA AUGCAGACU UUGACGGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUAUCAGGAGCU GGCCAUGGUUCCUCGCAUGAUUGUCACCCCCCAGAGCAAUCCGGCCUGUCAUGGGUAUUGUGCAGGAC ACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUGU GUGAUGGGGACAUGUUAUCUUAAGGGCUCACAGAGUAAGAUAUGACUAAACGCUUAAAGUUAUUC UUAUAUUGGGCGGUACGUAAAGAGACCAGUGGACAUAUCGACGGUAACAGCAUAUUCUAG	
5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAUCCCGAACCAAGCUCAGAAAUGGAGAAGGUGU	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE- 4-5 POLR2A precursor
8	AGAGACUGGAAGGCAGGUACCCUAAACAGUAAAGCUGAGGGUAAAGGGACAGUCCAGACCACAAACUG	
7	GUUAAUCCAGCGAGUAAACUGUAGUAGGUAAAGCAUAAACCCCGCAGCCAACUCUGCACA AAAUG	
	UCCAUGAUGGGGCAUCGGGUCCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGACAA CUCCGUACA AUGCAGACUUAUGACGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCG AGCAGAGAUCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCACCCCCCAGAGCAAUCCGGCCUGUC UUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGC GGUGGAAACGGCACAUUGUGAUGGGGACAUGUUAUUCUUAAGGGCUCACAGAGUAAGAUAUGACUA AACGCUUAAAGUUAUUCUUAUUAUGGGCGGUACGUAAAGAAACUUAAGUAUGUUAUCCUGUCAAAC UCGGGGZAGCCAAUAGCGCGGUAGACCAGUGGACAUAUCGACGGUAACAGCAUAUUCUAG	
5	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAAUCCAGGCAGUGAAAACUGUAGAUGGUAAAGCAU	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-9 POLR2A precursor
8	AACCCCGCGGACGCCAACUCUGCACA AAAUGUCCAUGAUGGGGCAUCGGGUCCGAUUCUCCCAUGGU	
8	CUACCUUUCGCUUGAAUCUUAUGUGACAACUCCGUACA AUGCAGACUUAUGACGGGAUGAGAUA	
	CUUGACCCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUAUCAGGAGCUGGCCAUGGUUCCUCGCAUG AUUGUCACCCCCCAGAGCAAUCCGGCCUGCAUUGGUUAUUGUGCAGGACACUCACAGCAGUGCGCA AAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUGUGUGAUGGGGACAUGUUAU CUUAAGGGGCUCAACGAGUAAGAUAUGACUAAACGCUUAAAGUUAUUCUUAUUAUGGGCGGUACGUA AAGAAACUUAUGUUAUCCUGUCAAAACUCGGGGAAGCCAAUAGCGCGUAUCCCGAACCAAGC UCCAGAAUUGGAGAAGGUGUAGAGACUGGAAGGCAGGUACCCUAAACAGUAAAGCUGAGGGUAAAGGG ACAGUCCAGACCACAACACUGGUUAGACCAGUGGACAUAUCGACGGUAACAGCAUAUUCUAG	
5	gggagacccUcgaCCGUCGAUUGUCCACUGGUCAAUCCGGUGAAGGUGUAGAGACUGGAUGGGCAGC	Agrobacteriu m- tumefaciens GROUP I INTRON-PIE-6 POLR2A precursor
8	ACCUAAAGGGCGCAAGGUCCACGGUGAAGGGACAGUCCAGACCAGAACGCCUUUACCGCAUUGAGA	
9	UGGGCGGGCGGCAAGCCGAAGCGGUAGaaUccUggggCCGGCAGCCAACUCUGCACA AAAUGUCC	
	AUGAUGGGGCAUCGGGUCCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGACAACUC CGUACA AUGCAGACUUAUGACGGGAUGAGAUAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGC AGAGAUCAGGAGCUGGCCAUGGUUCCUCGCAUGAUUGUCACCCCCCAGAGCAAUCCGGCCUGUCAUG GGUAUUGUGCAGGACACACUCACAGCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGG UGGAACGGCACAUGUGUGAUGGGGACAUGUUAUUCUUAACaggaUUccUAAGGUAAAUUGGGGUGU GCGCCCGAAACGACGCAUUCGAUCUGCUCAAAGUCGGGGAAGCUUCGCUUGUCUUCAGUAUGCC AAUCCGAGCCAAAGCUCGGAGACACAGUGGACAUAUCGACGGaUaacagcaUaUcUag	
5	AAAACCCGUUGACCCGCGAGCCAACUCUGCACA AAAUGUCCAUGAUGGGGCAUCGGGUCCGCAUUCUC	Anabaena pre-tRNA-Leu gene GROUP I INTRON POLR2A circRNA
9	CCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGACAACUCCGUACA AUGCAGACUUAUGACGGGAUG	
0	AGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUAUCAGGAGCUGGCCAUGGUUCC	
	UCGCAUGAUUGUCACCCCCCAGAGCAAUCCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCA GUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUGUGUGAUGGGGACA UUGUUAUUCUCAAAGACGCUACGACUU	

5 9 1	AAUCCGUUGGUGUCUGCCGGCAGCCAACUCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGCAUUCUCCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGUGACAACUCCGUACA AUGCAGACUUGACGGGG AUGAGAUAAUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGU UCCUCGCAUGAUUGUCACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACA GCAGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGG ACAUUGUUAUCUUAAGAGCAGCGGACUCAU	Azoarcus olearius BH72 GROUP I INTRON-PIE-2 POLR2A circRNA
5 9 2	AACUUUUAGUUAUCCGGCAGCCAACUCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGCAUUCU CCCAUGGUCUACCUUUCGCUUGAAUCUUAUGUGUGACAACUCCGUACA AUGCAGACUUGACGGGGAU GAGAUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUC CUCGCAUGAUUGUCACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGC AGUGCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGAC AUUGUUAUCUUAACUACUGAAAGCAU	Staphylococcus phage Twort GROUP I INTRON- PIE-2 POLR2A circRNA
5 9 3	CUGUUAACCGGACGCCAACUCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGCAUUCUCCCAUGG UCUACCUUUCGCUUGAAUCUUAUGUGUGACAACUCCGUACA AUGCAGACUUGACGGGGAUAGAGAU ACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAU GAUUGUCACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGCGC AAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGACAUUGUUA UCUCAAUAACAGGU	Scytalidium dimidiatum GROUP I INTRON-PIE-2 POLR2A circRNA
5 9 4	CAUAACCCCGGACGCCAACUCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGCAUUCUCCCAU GGUCUACCUUUCGCUUGAAUCUUAUGUGUGACAACUCCGUACA AUGCAGACUUGACGGGGAUAGAGAU GAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUCGCAU GAUGAUUGUCACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGC GCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGACAUUGU UAUCUUAAGGGCU	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-6 POLR2A circRNA
5 9 5	AAUCCUGGGGCGGACGCCAACUCUGCACAAAUGUCCAUGAUGGGGCAUCGGGUCCGCAUUCUCCC AUGGUCUACCUUUCGCUUGAAUCUUAUGUGUGACAACUCCGUACA AUGCAGACUUGACGGGGAUAGAG AUGAACUUGCACCUGCCACAGUCUCUGGAGACGCGAGCAGAGAUCCAGGAGCUGGCCAUGGUUCCUC GCAUGAUUGUCACCCCCCAGAGCAAUCGGCCUGUCAUGGGUAUUGUGCAGGACACACUCACAGCAGUGC GCGCAAAUUCACCAAGAGAGACGUCUUCUGGAGCGGGUGGAACGGCACAUUGUGAUGGGGACAUU GUUAUCUUAACAGGAUCCU	Agrobacteriu m- tumefaciens GROUP I INTRON-PIE-6 POLR2A circRNA
5 9 6	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUAACAUAUGAGACUUAACAUAUCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAACGUCAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCA AUGAGGACUAGCGAAAGCUGCAAGAGAAUGAAAUCUGUUAUAAACAGCCUGUGGGUUGCCACCC ACCCACAGGGGCCACUGGGCGUAGCACUCUGGUUAUCACGAUACCUUUGUGGCCUGUUUAUAUCC CCACCCCGAGUAAACGUUAGAAGUUAACGCAACCCCGAUCAUAGUAGGUUAGCACUCCAGCUGCAU CGAGAUCAAGCACUUCUGUCUCCCCGACCGAGUAUCAUAGACUGCUAACCGCGUUGAAGGAGAAA ACGUUCGUUAACCGGCCAAUUAUUCGAGAAGCCAGUAGUGCCGUGAAAGUUGCGGAGUGUUUCGC UCAGCACUUCUCCCCGUGUAGAUACAGGCUGAUGAGUACCCGGAUCCACAGGUGACUGUGCGGUG GCUUGCGUUGGGGGCUGCCUAUGGGGCAACCAUAGGACGCUCAUUAACAGACAUGGUGCGAAGAGC CUAUUGAGCUAAUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACUGCGGAGCAUGCCCC UCAAACAGGGGGUGGUGUGUCGUAACGGGUAAUCUCUGAGCGGAACCGACUACUUGGGUGUCCGU GUUUUUUUUAUUCUUAUAAUGGUGCUUAUGGUGACAAUUAAGAAUUGUUAACCAUAUAGCUAUUG GAUUGGCCAUCCGGUGACUAAACAAUUCGCUCAUUAUACAGUUGUUGUUUUGUCCCUUAUACAU ACAGCUCAUAACACCCUUAUUAUUAUACUAAUUAUAGCAAGAAUUGAGGACGCCAAGAACAU CAAGAAAGGGCCCGGCCUUCUACCCGUGGAGGACGGGACCGCCGGCAGCAGCUCCACAAGGCC AUGAAGCGGUACGCCUGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUAGAGUCCGACAUCA CCUACCGCGGAGUACUUCGAGAUAGCGUGCGCCUGGCCAGGCCAUGAAGCGGUACGGCCUGAACAC CAACACCGGAUCUGGUGUGUCUGGAGAACAGCCUGCAGUUCUUAUGCCGUGUGGCCGCCUUC UUCAUCGGCGUGGCCUGGCCCGGCCGGAACGACUUCUACAACGAGCGGGAGCUGUGAACAGCAUGG GGAUCAGCCAGCCAGCCUGGUGUUCUGAGCAAGAAAGGCCUGCAGAAUCCUGAACGUGCAGAA GAAGCUGCCCAUCAUCCAGAAGAUCAUCAUGGACAGCAAGACCGACUACAGGGCUUCCAGUCG AUGUACACGUUCUGUACCGAGCCACUCCCGCCGGGCUUACAAGAGUACGACUUCGUCCCGAGAGCU UCGACCGGGACAAGACCAUCGCCUGUAUGAAGACGAGCGGACGACCGGCCUGCCGAAGGGGU GGCCUGCCGACCGGACCGCCUGCGUGCGCUUCUGCAGCGCCGGGACCCAUUCGGCAACCCAG AUCUACCCGGACACCGCAUCCUGAGCGUGGUGCCGUUCCACACGGCUUCGGCAUGUUCACGACCC UGGGUACCUCAUCUGCGGCUUCCGGGUGGUCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG GAGCCUGCAGGACUACAAGAUCCAGAGCGCGUGUCUGGCGACCCUGUUCAGCUUCUUCGCCAAG AGCACCCUUAUCGACAAGUACGACCUUGUGCAACCGACGAGAUCCGACCGGGGGCGCCCGCUA GCAAGGAGGUGGGCGAGGCCUGGCCAAGCGGUUCCACUCCGGGCAUCCGCCAGGGCUACGGCCU GACCGAGACACGAGCGCAUCCUGAUCAACCCGAGGGGACGACAAGCCGGGCGCCUGGGCCAG GUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGACUGGACACCGGCAAGACCCUGGGGUGAACCCAG GGGGCGAGCGUUGCGUGGGGGCGGAUGAUCAUGAGCGGCUACGUGAACAAACCGGAGGCCACCA CGCCUUAUCGACAAGGACGGCUGGUGCACAGCGGCGACUCCCUACUGGGACGAGGACGAGCAC UUCUUAUCGUGCAGCCGUGAAGUCGUGUAUAGUACAGGGCUACAGGUGGGCGCCGGCCGAGC UGGAGAGCAUCCUGUCCAGCACCCCAACUUCGACGCGCGGUGGCGGCGUGCGGCGCGGACGACA CGCCGGCGAGCUGCCGGCGCGGUGGUGGUGGUGGAGCAGCGCAAGACCAUGACGGAGAAGGAGAU GUCGACUACGUGGCCAGCCAGGUGACCAACCGCAAGAGUGCGGGGCGGCGUGGUGUUCUGGACG	AnaX-CAV2- Luciferase precursor

	AGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGAAGAUCCGCGAGAUCUGAUCAAGGCCAA GAAGGGCGGCAAGAUCCGCCUGAGCUAAAGACGCUACGGACUUAUAAUUAUGAGCCUUAAGAAGAA AUUCUUUAAGUGGAUGCUCUCAAACUCAGGGAAACCUAAAUUCUAGUUUAUAGACAAGGCAAUCCUGAG CCAAGCCGAAGUAGUAAUUAAGUAAAGACCAGUGGACAAUCCGACGGAAACAGCAUAUCUAG	
5 9 7	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAAUGGAGAAGGUGUAGAGACUGGAAGGCGAGUAC CCUAACAGUAAAGCUGAGGCUAAAGGGACAGUCCAGACCACAAACUGGUUAAUCCAGCCAGUGAAAA CUGUAGAUUGGUAAGCAUAAACCUUAAACAGCCUGUGGGUUGCACCACCCACAGGGCCACUGGGC GCUAGCACUCUGGUUACAGAUACCUUUGUGCGCCUGUUUUUAUACCCACCCGAGUAAACGUUAG AAGUUAACGCAACCCCGAUCAAUAGUAGGUGUAGCACCAGCUGCAUCCGAGAUCAAGCACUUCUGUC UCCCCGACCGAGUAUCAAUAGACUGCUAACCGGUGUAGAGGAGAAACGUUCGUUACCCGGCCAAU UACUUCGAGAAGCCAGUAGUGCCGUGAAAGUUGCGGAGUGUUCGCUACAGCACUCCCCCGUGUAG AUCAGGCGUAGAGUACCCGCAUCCACAGGUGACUGUGCGGUGGCGUGCUUGGCGGCCUGCCU AUGGGGCAACCAUAGGACGCUUAUACAGACAUGGUGCGAAGAGCCUAUUGAGCUAAUUGGUAGU CCUCCGGCCCCUGAAUGCGGCUAAUCCUAAUCUGCGGAGCACAUCCCCUAAACAGGGGGUGUGUG UCGUAAACGGGUAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUUUAUUCUUAUA UGGCGUUAUUGGUGACAAUUAAGAAUUGUUAACCAUUAUAGCUAUGGAUUGGCCAUCCGGUGACUA ACAAAUCCGCUCAUAUACGUUUUGUUGUUUUGUCCCUUAUACAUACAGCUCAUAAACCCUCUU AUUAUUACUACAUAUUAAGAAUAGCAAGAGCCACCAUGGAGGAGCCAAAGAACUAAGAAGGGCCCG CGCCUUCUACCCCGUGGAGGACGGGACCCCGCGGAGCAGCUCCACAGGCCAUAGAAGCGGUACGC CCUGGUGCGGGGACGAUCGCCUUCACCGACGCCCAUACGAGGUGCGAUACCUACCGCGGAGUAC UUCGAGAUAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAACCCAGGUAUCG UGGUGUGCUCGGAGAACAGCCUGCAUUCUUAUGCCGGUGUGGGGCCCCUUCUACGCGGUGGC CGUCCGCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCUGAACAGCAUUGGGGAUCAGCCAGCCG ACCGUGUGUUCUGAGCAGAAGAGGGCCUGCAGAAGAUCCUGAACGUGCAGAGAAGCUGCCCAUA UCCAGAAGAUCAUCAUGGACAGCAAGACCGACUACAGGGCUUCCAGUCGAUGUACAGCUUCGU GACCCAGCCACUCCCGCCGGGUUCAAACGAGUACGACUUCGUCCCGGAGAGUUCGACCGGGCAAG ACCAUCCGCCUGAUCAUAGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGUGGCCUUGCCGACC GGACCGCCUGCGUGCGCUUCUGCAGCGCCGGGACCCCAUCUUCGGCAACGAUACUCCGGACAC GCCAUCCUGAGCGUGGUGCCGUUCCACACCGCUUCGGCAUGUUAACGACCCUGGGCUACCCUACU UGCGGCUUCCGGGUGGUCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCGGAGCCUGCAGGACU ACAAGAUCCAGAGCGCGUGCUGCGGACCCUGUUCAGCUUCUUCGCCAAGAGCACCUGAUCGA CAAGUACGACCCUGUCGAACCGACAGAGAUCCGAGCGGGGCGGCCCGUGAGCAAGGAGGUGGGC GAGGCCUGGCGCAAGCGUUCACCUCUCCGGCAUCCGCCAGGGCUACGGCCUGACCGAGACCACGA GGCGGAUCCUGAUACCCCGGAGGGGACGACAAGCCGGCGCCGUGGGCAAGGUGGUCCCGUUCUU CGAGGGCAAGGUGGUGGACUUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGC GUGCGGGGGCCGAUGAUCAUGAGCGGCUACGUGAACAACCCGAGGECACCAACGCCUUAUCGACA AGGACGGCUGGUGCAGCAGCGGACAUCCGCUACUGGGACGAGGACGAGCACUUCUUAUCUGUGCA CCGGCUGAAGUCGUGAUCAAGUACAAGGGCUACAGGUGGGCGCCGGCAGCUGGAGAGCAUCCUG UCGCGGUGAGCUGAUAAAGGCUCAACGAGUAGAUGACUAAACGCUUAAAGUUUAUUCUUAUCU GGCGGUAAGUAAAGAAACUUAUGUUAUCCUGUCAAACUCGGGGAGGCCAAUAGCGCGGUAAUC CCGAACCAAGCUCCAGAGACCAGUGGACAAUCGACGGAAACAGCAUAUCUAG	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-6 CAV2- Luciferase precursor
5 9 8	AAAAUCCGUUGAUUAAACAGCCUGUGGGUUGCACCCACCCACAGGGCCACUGGGCGCUAGCACUC UGGUUAUCACGAUACCUUUGUGCGCCUGUUUAUACCCACCCAGUAAACGUUAGAAGUACGCA ACCCCGAUCAUAGUAGGUGUAGCACCAGCUGCAUCGAGAUCAAGCACUUCUGUCUCCCGGACC GAGUAUCAUAGACUGCUAACCCTGUGUAGGAGAAAACGUUCGUUACCCGGCCAAUUAUCUUCGAG AGCCCAUAGUGCCGUGAAAGUUGCGGAGUGUUCGCUACGACUUCUCCCGUGUAGAUACGGCUGA UGAGUCACCGCGAUCCCAACAGGUGACUGUGCGGUGGCGUGCGUUGGGCGCCUGCCUAGGGGCAAC CCAUAGGACGCUUAAUACAGACAUGGUGCGAAGAGCCUUAUAGACUAAUUGGUAGUCCUCCGGCCC CUGAAUGCGGCUAAUCCUAAUCUGCGGAGCACAUGCCUCAAACAGGGGUGGUGUGUUGUAACGGG UAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUCUUAUUAUUAUUAUUGGUGCUUA UGGUGACAAUUAAGAAUUGUUAACCAUUAAGCUAUGGAUUGGCCAUCCGGUGACUAAACAAUCCGU CAUAUACCAUUGUUGUUGUUCUCCUUAUACCAUACAGCUCAUAAACCCUUAUUAUUAUUAUUAU CAAUUGAAUAGCAAGAAUUGGAGGACGCCAAGAACAUCAAGAGGGCCCGGCGCCUUAUACCCGUA GGAGGACGGGACCGCGCGGAGCAGUCCACAAGGCCAUGAAGCGGUACGCCUGGUGCCGGGACG AUCGCCUACACCGACGCCCAUCCGAGGUGGACAUACCUACCGGAGUACUUCGAGAUAGCGGUGC GCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAACCCGGAUCGUGGUGUGUCGCGGAGAA CAGCCUGCAGUUCUUAUGCCGGUGCUGGGCGCCUUCUUAUCGGCGUGGCCUGCCCGGCGAAC GACAUCUACAACGAGCGGAGCUGGAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGA GCAAGAAGGGCCUGCAGAAGAUCCUGAACGUGCAGAAGAAGCUGCCCAUACUCCAGAAGAUCAU CAUGGACAGCAAGACCGACUACCGAGGCUUCCAGUGCAUGUACAGUUCGUGACCAACCCUCCG CCGGGCUUAACAGAUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAUCA UGAACAGCAGCGGACGACCGGCCUGCCGAAGGGGUGGCCUUGCCGACCGGACCGCCUGCGUGCG UUCUCCGACGCCCGGACCCCAUUCGCGCAACCAAGAUCAUCCCGACACCGCAUCCUGAGCGUG GUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCUGGGCUACCUCAUCUGCGGCUUCGGGUGG UCCUGAUGUACGGGUCCGAGGAGGAGCUGUCCUGCGGAGCCUGCAGGACUACAAGUCCAGAGCGC	AnaX-CAV2- Luciferase circRNA

	GCUGCUCGUGCCGACCCUGUUCAGCUUCUUCGCCAAGAGCACCCUGAUCGACAAGUACGACCUUGUGG AACCCUGCACGAGAUCCGACGCGGGGGCGCCCCGCGUGAGCAAGGAGGUGGGCGAGGGCCUGGGCCAAAGC GGUUCACCCUCCCGGGCAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGGAUCCUGAUCAC CCCCGAGGGGGACGACAAGCCGGGGCGCCUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUG GACCCUGGACACCGGCAAGACCCUGGGCGUGAACCGCGGGCGAGCUGUGCGUGCGGGGGCCGAUGA UCAUGAGCGGCUACGUGAACAACCCGAGGGCCACCAACGCCCUCAUCGACAAGGACGGCUGGCUGCA CAGCGGGCGCAUCGCCUACUGGGGACGAGGACGAGCACUUCUUCUUCGUGGACCGGGCUGAAGUGCGUG AUCAAAGUACAAGGGCUACCAAGGUGGCGCGGCCGAGCUGGAGAGCAUCCUGUCCAGCACCCCAACA UCUUCGACGCGCGGCGUGGCCGGGCGGCCGAGCAGCAGCGCGCGAGCUGCCGGCCGCGGUGGUGGU GCUGGAGCACGGCAAGACCAUGACGGAGAAGGAGAUUCGUGACUACGUGGCCAGCCAGGUGACCACC GCCAAGAAGCUGCGGGGCGGCGUGGUGUUCGUGGACGAGGUCCGAAGGGCCUGACCGGGAAGCUCG ACGCCCGGAAGAUCGCGGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGAGCUAAG ACGCUACGGACUU	
5 9 9	CAUAACCCUUAAAACAGCCUGUGGGUUGCACCCACCCACAGGGCCCACUGGGCGCUAGCACUCUGGU AUCACGUAUACCUUUGUGCGCCUGUUUAUAUCCCCACCCCGAGUAAACGUAAGAUUACGCAACCC CGAUCAAUAGUAGGUGUAGCACUCCAGCUGCAUCGAGAUCAAGCACUUCUGUCUCCCGGACCGAGU AUCAAUAGACUGCUAACGCGGUUGAAGGAGAAAACGUUCGUUACCCGGCCAAUUAUUCGAGAAGCC CAGUAGUGCCGUGAAAGUUGCGGAGUGUUUCGUCAGCACUUCUCCCGUGUAGAUACGGCUGAUGAG UCACCGCGAUCCCCACAGGUGACUGUGGCGGUGGCGUGCGUUGCGGCCUGCCUAUGGGGCAACCCAU AGGACGCUUCAAUACAGACAUGUGGCGAAGAGCCUAUUGAGCUAAUUGGUAGUCCUCCGGCCCCUGA AUGCGGCUAAUCCUAACUGCGGAGCACAUGCCCUCAAACAGGGGGUGGUGUGUCGUAACGGGUAAAC UCUGCAGCGGAACCGACUACUUUGGGUGUCCGUGUUUUUAUUCUUAUAAUGGCUGCUUAGGU GACAUAUAAAGAAUUGUUACCAUAUAGCUAUUUGGAUUGGCCAUCCGGUGACUAACAAAUCGCUCAUA UACCAGUUUGUGUUUUUGUCCCUUAUCACAUACAGCUCAUAAACCCUCUUAUUAUUAUACUACAAU UGAAUAGCAAGAAGCCACCAUGGAGGACGCCAAGAACAUCAGAAGGGCCCGGCGCCUUUCUACCCG CUGGAGGACGGGACCGCCGCGGAGCAGUCCACAAGGCCAUGAAGCGGUACGCCUGGUGCCGGGCA CGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUACCUACGCGGAGUACUUCGAGAUAGAGCGU GCGCUGGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAACCACCGGAUCGUGGUGUGUCGGAG AACAGCCUGCAGUUCUUCUAGCGCGGUGCUGGGCGCCUUCUUCUUCGCGUGGCCGUCGCCCGGCGA ACGACAUCUACAACGAGCGGGAGCUGCUGAACGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGU GAGCAAGAAGGGCCUGCAGAAGAUCUGAACGUGCAGAAGAAGCUGCCAUCAUCCAGAAGAUAUC AUCAUGGACAGCAAGACCGACUACCAAGGGCUUCCAGUUCGAUGUACACGUUCGUGACCAGCCACCUCC CGCCGGGCUUCAACGAGUACGACUUCUGUCCGGAGAGCUUCGACCGGGACAAGACCAUCCGCCUGAU CAUGAACAGCAGCGGACGACCCGCCUGGCCAAGGGGGUGGGCCUGCCGACCCGACCGCCUGCGUG CGCUUCUCGCACGCCCGGGACCCCAUUCUGGCCAACAGAUCAUCCGGACACCGCCAUCCUGAGCG UGGUGCCGCUUCCACACGCGUUCGGCAUGUUACGACCCUGGGCUACCUCAUCUGCGGCUUCCGGGU GGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCGGAGCCUGCAGGACUACAAGAUCAGAGC GGCGUGCUGCGGCCACCCUGUUCAGCUUCUUCGCCAAGAGCACCCUGAUCGACAAGUACGACCUUGU CGAACCUUGCAGAGAUCCGACGCGGGGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAA GCGGUUCCACCUUCCCGGCAUCCGCCAGGGCUACGGCCUGACCGAGACCAGAGCGCGGAUCCUGAUC ACCCCCGAGGGGACGACAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGG UGGACCUUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCCGAU GAUCAUGAGCGGCUACGUGAACACCCGGAGGCCACCAACGCCCUCAUCGACAAGGACGGCUGCGUG CACAGCGGCGCAUCGCGUACUGGGACGAGGACGAGCACUUCUUCUACUGGUCGACCGGCUAAGUCGC UGAUAAGUACAAGGGCUACCAAGGUGGCGCGCGGCGAGCUGGAGAGCAUCCUGUCCAGCACCCCAA CAUCUUCGACGCGGCGUGGCCGGGCGGCCGAGCAGCAGCGCGGCGAGCUGCCGGCCGCGGUGGUG GUGCUGGAGCACGGCAAGACCAUGACGGAGAAGGAGAUUCGUGACUACGUGGCCAGCCAGGUGACCA CCGCCAAGAAGCUGCGGGGCGGCGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCU CGACGCCCGGAAGAUCGCGGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGAGCUGA UAAGGGCU	Scytonema- hofmani tRNA fMet gene GROUP I INTRON-PIE-6 CAV2- Luciferase circRNA
6 0 0	CCGAUUGUUGUAGUCAUCUAUUG	Ligand Intron
6 0 1	UCAACGGAUUUUAGUCCGUAGCGUCU	Ligand Feature
6 0 2	AUUUUAGUC	Ligand F10
6 0 3	AACGGAUUUUAGUCCGUAG	Ligand F20
6 0 4	CAACGGAUUUUAGUCCGUAGCG	Ligand F23
6 0 5	UCAACGGAUUUUAGUCCGUAGCGU	Ligand F25

6 0 6	UCAACGGAUUUUAGUCCGUAGCGUCU	Ligand F27
6 0 7	AUCAACGGAUUUUAGUCCGUAGCGUCU	Ligand F29
6 0 8	UGCGAGGAACCAUCCGGCCAGCUCUG	Ligand- Feature/liga se
6 0 9	GAUAGUCGUUAGGCAUUUAUGUAGA	Ligand-Td- Intron
6 1 0	UAUUAAACGGUAGACCCCAAGAAAACAU	Ligand-Td- Feature
6 1 1	GGGAGACCCUCGAGGCCCCUCUGCAGUGUCUGUUGUGGAGGCCAGAUACGUAAACUGCUUCCGCAACA AUAGAUGACUUAACAUAUUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAGAGACGAG GGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAACCC CCCCACCCACAGCAAGAAUGCCAUCAUCUGUCCUACCCCCAAUUUCCCUUUUCUCCUCCUGCAAC CAUUAACGUUACUCGCAUGGCAUUGAGUGGUGCAUGUGUUGAACAAACAGCUACACUCACAUGGGG GCGGGUUUCCCGCCCUACGGCCUCUCGCGAGGCCACCCUUCUCCUCCCUUAUAACUACAGUGCU UUGGUAGGUAAGCAUCCUGAUCCCCCGGGAAGCUCACGUGGCAACUGUGGGACCCAGACAGG UUAUCAAGGACCCCGGUCUUCGCGCUUCAGGAGUAUCCCUACUAGUGAAUUCUAGCGGGGUCUG CUUGGUGCCAAACUCCCCCAAUUGCGCGCUGCGGGAGUGUCUUCUCCCAACUACCCUAGUAUCCUC UCAUGUGUGUCUUGGUCAGCAUUCUGAGACGAUUGUCCGUGUCCAGACCAGUCCAGUAAUGGA CGGGCCAGUGCGUGUAGUCGUCUUCGCGCUUGUCCGGGGCAUGUUUGGUGAACCGGUGGGGUAAGGU UGGUGUGCCCAACGCGCGUACUUGGUGACACCUCAGACCCAGGAAUGCCAGGGAGGUACCCC ACUCUACGGUGGGUACUGACCCUGGGCUAAUUGUCUACGGUGGUUUCUUCUUGCUUCCACUUCUUCU UCUGUUCACGGCCACCAUGGAGGACGCCAAGAACAUAAGAAGGGCCCGCGCCCUUCUACCCGCG GAGGACGGGACCGCGCGAGCAGCUCACCAAGGCCAUGAAGCGGUACGCCCUGGUGCCGGGACGA UCGCCUUCACGACGCCACAUCGAGGUCGACAUCACCUACGCGGAGUACUUCGAGAUAGCGUGCG CCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAACCACCGGAUCGUGGUGUCUCGGAGAAC AGCCUGCAGUUCUUAUGCCGGUGUCUGGGCGCCUCUUAUCGCGGCGGCGUCCCGCGCGCAACG ACAUCUACAACGAGCGGGAGCUGCUGAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAG CAAGAAGGGCCUGCAGAGAUAUCUGAACGUGCAGAGAAGCUGCCCAUACUCCAGAGAUAUCAUAC AUGGACAGCAAGACCGACUACCGGGCUUCCAGUCGAUGUACAGUUCGUGACCGACCCACUCCCGC CGGGCUUACAAGAGUACGACUUCGUCUCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUUAUCAU GAACAGCAGCGGCAGCACCGGCCUGGCCAAGGGGUGGCCUUGCCGACCGGACCGCCUUCGCGUGCG UUCUCGCACGCGCGGACCCCAUCUUCGGCAACCAGAUACUCCGGACACCGCAUCCUGAGCGUGG UGCCGUUCCACACCGCUUCGGCAUGUUCACGACCCUGGGCUACCUCAUCUGCGGCUUCCGGGUGU CCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCGGAGCCUGCAGGACUACAAGAUCCAGAGCGG CUGCUCUGCCGACCCUUGUUCAGCUUUCUCCCAAGAGCACCCUGAUACGACAAGUACGACCUUGCGA ACUCGACGAGAUACGCCAGCGGGGGCGCCCGCUGAGCAAGGAGGUGGGCAGGGCCGUGGCCAAGCG GUUCCACCUCCCGGGCAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUACCC CCCGAGGGGACGACAAGCGGGCGCCUGGGCAAGGUGUCCCGUUCUUCGAGGCCAAGGUGGUGG ACCUGGACACCGCAAGACCCUGGGCGUAACCAGCGGGCGAGCUGUGCGUGCGGGGGCCGAUGAU CAUGAGCGGCUACGUGAACACCCGGAGGCCACCAACGCCUUAUCGACAAGGACGGCUGGCGUGCAG AGCGGCACAUCCGCUACUGGGACGAGGACGACACUUCUUAUCGUCGACCGGCGUAGAGUGCGUGA UCAAGUACAAGGGUACACAGGUGGCGCGGGCCGAGCUGGAGAGCAUCCUGUCCAGCACCCCAACAU CUUCGACGCCGCGUGGGCCGGGCGCCGACGACGCGCGGAGCUGCCGCGCGCGGUGGUGGUG CUGGAGCACGGCAAGACCAUGACGGAGAAGGAGAUUGUCGACUACGUGGCCAGCCAGGUGACCACCG CCAAGAAGCUGCGGGCGCGGUGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGGAAGCUCGA CGCCCGGAAGAUCCGCGAGAUCCUGAUCAAGGCCAAGAGGGCGGCAAGUCCGUGAGCUGAUAA CGUAAAUAUCACAAACACCAUUGUCACACUCCAACUUAUAUCCAGACGGGGGGGCUAAAUAUUAUGA GCCUUAAGAAGAAAUUCUUUAAGUGGAGUCUCUCAAACUACAGGAAACCUAAAUCUAGUUAUAGAC AAGGCAAUCCUGAGCCAAGCCGAAGUACUAAUAGUAGCGGAAGCAGUUAACGUUACUGGCCUCCAC AACGACACUCGAGAGGGGCCAUACAGCAUAUCUAG	circRNA-1X miRNA-122-5P
6 1 2	GGGAGACCCUCGAGGCCCCUCUGCAGUGUCUGUUGUGGAGGCCAGAUACGUAAACUGCUUCCGCAACA AUAGAUGACUUAACAUAUUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAGAGACGAG GGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAACCC CCCCACCCACAGCAAGAAUGCCAUCAUCUGUCCUACCCCCAAUUUCCCUUUUCUCCUCCUGCAAC CAUUAACGUUACUCGCAUGGCAUUGAGUGGUGCAUGUGUUGAACAAACAGCUACACUCACAUGGGG GCGGGUUUCCCGCCCUACGGCCUCUCGCGAGGCCACCCUUCUCCUCCCUUAUAACUACAGUGCU UUGGUAGGUAAGCAUCCUGAUCCCCCGGGAAGCUCACGUGGCAACUGUGGGACCCAGACAGG UUAUCAAGGACCCCGGUCUUCGCGCUUCAGGAGUAUCCCUACUAGUGAAUUCUAGCGGGGUCUG CUUGGUGCCAACUCCCCCAAUUGCGCGCUGCGGGAGUGUCUUCUCCCAACUACCCUAGUAUCCUC UCAUGUGUGUCUUGGUCAGCAUUCUGAGACGAUUGUCCGUGUCCAGACCAGUCCAGUAAUGGA CGGGCCAGUGCGUGUAGUCGUCUUCGCGCUUUCGCGGGCAUGUUUGGUGAACCGGUGGGGUAAGGU	circRNA-2X miRNA-122-5P

	UGGUGUGCCCAACGCCCGUACUUUGGUGACACCUCAGACCACCCAGGAAUGCCAGGGAGGUACCCC ACCUCACGGUGGGGAUCUGACCCUGGGCUAAUUGUCUACGGUGGUUCUUCUGCUUCCACUUCUUUCU UCUGUUCACGGCCACCAUGGAGGACGCCAAGAACAUCAGAAAGGGCCCGGCCCCUUCUACCCCGCUG GAGGACGGGACCGCCGGCAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCUUGGUGCCGGGCACGA UCGCCUUCACCGACGCCACAUCGAGGUGGACAUACCUACGCGGAGUACUUCGAGAUAGCGUGCG CCUGGCCAGGCCAUGAAGCGGUACGGCCUGAACACCAACCACCGAUCUGUGUGUCUGGAGAAC AGCCUGCAGUUCUUCAGCCGGUGUCUGGGCGCCCUUCUACUGCGCGUGGCCGUCGCCCGCGCAACG ACAUCUACAACGAGCGGGAGCUGCUGAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAG CAAGAAGGGCCUGCAGAAGAUCUGAACGUGCAGAAGAAGCUGCCCAUACUCCAGAAGAUAUCAUC AUGGACAGCAAGACCGACUACACGGGCUUCCAGUCCAGUACACGUUCGUGACCGACCCUCCCGC CGGGCUUCAAACGAGUACGACUUCGUCGCCGAGAGCUUCGACCGGGACCAAGACCAUCGCCUGAUCAU GAACAGCAGCGCGACACCGGCCUGCCGAAGGGGUGGCCUUGCCGACCCGACCGCCUUGCGUGCGC UUCUCGACGCGCCGGGACCCAUUCUGGCAACAGAUACUCCGGACACCGCCAUCCUGAGCGUGG UGCCGUUCCACACGGCUUCGGCAUGUUCACGACCCUGGGCUACCUAUCUGCGGGCUUCCGGGUGGU CCUGAUUACCGGUUCGAGGAGGAGCUGUUCUGCGGAGCCUGCAGGACUACAAGAUCAGAGCGCG CUGCUCUGGCGACCCUGUUCAGCUUCUUCGCCAAGAGCACCUGAUCGACAAGUACGACCUUGCGA ACCUGCAGGAGAUCCGACGCGGGGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCUGGCCAAGCG GUUCCACCUCCCGGCAUCGCGCAGGGCUACGGCCUGACCGAGACCAGAGCGCGAUCCUGAUCACCC CCGGAGGGGACGACAAGCCGGGCGCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGG ACCUGGACACCGGCAAGACCCUGGGCGUGAACCGCGGGCGAGCUGUGCGUGCGGGGGCCGAUGAU CAUGAGCGGCUACGUGAACACCCGGAGGCCACCAACGCCCUCAUCGACAAGGACGGCUGGCGUCAC AGCGGCCACAUCCGCUACUGGGAGCAGGACGAGCAGUUCUUAUCGUGCAGCGGCCUGAAGUCGUGA UCAAGUACAAGGGCUACCGGUGGCGCCGGCCGAGCUGGAGAGCAUCCUGCUCACGACCCCAACAU CUUCGACGCGCGCGUGGCGGGCGUGCCGACGACGACGCGCGGAGCUGCCGCGCGCGUGGUGGUG CUGGAGCAGCGCAAGACCAUGACGGAGAAGGAGAUUCGUGACUACGUGGCGAGCCAGGUGAGCCCG CCAAGAAGCUGCGGGGGCGCGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAGCUCGA CGCCCGGAAGAUCCGCGAGAUCCUGAUCAAGGCCAGAAGGGCGGCAAGAUCCCGUGAGCUGAUAA GAUUAUUUCGCAACACCAUUGUCACACUCCAACUAAAUCCCAACACCAUUGUCACACUCCAACU UAAAUCCAGACGGGGGGGGCUAAAUAUUGAGCCUUAAGAAGAAAUUCUUUAAGUGGAUGCUCUC AAACUCAGGGAACCUAAUUCUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAGUAGUAAUAG UAAGCGGAAGCAGUACGUUACUGGGCUCCACAACAGACACUGCAGAGGGGCCAUACAGCAUUCU AG	
6 1 3	GGGAGACCCUCGAGGCCCCUCUGCAGUGUCUUGUGGAGGCCAGAUACGUACUGCUUCCGCAACA AUAGAUGACUUAACUAAUUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAACGUCAAGACGAG GGUAAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAACCC CCCCACCCACAGCAAGAAGGCCAUCAUCUGUCCUACCCCCAAUUUUCCCUUUUCCUCCCCUGCAAC CAUUAACGCUUACUCGCAUGUGCAUUGAGUGGUGCAUGUGUUGAACAAACAGCUACACUACAUGGGG GCGGGUUUUCCCGCCUACGGCCUCUGCGAGGGCCACCCUUCUCCUCCUUAUAACUACAGUGCU UUGGUAGGUAAAGCAUCCUGAUCCCCCGGGAAGCUGCUCACGUGGCAACUGUGGGGACCCAGACAGG UUAUCAAAGGACCCGGUCUUUCCGCCUUCAGGAGUAUCCUACUAGUGAAUUCUAGCGGGGCUUCG CUUGGUGCCAACCUCCCCAAUUGCGCGCUGCGGGAGUGUCUUCUCCCAACUACCCUAGUAUCCUC UCAUGUGUGUGCUUGGUCAGCAUUCUGAGACGAUGUCCGCUUGUCCAGACACAGUCCAGUAAUGGA CGGGCCAGUGCGUGUAGUCGUCUUCCGGCUUGUCCGGGCAUGUUUGUGAACCUGUGGGGUAAGGU UGGUGUGCCCAACGCCCGUACUUUGGUGACACCUCAGACCACCCAGGAAUGCCAGGGAGGUACCCC ACCUCACGGUGGGAUUCGACCCUGGGCUAAUUGUCUACGUGGUUUCUUCUGCUUCCACUUCUUCU UCUGUUCACGGCCACCAUGGAGGACGCCAAGAACAUCAGAAGGGCCCGCGCCCUUCUACCCCGCUG GAGGACGGGACCGCCGCGAGCAGCUCCACAGGCCAUGAAGCGGUACGCCUUGGUGCGGGGACGA UCGCCUUCACCGACGCCACAUCGAGGUCGACAUACCUACGCGGAGUACUUCGAGAUAGCGUGCG CCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAACCAGGAUCGUGGUGUGCUCGGAGAAC AGCCUGCAGUUCUUCAGCCGGUGCUGGGCGCCCUUCUACUGGGCGUGGCCGUCGCCCGCGGCAACG ACAUCUACAACGAGCGGGAGCUGUGAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCUGGAG CAAGAAGGGCCUGCAGAAGAUCUGAACGUGCAGAGAAGCUGCCAUCAUCCAGAAGAUAUCAUC AUGGACAGCAAGACCGACUACACGGGCUUCCAGUCGAUGUACAGUUCGUGACCGACCCUCCCGC CGGGCUUCAAACGAGUACGACUUCGUCGCCGAGAGCUGACCGGGGACAGACCAUCGCCUGAUCAU GAACAGCAGCGCGACGACCGGCCUGCCGAAGGGGUGGCCUUGCCGACCGGACCGCCUUGCGUGCG UUCUCGACGCGCCCGGACCCCAUUCUGGCAACCAAGAUACUCCGGACACCGCCAUCCUGAGCGUGG UGCCGUUCCACACCGCUUCGGCAUGUUCACGACCCUGGGCUACCUAUCUGCGGCUUCCGGGUGGU CCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCGGAGCCUGCAGGACUACAAGAUCAGAGCGCG CUGCUCGUGCCGACCCUGUUCAGCUUCUUCGCCAAGAGCACCCUGAUCGACAAGUACGACCCUGCGA ACCUGCAGGAGAUCCGACGGGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCUGGCCAAGCG GUUCCACCUCCGGGCAUCCGCCAGGGCUACGGCCUGACCGAGACACGAGCGCGAUCCUGAUCACC CCGAGGGGGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGG ACCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCCGAUGAU CAUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCUUCUACGACAAGGACCGCUGGUGCAC AGCGGCGACAUCGCCUACUGGGACGAGGACGAGCAGUUCUUAUCGUGCAGCGGCGUGAAGUCGCUA UCAAGUACAAGGGCUACCGAGGUGCGCGCGGCGAGCUGGAGAGCAUCCUGUCCAGCACCCCAACAU CUUCGACGCCGCGUGGCGGGCGUGCCGGACGACGACGCGCGGAGCUGCGCGCGCGGUGGUGGUG CUGGAGCACGGCAAGACCAUGACGGAGAAGGAGAUUCGUGCAGUACGUGGCCAGCCAGGUGACACCG	circRNA-1X miRNA-671-5P

244

	UGCUGACUACAACUACAACUGCCUGAUGACUUCACAGGCGUGUGAUUGCCUGGAACAGCAACAC CUGGACAGCAAGGUGGGAGGCAACUACAACUACCUCUACAGACUGUUCAGGAAGAGCAACUGAAAC CAUUGAGAGGGACAUACGACAGAGAUUACCAGGCGUGGCAGCACACCAUGUAUUGGAGUGGAGGG CUUACAUGUUAUUCCACUCCAAUCCUAUGGCUUCCAACCAACCAUUGGAGUGGGCUACCAACCA UACAGGGUGUGUGUGUCUCCUUUGAACUGCUCCAUGCCCCUGCCACAGUGGAGCCUAGAGCAGCG ACAAGACCCACAUAGCCUCCAUUGUCCUGCCAGAACUGCUCGGAGGCCUUCUGUGUUCUGUU UCCUCCAAGGCCUAGGACACCCUGAUGAUCUCCAGAACACCCGAAGUGACCCUGGUGGUGGUGCAU GUGUCUACAGGAGCCUGAAGUGAAGUUAUUGGUACGUCGACGGCGUGGAAGUGCACACAGCCA AGACCAAGCCUAGAGAGGAACAGUACAACAGCACCUCACAGAGUGGUGUCCGUGCUGACCGUGCUGCA CCAGGAUUGGCGUAAACGGCAAGAGUACAAGUGCAAGGUGUCCAACAAGGCCUGCCAGCUCCUUAUC GAGAAAACCAUCAGCAAGGCCAAGGGCCAGCCUAGGGAACCCAGGUUUAACACACUGCCUCCAAGCC GGAAGAGAUAGCCAGAACACAGGUGUCCUGACCUGCCUGGUCAAGGGCUUUAUCCCUUCCGAUUAU CGCCUGGAAUUGGGAGAGCAAUGGCCAGCCUGAGAACAACUACAAGACCACACCUCUGUGCUGGAC AGCGACGGCUCAUUCUCCUGUACAGCAAGCUGACAGUGGACAAGAGCAGAUGGCAGCAGGGCAAGC UUCUACGUGCGAGCUGAUGCAGGCGCCUGGCACAACCACUACACCCAGAAGUCCUGAGCCUGGUC UCCUGGCAAAUGAUAAAGACGGGGGGGCUUAAUAAUUGAGCCUUAAGAAGAAAUUCUUUAAGUG GAGUCUCUACAACUCAGGGAACCUAAUUCUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAAGU AGUAAUAGUAAGACCAGUGGACAUCGACGGGACCAGUGGACAUCGACGGUAACAGCAUAUCUA G	
6 1 6	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUGCCGUCGAUUGUCCACUGGUAACAAUAGAUGACU UACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUAAGACGAGGGUAAAGAGA GAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAACCCCCCAAAAA AAAAAACAACAAAAAACAACAAAAAUAUUAUAGCUAACCCACAGCAAGAAUGCCAUCAU CUGUCCUACACCCCAUUAUUCUUCUCCUCCUGCAACCAUUAACGCUUACUCGCAUGUGCAUUGA GUGGUGCAUGUGUUAACAACACUACUACUACUACUACUACUACUACUACUACUACUACUACUACU GCGAGGCCACCCUCCUCCUCCUUAUAACUACAGUGCUUUGUAGUAAGCAUCCUGAUCUCCCG CGAAGCUGCUCACGUGGCAACUGUGGGGACCCAGACAGGUUAUCAAAGGCACCCGGUUCUUCGGC UUCAGGAGUAUCCUACUAGUGAAUUCUAGCGGGGCUUCGCUUGGUGCCAACCUCCCCCAAUGCGC GCUGCGGAGUGCUCUCCCAACUCACCCUAGUAUCCUCUACUAGUGUGCUGUGGUCAGCAUAUCU GAGACGUGUUCUGCUGUCCAGACAGUCCAGUAUUGGACGGGCCAGUGCGUGUAGUCGUCUCCG GCUUGUCCGGGGCAUGUUGGUGAACCUGGUGGGUAAAGGUUGUGUGGCCAACGCCCGUACUUGGU GACACCUCAAGACCACCCAGGAUUGCCAGGGAGGUACCCACCUCACGGUGGGAUUCGACCCUGGGC UAUUGUCUACGGUGGUUCUUCUUGCUUCCAUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCU AGCCCAUGUUCUACUGUGGGGCGUGCUGCUUUGGUGGCCUGGCGUAAAUGCAACUACCAAC CUGUGUCCAUUUGGAGAGGUGUUAUUGCCACCCAGCUUUGCCUUCUGCUAUGCCUGGAACAGGAAG GGAUUAAGCAACUGUGUGGUGACUACUACUACUACUACUACUACUACUACUACUACUACUACUACU UUAUGGAGUGAGCCCAACCAACUGAAGUACUGUGUUAUACCAUUGUUAUGUACUACUACUACU AUUAGGGGAGUGAGGUGAGACAGAUUGCCCGUGGACAAACAGGCAAGAUUGCUGACUACUACUACA AACUGCCUGAUGACUACAGGCGUGUGAUGGCCUGGAACAGCAACAACUGGACAGCAAGGUGGG AGGCAACUACACUACCUUACAGACUGUUCAGGAAGAGCAACCUGAACCAUUAUGAGAGGGACAU AGCACAGAGAUUUAACAGGCGUGGCAGCACACCAUGUAUUGGAGUGGAGGGCUUACUUGUUAUUC GACUCCAAUCCUUAUGGCUUCCAAACCAACCAUUGGAGUGGGCUACCAACCAUACAGGGUGGUGGCU GUCCUUGAAGUGCUCCAUUGCCCGUGCCACAGUGGAGCCUAAGAGCAGCGACAAGACCCACACAUGC CCUCCAUGUCCUGCUCCAGAACUGCUCGGAGGCCUUCUGUGUUCUUGUUCUCCAAAGCCUAAAGG ACACCCUGAUGAUCUCCAGAACACCCGAAGUGACCUUGCUGGUGGUGCAUGUGUUCACAGGAGCC UGAAGUGAAGUUAUUGGUACGUCGACGGCGUGGAAGUGCACAACGCCAAGACCAAGCCUAGAGAG GAACAGUACAACAGCACCUCACAGAGUGGUGUCCGUGCUGACCUGUGCAGCAGGAUUGGCUAAGC GCAAGAGUACAAGUGCAAGGUGUCCAACAGGCCUGCCAGCUCCUUAUCGAGAAAACCAUCAGCAA GGCCAAAGGGCCAGCCUAGGGAACCCAGGUUAACACACUGCCUCCAAAGCCGGGAAGAGAUAGCCAA AACCAGGUGUCCUGACCUUGGUGUACAGGCUUUAUCCUCCGAUUAUCCCGUGGAUUGGGAGA GCAUUGGCCAGCCUGAGAAACAACUACAAGACCACACCUCCUGUGGACAGCGACGGCUCAUUCU CCUGUACAGCAAGCUGACAGUGGACAAGAGCAGAUUGGCAGCAGGGCAACGUGUACGUGCAGCGUG AUGCACGAGGGCCUGCACAAACUACACCCAGAGUCCUGAGCCUGUCUCCUGGCAAAUGAUAAA GACGGGGGGGCUUAAUAAUUGAGCCUUAAGAGAGAAUUCUUUAAGUGGAGUCUCUACAACUCAG GGAAACUAAUACUAGUUAUAGACAAGGCAUCCUGAGCCAAGCCGAAGUAGUAAUUAAGUAGACCA GUGGACAAUACAGGGACAGUGGACAUCGACGGGAUACAGCAUAUCUAG	circRBD-UTR-1
6 1 7	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUGCCGUCGAUUGUCCACUGGUAACAAUAGAUGACU UACAACUAAUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUAAGACGAGGGUAAAGAGA GAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAACCCCCAGGGAA AUAGAGAGAAAAGAGUAAGAAAGAAUUAAGACCCCGGCCCCCCACAGCAAGAAUGCCAUCA UCUGUCCUACACCCCAUUAUUCUUCUCCUCCUGCAACCAUUAACGCUUACUCGCAUGUGCAUUG AGUGGUGCAUGUGUUAACAACAGCUACACUACUACUACUACUACUACUACUACUACUACUACUACU CGCGAGGCCACCCUCCUCCUCCUUAUAACUACAGUGCUUUGGUAGGUAAAGCAUCCUGAUCUCC GCGGAAGCUGCUCACGUGGCAACUGUGGGGACCCAGACAGGUUAUCAAAGGCACCCGGUCUUCGGC CUUACAGGUAUCCUACUAGUAAUUCUAGCGGGGCUUCGCUUGGUGCCAACCUCCCCAAUUGCG CGCUGCGGGAGUGCUUCCCAACUCACCCUAGUAUCCUCUACUGUGUGGUGUUGGUCAGCAUAUC UGAGACGAUGUCCGUGUCCAGACCAGUCCAGUAUUGGACGGGCCAGUGCGUGAUGUCUUGCC	circRBD-UTR-2

	GGCUUGUCCGGGCAUGUUUGGUGAACCGUGGGGUAAGGUUGGUGGCCAACGCCCGUACUUUGG UGACACCUCUAGACCACCCAGGAUUGCCAGGGAGGUACCCACCUCACGGUGGGAUCUGACCCUGGG CUAAUUGUCUACGGUGGUUCUUCUUGCUUCCACUUCUUCUUCUGUUCACGGCCACCAUGGAUAUGA GAGCCCAUGUUAUCUGCUGGGCCUGCUGCUGCUUUGGCGUCCUGGCGCUAAAUGCAACAUCACCAA CCUGUGUCCAUUUGGAGAGGUGUUAUUGCCACCAGGUUUGCCUCUGUCUAUGCCUGGAACAGGAAG AGGAUUAAGCAACUGUGUGCUGACUACUCUGUGCUUACAACUCUGCCUCCUUCAGCACCUCUAGU GUUAUGGAGUGAGCCCAACCAACUGAAUGACCUGUUAUACCAUUGUCUAUGCUGACUCCUUUGU GAUUAAGGGGAGAUGAGGUGAGACAGAUUGCCCGGACAAACAGGCAAGAUUGCUGACUACAACUAC AAACUGCCUGAUGACUUCACAGGCGUGUGAUUGCCUGGAACAGCAACAACUGGACAGCAAGGUGG GAGGCAACUACAACUACCUCUACAGACUGUUCAGGAAGAGCAACCUGAAACCAUUUGAGAGGGACA CAGCACAGAGAUUUAACAGGCGGCGAGCACACCAUGUAAUGGAGUGGAGGGCUUACACUGUUAUUU CCACUCCAAUCCUAUGGCUUCCAAACCAACCAUUGGAGUGGGCUACCAACCAUACAGGGUGGUGGUC UGUCCUUUGAACUGUCCUAGGCCUGCCACAGUGGAGCCUAAAGAGCAGCGACAAGACCCACACAU CCUCCUAGUCCUGUCCAGAACUGCUGGAGGGCCUUCGUGUUCUGUUCUCCUCAAAGCCUAAAG GACACCUGAUGAUUCCAGAACACCCGAAGUGACCUGCGUGGUGUGUGCAUGUGUCUACAGGACC CUGAAGUGAAGUUAUUGGUACGUGCAGCGGCGUGGAAGUGCACAAACGCCAAGACCAAGCCUAGAGA GGAACAGUACACAGCACCUCACAGAGUGGUGUCCGUGCUGACCGUGCAGCAGGAGUUGGCGUAC GGCAAAGAGUACAAGUGCAAGGUGUCCAAACAGGCCUGCCAGCUCCUACUGAGAAAACCAUCAGCA AGGCCAAGGGCCAGCCUAGGGAACCCAGGUUUAACACUGCCUCCAAAGCCGGGAAGAGAUAGACAA GACACAGGUGUCCUGACUGCCUGGUCAAGGGCUUCUACCCUUCGGAUUAUGCCGUGGAUUGGAG AGCAAUGGCCAGCCUGAGAACAACUACAAGACCACACCUCUGUGCUGGACAGCGACGGCUCAUUCU UCCUGUACAGCAAGCUGACAGUGGACAAGAGCAGAUUGCCAGCAGGGCAACGUGUUCAGCUGCAGCG GAUGCAGGAGGCCUGCACAAACCAUACACCCAGAAGUCCUGAGCCUGUCUCCUGGCAAAUGAUAA AGACGGGGGGGCUUAAAUAAUUGAGCCUAAAGAGAGAAUUCUUUAAUGUGGAUGCUCUCAAACUCA GGGAAACCUAAAUUAUAGUUAUAGACAAGGCAAUCCUGAGCCAAGCCGAAGUAUUAUAGUAGACC AGUGGACAAUCGACGGGACCAGUGGACAAUCGACGGUAACAGCAUAUCUAG	
6 1 8	GGGAGACCCUGACCGUCGAUUGUCCACUGGUCCCCGUCGAUUGUCCACUGGUCACAAUAGAUGACU UACAACUAAUCCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAAACGUCAAGACGAGGGUAAAGAGA GAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAACCCCCCAGAGAA UAAACUAGUAUUCUUGUGUCCCCACAGACUCAGAGAGAACCCCCACAGCAAGAAUGCCAUCAUCU GUCCUACCCCCCAAUUUUCCUUCUUCUCCUGCAACCAUUAACGCUUACUGCAUGUGCAUUGGAG GGUGCAUGUGUUGAACAAACAGCUACACUACAUGGGGGCGGGUUUUCCCGCCCUACGGCCUUCGCG GAGGCCACCCCUUCCUCCUUAUUAACUACAGUCUUUGGUAGGUAAAGCAUCCUGAUUCCCGCG GAAGCUGCUCACGUGGCAACUGUGGGAGCCAGACAGGUUAUCAAGGCACCCGGUUCUCCGCCU CAGGAGUAUCCCUACUAGUGAAUUCUACGGGGGCUUCGCUUGGUGCCAAACUCCCCCAAUUGCGCG UGCGGGAGUGCUCUCCCAACUACCCUAGUAUCCUCUCAUGUGUGUGCUUGGUCAGCAUAUCUGA GACGAUGUUCGCGUUCUCCAGACCAGUCCAGUAAUGGACGGGCCAGUGCGUGUAGUGCUCUCCGCG UUGUCCGGGGCAUUGUUGGUGAACCGGUGGGGUAAGGUUGGUGGCCAACGCCCGUACUUGUGUGA CACCUCAGACACCACCGAAUUGCCAGGGAGGUACCCACCUCACGGUGGGAUCUGACCCUGGGCUA AUUGUCUACGGUGGUUCUUCUUGCUUCCACUUCUUCUUCUGUUCACGGCCACCAUGGAUAUGAGAG CCCAUGUUAUCUGCUGGGCCUGCUGCUGCUUUGGUGCCUGGCGCUAAAUGCAACAUCACCAACCU GUGUCCAUUUGGAGAGGUGUUAUUGCCACCAAGGUUUGCCUCUGUCUAUGCCUGGAACAGGAAGAG AUUAGCAACUGUGUGGUGACUACUCUGUGCUUACAACUCUGCCUCCUUCAGCACCUCUAGUGUU AUGGAGUGAGCCCAACCAACUGAAUAGACCUGUGUUAACCAUUGUCUAUGCUGACUCCUUGUGAU UAGGGGAGAUAGGUGAGACAGAUUGCCCGGACAAACAGGCAAGAUUGCUGACUACAACUACAAA CUGCCUGAUGACUUCACAGGCGUGUGAUUGCCUGGAACAGCAACAACCCUGGACAGCAAGGUGGGAG GCAACUACAACUACCUCUACAGACUGUUCAGGAAGAGCAACCUGAAACCAUUGAGAGGGGACAUCAG CACAGAGAUUUAACAGGCGUGGACGACACCAUGUAAUGGAGUGGAGGCUUACUUGUUAUCCUCC CUCCAAUCCUUAUGGCUUCCAAACCAACCAUUGGAGUGGGCUACCAACCAUACAGGGUGGUGGUGCUG CCUUGAACUGCUCUAGUCCCGGCCACAGUGGAGCCUAAAGAGCAGCGACAAGACCCACACAUGCCC UCCAUGUCCUGCUCAGAACUGCUGGAGGGCCUUCGUGUUCUGUUCUCCUCAAAGCCUAAAGGAC ACCCUGAUGAUUCCAGAACACCCGAAGUGACCUGCGUGGUGUGUGCAUGUGUCUACGAGGACCCUG AAGUGAAGUUAUUGGUACGUCGACGGCGUGGAAGUGCACAACGCCAAGACCAAGCCUAGAGAGGA ACAGUACAACAGCACCUCACAGAGUGGUCUCCUGCUGACCGUGCUGACCAAGAUUGGCGUACGGC AAAGAGUACAAGUGCAAGGUGUCCAAACAGGCCUCCAGCUCUUAUCGAGAAAAACCAUCAGCAAGG CCAGGGCCAGCCUAGGGAACCCAGGUUAACACACUGCCUCCAAAGCCGGGAAGAGAUAGCCAAAGAA CCAGGUGUCCUGACUGCCUGGUCAAGGGCUUCUACCCUUCGUAUUCGCCUGGAUUGGGAGAGC AAUGGCCAGCCUGAGAACAACUACAAGACCACACCUCUGUGCUGGACAGCGACGGCUCAUUCUCC UGUACAGCAAGCUGACAGUGGACAAGAGCAGAUUGGCAGCAGGGCAACGUGUUCAGCUGCAGCGUGAU GCACGAGGCCUGCACAAACCAUACACCCAGAAGUCCUGAGCCUGUCUCCUGGCAAAUGAUAAAGA CGGGGGGGGCUUAAAUAAUUGAGCCUAAAGAGAGAAUUCUUUAAUGUGGAUGCUCUCAAACUCAGGG AAACCUAAAUUAGUUAUAGACAAGGCAAUCCUGAGCCAAGCCGAAGUAUUAUAGUAAAGACCAGU GGACAAUCGACGGGACCAGUGGACAAUCGACGGUAACAGCAUAUCUAG	circRBD-UTR- 3
6 1 9	AAAAAAAAAAACCAAAAAAAAAAAACAAAAAAAAAAAUAAUUGACUAA	UTR-1

6 2 0	GGGAAAUAGAGAGAAAAGAAAGAGUAAGAAGAAAUUAAGACCCGGCGCCGCCACC	UTR-2
6 2 1	GAGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACC	UTR-3
6 2 2	CAAACACCAUUGUCACACUCCA	hsa-miR-122-5p
6 2 3	CUCCAGCCCCUCCAGGGCUUCCU	hsa-miR-671-5p
6 2 4	AAAUAGAGAGAAAAGAAAGAGUAAGAAGAAAUUAAGAGCCACCAUGGAGGACGCCAAGAACAUCAA GAAGGGCCCGCGCCCUUCUACCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUG AAGCGGUACGCCUUGGUGCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUACCU ACGCGGAGUACUUCGAGAUAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCUUGCCGGUGCUGGGCGCCUUCUUC AUCGGCGUGGCCGUCGCCCGGGCGAACGACAUCAACAGAGCGGGAGCUGUGAACAGCAUGGGGA UCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAAGAUCUGAACGUGCAGAAGAA GCUGCCCAUCAUCCAGAAGAUCAUCAUGGACAGCAAGACCGACUACCGAGGCCUCCAGUCGAUG UACACGUUCGUGACAGCCACCUCGCCCGGGCUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCG ACCGGGACAAGACCAUCGCCUGAUCAGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGUGGC CCUGCCGCACCGGACCGCCUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAU AUCCCGGACACCGCAUCCUGAGCGUGGUGCCGUUCCACACGGCUUCGGCAUGUUCACGACCCUGG GCUACCUCAUCUGCGGCUUCCGGGUGGUCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCGGAG CCUGCAGGACUACAAGAUCACAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCUUCGCCAAGAGC ACCUGAUCGACAAAGUACGACCUUGUCGAACCGCAGCAGAUCCGAGCGGGGGCGCCCGCUGAGCA AGGAGGUGGGCGAGGCCUGGCCAAGCGGUUCCACCUCCGGGCAUCCGCCAGGGCUACGGCCUGAC CGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGGGGACGACAAGCCGGGCGCCGUGGGCAAGGUG GUCCCGUUCUUCGAGGCCAAGGUGGUGGACCGGACACCGGCAAGACCCUGGGCGUGAACAGCGGG GCGAGCUGUGCGUGCGGGGCGGUAUGAUCAGCGGCUACGUAACAACCGGAGGCCACCAACGC CCUCAUCGACAAGGACGGCUGGCGACAGCGGGGACAUCCGUACUGGGACGAGGACGAGCACUUC UUCAUCGUCGACCGGCUAGAGUCGCUAUCAGUACAAGGCUACCGAGGUGGCGCGGCCGAGCUGG AGAGCAUCCUGUCAGCACCCCAACAUUCGACGCCGGCGUGGCCGGGCUUGCCGACGACGAGCGC CGGCGAGCUGCCGGCCGGUGGUGGUGGUGGAGCAGCGCAAGACCAUGACGGAGAGGAGAUUGUC GACUACGUGGCCAGCCAGGUGACCAACGCCAAGAAGCUGCGGGGCGGCGUGGUGUUCGUGGACGAGG UCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAUCGCGGAGAUCCUGAUCAGGCCAAGAA GGGCGGCAAGAUCCGGUGAGCUAAGCUGCCUUCUGCGGGGCUUGCCUUCUGGCCAUGCCCUUCUUC UCUCUUUGCACCUGUACCUUGGUCUUUGAAUAAAGCCUGAGUAGGAAGUAUCCCAAUGGGCGGC CGAGCUUGGCAA AA	Luciferase mRNA

CLAIMS

We claim:

1. A circular RNA precursor comprising from 5' to 3' direction:

- a) a 3' self-splicing intron fragment; and
- b) a 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA through the self-splicing of the circular RNA precursor, and

wherein the self-splicing intron is selected from Group I introns and Group II introns.

2. The circular RNA precursor of claim 1, wherein the self-splicing intron is selected from the following group: cyanobacterium *Anabaena* Group I Intron such as AnaX, *Azoarcus* Group I Intron (Azo), *Scytalidium dimidiatum* Group I Intron (Sd), *Staphylococcus* phage Twort Group I Intron (Twort), *Scytonema-hofmani* tRNA fMet group I intron (Sh), or *Agrobacterium-tumefaciens* group I intron (At), for example, the self-splicing intron has a nucleotide sequence selected from any one of SEQ ID NO: 1-6, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 1-6.

3. The circular RNA precursor of claim 1 or 2, wherein the 3' self-splicing intron fragment is the 3' terminal portion of the self-splicing intron starting from an intron internal split site I to the 3' end of the self-splicing intron, and the 5' self-splicing intron fragment is the 5' terminal portion of the self-splicing intron starting from the 5' end of the self-splicing intron to an intron internal split site II.

4. The circular RNA precursor of any one of claims 1-3, wherein the self-splicing intron is cyanobacterium *Anabaena* Group I Intron such as AnaX having a nucleotide sequence as set forth in SEQ ID NO. 1, and the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 115 and nucleotide A at position 116 from 5' end of SEQ ID NO. 1.

5. The circular RNA precursor of any one of claims 1-3, wherein the self-splicing intron is *Azoarcus* Group I Intron (Azo) which has a nucleotide sequence as set forth in SEQ ID NO. 2, and wherein

- 1) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 24 and nucleotide A at position 25 from the 5' end of SEQ ID NO. 2; or
- 2) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 85 and nucleotide A at position 86 from the 5' end of SEQ ID NO. 2; or

3) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide C at position 71 and nucleotide C at position 72 from the 5' end of SEQ ID NO. 2; or

4) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide C at position 108 and nucleotide U at position 109 from the 5' end of SEQ ID NO. 2; or

5) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 154 and nucleotide A at position 155 from the 5' end of SEQ ID NO. 2; or

6) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide C at position 198 and nucleotide A at position 199 from the 5' end of SEQ ID NO. 2.

6. The circular RNA precursor of any one of claims 1-3, wherein the self-splicing intron is Staphylococcus phage Twort Group I Intron (Twort) which has a nucleotide sequence as set forth in SEQ ID NO. 3, and

1) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide U at position 70 and nucleotide A at position 71 from the 5' end of SEQ ID NO. 3; or

2) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide U at position 103 and nucleotide A at position 104 from the 5' end of SEQ ID NO. 3; or

3) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 132 and nucleotide U at position 133 from the 5' end of SEQ ID NO. 3; or

4) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide U at position 173 and nucleotide C at position 174 from the 5' end of SEQ ID NO. 3; or

5) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 206 and nucleotide A at position 207 from the 5' end of SEQ ID NO. 3; or

6) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide G at position 226 and nucleotide U at position 227 from the 5' end of SEQ ID NO. 3; or

7) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide U at position 23 and nucleotide A at position 24 from the 5' end of SEQ ID NO. 3; or

8) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 82 and nucleotide A at position 83 from the 5' end of SEQ ID NO. 3.

7. The circular RNA precursor of any one of claims 1-3, wherein the self-splicing intron is Scytalidium dimidiatum Group I Intron (Sd) which has a nucleotide sequence as set forth in SEQ ID NO. 4, and wherein

1) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 88 and nucleotide G at position 89 from the 5' end of SEQ ID NO. 4; or

2) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide U at position 173 and nucleotide G at position 174 from the 5' end of SEQ ID NO. 4; or

3) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 204 and nucleotide G at position 205 from the 5' end of SEQ ID NO. 4; or

4) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide C at position 252 and nucleotide G at position 253 from the 5' end of SEQ ID NO. 4; or

5) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide C at position 370 and nucleotide G at position 371 from the 5' end of SEQ ID NO. 4.

8. The circular RNA precursor of any one of claims 1-3, wherein the self-splicing intron is *Scytonema-hofmani* tRNA fMet group I intron (Sh) which has a nucleotide sequence as set forth in SEQ ID NO. 5, and

1) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 129 and nucleotide A at position 130 from the 5' end of SEQ ID NO. 5; or

2) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 60 and nucleotide A at position 61 from the 5' end of SEQ ID NO. 5; or

3) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 109 and nucleotide A at position 110 from the 5' end of SEQ ID NO. 5; or

4) the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 213 and nucleotide A at position 214 from the 5' end of SEQ ID NO. 5.

9. The circular RNA precursor of any one of claims 1-4, wherein the self-splicing intron is *Agrobacterium-tumefaciens* group I intron (At), and the intron internal split site II and the intron internal split site I respectively correspond to nucleotide A at position 107 and nucleotide A at position 108 from the 5' end of the At intron having a nucleotide sequence as set forth in SEQ ID NO. 6.

10. The circular RNA precursor of claim 1, wherein the self-splicing intron is the cyanobacterium *Anabaena* Group I Intron such as AnaX, and the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 7 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100%

identity to SEQ ID NO: 7, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 8 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 8.

11. The circular RNA precursor of claim 1, wherein the self-splicing intron is the Azoarcus Group I Intron (Azo), and the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 9, 11, 13, 15, 17 or 19 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 9, 11, 13, 15, 17 or 19, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 10, 12, 14, 16, 18 or 20 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 10, 12, 14, 16, 18 or 20.

12. The circular RNA precursor of claim 1, wherein the self-splicing intron is the Scytalidium dimidiatum Group I Intron (Sd), and the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 37, 39, 41, 43 or 45 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 37, 39, 41, 43 or 45, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 38, 40, 42, 44 or 46 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 38, 40, 42, 44 or 46.

13. The circular RNA precursor of claim 1, wherein the self-splicing intron is the Staphylococcus phage Twort Group I Intron (Twort), and the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 21, 23, 25, 27, 29, 31, 33 or 35, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 22, 24, 26, 28, 30, 32, 34 or 36.

14. The circular RNA precursor of claim 1, wherein the self-splicing intron is the *Scytonema hofmani* tRNA fMet group I intron (Sh), and the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 47, 49, 51 or 53 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 47, 49, 51 or 53 and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 48, 50, 52 or 54 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 48, 50, 52 or 54.

15. The circular RNA precursor of claim 1, wherein the self-splicing intron is the *Agrobacterium tumefaciens* group I intron (At), and the 3' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 55 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 55, and the 5' self-splicing intron fragment comprises or consists of a nucleotide sequence of SEQ ID NO: 56 or a nucleotide sequence having at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% identity to SEQ ID NO: 56.

16. The circular RNA precursor of any one of claims 1-15, which further comprises from 5' to 3' direction: optionally a first residual circularizing element; a nucleotide sequence of interest; and optionally a second residual circularizing element, located between a) and b), wherein the total length of the first residual circularizing element and the second residual circularizing element is about 0 to about 500 nucleotides; and/or

the circular RNA precursor further comprises a homology arm sequence I upstream of the 3' self-splicing intron fragment and a homology arm sequence II downstream of the 5' self-splicing intron fragment.

17. The circular RNA precursor of claim 16, comprising from 5' to 3' direction:

- i) the 3' self-splicing intron fragment;
- ii) the first residual circularizing element;
- iii) the nucleotide sequence of interest;
- iv) the second residual circularizing element; and
- v) the 5' self-splicing intron fragment.

18. The circular RNA precursor of claim 16 or 17, wherein the first residual circularizing element comprises or consists of from 5' to 3' direction a 3' exon region and optionally a first spacer, the second residual circularizing element comprises or consists of from 3' to 5' direction a 5' exon region and optionally a second spacer.

19. The circular RNA precursor of claim 18, wherein the 3' exon region is derived from the native 3' exon of the self-splicing intron, the 5' exon region is derived from the native 5' exon of the self-splicing intron, and the 3' exon region and 5' exon region can be recognized and/or spliced by the self-splicing intron or a combination of the 3' self-splicing intron fragment and the 5' self-splicing intron fragment.

20. The circular RNA precursor of claim 19, wherein 3' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to the native 3' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 5' terminal nucleotide of the native 3' exon, and

the 5' exon region is a sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to the native 5' exon or a contiguous fragment of about 1- about 50 nucleotides starting from the 3' terminal nucleotide of the native 5' exon.

21. The circular RNA precursor of any one of claims 16-20, wherein the nucleotide sequence of interest comprises at least one protein-coding sequence and a translation initiation element such as an internal ribosome entry site (IRES) operably linked thereto.

22. The circular RNA precursor of claim 21, wherein the translation initiation element is an internal ribosome entry site (IRES), wherein said IRES is selected from the following IRES

sequences: Taura syndrome virus, blood-sucking bug virus, Tyler's encephalomyelitis virus, simian virus 40, red fire ant virus 1, cereal constrictor virus, reticulovirus Endothelial hyperplasia virus, Forman poliovirus 1, soybean inchworm virus, Kashmir bee virus, human rhinovirus 2, glass leafhopper virus-1, human immunodeficiency virus type 1, glass leafhopper virus-1, lice P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinovirus, Tea inchworm-like virus, Encephalomyocarditis virus (EMCV), Drosophila C virus, Cruciferous tobacco Virus, cricket paralysis virus, bovine viral diarrhea virus 1, black queen cell virus, aphid lethal paralysis virus, avian encephalomyelitis virus, acute bee paralysis virus, hibiscus yellow ring spot virus, swine fever virus, human FGF2, human SFTPA1, Human AML1/RUNX1, Drosophila Antennae, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 α , Human n.myc, mouse Gtx, human p27kipl, human PDGF2/c-sis, human p53, human Pim-1, mouse Rbm3, Drosophila reaper, canine Scamper, Drosophila Ubx, human UNR, mouse UtrA, human VEGF-A, human XIAP, Drosophila hairless, Saccharomyces cerevisiae TFIID, Saccharomyces cerevisiae YAP1, human c-src, human FGF-1, simian picornavirus, turnip crepe disease virus, eIF4G aptamer, Coxsackie Virus B3 (CVB3) or Coxsackie virus A (CVA1/2), preferably, the IRES is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2.

23. The circular RNA precursor of claim 22, wherein the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to one of SEQ ID NOs: 120-526.

24. The circular RNA precursor of any one of claims 21-23, wherein the at least one protein-coding sequence encodes a protein of eukaryotic, prokaryotic or viral origin, for example, a protein for therapeutic or diagnostic use;

for example, the at least one protein-coding sequence encodes IL-2, or an aquaporin, such as AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

25. The circular RNA precursor of any one of claims 21-24, wherein the at least one protein-coding sequence comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least

89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

26. The circular RNA precursor of any one of claims 16-23, wherein the nucleotide sequence of interest is a non-coding sequence, for example, the non-coding sequence can be an antisense RNA, aptamer, guide RNA, or non-protein-coding RNA existing in any organism.

27. The circular RNA precursor of any one of claims 1-15 comprising from 5' to 3' direction: the 3' self-splicing intron fragment; a nucleotide sequence of interest; and the 5' self-splicing intron fragment;

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA which is formed by the nucleotide sequence of interest through the self-splicing of the circular RNA precursor;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a circularizing sequence fragment II comprising a ribozyme recognition site II at its 5' end and a circularizing sequence fragment I comprising a ribozyme recognition site I at its 3' end,

wherein the circularizing sequence fragment I is the 5' terminal portion of a circularizing sequence, and the circularizing sequence fragment II is the 3' terminal portion of a circularizing sequence, and

wherein the ribozyme recognition site I is connected with the ribozyme recognition site II in the generated circular RNA, and thus the circularizing sequence fragment I is connected with the circularizing sequence fragment II to form a circularizing sequence.

28. The circular RNA precursor of claim 27, wherein the circularized nucleotide sequence of interest comprises one or more sequences selected from the following group: at least one protein coding sequences, a non-coding sequence, translation initiation element (TIE), a non-TIE functional element or any combination thereof.

29. The circular RNA precursor of claim 27 or 28, wherein the circularizing sequence is selected from at least one protein-coding sequence, a non-coding sequence, a translation initiation element (TIE), or a non-TIE functional element.

30. The circular RNA precursor of any one of claims 1-15, comprising from 5' to 3' direction: the 3' self-splicing intron fragment; a nucleotide sequence of interest; and the 5' self-splicing intron fragment,

wherein the 3' self-splicing intron fragment and the 5' self-splicing intron fragment in combination retain the self-splicing activity of a self-splicing intron and allow generation of a circular RNA which is formed by the nucleotide sequence of interest through the self-splicing of the circular RNA precursor and which comprises at least one protein-coding sequence;

wherein the nucleotide sequence of interest comprises from 5' to 3' direction: a protein-coding sequence fragment II, a translation initiation element such as an internal ribosome entry site (IRES), and a protein-coding sequence fragment I;

wherein the protein-coding sequence fragment I is the 5' terminal portion of the at least one protein-coding sequence, and the protein-coding sequence fragment II is the 3' terminal portion of the at least one protein-coding sequence, and

the protein-coding sequence fragment I comprises a ribozyme recognition site I located at its 3' end, the protein-coding sequence fragment II comprises a ribozyme recognition site II located at its 5' end, wherein in the generated circular RNA the ribozyme recognition site I is connected with the ribozyme recognition site II, and thus the protein-coding sequence fragment I is connected with the protein-coding sequence fragment II to form the at least one protein-coding sequence.

31. The circular RNA precursor of claim 30, wherein the generated circular RNA comprises at least two protein-coding sequences, and a further translation initiation element, 2A sequence or spacer may be linked between any two adjacent protein-coding sequences, and the translation initiation element comprises: IRES sequence, 5' UTR sequence, Kozak sequence, a sequence containing m6A, a sequence complementary to ribosome 18S rRNA, or any combination thereof.

32. The circular RNA precursor of any one of claims 29-31, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured such that the circular RNA precursor comprising them has a high circularization efficiency.

33. The circular RNA precursor of any one of claims 29-32, wherein the total length of the ribozyme recognition site I and the ribozyme recognition site II is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100 nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

34. The circular RNA precursor of any one of claims 29-33, wherein the total length of the ribozyme recognition site I is about 0 to about 100 nucleotides, about 0 to about 90 nucleotides, about 0 to about 8 nucleotides, about 0 to about 70 nucleotides, about 0 to about 60 nucleotides, about 0 to about 50 nucleotides, about 0 to about 40 nucleotides, about 0 to about 35 nucleotides, about 2 to about 100 nucleotides, about 5 to about 100 nucleotides, about 10 to about 100

nucleotides, about 20 to about 100 nucleotides, about 30 to about 100 nucleotides, or about 30 to about 100 nucleotides, such as 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, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides in length.

35. The circular RNA precursor of any one of claims 29-34, wherein the ribozyme recognition site I and the ribozyme recognition site II are configured to be capable of forming a stem-loop structure, e.g., upon self-splicing for circularization.

36. The circular RNA precursor of claim 35, wherein the loop of the stem-loop structure comprises the splicing junction.

37. The circular RNA precursor of any one of claims 29-36, wherein the ribozyme recognition site I comprises the sequence structure of the following formula: 5'-third non-pairing sequence-third pairing sequence-third loop sequence-3'; and the ribozyme recognition site II comprises the sequence structure of the following formula: 5'-fourth loop sequence-fourth pairing sequence-fourth non-pairing sequence-3',

wherein the third non-pairing sequence or the fourth non-pairing sequence may be independently present or absent,

the third pairing sequence and the fourth pairing sequence can complementarily pair to each other to form the stem of the stem-loop structure, wherein the third loop sequence and the fourth loop sequence can form the loop of the stem-loop structure, e.g., through self-splicing for circularization.

38. The circular RNA precursor of any one of claims 29-37, wherein the ribozyme recognition site II can be selected from various sequence.

39. The circular RNA precursor of claim 37 or 38, wherein the fourth loop sequence comprises or consists of one or more nucleotides which can pair with the P1 region of the corresponding self-splicing intron to form a P10 duplex region during the circularization.

40. The circular RNA precursor of any one of claims 37-39, wherein the fourth loop sequence consist of a nucleotide sequence of (N)_n, wherein N represents any nucleotides (A, G, U, or C), n represents an integer from 1-20, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

41. The circular RNA precursor of any one of claims 37-40, wherein the fourth loop sequence comprises or consists of the sequence of AAAA, AGAA, AA, UUUU, UAAA, CAAA, or GAAA, preferably AAAA.

42. The circular RNA precursor of any one of claims 37-41, wherein the ribozyme

recognition site I comprises a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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 100% sequence identity to a nucleotide sequence of any one of SEQ ID NO: 95-116.

43. The circular RNA precursor of any one of claims 37-42, wherein the third loop sequence comprises or consists of one or more nucleotides which can pair with the internal guide sequence (IGS) of the corresponding self-splicing intron to form a P1 duplex region during the circularization.

44. The circular RNA precursor of any one of claims 37-43, wherein the third loop sequence comprises or consists of a nucleotide sequence of any one of SEQ ID NO: 95-116, such as CUU, CAU or CUC, preferably CUU.

45. The circular RNA precursor of any one of claims 37-44, wherein the loop of the stem-loop structure has a sequence of CUUAAAA, CUUUUUU, CUUAA, CUUGAAA, CUUUAAA, CUUAGAA, CUUCAAA or CUCAAAA, preferably CUUAAAA.

46. The circular RNA precursor of any one of claims 29-45, wherein the translation initiation element is an internal ribosome entry site (IRES) which is selected from the following IRES sequences: Taura syndrome virus, blood-sucking bug virus, Tyler's encephalomyelitis virus, simian virus 40, red fire ant virus 1, cereal constrictor virus, reticulovirus Endothelial hyperplasia virus, Forman poliovirus 1, soybean inchworm virus, Kashmir bee virus, human rhinovirus 2, glass leafhopper virus-1, human immunodeficiency virus type 1, glass leafhopper virus-1, lice P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinovirus, Tea inchworm-like virus, Encephalomyocarditis virus (EMCV), Drosophila C virus, Cruciferous tobacco Virus, cricket paralysis virus, bovine viral diarrhea virus 1, black queen cell virus, aphid lethal paralysis virus, avian encephalomyelitis virus, acute bee paralysis virus, hibiscus yellow ring spot virus, swine fever virus, human FGF2, human SFTP1, Human AML1/RUNX1, Drosophila Antennae, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 α , Human n.myc, mouse Gtx, human p27kipl, human PDGF2/c-sis, human p53, human Pim-1, mouse Rbm3, Drosophila reaper, canine Scamper, Drosophila Ubx, human UNR, mouse UtrA, human VEGF-A, human XIAP, Drosophila hairless, Saccharomyces cerevisiae TFIID, Saccharomyces cerevisiae YAP1, human c-src, human FGF-1, simian picornavirus, turnip crepe disease virus, eIF4G aptamer, Coxsackie Virus B3 (CVB3) or Coxsackie virus A (CVA1/2), preferably, the IRES is selected from CVB3, HRV-B3, BRAV-1_L, PV1_L, CAV2_L, BRAV-1, PV1, or CAV2.

47. The circular RNA precursor of any one of claims 29-46, wherein the IRES comprises a nucleotide sequence set forth in one of SEQ ID NOs: 120-526, or comprises a nucleotide sequence having at least 75%, e.g., at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to one of SEQ ID NOs: 120-526.

48. The circular RNA precursor of any one of claims 29-47, wherein the at least one protein-coding sequence encodes a protein of eukaryotic, prokaryotic or viral origin, for example, a protein for therapeutic or diagnostic use, in particular a protein of mammalian or human origin.

49. The circular RNA precursor of any one of claims 29-48, wherein the at least one protein-coding sequence encodes IL-2, or an aquaporin.

50. The circular RNA precursor of claim 49, wherein the aquaporin is AQP-1 protein or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

51. The circular RNA precursor of any one of claims 29-50, wherein the at least one protein-coding sequence comprises a nucleotide sequence of any one of SEQ ID NO: 117-118, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 117-118.

52. The circular RNA precursor of any one of claims 29-51, which comprises a nucleotide sequence selected from any one of SEQ ID NO: 530, 531 and 533-534, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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% , or at least 99% sequence identity to any one of SEQ ID NO: 530, 531 and 533-534.

53. A nucleic acid vector for generating a circular RNA molecule, said vector comprises a coding sequence of the circular RNA precursor of any one of claims 1-52.

54. The nucleic acid vector of claim 53, which further comprises an RNA polymerase promoter sequence operably linked to the coding sequence of the circular RNA precursor.

55. The nucleic acid vector of claim 53 or 54, wherein the promoter is a T7 RNA polymerase promoter, a T6 viral RNA polymerase promoter, a SP6 viral RNA polymerase promoter, a T3 viral RNA polymerase promoter or a T4 viral RNA polymerase promoter, preferably a T7 RNA polymerase promoter.

56. A circular RNA, which is prepared from the circular RNA precursor of any one of claims 1-52, preferably said circular RNA has a duration of in vivo expression of at least 24 hours, for example, 36 hours, 48 hours, 72 hours, 96 hours, 5 days, 10 days, 15 days, 16 days, 17 days, 18 days, 19 days, 20 days or more.

57. The circular RNA of claim 56, which comprises or has a nucleotide sequence of any one of SEQ ID NO: 543-558, or a nucleotide sequence having at least 50%, at least 60%, at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity to any one of SEQ ID NO: 543-558.

58. Use of the circular RNA precursor of any one of claims 1-52 and/or the circular RNA of 56 or 57 as an expression vector.

59. A pharmaceutical composition comprising the circular RNA precursor of any one of claims 1-52 and/or the nucleic acid vector of any one of claims 53-55 and/or the circular RNA of claim 56 or 57, and a pharmaceutically acceptable carrier.

60. A product expressed by the circular RNA of claim 56 or 57.

61. The product of claim 60, which comprises IL-2, POLR2A, or an aquaporin, such as AQP-1 protein having a sequence as set forth in SEQ ID NO: 118 or a protein having the same activity as the AQP-1 protein and derived from the AQP-1 protein through substitution, deletion or addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in the amino acid sequence of the AQP-1 protein.

62. A method for preventing or treating a disease or condition in a subject, comprising administering to a subject the circular RNA of any one of claims 56-57, or the pharmaceutical composition of claim 59.

63. The method of claim 62, wherein the disease or condition comprises but not limited to xerostomia, xerophthalmia, Sjogren's syndrome or the like.

64. Use of the circular RNA of claim 56 or 57 or the pharmaceutical composition of claim 59 in the manufacture of a medicament for preventing or treating a disease or condition in a subject.

65. The use of claim 64, wherein the disease or condition is xerostomia, xerophthalmia, or Sjogren's syndrome.