



(51) International Patent Classification:

C12N 15/67 (2006.01) C12N 15/113 (2010.01)

(21) International Application Number:

PCT/EP2024/065837

(22) International Filing Date:

07 June 2024 (07.06.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2308675.4 09 June 2023 (09.06.2023) GB

(71) Applicant: UNITED KINGDOM RESEARCH AND IN-
NOVATION [GB/GB]; Polaris House, North Star Avenue,
Swindon SN2 1FL (GB).

(72) Inventors: DU, Yifei; c/o MRC Laboratory of Molecu-
lar Biology, Cambridge, Cambridgeshire, CN2 0QH (GB).
RAMAKRISHNAN, Venkatraman; c/o MRC Laborato-
ry of Molecular Biology, Cambridge, Cambridgeshire, CN2
0QH (GB).

(74) Agent: MEWBURN ELLIS LLP; Aurora Building, Coun-
terslip, Bristol BS1 6BX (GB).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, CV,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,
SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ,
RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ,
DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT,
LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,
SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,
GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished
upon receipt of that report (Rule 48.2(g))
- in black and white; the international application as filed
contained color or greyscale and is available for download
from PATENTSCOPE

(54) Title: CIRCULAR RNAS AND METHODS FOR MAKING THE SAME

(57) Abstract: This invention relates to recombinant nucleic acid molecule for making a circular RNA, comprising in the 5' to 3' direction: a) an internal guide sequence (IGS), b) a ribozyme, c) a first portion of an extended anticodon arm (eACA) sequence, d) gene of interest, and e) a second portion of the eACA sequence. A nucleotide in the second portion of the eACA sequence forms a wobble base pair with a nucleotide in the IGS. Recombinant nucleic acid molecules, circular RNA, and methods of producing and using circular RNA are provided.



CIRCULAR RNAS AND METHODS FOR MAKING THE SAME

This application claims priority from GB2308675.4 filed 09 June 2023, the contents and elements of which are herein incorporated by reference for all purposes.

FIELD OF THE INVENTION

The invention relates to recombinant nucleic acid molecules for making circular RNAs, methods for making circular RNAs, and circular RNAs which do not comprise exogenous exon sequences.

BACKGROUND

Circular RNAs (circRNAs) are covalently closed single-stranded RNA molecules. Over the last decade, tens of thousands of circRNAs have been identified from viruses to humans. Within human cells, they are mainly produced by a back splicing process. Due to their ubiquitous distribution, there is an increasing interest in them. Although their natural function is not yet well determined, some hypotheses are that they act as sponges for microRNAs and proteins, or as templates to produce peptides or proteins. At the same time, circRNAs are now becoming the basis of next-generation mRNA therapeutics, mainly because of their high stability.

RNA *in vitro* circularization can be achieved by chemical or enzymatic ligation, but such strategies are inefficient in producing long circRNAs and prone to forming intermolecular ligations (Petkovic, S. & Muller, S. RNA circularization strategies in vivo and in vitro. *Nucleic Acids Res* **43**, 2454-2465, (2015)). An alternative method, the so-called permuted intron–exon method (PIE) is efficient in circularizing long RNAs (Puttaraju, M. & Been, M. D. Group I permuted intron-exon (PIE) sequences self-splice to produce circular exons. *Nucleic Acids Res* **20**, 5357-5364 (1992); Wesselhoeft, R. A., Kowalski, P. S. & Anderson, D. G. Engineering circular RNA for potent and stable translation in eukaryotic cells. *Nat Commun* **9**, 2629, doi:10.1038/s41467-018-05096-6 (2018); WO2023/046153). The PIE method takes advantage of the self-splicing property of the group I intron, a well-studied ribozyme that can join flanking exons by a two-step transesterification reaction in the absence of any protein. The PIE method moves the 5' half of the intron to the 3' end of the RNA and thereby connects flanking exons, so that the splicing product is a circular form of the preconnected exons.

Although the PIE method is an attractive one-step method to produce circRNA, it has some drawbacks. The key for PIE to work is the formation of the ribozyme core by the two intron halves that are separated by the preconnected exon. The long preconnected exon has the potential to interact with adjacent intron halves and could block the formation of the functional ribozyme core. Also, circRNAs produced by PIE contain unwanted sequences which have been reported to induce immune responses (Liu, C. X. *et al. Mol Cell* **82**, 420-434 e426 (2022)). Group I intron can also act as a trans ribozyme to target a second molecule (Inoue T *et al*, Intermolecular exon ligation of the rRNA

precursor of Tetrahymena: Oligonucleotides can function as 5' exons. *Cell*, vol 43(2) (1985); Been, MD et al, One binding site determines sequence specificity of Tetrahymena pre-rRNA self-splicing trans-splicing and RNA enzyme activity. *Cell*, vol. 47(2) (1986)). Many efforts have been spent in engineering the trans ribozyme to deliver correct sequences to defective mRNAs *in vivo* (Sullenger BA et al, Ribozyme-mediated repair of defective mRNA by targeted trans-splicing *Nature*, vol. 371(6498) (1994); Lan, N., Howrey, R. P., Lee, S. W., Smith, C. A. & Sullenger, B. A. Ribozyme-mediated repair of sickle beta-globin mRNAs in erythrocyte precursors. *Science* 280, 1593-1596 (1998). Jones JT et al, Tagging ribozyme reaction sites to follow trans-splicing in mammalian cells *Nat Med*, vol. 2(6) (1996). Kohler U et al, Trans-splicing Ribozymes for Targeted Gene Delivery, *J Mol Biol* vol. 285(5), (1999)), but unfortunately the efficiency is low.

Various attempts have been made to provide circular RNAs which do not comprise unwanted sequences, for example, which do not comprise spacer sequences or native exon sequences such as those introduced with the PIE method. One such approach is described in WO2022/191642, in which the internal guide sequence (IGS) of the ribozyme is reverse complementary to a target site region in the gene of interest, in order to self-splice and circularize. The IGS region and the target site region flank the ribozyme and a gene of interest, respectively, and a guanine at the 5' end of the IGS region forms a wobble base pair with an uracil at the 3' end of the target site region. However, such approaches are limited by poor efficiency in converting precursor to circular RNAs and produce high levels of nicking with low yield.

Accordingly, there is a need in the art for more efficient methods of RNA circularization which also enable circular RNAs to be produced without unwanted sequences.

SUMMARY

In a first aspect, the invention provides a recombinant nucleic acid molecule for making a circular RNA, comprising in the 5' to 3' direction:

- a) an internal guide sequence (IGS),
- b) a ribozyme,
- c) a first portion of an extended anticodon arm (eACA) sequence,
- d) a gene of interest, and
- e) a second portion of the eACA sequence,

wherein a nucleotide in the second portion of the eACA sequence forms a wobble base pair with a nucleotide in the IGS.

A gene of interest in a recombinant nucleic acid molecule of the first aspect may comprise a coding sequence. The gene of interest may further comprise an internal ribosome entry site (IRES) sequence.

In some embodiments, a recombinant nucleic acid molecule of the first aspect may be devoid of stop codons in frame with the coding sequence.

In a second aspect, the invention provides the use of a recombinant nucleic acid molecule as described herein in a method of making a circular RNA.

In a third aspect, the invention provides a method for producing a circular RNA, comprising:

- a) providing a recombinant nucleic acid molecule as described herein, and
- b) circularizing the recombinant nucleic acid molecule.

In a fourth aspect, the invention provides a method for producing a circular RNA, comprising:

- a) providing a recombinant nucleic acid molecule as described herein,
- b) transcribing the recombinant nucleic acid molecule to produce an RNA precursor, and
- c) circularizing the recombinant nucleic acid molecule.

In a fifth aspect, the invention provides a method for producing a circular RNA, comprising:

- a) identifying a gene of interest comprising a sequence capable of forming an eACA stem-loop,
- b) preparing a recombinant nucleic acid molecule comprising, in a 5' to 3' direction: an internal guide sequence (IGS) – a ribozyme – a sequence encoding the gene of interest, and
- c) circularizing the recombinant nucleic acid molecule.

In a sixth aspect, the invention provides a method for producing a circular RNA, comprising:

- a) providing a recombinant nucleic acid molecule as described herein, and
- b) transcribing and circularizing the recombinant nucleic acid molecule, wherein circularization occurs co-transcriptionally.

In a seventh aspect, the invention provides a circular RNA comprising a sequence encoding a gene of interest, wherein the circular RNA does not comprise exogenous splicing sequences.

In an eighth aspect, the invention provides a circular RNA comprising a sequence encoding a gene of interest, wherein the circular RNA does not comprise RNA from a circularization agent.

In a ninth aspect, the invention provides a circular RNA obtainable by the methods described herein, wherein the circular RNA is less immunogenic compared to a circular RNA which comprises exogenous exon sequences.

In a tenth aspect, the invention provides a circular RNA obtainable by the methods described herein, wherein the circular RNA does not comprise exogenous exon sequences.

A circular RNA of any of the seventh to the tenth aspects may be produced by a method of any of the third to the sixth aspects.

- 5 In an eleventh aspect, the invention provides the use of a circular RNA as described herein in an *in vitro* method of expressing a protein in a cell.

In a twelfth aspect, the invention provides a method for expressing a gene of interest in a cell, the method comprising (a) circularizing a recombinant nucleic acid molecule as described herein to
10 provide a circular RNA comprising the gene of interest, and (b) administering the circular RNA to the cell.

In a thirteenth aspect, the invention provides a method of treating a disease in a subject, the method comprising (a) circularizing a recombinant nucleic acid molecule as described herein to provide a
15 circular RNA, and (b) administering the circular RNA to the subject.

In a fourteenth aspect, the invention provides a recombinant nucleic acid molecule as described herein, for use as a medicament.

- 20 In a fifteenth aspect, the invention provides a recombinant nucleic acid molecule as described herein, for use in a method of treating a disease in a subject.

BRIEF DESCRIPTION OF THE DRAWINGS

25 **Figure 1:** Comparison of prior art approaches and trans ribozyme-based circularization (TRIC). (A) Schematic depicting splicing by the group I intron from cyanobacterium *Anabaena* (Ana), resulting in a linear molecule comprising joined exons. (B) Schematic depicting circularization of a gene of interest by the permuted intron-exon (PIE) method. (C) Schematic depicting splicing by a group I intron acting as a *trans* ribozyme, in which a target sequence is edited. (D) Schematic depicting circularization of a
30 gene of interest using the TRIC approach.

Figure 2: (A) Schematic of the Ana tRNA-intron. (B) Initial version of TRIC (V0). The tRNA sequence of Ana has been retained (L15/R30). Insert: reaction steps of TRIC-V0. (C) (right of Figure) 12% and 6% denaturing PAGE to identify circular 3*Flag produced by TRIC-V0. (D) RT-PCR of *in vitro*
35 transcription product, band I product and band II product. (E) Sanger sequencing of RT-PCR product.

Figure 3: (A) Schematic of TRIC-V1 (L15/R30) in which EGS and loop regions are present. (B) 6% denaturing PAGE was run with product of *in vitro* transcription of V1 variants. (C) RNase R digestion of the *in vitro* transcription product of TRIC-V1.

40

Figure 4: (A) Relative lengths of long genes of interest. (B) 0.8% native agarose gel run with in vitro transcription samples and circularized samples of TRIC-V1. Expected band positions for circular products are labelled with black circles on the gel.

Figure 5: (A) Schematic of DNA template, RNA precursor and resultant circular RNA of the PIE-Ana approach. Sequence of the linear precursor is shown on the right. (B) Schematic of DNA template, RNA precursor and resultant circular RNA of the TRIC-V1 approach. Sequence of the linear precursor is shown on the right.

Figure 6: (A) Results of in vitro transcription performed at varying concentrations of Mg²⁺. (B) Full length precursors of the TRIC-V1 were subject to circularization for 10, 20, 40 min and separated in a 0.8% native agarose gel. (C) Circularization efficiency of TRIC-V1 for long GOIs. (D) Comparison of TRIC-V1 to the PIE-Ana method. Expected band positions for circular products are labelled with black circles on the gels.

Figure 7: (A) Schematic of TRIC-V1 with a short GOI. (B) Upper panel and middle table show sequence information for TRIC-V1. Lower table shows the right and left arm configurations of the tRNA sequence in TRIC variants 1.30-1.39.

Figure 8: (A) Schematic of TRIC-V2. (B) 6% denaturing PAGE showing circularization with V1.30-V1.39 variants and with V2.0, 2.1 and 2.2 variants. L/R configurations are shown under the gel. (C) The extended anticodon arm (eACA). As long as a stem loop structure can be found where the loop contains a uracil in 3rd position and the stem is ≥5bp, a circularization site can be assembled. Thus, circularization sites can be either placed in UTR or CDS. (D) Circularization of RNAs with the eACAs in CDS. Expected band positions for circular products are labelled with black circles on the gels.

Figure 9: (A) Comparison between TRIC-V1 and TRIC-V2 on 2% native agarose gel, with varying EGS and eACA stem lengths. (B, C) Comparison between PIE and TRIC-V2 on CVB3-EGFP, CVB3-Spike-EGFP and CVB3-Cas9-EGFP. (D) Michaelis-Menten fitting of the TRIC-V2 and the PIE for RNA circularization. Data presented as means + SDs; n = 3; **p < 0.01. (E) Comparison of efficiency, yield and nicking between TRIC-V1, TRIC-V2 and PIE. Data presented as means + SDs; n = 3; **p < 0.01.

Figure 10: (A) Schematic overview of TRIC-V2 approach. (B) Legend for components of TRIC-V2.

Figure 11: (A) Schematic of circular RNAs produced by TRIC-V2, including position of siRNA targeting sites (msiTS). Left: arrangement of the resultant circ GOIs: eACA (UTR)-msiTS-IRES-start codon-CDS-stop codon-PolyAC-eACA (UTR). The eACA is assembled in UTR of circ. GOIs. Right: arrangement of the resultant circ GOIs: eACA (CDS)-stop codon-PolyAC-msiTS-IRES-start codon-eACA (CDS). The eACA is assembled in CDS of circ. GOIs. A multiple siRNA target site (msiTS) is

introduced for siRNA targeting and to separate IRES from eACA. (B) Upper panel: HPLC profile of a circularized RNA sample. Lower panel: HPLC fractions were resolved in a 0.8% native agarose gel. (C) Two-step purified (HPLC and RNase R digestion) circular RNAs on a 6M-1.5% urea-agarose gel. (D-F) Expression of INF-beta (D), IL6 (E) and CCL5 (F) by A549 cells. Data presented as means + SDs; n = 3; *p<0.05, **p < 0.01, ***p<0.001. (G) Expression of Nluc in HEK293 cells following transfection with linear or circular constructs. Data presented as means + SDs; n ≥ 3; *p<0.05, **p < 0.01. (H) Expression of Nluc following siRNA treatment. Data presented as means + SDs; n = 6; **p < 0.01.

Figure 12: Schematic depicting how a circular RNA can be provided from a gene of interest which comprises an extended anticodon arm sequence.

Figure 13: Schematic showing the positioning of circularization site in either UTR or CDS. (A) Overview for TRIC-V2, showing resulting circular constructs. (B) Exemplary sequences with eACA placed in UTR (SEQ ID NO: 61) or placed in CDS (SEQ ID NO: 62).

Figure 14: Schematic showing the different recombinant nucleic acid constructs used in the experiments described herein.

Figure 15: (A) schematic illustrating that when circRNAs contain only a CDS but no stop codons, ribosomes can translate such circRNA indefinitely and produce a polyprotein, a process known as the rolling circle translation (RCT). (B) schematic illustrating that when a 2A sequence is present, such as T2A or P2A, the polyprotein will be converted to protein monomers. (C) Secondary structure of the 373 nt CSFV IRES. Multiple stop codons are present in all possible frames: 4, 10, and 8 in frames 1-3, respectively. Frame 1 was cleaned of stop codons by three mutations (U52C, A171G, and A191G), and one deletion (d349U) to align with the CDS. (D) Equimolar amounts of circOR4F17-Nluc-RCT (SEQ ID NO: 78), circCSFV-Nluc (SEQ ID NO: 74), circCSFV-Nluc-RCT (SEQ ID NO: 75), V2 circNluc, and N¹Ψ-modified linear Nluc were transfected into HEK 293F cells. 24 hours later, Nluc expression was monitored. (E) Western blot analysis of Nluc expression from either single shot translation or RCT using CSFV IRES. Actin was used as internal reference. Polyproteins were observed. Data in (D) are mean ± SD for 4 biological replicates. *p < 0.05, **p = 0.01 and ***p < 0.001, unpaired two-tailed t-test

Figure 16: (A) schematic showing the original eACA that contains a 7 nt loop with a U (T) at the third position and a stem structure (a-0) and new constructs with different Loop length and U positions (b-1 to m-12). (B) Precursor RNAs were diluted to 400 ng/ul, denature at 95 °C for 2 min and placed on ice for at least 3 min. Then 1 ul of 10X splicing buffer was added to each tube contains 9 ul of refolded RNA precursors. 1 or 3 min later, add 2 ul 100 mM EDTA to each tube to stop the reaction. Then ~500 ng of each RNA was loaded to a 1.5% native agarose gel and run at 25W for 30min.

Figure 17: (A) Full length precursors of TRICv2 constructs with GU, or CA base pairs were circularized for 1 to 20min and analysed on a 1.2% native agarose gel. (B) Circularization of TRICv2 (C●A) was confirmed by Urea-Agarose gel.

Figure 18: (A-B) Comparison between V2 using either Ana group I intron or Tetrahymena thermophila group I intron (Tetra) and the Tetra-STS and the Tetra-Rzy constructs. CVB3 (coxsackievirus B3)-EGFP was cloned to the best Tetra-STS (AU-rich no. 16) and Tetra-Rzy (CVB3IRES-GFP) constructs. FL were circularized for designated time and analyzed on native agarose gels (a) or urea-agarose gels (b).

DETAILED DESCRIPTION

The inventors have determined that if a 5' end of an RNA molecule is connected to a 3' end of a trans-splicing ribozyme, the splicing product is a circular RNA. The approach described herein is named "Trans Ribozyme-based Circularization", or TRIC. TRIC uses an intact intron, joined upstream of the sequence to be circularized. This approach better enables the folding of the catalytic structural core compared to the PIE method, in which the intron is split in two. This in turn permits the circularization of longer sequences or genes of interest because there is less chance of interference between the gene of interest and the intron.

An important feature of the invention is the presence of an extended anticodon arm (eACA) sequence provided in the recombinant nucleic acid molecule to be circularized. This sequence provides a structure that is structurally similar to the anticodon arm found in *Ana* tRNA, comprising a stem and a loop, and enables circularization with greater efficiency compared to existing PIE methods and methods such as those described in WO2022/191642. Because the advantages of the eACA are dependent on structure, rather than specific nucleotide sequence, eACA sequences (that is, sequences capable of forming a stem-loop structure as described herein) may be readily identified in genes of interest so that they may be circularized without the need to add additional sequence or perform extensive mutations. Consequently, the recombinant nucleic acid molecules described herein may produce circular RNAs comprising only the gene of interest, without exogenous splicing sequences or unwanted spacer sequences, and may do so in a highly efficient way by utilising the propensity of the eACA sequence to form a stem-loop structure.

Abbreviations

ACA – anticodon arm

CDS – coding sequence

EGS – extended guide sequence

eIGS – extended internal guide sequence

eACA – anticodon arm-like structure

GOI – gene of interest

IGS – internal guide sequence

IVT – in vitro transcription

PIE – permuted intron-exon

RNA – ribonucleic acid

5 TRIC – trans ribozyme based circularization

tRNA – transfer RNA

UTR – untranslated region

RCT – rolling circle translation

10 Recombinant nucleic acid molecules are provided herein for making circular RNAs. The nucleic acid molecules are generally linear prior to circularization. In general, the nucleic acid molecule for circularization comprises the gene of interest (GOI), which is the gene to be circularized, a ribozyme, capable of performing the circularization, and an extended anticodon arm (eACA) sequence, split into two portions.

15

The ribozyme may be any ribozyme capable of acting as a trans-splicing ribozyme. In some cases, the ribozyme is derived from or is a group I intron. Examples of suitable ribozymes are the *Tetrahymena* ribosomal intron, T4 phage thymidylate synthase intron, *Anabaena* (Ana) pre-tRNA intron, Azoarcus sp. BH72 Ile tRNA intron, and Staphylococcus phage Twort ribonucleotide reductase intron. Sequences of these ribozymes are shown below, with the internal guide sequence (IGS) shown with underlined, shaded letters.

Ana group I intron:

25 AATAATTGAGCCTTAGAGAAGAAATTCCTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCT
AGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTT
ACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA
GAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATG (SEQ ID NO: 1)

T4 phage group I intron:

30 AATTGAGGCCTGAGTATAAGGTGACTTATACTTGTAATCTATCTAAACGGGGAACCTCTCTAGTAG
ACAATCCCGTGCTAAATTGTAGGACTTGCCCTTTAATAAACTTCTATATTTAAAGAGGTATTTAT
GAAAAGCGGAATTTATCAGATTAAAAATACTTTAAACAATAAAGTATATGTAGGAAGTGCTAAAGAT
TTTGAAAAGAGATGGAAGAGGCATTTTAAAGATTTAGAAAAAGGATGCCATTCTTCTATAAACTT
35 CAGAGGTCTTTTAAACAACATGGTAATGTGTTTGAATGTTCTATTTTGAAGAAAATTCATATGAGA
AAGATTTGATTATTGAACGAGAGAAAATTTTGGATTAAAGAGCTTAATTCTAAAAATTAATGGATACAA
TATTGCTGATGCAACGTTTGGTGATACATGTTCTACGCATCCATTAAAAGAAGAAATTATTAAGAAA
CGTTCTGAAACTGTTAAAGCTAAGATGCTTAAACTTGGACCTGATGGTCGGAAAGCTCTTTACAGT
AAACCCGGAAGTAAAAACGGGCGTTGGAATCCAGAAACCCATAAGTTTTGTAAGTGCGGTGTTTCG
CATACAAACCTTCTGCTTATACTTGTAAGTAAATGCAGAAATCGTTCAGGTGAAAATAATTCATTCTTT
40 AATCATAAGCATTGACATAACTAAATCTAAATATCAGAAAAGATGAAAGGTAAAAAGCCTAGT
AATATTAATAAGATTTTCATGTGATGGGGTTATTTTGAATGTTGTCAGCAGATGCAGCTAGACATTTTA
AAATTTCTGCTGGATTAGTTACTTATCGTGTAATCTGATAAATGGAATTGGTTCTACATAAATGC
CTAACGACTATCCCTTTGGGGAGTAGGGTCAAGTGACTCGAAACGATAGACAACCTTGCTTTAACA
AGTTGGAGATATAGTCTGCTCTGCATGGTGACATGCAGCTGGATATAATTCCGGGGTAAGATTAA
45 CGACCTTATCTGAACATAATG (SEQ ID NO: 2)

Tetrahymena ribosomal intron:

AAATAGCAATATTTACCTTTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAAC
 CAATAGATTGCATCGGTTTAAAAGGCAAGACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTT
 AGTACCAAGTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGA
 CATGGTCCTAACCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCA
 5 CAGACTAAATGTCGGTTCGGGAAGATGTATTCTTCTCATAAGATATAGTCGGACCTCTCCTTAATG
 GGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGAAAGTAT
 ATTGATTAGTTTTGGAGTACTCG (SEQ ID NO: 3)

Azoarcus sp. BH72 Ile tRNA intron:

10 ATTTTCGATGTGCCTTGCGCCGGGAAACCACGCAAGGGATGGTGTCAAATTCGGCGAAACCTAAG
 CGCCCGCCCGGGCGTATGGCAACGCCGAGCCAAGCTTCGGCGCCTGCGCCGATGAAGGTGTAG
 AGACTAGACGGCACCCACCTAAGGCAAACGCTATGGTGAAGGCATAGTCCAGGGAGTGGCGAAA
 GTCACACAAACCGG (SEQ ID NO: 4)

Staphylococcus phage Twort ribonucleotide reductase intron:

15 AAACAAATTCGCCCCCTATATTAGTAATAGTGTAGGTTAAAAAACTTCCTTAATTCATGGGAAATCT
 CCCTGACTTTTATTATAAATTTGGTATAGTAAATGAGAAGGAGATAATCATGAGCAAGATAACCA
 ATAGCAAAGAAACCCAAAAAGTACCTAACGCTACTAAAAGTATTTACCACCATATAAAAAGTAAAA
 GAAGGATGGAAGTCATTAAATCACTTAATGAATTGGTAATTATCTTGTGCAACGACTAGAGAAAAG
 20 ATAGTTTATTGTTACAGGCAGTAAATGAAGACTGAGTATCGTACACACAAGTGAGTGGAACAGG
 AAGTATCCTAGAGTAACGACTAGGATAATGATATAGTCTGAACATTGTAGGTGACTACAAGAAGGT
 AAGGAGTAACGAACCTTATCGTAACATAATTG (SEQ ID NO: 5)

Recombinant nucleic acid molecules described herein also generally comprise an internal guide

25 sequence (IGS). The IGS may be part of the ribozyme, such as part of the group I intron. Generally,
 this enables the TRIC approach to use an intact intron. However, it is possible to use a truncated
 ribozyme sequence in the TRIC method, with the native IGS removed and replaced by a different IGS
 to that which would normally be present.

30 The function of the IGS is to base pair with the end regions of the gene of interest, in order to bring
 them into proximity with each other so that circularization may occur. The IGS binds through
 complementary base pairing to both first and second portions of the eACA sequence, which are
 located at either end of the GOI to be circularised.

Extended anticodon arm (eACA) sequence

35 An extended anticodon arm (eACA) sequence is one which is capable of forming a stem-loop
 structure (also known as a hairpin or hairpin loop). Stem-loop structures form when two regions of
 single-stranded RNA which are generally complementary to each other (when read in opposite
 directions) base-pair with each other. The base-pairing results in a double helix structure ending in an
 40 unpaired loop. The natural propensity of eACA sequences to form stem-loop structures may be
 utilised to enable circularization of a gene of interest, as shown in Figure 12.

Prior to circularization, the linear recombinant nucleic acid molecule comprises an eACA sequence in
 two separate portions. A first portion of the eACA sequence is positioned at or near the 5' end of the
 45 gene of interest, and a second portion of the eACA sequence is positioned at or near the 3' end of the
 gene of interest, as shown in Figure 10A (top panel). During circularization, splicing by the trans-

ribozyme causes the first and second portions of the eACA sequence to be covalently joined, in order to create a circular version of the gene of interest. In the resulting circular molecule, the first and second portions are joined to form the eACA sequence, which is generally capable of forming a stem-loop structure as shown in Figure 8C, Figure 10A, and Figure 12.

5

The first portion of the eACA sequence may comprise a first eACA stem portion and a first eACA loop portion. The second portion of the eACA sequence may comprise a second eACA stem portion and a second eACA loop portion. By “stem portion” it is meant a part of the first (or second) portion of the eACA sequence that is capable of forming the stem of a stem-loop structure. By “loop portion” it is meant a part of the first (or second) portion of the eACA sequence that is capable of forming the loop of a stem-loop structure. Figure 12 shows how a stem-loop forming structure can be identified in a gene of interest, and subsequently used for circularizing the RNA. Thus, the stem and loop portions of the eACA sequence are capable of forming a stem-loop structure in the recombinant nucleic acid molecules described herein.

15

The specific nucleotide sequence of the eACA sequence is not important for TRIC and does not determine whether circularization will occur. Rather, it is the structure as opposed to the sequence of the eACA that is important. Consequently, the eACA may be of any nucleotide sequence, provided that the last nucleotide in the second eACA loop portion is one which can form a wobble base pair with a corresponding nucleotide in the internal guide sequence described herein. The last nucleotide in the second eACA loop portion may be a uracil that forms a wobble base pair with a corresponding guanine in the internal guide sequence. Alternatively, the last nucleotide in the second eACA loop portion may be a cytosine that forms a wobble base pair with a corresponding adenine in the internal guide sequence.

25

The first and second eACA stem portions may be complementary to each other, though this is not necessary.

The first and second eACA stem portions are generally each at least 5 nucleotides in length but may be as short as 1 nucleotide in length each. The stem portion lengths may be adapted depending on the gene of interest to be circularized. For example, longer stem lengths (such as lengths greater than 15 nt) may be advantageous if circularizing long (>500 nt) genes of interest.

35

Thus, the first and second stem portions may each be at least 1, at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, at least 25 or at least 30 nucleotides in length. In particular, the first and second stem portions may each be at least 15 or at least 25 nucleotides in length. For example, the first and second eACA stem portions may be 1 to 50 nucleotides in length, for example 5 to 40 nucleotides.

The first and second stem portions need not be the same length, for example one stem portion may be one or two nucleotides shorter than the other, provided a stem-loop structure can still be formed.

The anticodon arm loop of the *Ana* tRNA group I intron is naturally 7 nucleotides in length.

5 Consequently, in the circular RNAs described herein, the loop of the stem-loop structure may be 7 nucleotides in length, particularly when the ribozyme used is, or is derived from, the *Ana* group I intron. If other group I introns are used, the loop of the stem-loop structure may have a different nucleotide length. The loop of the stem-loop structure may be between 3 to 40 nucleotides in length and is generally between 3 and 10 nucleotides in length. In many cases, the loop of the stem-loop structure is at least 3 nucleotides in length, at least 4 nucleotides in length or at least 5 nucleotides in length, particularly if the ribozyme is or is derived from the *Ana* group I intron. In some embodiments, the loop of the stem-loop structure is generally between 5 and 11 nucleotides in length.

15 Generally, if the loop of the stem-loop structure comprises 7 nucleotides, then the first eACA loop portion comprises 4 nucleotides, and the second eACA loop portion comprises 3 nucleotides.

Accordingly, the first and second eACA stem portions may each be at least 15 nucleotides in length, the first eACA loop portion may be 4 nucleotides in length, and the second eACA loop portion may be 3 nucleotides in length. In the resulting circular RNA molecule, this produces an eACA sequence which is capable of forming a stem-loop structure, wherein the stem comprises at least 15 base pairs, and the loop is 7 nucleotides in length.

In total, the first portion of the eACA sequence may comprise as few as 5 nucleotides (for example, 1 stem nucleotide and 4 loop nucleotides). The second portion of the eACA sequence may comprise as few as 4 nucleotides (for example, 1 stem nucleotides and 3 loop nucleotides). Alternatively, the first portion of the eACA sequence may comprise, for example, 19 nucleotides (e.g. 15 stem nucleotides and 4 loop nucleotides), or 29 nucleotides (e.g. 25 stem nucleotides and 4 loop nucleotides). The second portion of the eACA sequence may comprise, for example, 18 nucleotides (e.g. 15 stem nucleotides and 3 loop nucleotides), or 28 nucleotides (e.g. 25 stem nucleotides and 3 loop nucleotides).

An exemplary first portion of the eACA sequence comprising a 1 nucleotide stem and a 4 nucleotide loop may comprise the nucleotide sequence 5'-NNNNN-3', wherein N is any nucleotide, and an exemplary second portion of the eACA sequence comprising a 1 nucleotide stem and a 3 nucleotide loop may comprise the nucleotide sequence 5'-NNNU-3', wherein N is any nucleotide. As an example, in the case of Nano Luciferase (NLuc), a possible ACA sequence is **GATCACCAC**TTTAAGGTGATC (SEQ ID NO: 58). For circularization, the NLuc GOI is rearranged such that the ACA sequence is provided in two portions (a 5' first portion and a 3' second portion). The first portion of the eACA sequence comprises the sequence TTAAGGTGATC (SEQ ID NO: 59). The second portion of the eACA sequence comprises the sequence GATCACCAC (SEQ ID NO: 60).

It will be realised that the provision of a thymine in a DNA precursor molecule will result in a uracil in the translated RNA molecule for circularization.

The first and second eACA loop portions base pair with the internal guide sequence (IGS) to form the P1 and P10 regions, which are critical for ribozyme activity. The first eACA loop portion, positioned towards the 5' end of the gene of interest, base pairs with the IGS to form the P10 region. It is not necessary for all the nucleotides in the first eACA loop portion to form the P10 region, and in some cases only two nucleotides of the first eACA loop portion form the P10 region. The second eACA loop portion, positioned towards the 3' end of the gene of interest, base pairs with the IGS to form the P1 region.

As described above, in order for circularization to occur, the last nucleotide of the second eACA loop portion (i.e. the nucleotide at the 3' end of the second eACA loop portion) may form a wobble base pair with a corresponding nucleotide in the IGS. Generally, the wobble base pair is a GU wobble base pair, with G in the IGS and U in the second eACA loop portion or an AC wobble base pair with A in the IGS and C in the second eACA loop portion. The wobble base pair provides the circularization site, such that once circularized, the nucleotide at the 3' end of the second eACA loop portion forms the third nucleotide in the loop of the eACA stem-loop structure. This is depicted in Figure 8C.

The P1 region may also be formed by base pairing of the IGS with a region adjacent to the second eACA loop portion in the 3' direction, known as the "P1 extension". If present, the P1 extension typically comprises between 2 and 4 nucleotides, which base pair with the IGS. The P1 region may therefore be formed by the P1 extension and second eACA loop portion base pairing with the IGS, as shown in Figure 10A. Consequently, in some embodiments, the second portion of the eACA sequence and the P1 extension together are capable of forming a P1 region. If the P1 extension is not present, the P1 region is formed by only the second eACA loop portion base pairing with the IGS. P1 extensions have been described in Olson & Muller (2012) *RNA* 18:581-589. The contents of which are incorporated herein by reference. Generally, if an extended guide sequence (EGS) is used, the P1 extension region will be present.

A particular advantage of the TRIC method is that it may utilise eACA sequences which are already present in the gene of interest. For example, if a stem-loop forming eACA sequence can be found in a gene of interest, this gene may be circularized efficiently without introducing any additional sequences. This in turn means that the resulting circular RNA is far less likely to be immunogenic. This is depicted in Figure 12, which shows first the identification of an eACA sequence (i.e. a stem-loop forming structure) in a gene of interest. Subsequently, the gene of interest is rearranged such that the eACA sequence is split into two portions, one at each end of the gene of interest. This rearranged gene is then cloned into the TRIC construct for circularization. An example of this is the protein coding circular RNA T2A Nano Luciferase. This circular RNA already comprises an eACA sequence in its natural

sequence. This means a circularization site can be introduced using the naturally occurring eACA sequence, without the need to perform mutations or introduce additional sequence.

If the natural coding sequence (CDS) does not contain an eACA sequence, it is possible to perform selective mutations to create one. Codon redundancy means that mutations may be made to the nucleotide sequence of the GOI without affecting the resulting peptide sequence. Consequently, an eACA sequence may be provided in the GOI without requiring the introduction of additional sequences. Instead, only selective mutation of the existing sequence is needed, following the rules of codon redundancy. Examples of this are the T2A-EGFP and circZNF609 circular RNAs described in Example 5, in which mutations are introduced based on codon redundancy to introduce the circularization site.

If there is no eACA in a non-coding RNA, or if the circularization site is placed in an untranslated region (UTR) of a protein-coding RNA, the circularization site may be created by introducing additional nucleotides. For example, as shown in Figure 8C, 5 nucleotides (light grey nt) may be introduced to create a stem portion of the eACA sequence, using the existing sequence (black nt) of the GOI to provide the remainder of the eACA.

Accordingly, in the recombinant nucleic acids described herein, the first and/or second portions of the eACA sequence may naturally occur in the gene of interest. In other words, they may be part of the gene of interest and so are present without having to mutate the existing sequence or introduce additional sequence.

Alternatively, in the recombinant nucleic acids described herein, all or part of the eACA sequence may be derived from human ribosomal RNA (rRNA). The use of human rRNA has the potential to provide circular RNAs which are less immunogenic.

The location of the first and second portions of the eACA sequence in the gene of interest has little impact on circularization. One portion of the eACA sequence may be identified in or placed in a coding sequence, whilst the other may be identified in or placed in an untranslated region. It is not necessary for both portions to be in the coding sequence, for example, or for both portions to be in the untranslated region.

Extended guide sequence (EGS)

The recombinant nucleic acids described herein may further comprise an extended guide sequence (EGS), in particular a first EGS and a second EGS which are capable of complementary base pairing to each other. The function of the EGS is to increase the length of the complementary base-pairing region at the two ends of the recombinant nucleic acid molecule. In this way, the EGS may be included to compensate for a shorter IGS, in particular when longer (>500 nt) GOIs are circularized.

The recombinant nucleic acids described herein may comprise a first EGS positioned 5' of the IGS. The recombinant nucleic acids described herein may comprise a second EGS positioned 3' to the second portion of the eACA sequence. Generally, there is a loop sequence situated between the first EGS and the IGS, as described in more detail below. Similarly, there is also generally a loop sequence situated between the second EGS and the second portion of the eACA sequence.

The first and second EGS may be partly or fully complementary to each other. Generally, mismatches are tolerated well and do not materially affect circularization. In the recombinant nucleic acids described herein, the first EGS may be at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% complementary to the second EGS.

If present, the first and second EGS may each be between 1 and 500 nucleotides in length. For example, the first and second EGS may each be between 10 and 50 nucleotides in length. The first and second EGS may each be 20, 30, or 40 nucleotides in length.

An exemplary first EGS sequence is GGUCAAUCGGUUGGCUUCCG (SEQ ID NO: 56).

An exemplary second EGS sequence is CGGAAGCCAACCGAUUGACC (SEQ ID NO: 57).

Loops

Lengthening the base paired P1 region may have a deleterious effect on circularization. To avoid this, the recombinant nucleic acids described herein may further comprise loop sequences, such as a first loop sequence and a second loop sequence.

The first and second loops may act as spacers between, at the 5' end, the internal guide sequence (IGS) and the first extended guide sequence (EGS), and at the 3' end, the P1 region and the second EGS. The loop sequences are preferably not complementary to each other, such that there is little or no base pair interaction between the first and second loop sequences. Because of the low or non-complementarity between the two loop sequences, the base-paired P1 region remains at a fixed length. The first loop may alternatively be described herein as "left loop". The second loop may alternatively be described herein as "right loop".

The first and second loop sequences may each be between 1 and 10 nucleotides in length. It is not necessary for the first and second loop sequences to have the same number of nucleotides, and in fact the TRIC method works well when the first and second loop sequences are different lengths. The first loop sequence may be 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides in length. The second loop sequence may be 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides in length. A preferred combination is a 6 nucleotide first loop sequence and a 5 nucleotide second loop sequence. Another preferred combination is a 3 nucleotide first loop sequence and a 2 nucleotide second loop sequence.

The first loop sequence is positioned 3' to the first EGS and 5' to the IGS, in other words, in between the first EGS and the IGS. The second loop sequence is positioned 3' to the second portion of the eACA sequence, and 5' to the second EGS, in other words, between the second portion of the eACA sequence and the second EGS. If the P1 extension region is present, the second loop sequence is 3' to the P1 extension.

It is not always necessary to have two loop sequences. The recombinant nucleic acids described herein may comprise only a first loop sequence, positioned in between the first EGS and the IGS, without a second loop sequence. Alternatively, the recombinant nucleic acids described herein may comprise only the loop sequence positioned between the second portion of the eACA sequence (or, if present, P1 extension) and the second EGS.

An exemplary sequence for the first loop sequence is AAATAA (SEQ ID NO: 54).

An exemplary sequence for the second loop sequence is ACACC (SEQ ID NO: 55).

Gene of interest (GOI)

The gene of interest (GOI) refers to the sequence which is to be circularized. The GOI may comprise a coding sequence, coding for a peptide or protein, or may be a noncoding sequence. The GOI may also comprise a combination of coding and noncoding sequence. It will also be understood that, as used herein, the term "gene of interest" encompasses sequences which include additional sequence elements, for example, a translation initiation element, such as an internal ribosome entry site (IRES) sequence, multiple siRNA target sites (msiTS), spacer sequences such as polyAC sequences, start codons, stop codons, and any other sequence elements known to be useful in the art for producing circular RNA. For example, where an eACA is located in a coding sequence to be circularized, the GOI may comprise, in the 5' to 3' direction: a stop codon, a polyAC sequence, multiple siRNA target sites (msiTS), an IRES, a start codon, and the coding sequence including the eACA. Where an eACA is placed in an untranslated region of a sequence to be circularized, the GOI may comprise, in the 5' to 3' direction: multiple siRNA target sites (msiTS), an IRES, a start codon, a coding sequence, a stop codon, a polyAC sequence, and the eACA. See, for example, Figures 10A and 10B.

Suitable translation initiation elements include internal ribosome entry site (IRES) sequences. IRESs for use in the invention may include viral IRESs, such as the Coxsackievirus B3 (CVB3), Cafeteria roenbergensis Virus (CroV), or Classical Swine Fever Virus (CSFV) IRES, the DNA sequences of which are set out below.

CVB3-IRES

TTAAACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACG
GTACCTTTGTGCGCCTGTTTTATACCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGGATC
AACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATC
AATAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTTCGTTATCCGGCCAACTACTTCGAAAAACCT

AGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCCAGTGTAGATCAGGTCGATGA
 GTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAAA
 CCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCG
 GCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCG
 5 TAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG
 GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTA
 ATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTAAAACATTACA
 ATTCATTGTTAAGTTGAATACAGCAAA (SEQ ID NO: 71)

10 CroV-IRES:

GTATAAGAGACAGGTGTTTGCCTTGTCTTCGGAAGTGGCATCTTGGGACCAACCCCCCTTTTCCCC
 AGCCATGGGTAAATGGCAATAAAGGACGTAACAACTTTGTAACCATTAAGCTTTGTAATTTTGT
 ACCACTAAGCTTTGTGCACATAATGTAACCATCAAGCTTGTAGTCCCAGCAGGAGGTTTGCATG
 CTTGTAGCCGAAATGGGGCTCGACCCCCCATAGTAGGATACTTGATTTTGCATTCCATTGTGGAC
 15 CTGCAAACTCTACACATAGAGGCTTTGTCTTGCATCTAAACACCTGAGTACAGTGTGTACCTAGAC
 CCTATAGTACGGGAGGACCGTTTGTTCCTCAATAACCCTACATAATAGGCTAGGTGGGCATGCC
 CAATTTGCAAGATCCCAGACTGGGGGTGCGTCTGGGCAGGGTTAGATCCCTGTTAGCTACTGCC
 TGATAGGGTGGTGCTCAACCATGTGTAGTTTAAATTGAGCTGTTTCATATACC (SEQ ID NO: 72)

20 CSFV-IRES:

GTATACGAGGTTAGTTCATTCTCGTATACACGATTGGACAAATCAAAATTATAATTTGGTTCAGGG
 CCTCCCTCCAGCGACGGCCGAAGTGGGCTAGCCATGCCCATAGTAGGACTAGCAAACGGAGGG
 ACTAGCCGTAGTGGCGAGCTCCCTGGGTGGTCTAAGTCCTGAGTACAGGACAGTCGTACGTAGT
 TCGACGTGAGCAGAAGCCCACCTCGAGATGCTACGTGGACGAGGGCATGCCCAAGACACACCTT
 25 AACCCTAGCGGGGGTCTGCTAGGGTGAAATCACACCACGTGATGGGAGTACGACCTGATAGGGC
 GCTGCAGAGGCCCACTATTAGGCTAGTATAAAAATCTCTGCTGTACATGGCAC (SEQ ID NO: 73)

In some preferred embodiments, a viral IRES may be modified to remove stop codons in open reading
 frames. Suitable modified viral IRESs include modified CSFV IRESs. A modified CSFV IRES may for
 30 example comprise the DNA sequence of SEQ ID NO: 96; nucleotides 324 to 696 of SEQ ID NO: 75; or
 nucleotides 596 to 968 of SEQ ID NO: 77. Modified IRESs may be useful in the rolling circle
 translation of the circular RNA as described herein.

The TRIC method is suitable for genes of interest of any length. TRIC is particularly suitable for long
 35 genes of interest. In this context, "long" is generally considered to mean a sequence of at least 500
 nucleotides. Thus, the gene of interest may be at least 100, at least 250, at least 500, at least 1000,
 at least 2000, at least 3000, at least 4000, at least 5000, at least 6000, at least 7000, or at least 8000
 nucleotides in length.

Thus, recombinant nucleic acid molecules for making a circular RNA as described herein may comprise, in the 5' to 3' direction:

- a) an internal guide sequence (IGS),
- b) a ribozyme,
- 5 c) a first portion of an extended anticodon arm (eACA) sequence,
- d) a gene of interest, and
- e) a second portion of the eACA sequence,

wherein a nucleotide in the second portion of the eACA sequence forms a wobble base pair with a nucleotide in the IGS.

In some embodiments, recombinant nucleic acids described herein may comprise, in a 5' to 3' direction:

- a) a first EGS,
- b) a first loop sequence,
- 15 c) an internal guide sequence (IGS),
- d) a ribozyme,
- e) a first portion of an extended anticodon arm (eACA) sequence,
- f) a gene of interest,
- g) a second portion of the eACA sequence,
- 20 h) a P1 extension,
- i) a second loop sequence, and
- j) a second EGS.

Where the gene of interest to be circularized is long (>500 nt), the first and second eACA stem portions may be 15 or 25 base pairs and the loop of the resulting stem-loop structure may be 7 nucleotides.

Further components

The recombinant nucleic acids described herein may further comprises elements to facilitate the transcription or circularization process. For example, the recombinant nucleic acids described herein may further comprise a T7 high efficiency sequence, one or more restriction enzyme cleavage sites, and/or a poly(A) tail. If the recombinant nucleic acid is a DNA template, it may further comprise a T7 promoter sequence. A recombinant nucleic acid may further comprise a nucleotide sequence encoding a self-cleaving peptide to ensure the production of monomeric protein during rolling circle amplification. Such additional elements are known in the art.

In other embodiments, the recombinant nucleic acids described herein may lack stop codons. For example the recombinant nucleic acids may be engineered or modified to remove stop codons in open reading frames. This may facilitate rolling circle amplification.

Methods for producing a circular RNA

Generally, methods for producing a circular RNA comprise the provision of a linear recombinant nucleic acid molecule, such as those described herein, and the splicing of that molecule to generate a circular RNA.

5

Thus, methods for producing a circular RNA may comprise:

- a) providing a recombinant nucleic acid molecule as described herein, and
- b) circularizing the recombinant nucleic acid molecule.

10 It will be appreciated that the term “recombinant nucleic acid molecule” may refer to either a DNA template molecule or an RNA precursor.

Suitable protocols for circularization are known in the art, for example as described in Puttaraju & Been (1992) *Nucleic Acids Res* 20, 5357-5364; and Wesselhoeft, Kowalski & Anderson (2018) *Nat Commun* 9, 2629. The contents of which are incorporated herein by reference.

15

In some cases, the recombinant nucleic acid molecule is a linear DNA template molecule. In a first part of the method, *in vitro* transcription is performed using the DNA template molecule to obtain a linear RNA precursor molecule. Subsequently, a circularization/splicing step is performed with the RNA precursor, to generate a circular RNA molecule.

20

Thus, methods for producing a circular RNA may comprise:

- a) providing a recombinant nucleic acid molecule as described herein,
- b) transcribing the recombinant nucleic acid molecule to produce an RNA precursor, and
- c) circularizing the RNA precursor.

25

In the above methods, some splicing (and consequently circularization) may occur during transcription, known as co-transcriptional splicing. In some cases, it is preferable to suppress splicing during the transcription step, so that splicing and circularization occurs only post-transcriptionally.

30

Thus, the methods described herein may additionally comprise the suppression of splicing (and circularization) during step (b) (the transcription step).

Although magnesium ions are generally essential for *in vitro* transcription and splicing to occur, co-transcriptional splicing/circularization can be suppressed by performing the transcribing step at low concentrations of magnesium ions and with an excess of nucleoside triphosphates (NTPs). NTPs can act to chelate Mg²⁺ during transcription. Thus, the methods described herein may comprise performing step (b) in the presence of NTPs at a concentration of approximately 24 mM, and Mg²⁺ at a concentration less than 18 mM, less than 16 mM, less than 14 mM, or less than 12 mM. In some cases, the transcribing step is performed in the presence of nucleoside triphosphates (NTPs) at a

35

concentration of approximately 24 mM and Mg²⁺ at a concentration of 16 mM or less. If the concentration of NTPs is more or less than 24 mM, the concentration of Mg²⁺ may also vary.

Methods for producing a circular RNA may therefore comprise:

- 5 a) providing a recombinant nucleic acid molecule as described herein,
 b) transcribing the recombinant nucleic acid molecule to produce an RNA precursor, and
 c) circularizing the RNA precursor;
 wherein step (b) is performed in the presence of Mg²⁺ at a concentration of 16 mM or less and
 NTPs at a concentration of approximately 24 mM.

10

In other cases, suppression of splicing during transcription is not necessary to produce circular RNAs. Thus, described herein are methods for producing a circular RNA, comprising:

- a) providing a recombinant nucleic acid molecule as described herein, and
 b) transcribing and circularizing the recombinant nucleic acid molecule,
15 wherein splicing occurs co-transcriptionally.

20

As described above, an advantage of the TRIC method is that it can be used to circularise genes of interest that naturally comprise a sequence that is capable of forming an eACA stem-loop. This enables a circular RNA to be created that does not contain exogenous sequence material (such as exon sequences), which are frequently immunogenic, and so provides clear benefits over existing methods such as PIE.

Thus, methods for producing a circular RNA may comprise:

- 25 a) identifying a gene of interest comprising a sequence capable of forming an eACA stem-loop,
 b) preparing a recombinant nucleic acid molecule comprising, in a 5' to 3' direction: an internal guide sequence (IGS) – a ribozyme – a sequence encoding the gene of interest, and
 c) circularizing the recombinant nucleic acid molecule.

30

Step (b) above may comprise rearranging the gene of interest to place a first portion of the sequence capable of forming an eACA stem-loop at the 5' end, and a second portion of the sequence capable of forming an eACA stem-loop at the 3' end.

35

Suitable genes of interest comprising a sequence capable of forming an eACA stem-loop may be identified by the skilled person using approaches known in the art, including the use of software such as RNAFold (University of Vienna).

Also described herein is the use of any of the recombinant nucleic acid molecules described herein in methods for producing a circular RNA.

Methods for producing a circular RNA may also comprise the expression of the recombinant nucleic acid molecule in a cell, with subsequent circularization being performed in the cell.

Circular RNAs

- 5 Circular RNAs can be generated using the recombinant nucleic acid molecules described herein. The circular RNAs comprise a sequence encoding a gene of interest. One advantage of the described circular RNAs is that in some cases, they do not comprise exogenous splicing sequences, such as the circular RNAs obtained with the PIE method, which typically comprise exogenous exon sequences.
- 10 By “exogenous”, it is meant any sequence information that is not naturally present in the gene of interest. Circular RNAs described herein may not comprise RNA from a circularization agent. For example, circular RNAs described herein may not contain any RNA from a ribozyme, in particular from a group I intron. In some cases, the circular RNAs described herein may comprise only the gene of interest. This in turn reduces the immunogenicity of the circular RNAs compared to circular RNAs
- 15 generated using methods of the art (such as the PIE method). The circular RNAs described herein may be less immunogenic compared to circular RNAs (e.g. those encoding the same GOI) which comprise exogenous exon sequences. By less immunogenic, it is meant that transfection with the circular RNA produces less of an immune response by the host cell, for example, reduced production of cytokines, chemokines or other immune signalling molecules. Suitable methods for determining
- 20 immunogenicity are known in the art. For example, immunogenicity can be determined by measuring the production of immune factors (cytokines, chemokines, etc) in transfected cells for a period of time following transfection.

In the circular RNAs described herein, the sequence encoding the gene of interest may comprise a

25 sequence capable of forming an eACA stem-loop structure as described herein. For example, a sequence capable of forming an eACA stem-loop structure with a stem of at least 5 base pairs, and a loop of 7 nucleotides, wherein the third nucleotide of the loop in the 5' to 3' direction is uracil. The sequence capable of forming an eACA stem-loop structure may be naturally occurring in the gene of interest, or may have been introduced prior to circularization.

30

Also described herein are circular RNAs obtainable by the methods disclosed herein. For example, described herein is a circular RNA obtainable by a method comprising:

- a) providing a recombinant nucleic acid molecule as described herein,
- b) transcribing the recombinant nucleic acid molecule to produce an RNA precursor, and
- 35 c) circularizing the RNA precursor.

Uses of circular RNAs/precursors

The recombinant nucleic acids described herein may be useful in therapeutic methods or as a medicament. Circular RNAs obtained from the recombinant nucleic acids described herein may be

40 used in various ways to exert a therapeutic effect. For example, circular RNAs can act as a “sponge” for micro RNAs (miRNA), in turn preventing or reducing the degradation of a target mRNA by the

miRNA. Circular RNAs may also influence protein trafficking and subcellular protein localisation. The specific therapeutic effect will be determined by the gene of interest, including whether the gene of interest is a coding or noncoding sequence.

- 5 In addition, circular RNAs can also be used to drive expression of a gene of interest in vivo or in vitro. Thus, described herein are methods for expressing a gene of interest in a cell, in which a recombinant nucleic acid molecule as described herein is circularized to provide a circular RNA comprising the gene of interest, and the circular RNA is administered to the cell.

10 Rolling Circle Translation

- Circular RNAs obtained from the recombinant nucleic acid molecules as described herein may be used to drive expression of a gene of interest through rolling circle amplification. Suitable circular RNAs may lack in-frame stop codons. For example, a circular RNA may comprise a modified viral IRES as described herein which lacks in frame stop codons. In some embodiments, the circular RNA may be translated continuously multiple times by a ribosome to generate polyproteins. In other
15 embodiments, a suitable circular RNA may further comprise a DNA sequence encoding a self-cleaving peptide. The self-cleaving peptide causes ribosomal skipping during translation of the circular RNA by a ribosome to generate monomeric proteins. Suitable self-cleaving peptides are well-known in the art and include 2A peptides, such as T2A, P2A, E2A and F2A.

20 Methods of Treatment

- Described herein are methods for treating a disease in a subject, the methods comprising (a) circularizing a recombinant nucleic acid molecule as described herein to provide a circular RNA, and
(b) administering the circular RNA to the subject.

- 25 Also described herein are recombinant nucleic acid molecules such as those described herein for use as a medicament; recombinant nucleic acid molecules as described herein for use in a method of treating a disease in a subject; circular RNA obtained from a recombinant nucleic acid molecule as described herein for use as a medicament; and circular RNA obtained from a recombinant nucleic acid
30 molecule as described herein for use in a method of treating a disease in a subject

Also described herein are pharmaceutical compositions comprising a circular RNA or a recombinant nucleic acid molecule as described herein, and a pharmaceutically acceptable excipient.

- 35 The terms "comprising" or "comprises" may be substituted with the terms "consisting of", "consists of", "consisting essentially of" or "consists essentially of" and vice versa, wherever they occur herein.

It is to be understood that the application discloses all combinations of any of the above aspects and embodiments described above with each other, unless the context demands otherwise. Similarly, the

application discloses all combinations of the preferred and/or optional features either singly or together with any of the other aspects, unless the context demands otherwise.

Modifications of the above embodiments, further embodiments and modifications thereof will be apparent to the skilled person on reading this disclosure, and as such, these are within the scope of the present invention.

All documents and sequence database entries mentioned in this specification are incorporated herein by reference in their entirety for all purposes.

“and/or” where used herein is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example, “A and/or B” is to be taken as specific disclosure of each of (i) A, (ii) B and (iii) A and B, just as if each is set out individually herein.

EXAMPLES

The invention will now be described with reference to the following non-limiting examples.

Materials & Methods

In vitro transcription (IVT) protocol:

DNA templates were linearized and cleaned by phenol:chloroform:isoamyl alcohol extraction. IVT was performed at 50 ng/ul of DNA template, 14 ug/ul of homemade T7 polymerase, 0.04 U/ul of RNase inhibitor (Promega), 6 mM of each NTPs and 1X IVT buffer. For IVT where co-transcriptional splicing is allowed, 1X IVT buffer contains 80 mM HEPES-K (pH7.5), 2 mM spermidine, 40 mM DTT and 24 mM MgCl₂. When co-transcriptional splicing is to be suppressed, the concentration of MgCl₂ in 1X IVT buffer is 14 mM. IVT reactions were incubated at 37 °C for 3-5h and then digested by RNase-free DNase I for 20min. Afterwards, 100 mM EDTA was added to a concentration of 25 mM to clear any precipitation. Then equal volume of 7.5 M lithium chloride is added to precipitate RNAs for 30 min to overnight at -20 °C. Then precipitations were spun down at 13000 rpm/min for at least 20min. RNA pellets were washed by 75% alcohol, air dried and dissolved in DEPC treated H₂O.

Circularization protocol:

3*Flag circular RNAs were circularized during IVT. For co-transcriptional circularization of EGFP, Fluc1, Spike, Cas9 and Factor 8, DNA template digested IVT reactions were supplied with extra GTP at 2 mM concentration and heated at 55 °C for 20min. For full-length precursors, 9 ul of RNAs were firstly denatured at 95 °C for 2 min and then annealed on ice for 3 min. Then annealed RNAs were supplied with 1ul of 10X circularization buffer (500 mM Tris-HCl, pH=7.4, 100 mM MgCl₂, 10 mM DTT, 20 mM GTP) and heated at 55 °C for designated time.

Michaelis-Menten fitting:

Concentration of ribozyme was fixed at 1 μ M and concentration of GTP was varied from 1 μ M to 2000 μ M. At each concentration of GTP, time course of circularization of TRIC-V2 and PIE were monitored. Then time for 50% circularization completeness ($t_{1/2}$) for each sample was calculated and used for estimation of initial circularization speed of each construct at each GTP concentration (V_{obs}). V_{obs} was then plotted against GTP concentration to calculate kinetics parameters of the TRIC-V2.

Calculation of circularization efficiency:

To measure efficiency, ratio of yield and nicking, samples were loaded to native agarose gels, and intensity of full length, circular RNA and nicked RNAs were measured in image J. Efficiency equals the percentage of full-length precursor that has been converted to circular RNAs. Yield is calculated by dividing total RNA with circular RNA. Nicking rate equals the ratio between nicked RNA and circular RNA.

RNase R digestion :

To do RNase R digestion, IVT of TRIC-V1.0 was precipitated by 7.5 M LiCl and dissolved in DEPC treated water. Then 5 μ g of RNA was digested with 10 U of RNase R (antibodies-online) at 37 °C for 15min. Digested RNAs were loaded to denaturing PAGE.

RT-PCR:

Reverse transcriptase and DNA polymerase used here are the SuperScrip IV Reverse Transcriptase (Thermo Fisher) and the Q5 High-Fidelity DNA Polymerase (NEB). Manufacturer's manuals were followed for reverse transcription and PCR. The IVT sample, RNA I and II were used as templates for reverse transcription and PCR using the using primers indicated in Figure 2B.

Circular RNA purification:

A two-step strategy was used to purify circular RNAs. First, circular RNA samples were injected into an SRT-2000 SEC column (Sepax) operated on an AKTA pure system (Cytiva). The running buffer contained 10 mM Tris-HCl (pH=6.5) and 1 mM EDTA. Then fractions containing circular RNAs were pooled, digested by RNase R (antibodies-online, 0.5U/ μ g) at 37 °C for 1 hour and cleaned using the RNA clean & concentrator kit (ZYMO RESEARCH).

Cell transfection, RT-qPCR, Nluc expression and siRNA knocking down:

A549 and HEK 293F cells were cultured in DMEM (with 10% FBS, High Glucose GlutaMAX, Life Technologies Ltd) and Freestyle (Gibco) medias, respectively.

For immunogenicity studies, 200 ng of each RNA was transfected into 100,000 A549 cells in 24-well plates using the MessengerMax transfection reagents (Invitrogen). 24 hours later, cells were harvested, and total RNAs were extracted by TRIzol (Invitrogen) and the RNA clean & concentrator kit. Then, 300 ng of each total RNA was reverse transcribed by the Superscript IV reverse

transcriptase (Invitrogen) using random hexamer plus 20 mer oligo dT (Invitrogen). The reverse transcription products were 20 times diluted and used as templates for qPCR, primers are listed in Table 1 (ViiA 7, Thermo Fisher).

- 5 For Nluc expression and siRNA knocking down studies, 50 ng of circular Nluc and an equal molar amount of linear mRNA were transfected into 10,000 cells in 96-well plates using the MessengerMax transfection reagent. For protein expression, 5ul of cells were taken for luciferase assay using the Nano-Glo® Luciferase Assay System (Promega). For siRNA knocking down, siRNAs were transfected into cells using the RNAiMax (Invitrogen) on day 3. 5 hours after siRNA transfection, expression of
- 10 Nluc was measured using the above kit.

Table 1 Primers used in qPCR

qGAPDH-F	AGCTCATTTCTGGTATGACA	SEQ ID NO: 6
qGAPDH-R	AGGGGAGATTCAGTGTGGTG	SEQ ID NO: 7
qIL6-F	AGCCACTCACCTCTTCAGAAC	SEQ ID NO: 8
qIL6-R	GCCTCTTTGCTGCTTTTCACAC	SEQ ID NO: 9
qCCL5-F	CACAGGTACCATGAAGGTCTC	SEQ ID NO: 10
qCCL5-R	GGTGTCCGAGGAATATGGG	SEQ ID NO: 11
qIFN β -F	TCTAGCACTGGCTGGAATGAG	SEQ ID NO: 12
qIFN β -R	GTTTCGGAGGTAACCTGTAAG	SEQ ID NO: 13

Statistical analysis

- 15 The t-Test-Two-Sample Assuming Unequal Variances model was used for statistical analysis. Details can be found in Figure legends.

Example 1 - TRIC

- 20 The inventors selected the tRNA^{Leu} intron from the cyanobacterium *Anabaena* (Ana) for the initial test of the TRIC approach. This first construct was designated TRIC-V0.

- Ana is short (249 nt) but highly active. The Ana intron divides the leucine transfer RNA (tRNA) into a 34 nt left half and a 51 nt right half (L34/R51) in the anticodon arm (ACA) (Figure 2A-insert labelled "tRNA"). The inventors set out to determine whether using some of the leucine tRNA sequence would
- 25 enable circularization. Of the Leu tRNA anticodon arm, a L15/R30 portion was reserved and joined on either side of a gene of interest (in this case a 3*Flag coding sequence) for circularization. The construct is shown in Figure 2B (SEQ ID NO: 14).

- In vitro transcription was performed, and the samples were loaded on a 12% denaturing
- 30 polyacrylamide gel (PAGE) as described in Puttaraju & Been (1992). The results of the gel are shown in Figure 2C (left column labelled IVT).

Consistent with the high activity of Ana, splicing occurs co-transcriptionally and multiple RNA species could be seen in the 12% gel. Full-length precursor (403 nt) and the spliced linear intron (257 nt) were identified, but not the circular gene of interest (3*Flag). The linear intron and two other major species (I and II) were electroeluted from the 12% gel and loaded onto a 6% PAGE (Figure 2C, right column labelled IVT). Both the major species I and II contained a fast-moving minor species, and both run faster than themselves in the 12% PAGE, indicating that they are circular RNAs. Since the minor bands in I and II run at the same place as linear intron and nicking 3*Flag, respectively, the inventors concluded that species I is circular intron and species II is circular 3*Flag (the gene of interest). To further confirm the circular identity of 3*Flag, the inventors performed reverse transcription followed by PCR (RT-PCR) on the IVT sample and species I and II (Figure 2D). As expected, a 109 bp DNA product was produced from IVT and II but not I. The inventors then cloned this PCR product to a sequencing vector and performed Sanger sequencing. The sequencing result clearly shows that circularization of the 3*Flag gene of interest happened as expected (Figure 2E). These results demonstrate that even with a short IGS, TRIC-V0 is able to circularize gene of interest highly efficiently during the IVT process (co-transcriptional circularization).

Example 2 - TRIC V1

The internal guide sequence (IGS) of the Ana group I intron is short (5 nucleotides), and the inventors speculated that a short IGS might make it difficult to circularize long genes of interest. To compensate for the short IGS, the inventors introduced an extended guide sequence (EGS) at the 5' end of the TRIC, which could form a 20 nucleotide base-paired structure with a corresponding region at the 3' end (Figure 3A). This construct was designated TRIC-V1.

The sequence connecting the IGS and EGS constitutes an internal loop, which the inventors also optimised in TRIC-V1. Three constructs were created, each containing a different loop configuration. In TRIC-V1.0 (SEQ ID NO: 15), the 6 nt sequence upstream of the Ana IGS was retained in the loop sequence, while for V1.1 (SEQ ID NO: 16) and V1.2 (SEQ ID NO: 17) the length of the loop was reduced by introducing base pair interactions between 3 nt at the 3' end of the construct and 3 nt at the 5' end of the loop. This effectively extended the EGS to 23 nt in length (Figure 3A), and reduced the 5' loop length to 3 nucleotides. The 3' loop length was 3 nucleotides in V1.1 and V1.2, whilst the 3' loop length was 5 nucleotides in V1.0.

In vitro transcription was performed for the three variants V1.0, V1.1 and V1.2, and the products were loaded onto a 6% PAGE together with purified circular 3*Flag (the species II from earlier experiments). All three variants were found to circularize the gene of interest (3*Flag) highly efficiently (Figure 3B). The inventors subsequently performed digestion of the in vitro transcription product from V1.0 by an RNase R. As expected, only the circular gene of interest (3*Flag) is resistant to RNase R digestion, while the linear intron is digested (Figure 3C). Compared to the TRIC-V0 variant, the efficiency (that is, the ratio of full-length precursor to circular GOI) of the three V1.1 variants was similar, since full length precursors mostly converted to circular RNAs during in vitro transcription. One advantage of

the V1 variants compared to V0 is that the amount of circular intron is reduced. Reduction of circular introns is beneficial because these circular introns cannot be removed by RNase R digestion. Of the three V1 variants, V1.1 gave the highest ratio of circular 3*Flag to linear intron, so this variant was taken forward for further investigation.

5

The 3*Flag sequence (SEQ ID NO: 18) is 141 nucleotides in length. The inventors also determined the capacity of TRIC-V1 for circularization of long genes of interest. Five new constructs were created with the aim of producing circular CVB3-EGFP (EGFP, 1638 nt) (SEQ ID NO: 19), CVB3-Firefly luciferase (Fluc, 2601 nt) (SEQ ID NO: 20), CVB3-Spike protein of SARS-CoV 2-EGFP (Spike, 5469 nt) (SEQ ID NO: 21), CVB3-spCas9-EGFP (Cas9, 5757 nt) (SEQ ID NO: 22) and CVB3-Factor 8-EGFP (Factor 8, 8706 nt) (SEQ ID NO: 23) (Figure 4A). Longer tRNA sequence (L15/R51), internal homology arms (iHRs) and spacer sequences (polyAC) were used in the TRIC V1.0 construct to produce an identical circular RNA as would be obtained with the PIE Ana construct (Figure 5A, 5B). The sequences of each of the GOIs are shown below.

15

3*Flag:

AAAATCCGTTGACCTTAAACGGTCGTGTGGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTT (SEQ ID NO: 18)

20

EGFP:

AAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGAAACGCAATAGCCGAAAAACAA
 AAAACAAAAAAACAAAAAAACCAAAAAACAAAAACACATTAAAAACAGCCTGTGGGTTGATCCACCCACAGGCCCA
 TTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACTGTAAGTAAAGTAA
 CACACACCGATCAACAGTCAGCGTGGCAGCAGCCACGTTTGGTCAAGCACTTCTGTTACCCCGGACTGAGTATCAA
 TAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAG
 TTGCAGAGTGTTTCGCTCAGCACTACCCAGTGATGAGTCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGG
 CGGTGGCTGCGTTGGCGGCCTGCCATGGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTG
 AGCTAGTTGGTAGTCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGT
 GTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTGGCTGCTTA
 TGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTG
 TTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACATTACAATTCAATTGTTAAGTTGAATACAGCAAAATGGTGAGCAA
 GGGCGAGGAGCTGTTACCCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCG
 TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCC
 GTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAG
 CACGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCCTCTTCTTCAAGGACGACGGCAACTAC
 AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGA
 GGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
 GAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCA
 GAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAG
 ACCCAACGAGAAGCGGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAG

40

AAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGGAACGCAATAGCCGAAAAACAA
AAAACAAAAAAACAAAAAAACCAAAAAACAAAAACACATTAAACAGCCTGTGGGTTGATCCCACCCACAGGCCCA
TTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACTGTAACCTAGAAGTAA
CACACACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAA
5 TAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAG
TTGCAGAGTGTTCGCTCAGCACTACCCCACTGTAGATCAGGTGCGATGAGTCACCGCATTCCCCACGGGCGACCGTGG
CGGTGGCTGCGTTGGCGGCCTGCCATGGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTG
AGCTAGTTGGTAGTCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGT
GTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTGGCTGCTTA
10 TGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAATAAGAGCTATTATATATCCCTTTG
TTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAACATTACAATTCAATTGTTAAGTTGAATACAGCAAAATGGGGATTCT
GCCTAGCCCCGGCATGCCCGCCCTCCTGAGCCTGGTGAGCCTGCTGAGCGTCCTGTTGATGGGGTGCGTGGCCGAGA
CCGGCACACAGTGCCTGAACCTGACAATAAGACAGCTGCCCCAGCCTACACCAACAGCTTTACTAGGGGAGTGT
ACTACCCTGATAAGTGCTTCAGGAGCAGTGTGCTGCACAGCACCCAGGACCTCTTTTTGCCCTTTTTCAGTAACGTGAC
15 CTGGTTCACGCTATCCACGTGAGCGGCACAAACGGCACAAACGGTTCGATAACCCAGTGTGCCCTTCAACGACGG
CGTCTACTTCGCCAGCACTGAGAAGTCCAACATTATCAGGGGGTGGATCTTCGGGACAACCCTGGACAGCAAGACACA
GAGCCTCCTGATTGTGAACAACGCCACCAACGTGGTCATTAAGGTCTGCGAGTTCAGTTCTGCAACGACCCCTTTCTG
GGGGTGTACTACCACAAGAACAACAAGTCTGGATGGAGTCCGAGTTTCGGGTGTACTCCAGCGCCAACAACCTGCACA
TTTGAGTACGTGAGCCAGCCCTTCTTGATGGATCTGGAGGGCAAGCAGGGGAACTTTAAGAACCTTCGGGAGTTTCGTCT
20 TTAAGAACATCGACGGGTACTTTAAGATCTACAGCAAGCACACCCCTATTAACCTGGTCCGCGATCTGCCCCAGGGCTT
CAGCGCCCTGGAGCCACTGGTGGACCTCCCCATCGGGATCAACATTACTAGGTTCCAGACACTCCTGGCCCTGCACCG
GAGCTACCTGACTCCCGCGATAGCTCCAGCGGGTGGACTGCCGGGGCCCGCCCTACTACGTGCGCTACCTGCAGC
CCAGGACATTCTGCTCAAGTACAACGAGAACGGCACCATCACTGATGCCGTGGACTGCGCTCTCGATCCCCTGAGCG
AGACAAAGTGACCCCTGAAGTCCTTACCCTGGAGAAGGGGATCTACCAGACATCTAATTCAGGGTGCAGCCCACTG
25 AGTCCATTGTGAGATTTCTAACATCACCAACCTGTGCCCTTCGGCGAGGTCTTTAACGCCACACGGTTCGCCAGCGT
CTACGCTTGAACAGGAAGCGGATCTCAACTGCGTGGCCGACTACAGCGTGCTGTACAACAGCGCCTCCTTCAGCAC
CTTTAAGTGCTACGGCGTGAGCCCCACTAAGCTGAACGACCTGTGCTTCACTAACGTGTACGCTGATAGCTTCGTGATT
AGGGGCGATGAGGTCCGGCAGATTGCCCCGGCCAGACAGGGAAGATCGCTGACTACAACATAAGCTGCCCCGATGA
CTTCACAGGGTGTGTGATCGCCTGGAACAGCAACAACCTGGATGCAAGGTGCGGGGGAACATAACTACCTGTACAG
30 GCTGTTCAAAAAGTCCAACCTGAAGCCCTTCGAGCGGGATATCAGCACTGAGATCTACCAGGCCGGGAGCACCCCTG
CAACGGCGTCGAGGGCTTTAACTGCTACTTTCCCTCCAGAGCTACGGCTTCAACCCACCAACGGCGTGGGGTACCA
GCCCTACCGGGTCGTGGTGCTTAGCTTCGAGCTCCTGCACGCCCCTGCAACCGTGTGCGGCCCCAAGAAGAGCACAAA
CCTCGTGAAGAACAAGTGCGTGAACCTCAACTTAACGGGCTGACCGGCACCGGCGTGCTGACTGAGAGTAACAAGAA
GTTTCTCCCCTTCAGCAGTGCGGGAGGGATATTGCCGACACTACTGATGCCGTGCGGACCCCTCAGACCCTTGAGAT
35 TCTGGACATCACTCCCTGCTCCTTCGGCGGAGTGTCCGTGATCACACCAGGGACCAACACTAGCAACCAGGTCGCCGT
CTTGATACAGGACGTGAACCTGCACCGAGGTGCCCGTGCCATCCACGCCGATCAGCTGACCCCAACTTGGCGGGTGTA
CTCCACCGGGAGCAACGTGTTTCAGACACGGGCCGGGTGCCTGATCGGGGCTGAGCATGTCAACAACAGCTACGAGT
GCGATATTCTATCGGCGCCGGGATTTGCGCCAGCTACCAGACACAGACAAACAGCCCTGGCTCGGCCAGCTCCGTGCG
CCAGCCAGAGCATCATCGCCTACACAATGTCTTGGGGGCCGAGAACAGCGTCGCCTACAGCAACAACAGCATTGCCA
40 TCCCCACCAACTTCACAATCAGCGTCACAACCGAGATCCTGCCCCTGAGCATGACTAAGACATCCGTGGATTGCACAAT
GTACATCTGCGGCGACTCCACAGAGTGCAAGTAACCTGCTGCTCCAGTACGGGTCTTTTGTACCCAGCTGAACAGAGC

CTTGACCGGCATTGCCGTGGAGCAGGACAAGAACACCCAGGAGGTGTTTGGCCAGGTGAAGCAGATCTACAAGACACC
CCCCATAAAGGACTTCGGGGGCTTTAACTTCTCCAGATCTTGCCCGACCCAGCAAGCCCTCCAAGCGGTCATTCATT
GAGGACCTGCTGTTTAAACAAGGTGACCCTCGCTGATGCAGGGTTTATCAAGCAGTACGGCGATTGCCTGGGGGACATT
GCCGCTAGGGACTTGATCTGCGCCCAGAAGTTCAACGGCCTGACCGTGCTCCCCCCCCCTGTTGACAGACGAGATGATC
5 GCCCAGTACACTAGCGCTCTGCTGGCCGGCACTATCACCAGCGGCTGGACCTTTGGGGCCGGCGCCGCCCTCCAGAT
TCCATTGCGCATGCAGATGGCCTACCGCTTCAACGGCATCGGCGTGACTCAGAACGTGCTGTACGAGAACCAGAAAGCT
GATTGCCAACCAAGTTCAACTCCGCCATCGGCAAGATCCAGGACAGCCTGTCCAGCACAGCATCCGCCCTGGGGAAGTT
GCAGGATGTGGTGAACCAGAACGCCAGGCCCTGAACACACTGGTGAAGCAGCTGAGCAGCAACTTTGGCGCCATCAG
TAGCGTCCTGAACGACATCCTGTCACGGCTTGATCCACCAGAGGCCGAGGTCCAGATCGATCGGCTGATTACAGGCCG
10 GCTGCAGAGTCTCCAGACCTACGTACCCAGCAGCTGATCCGGGCTGCCGAGATTAGAGCCTCCGCCAATCTCGCCGC
CACAAAGATGAGTGAATGCGTCCTGGGCCAGAGCAAAAGGTCGACTTTTGCGGGAAGGGGTACCACCTGATGTCCTT
TCCCCAGTCCGCCCTCACGGAGTGGTGTTCCTCCATGTCACATACGTGCCTGCCAGGAGAAGAACTTCACTACAGC
CCCCGCCATTTGCCACGATGGGAAGGCCCATTTCCCTAGAGAGGGCGTGTTCTGTCCAACGGGACCCACTGGTTCTGT
GACTCAGCGGAACTTTTATGAACCCAGATTATCACAACCGATAACACATTTGTCTCCGGCAACTGTGATGTGGTCATCG
15 GCATCGTGAACAACACCGTGTACGACCCACTGCAGCCTGAGCTTGACAGCTTTAAGGAGGAGCTGGACAAGTACTTTAA
GAACCACACCAGCCCTGATGTGGACCTTGGCGACATCAGCGGCATTAACGCCAGCGTGGTGAACATTCAGAAGGAGAT
CGATAGACTGAACGAGGTGGCCAAGAAGTTGAACGAGTCACTGATTGACCTGCAGGAGCTGGGGAAGTACGAGCAGTA
CATTAAGGGGTCCGGGAGGGAGAACCTGTACTTTAGGGGGGGGGGGGTCCGGCTACATTCGAGGCCCAAGG
GATGGCCAGGCCTACGTGAGGAAGGATGGGGAGTGGGTGCTGCTGAGCACCTTCTTGGGCAGCGGCAGCAGTATGGT
20 GAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT
TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAG
CTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCACTGCTTACGCCGCTACCCGACCATG
AAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTT
25 CAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAA
GCAGAAGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA
CCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGA
GCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATG
GACGAGCTGTACAAGTAAAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCGAGACGCTACGGACTT (SEQ
30 ID NO: 26)

spCas9-EGFP

AAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGAAACGCAATAGCCGAAAAACAA
AAAACAAAAAAACAAAAAAACCAAAAAACAAAAACATTAACACAGCCTGTGGGTTGATCCACCCACAGGCCCA
35 TTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACTGTAACCTAGAAGTAA
CACACACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAA
TAGACTGCTCAGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAG
TTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTGATGAGTACCGCATTCCCCACGGGCGACCGTGG
CGGTGGCTGCGTTGGCGGCCTGCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTG
40 AGCTAGTTGGTAGTCCTCCGGCCCCGTAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGT
GTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTGGCTGCTTA

TGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTG
TTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGATAAGAA
ATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGCTAAAAA
AGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGACAGTGGAGAG
5 ACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAGG
AGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGAAGAAGACA
AGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACTATCTATCATC
TGCGAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTGGCCTTAGCGCATATGATTAAGTTTC
GTGGTCATTTTTTATTGATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCT
10 ACAATCAATTATTTGAAGAAAACCTATTAACGCAAGTGAGTAGATGCTAAAGCGATTCTTCTGCACGATTGAGTAAAT
CAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCA
TTGGGTTTGACCCCTAATTTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAGATACTTACGATG
ATGATTTAGATAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTGGCAGCTAAGAATTTATCAGATGCTAT
TTACTTTTACAGATATCCTAAGAGTAAATACTGAAATACTAAGGCTCCCCTATCAGCTTCAATGATTAACGCTACGATGA
15 ACATCATCAAGACTTGACTCTTTTAAAGCTTTAGTTGACAACAACCTCCAGAAAAGTATAAAGAAATCTTTTTTATCAA
TCAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAA
AAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACG
GCTCTATTCCCATCAAATTCATTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAG
ACAATCGTGAGAAGATTGAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCGTT
20 TTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCT
CAATCATTTATTGAACGCATGACAACTTTGATAAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGTTTGCTTTATG
AGTATTTTACGTTTTATAACGAATTGACAAAGTCAAATATGTTACTGAAGGAATGCGAAAACAGCATTCTTTTACGGTG
AACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCA
AAAAATAGAATGTTTTGATAGTGTGAAATTTGAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATGATTT
25 GCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACC
TTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAGCTT
AAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGCAATCTGGCAAAAC
AATATTAGATTTTTTGAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAA
GAAGACATTCAAAAAGCACAGTGTCTGGACAAGGCGATAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGC
30 TATTA AAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAAATAT
CGTTATTGAAATGGCACGTGAAAATCAGACAACCTCAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAA
GAAGGTATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAATGAAAAGCTCTAT
CTCTATTATCTCCAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAAGTATTATGATGTCGAT
CACATTGTTCCACAAAGTTTCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAAATCGTGGTAAA
35 TCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACTATTGGAGACAACCTCTAAACGCCAAGTTAATCAC
TCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCC
AATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAAAT
GATAAACTTATTGAGAGGTTAAAGTGATTACCTTAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATA
AAGTACGTGAGATTAACAATTACCATCATGCCATGATGCGTATCTAAATGCCGTCGTTGGAACCTGCTTTGATTAAGAAAT
40 ATCCAAAACCTGAATCGGAGTTTGCTATGGTGATTATAAAGTTTATGATGTTTCGTAAAATGATTGCTAAGTCTGAGCAAG
AAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATTACACTTGCAAATGG

AGAGATTGCGAAACGCCCTCTAATCGAACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCC
ACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGGATTCTCCAAGG
AGTCAATTTTACCAAAAAAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTG
ATAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGAAATCGAAGAAGTTAAATCCGTAAA
5 GAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAG
GAAGTTAAAAAAGACTTAATCATTAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCT
AGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTATTTAGCTAGTCATTAT
GAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGAT
TATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAA
10 CATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTACGTTGACGAATCTTGGAGCTCCCGCTGCT
TTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAAT
CCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCAGCTAGGAGGTGACGGCAGCGGCAGCAGTATGGTGAGCAA
GGGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG
TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCC
15 GTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCACTGCTTACGCCGCTACCCCGACCACATGAAGCAG
CACGACTTCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGA
GGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCA
20 GAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCAGTCCGCCCTGAGCAAAAG
ACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAG
CTGTACAAGTAAAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCGAGACGCTACGGACTT (SEQ ID NO:
27)

25 **Factor 8-EGFP**
AAAATCCGTTGACCTTAAACGGTCGTGTGGGTCAAGTCCCTCCACCCCCACGCCGAAACGCAATAGCCGAAAAACAA
AAAACAAAAAAACAAAAAAACCAAAAAACAAAAACACATTAAAAACAGCCTGTGGGTTGATCCACCCACAGGCCCA
TTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACTGTAACCTAGAAGTAA
CACACACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAA
30 TAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAG
TTGCAGAGTGTTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGG
CGGTGGCTGCGTTGGCGGCCTGCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTG
AGCTAGTTGGTAGTCCTCCGGCCCCGTAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGT
GTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTA
35 TGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTG
TTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAACATTACAATTATTGTTAAGTTGAATACAGCAAAATGCAATAGA
GCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCATTCTGCTTTAGTGCCACCAGAAGATACTACCTGGGTGCAGTGGAAC
TGTCATGGGACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTCCTCCTAGAGTGCCAAAATCTTTT
CCATTCAACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAATTCACGGATCACCTTTTCAACATCGCTAAGCCAAG
40 GCCACCCTGGATGGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTTATGATACAGTGGTCATTACACTTAAGAACATG
GCTTCCCATCCTGTCAGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGAC

CAGTCAAAGGGAGAAAAGATGATAAAGTCTTCCCTGGTGGAAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAAT
GGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGG
CCTCATTGGAGCCCTACTAGTATGTAGAGAAGGGAGTCTGGCCAAGGAAAAAGACACAGACCTTGCACAAATTTATACTA
CTTTTGTCTGTATTTGATGAAGGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATC
5 TGCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCAGGTCTGATTGGATGCCACAGG
AAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCTGAAGTGCCTCAATATTCCTCGAAGGTCACACATTTCT
TGTGAGGAACCATCGCCAGGCGTCCTTGAAATCTCGCCAATAACTTTCTTACTGCTCAAACACTCTTGATGGACCTTG
GACAGTTTCTACTGTTTTGTATATCTCTTCCACCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCA
GAGGAACCCCACTACGAATGAAAAATAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGT
10 GGTGAGGTTTGTATGATGACAACTCTCCTTCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCCTAAAACCTGGGTAC
ATTACATTGCTGCTGAAGAGGAGGACTGGGACTATGCTCCCTTAGTCTCGCCCCGATGACAGAAGTTATAAAAGTCA
ATATTTGAACAATGGCCCTCAGCGGATTGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTA
AGACTCGTGAAGCTATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTTGATT
ATATTTAAGAATCAAGCAAGCAGACCATATAACATCTACCTCACGGAATCACTGATGTCCGTCCTTTGTATTCAAGGAG
15 ATTACCAAAGGTGTAAACATTTGAAGGATTTTCCAATTCTGCCAGGAGAAATATTCAAATATAAATGGACAGTGACTGT
AGAAGATGGGCCAACTAAATCAGATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAG
CTTCAGGACTCATTGGCCCTCTCCTCATCTGCTACAAAGAATCTGTAGATCAAAGAGGAAACCAGATAATGTCAGACAAG
AGGAATGTCATCCTGTTTTCTGTATTTGATGAGAACCGAAGCTGGTACCTCACAGAGAATATACAACGCTTTCTCCCCAA
TCCAGCTGGAGTGCAGCTTGAGGATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTTGATA
20 GTTTGCAGTTGTCAGTTTGTTCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCACAGACTGACTTCTTTCT
GTCTTCTTCTGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTCACCTATTCCCATTTCTCAGGAGAACT
GTCTTCATGTGATGAAAACCCAGGTCTATGGATTCTGGGGTGCCACAACCTCAGACTTTCCGAACAGAGGCATGACCG
CCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGATATTTAGCATACTTGC
TGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCAGAATTCAGACACCCCTAGCACTAGGCAAAAGCAATTTAAT
25 GCCACCACAATTCCAGAAAATGACATAGAGAAGACTGACCCTTGGTTTGCACACAGAACACCTATGCCTAAAATACAAA
TGTCTCCTCTAGTGATTTGTTGATGCTCTTGCGACAGAGTCCTACTCCACATGGGCTATCCTTATCTGATCTCCAAGAAG
CCAAATATGAGACTTTTTCTGATGATCCATCACCTGGAGCAATAGACAGTAATAACAGCCTGTCTGAAATGACACACTTC
AGGCCACAGCTCCATCACAGTGGGGACATGGTATTTACCCCTGAGTCAGGCCTCCAATTAAGATTAAATGAGAACTGG
GGACAACCTGCAGCAACAGAGTTGAAGAAACTTGATTCAAAGTTTCTAGTACATCAAATAATCTGATTTCAACAATTCCAT
30 CAGACAATTTGGCAGCAGGTAAGTACTGATAATACAAGTTCTTAGGACCCCCAAGTATGCCAGTTCATTATGATAGTCAATTA
GATACCACTCTATTTGGCAAAAAGTCATCTCCCTTACTGAGTCTGGTGGACCTCTGAGCTTGAAGAAAAATAATGA
TTCAAAGTTGTTAGAATCAGGTTAATGAATAGCCAAGAAAGTTCATGGGGAAAAAATGTATCGTCAACAGAGAGTGGTA
GGTTATTTAAAGGGAAAAAGAGCTCATGGACCTGCTTTGTTGACTAAAGATAATGCCTTATTCAAAGTTAGCATCTCTTTGT
TAAAGACAAACAAAACCTTCCAATAATTCAGCAACTAATAGAAAGACTCACATTGATGGCCCATCATTATTAATTGAGAATA
35 GTCCATCAGTCTGGCAAAATATATTAGAAAGTGACACTGAGTTAAAAAAGTGACACCTTTGATTCATGACAGAATGCTTA
TGGACAAAAATGCTACAGCTTTGAGGCTAAATCATATGTCAAATAAACTACTTCATCAAAAAACATGGAAATGGTCCAAC
AGAAAAAGAGGGCCCCATTCCACCAGATGCACAAAATCCAGATATGTCGTTCTTTAAGATGCTATTCTTGCCAGAATCA
GCAAGGTGGATACAAAGGACTCATGGAAAGAAGTCTCTGAACTCTGGGCAAGGCCCCAGTCCAAAGCAATTAGTATCCT
TAGGACCAGAAAAATCTGTGGAAGGTCAGAATTTCTTGTCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTTACA
40 AAGGACGTAGGACTCAAAGAGATGGTTTTTCCAAGCAGCAGAAACCTATTTCTTACTAAGTTGGATAATTTACATGAAAT
AATACACACAATCAAGAAAAAAAATTCAGGAAGAAATAGAAAAAGAGGAAACATTAATCCAAGAGAATGTAGTTTTGCCT

CAGATACATACAGTGACTGGCACTAAGAATTTTCATGAAGAACCTTTTCTTACTGAGCACTAGGCCAAAATGTAGAAGGTTCC
ATATGACGGGGCATATGCTCCAGTACTTCAAGATTTTAGGTCATTAAATGATTCAACAAATAGAACAAAGAAACACACAG
CTCATTTCTCAAAAAAGGGGAGGAAGAAAACCTTGAAGGCTTGGGAAATCAAACCAAGCAAATTGTAGAGAAATATGCA
TGCACCACAAGGATATCTCCTAATACAAGCCAGCAGAATTTTGTACGCAACGTAGTAAGAGAGCTTTGAAACAATTACAG
5 ACTCCCACTAGAAGAAACAGAACTTGAAAAAGGATAATTGTGGATGACACCTCAACCCAGTGGTCCAAAAACATGAAAC
ATTTGACCCCGAGCACCTTCACACAGATAGACTACAATGAGAAGGAGAAAGGGGCCATTACTCAGTCTCCCTTATCAGA
TTGCCTTACGAGGAGTCATAGCATCCCTCAAGCAAATAGATCTCCATTACCCATTGCAAAGGTATCATCATTTCCATCTAT
TAGACCTATATATCTGACCAGGCTCCTATTCCAAGACAACCTTCTCATCTTCCAGCAGCATCTTATAGAAAGAAAGATTCC
TGGGGTCCAAGAAAGCAGTCATTTCTTACAAGGAGCCAAAAAAATAACCTTTCTTTAGCCATTCTAACCTTGGAGATGA
10 CTGGTGATCAAAGAGAGGTTGGCTCCCTGGGGACAAGTGCCACAAATTCAGTCACATACAAGAAAGTTGAGAACACTGT
TCTCCCGAAACCAGACTTGCCCAAACATCTGGCAAAGTTGAATTGCTTCCAAAGTTCACATTTATCAGAAGGACCTAT
TCCCTACGGAACTAGCAATGGGTCTCCTGGCCATCTGGATCTCGTGGAAGGGAGCCTTCTTCAGGGAACAGAGGGAG
CGATTAAGTGGAATGAAGCAAACAGACCTGGAAAAGTTCCCTTTCTGAGAGTAGCAACAGAAAGCTCTGCAAAGACTCC
CTCCAAGCTATTGGATCCTCTTGCTTGGGATAACCACTATGGTACTCAGATACCAAAGAAGAGTGGAATCCCAAGAGA
15 AGTCACCAGAAAAACAGCTTTTAAAGAAAAGGATACCATTTTGTCCCTGAACGCTTGTGAAAGCAATCATGCAATAGCA
GCAATAATGAGGGACAAAATAAGCCCCGAAATAGAAGTCACCTGGGCAAAGCAAGGTAGGACTGAAAGGCTGTGCTCTC
AAAACCCACAGTCTTGAAACGCCATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTAT
GATGATACCATATCAGTTGAAATGAAGAAGGAAGATTTTGACATTTATGATGAGGATGAAAATCAGAGCCCCCGCAGCTT
TCAAAAGAAAACACGACACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCC
20 TAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACTGATGGCTCCTTTACT
CAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGGCCATATATAAGAGCAGAAGTTGAAGATAATAT
CATGGTAACTTTAGAAATCAGGCCTCTCGTCCCTATTCTTCTATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCA
AGGAGCAGAACCTAGAAAAAATTTGTCAAGCCTAATGAAACCAAACCTTACTTTTGAAAGTGCAACATCATATGGCAC
CCACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCCTG
25 ATTGGACCCCTTCTGGTCTGCCACACTAACCACTGAACCCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTC
TGTTTTTACCATCTTTGATGAGACCAAAGCTGGTACTTCACTGAAAATATGGAAAGAACTGCAGGGCTCCCTGCAAT
ATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGG
CTTAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATCCATTCTATTCAAT
TCAGTGGACATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTATCCAGGTGTTTTTGAG
30 ACAGTGGAATGTTACCATCCAAAGCTGGAATTTGGCGGGTGAATGCCTTATTGGCGAGCATCTACATGCTGGGATGA
GCACACTTTTTCTGGTGTACAGCAATAAGTGTGAGTCTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAAGATT
ACAGCTTCAGGACAATATGGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCCGGATCAATCAATGCCTGGAGCA
CCAAGGAGCCCTTTTCTGGATCAAGGTGGATCTGTTGGCACCAATGATTATTCAGGCATCAAGACCCAGGGTGCCCCG
TCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAG
35 GAAATCCACTGGAACCTTAATGGTCTTCTTTGGCAATGTGGATTCTCTGGGATAAAACACAATATTTTAAACCCTCCAA
TTATTGCTCGATACATCCGTTTGCACCCAACTCATTATAGCATTTCGAGCACTCTTCGCATGGAGTTGATGGGCTGTGAT
TTAAATAGTTGCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCCTACTTTAC
CAATATGTTTGCACCTGGTCTCCTTCAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATGCCTGGAGACCTCAGGTG
AATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACAGGAGTAACTACTCAGGGAGTAAAT
40 CTCTGCTTACCAGCATGTATGTGAAGGAGTTCCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTTTCAG
AATGGCAAAGTAAAGGTTTTTTCAGGGAAATCAAGACTCCTTCACACCTGTGGTGAAGTCTCTAGACCCACCGTTACTGAC

TCGCTACCTTCGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCACA
 GGACCTCTACGGCAGCGGCAGCAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGGCCATCCTGGTCG
 AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG
 ACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCCGTGCCCTGGCCACCCTCGTGACCACCTGACCTACGGCGTG
 5 CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAG
 GAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
 GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAA
 CAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAG
 GACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
 10 CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCTGT
 GACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAACAACAAACGGCTATTATGCGTT
 ACCGGCGAGACGCTACGGACTT (SEQ ID NO: 28)

After in vitro transcription, some samples were subject to an additional 20 minute post-transcriptional
 15 circularization protocol. Extra GTP was supplied to each reaction mixture at a final concentration of 2
 mM. The reaction mixtures were heated at 55 °C for 20 minutes to initiate circularization as previously
 described (Wesselhoeft, R. A., Kowalski, P. S. & Anderson, D. G. *Nat Commun* 9, 2629, (2018)). All
 the samples were loaded onto a 0.8% agarose gel and run for 1 hour 45 minutes. Results are shown
 in Figure 4B, with expected band positions for circular products labelled with black circles.

20 TRIC V1 can circularize long genes of interest co-transcriptionally, as indicated by the multiple RNA
 species observed for each construct in Figure 4B, without the post-transcriptional circularization
 protocol. During in vitro transcription, TRIC-V1 produces more splicing products than the PIE
 approach for all the genes of interest, suggesting that the co-transcriptional circularization efficiency of
 25 TRIC-V1 is higher than PIE. However, co-transcriptionally produced circular RNAs are heavily nicked,
 since circular RNAs for EGFP and Fluc1 are nicked at over 50% and circular RNAs for Spike, Cas9
 and Factor 8 were difficult to detect. The additional 20 minute post-transcriptional circularization did
 increase the amount of splicing product, but mainly by increasing the amount of nicking. Co-
 transcriptional nicking for the short GOI 3*Flag, but fatal for the long GOIs used in this experiment.
 30 Thus, the inventors sought to identify a protocol that could be used to produce long circular GOIs.

Example 3 - Post-transcriptional circularization

Mg²⁺ is believed to be the major source of circular RNA nicking (Wesselhoeft et al. (2018)). Since
 Mg²⁺ is essential for both in vitro transcription and circularization, the inventors speculated that
 35 reasonable ways to reduce nicking of the circular RNAs would be to suppress co-transcriptional
 circularization and to increase the speed of post-transcriptional circularization.

The inventors tested whether an excess of nucleoside triphosphates (NTPs) could be used to chelate
 Mg²⁺ during transcription in order to suppress co-transcriptional splicing. The concentration of NTPs
 40 was fixed at 24 mM and the concentration of Mg²⁺ was varied from 6 mM to 24 mM. As shown in

Figure 6A, when the concentration of Mg^{2+} was lower than 16 mM, the in vitro transcription products were mainly full length precursors, suggesting that co-transcriptional circularization was indeed suppressed. Consequently, for subsequent studies in vitro transcription was performed using 24 mM NTPs and 14 mM Mg^{2+} . Following transcription, full length precursors were subject to circularization for 0, 10, 20 or 40 minutes, and the reaction products analysed in a 0.8% agarose gel.

As shown in Figure 6B, using the post-transcriptional circularization protocol circular RNAs were now obtained for Spike, Cas9 and Factor 8 (as well as for EGFP and Fluci), and the splicing products were much less complex. Longer circularization times did increase the conversion from full length precursor to circular/nicked GOIs, but didn't always increase the amount of circular RNA, particularly for over 20 minutes of circularization. Consequently, the 20 minute circularization protocol was chosen to measure the efficiency of TRIC-V1 constructs. As shown in Figure 6C, the efficiency of TRIC-V1 was higher than 80% for the long GOIs, demonstrating that TRIC-V1 can efficiently circularize GOIs as long as >8000 nt. In addition, the efficiency of TRIC-V1 was at least as high as that observed for the PIE Ana approach (Figure 6D).

Example 4 - TRIC-V2

Both the TRIC-V1 and the PIE rely on native exon sequences, namely the tRNA sequence in the case of the Ana, to work. These native exon sequences will unavoidably be part of resulting circular GOIs (as shown in Figures 5A and 5B), although their inclusion is associated with certain drawbacks. Firstly, the native exon sequences are immunogenic (Liu, C. X. et al. Mol Cell 82, 420-434 e426 (2022)), which hinders the biomedical applications of circular RNAs produced by group I intron-based methods such as PIE. Secondly, the inclusion of the native exon sequences means that circularization sites must be positioned at untranslated regions (UTRs) of protein coding circular RNAs. Since UTRs are normally highly structured internal ribosome entry sites (IRESs), extra unwanted spacer sequences are needed to ensure the group I intron and remaining exons are separated from the IRESs. Finally, tens of thousands of natural circular RNAs exist whose functions are waiting to be deciphered, but group I intron-based methods are not favourable for this field because of the unwanted spacer and native exon sequences. To overcome these limitations, the inventors sought to provide an improved version of TRIC.

To reduce the dependence of circularization on tRNA sequences, length of left arm and right arm of Ana tRNA of the TRIC-V1.0 was systematically tested (Figure 7). The inventors first reduced the tRNA sequence of TRIC to just 5 nucleotides – 3 in the left arm and 2 in the right (L3/R2). This construct was designated TRIC-V1.30 (SEQ ID NO: 29) (see Figure 7B for an overview of the L/R configuration of the constructs used in this experiment). The aim of reducing the tRNA sequence to such a short length was to only allow the formation of the P1 and P10 structures. Following in vitro transcription, the reaction product was loaded to a 6% denaturing PAGE. The intron band was strongly detected but the circular 3*Flag band was very weak, indicating that circularization was not the major reaction (Figure 8B). Since the function of the native tRNA exon sequences are to bring RNA sequences into

proximity around the circularization site, the inventors sought to provide an interaction between the sequence 3' of P10 and the EGS (V1.31, SEQ ID NO: 30). The circular 3*Flag band was clearly detected with V1.31, but the circularization efficiency was not comparable with TRIC V1.0 (Figure 8B).

To improve the circularization efficiency, the inventors restored the left arm (L15/R2 - V1.32 (SEQ ID NO: 31)) or the right arm (L3/R30 - V1.33 (SEQ ID NO: 32)) of the tRNA sequence. Restoration of the right arm (V1.33) restored most of the circularization efficiency. Starting from this L3/R30 configuration in V1.33, the length of the right arm was reduced gradually in constructs V1.34 (L3/R25) (SEQ ID NO: 33), V1.35 (L3/R16) (SEQ ID NO: 34) and V1.36 (L3/R9) (SEQ ID NO: 35). The minimum length for the right arm was identified as 9 nucleotides (V1.36, L3/R9). Subsequently, the right arm was fixed at 9 nucleotides and an 8 nt left arm was tested (V1.39, L8/R9 (SEQ ID NO: 38)). The L8/R9 construct fully restored circularization efficiency to the V1.0 level. The requirement of the tRNA sequence in V1.39 has therefore been reduced to a 17-nucleotide structure which mimics the anticodon arm (ACA) found in the Ana tRNA.

The inventors sought to determine whether the group I intron activity could be replicated based on structure, rather than specific nucleotide sequences. To that end, the inventors reversed the sequence of the ACA stem, whilst keeping the sequence of the ACA loop the same (construct designated V2.0 (SEQ ID NO: 39)) and found the circularization efficiency is at least as high as V1.0 and V1.39. With V2.0, the requirement on tRNA sequence is reduced to 5nt (L-CTT/R-AA). A 5 nt sequence cannot be specific to any species, therefore TRIC-V2 doesn't rely on the original bacterial tRNA sequence.

Subsequently, the inventors also identified that the stem length could be increased to 15 base pairs (V2.1 (SEQ ID NO: 40)) without abolishing circularization efficiency. In this construct, the only tRNA sequence present is the L-CTT/R-AA sequence, and the remaining base pairs were not derived from the tRNA sequence. The longer stem is advantageous since it may enable better circularization of longer GOIs.

The inventors subsequently reversed the sequence of the P1 and P10 regions (from L-CTT/R-AA to L-GAT/R-TT), except for the uracil which forms a wobble base pair with the internal guide sequence (construct designated V2.2 (SEQ ID NO: 41)). Again, circularization efficiency was found to be unaffected. In this way, the inventors have shown that the IGS requires only a GU wobble base pair for circularization and is not otherwise dependent on sequence.

Based on the above, the inventors have determined that if an anticodon arm-like structure (eACA) which consists of a 7 nt loop with a uracil as the third nucleotide and a ≥ 5 base pair stem can be found in a gene of interest, this gene of interest can be circularized efficiently without introducing any unwanted sequence (Figure 8). If the circularization site is placed in a coding sequence, the site of choice can be extended because of codon redundancy. Potentially, only a few nt mutations would

need to be made without affecting the peptide sequence. If the circularization site is placed in a noncoding RNA or UTR of a protein coding circular RNA, the circularization site can be assembled by introducing as few as 5 extra nucleotides to create an eACA.

5 Example 5 - TRIC-V2 and long GOIs

The TRIC-V2 construct was then tested on three protein coding circular RNAs, including circular T2A-EGFP, T2A-Nano Luciferase, and circular Znf609. The circularization site was positioned at the CDS. For all the circular RNAs tested, multiple circularization sites in the CDS were found. Full-length precursors of those three RNAs were produced and circularized for 20 minutes as already described and shown in Figure 12. The sequences of the full length precursors for T2A-EGFP, T2A-Nano Luciferase and Znf609 are set forth in SEQ ID NOs: 42, 43 and 44 respectively. As shown in Figure 8D (urea-agarose gel), these long circular RNAs were all produced efficiently with TRIC-V2.

The results of this experiment suggested that longer eACA stems resulted in better circularization efficiency. To investigate this further, the CVB3-EGFP was cloned to the TRIC-V2 construct and two stem lengths were tested: 15 nt (SEQ ID NO: 45) and 25 nt (SEQ ID NO: 46). Separately, the effect of increasing the length of the extended guide sequence (EGS) was determined by providing a TRIC-V1 construct (that is, one having L15/R51 of the tRNA sequence) with a 40 nt EGS (SEQ ID NO: 47). Full length precursors of all these constructs were produced and circularization was performed for 4 minutes.

As shown in Figure 9A, with 4 minutes of circularization, TRIC-V1.0 converts ~50% of full length precursor to circular RNA. Increasing the EGS from 20 to 40 did not substantially increase circularization efficiency.

On the other hand, increasing the stem length to either 15 or 25 nt in TRIC-V2 did increase circularization efficiency. Increasing the EGS of TRIC-V2 from 20 to 40 nt did not substantially increase circularization efficiency.

TRIC-V2 was then compared with the PIE method on circularization of long GOIs EGFP, Spike and Cas9, for 1 or 3 minutes of circularization. At 1 minute of circularization, PIE converts a small number of EGFP precursors to circular RNAs, while TRIC-V2 converts many more. For example, when the eACA stem equals 25 nt, more than 50% of TRIC-V2 precursors converted to circular RNAs within 1 min (Figure 9B, 2% native agarose gel). At 3 minutes of circularization, PIE converts less than 50% of full-length EGFP precursor to circular RNAs but TRIC-V2 converts most of full-length precursors to circular RNAs. Moreover, the PIE produces a visible number of concatenations, which are largely absent with TRIC-V2 (Figure 9B). These differences between PIE and TRIC-V2 are also seen for Spike and the Cas9, suggesting that TRIC-V2 is much faster and cleaner than the PIE method for circularization of the long GOIs (Figure 9C, 0.8% native agarose gel).

Ribozyme kinetics were then measured using the Michaelis–Menten model. TRIC-V2 with a 25 nt stem is 3.7 times faster than PIE in producing the CVB3-EGFP circular RNA (Figure 9D).

Circularization efficiency, circular RNA yield (ratio between circular RNAs and total RNAs) and nicking ratio were then calculated using the CVB3-EGFP as an example. As shown in Figure 9E, for TRIC-V1, 20 min of circularization gives an efficiency of 95.6% with yield and nicking rates of 58.2% and 26.7%. For TRIC-V2, 3 min of circularization gives an efficiency of 89.4% with yield and nicking rates of 70.6% and 4.6%. Further circularization to 8 min of TRIC-V2 gives higher efficiency and yield (97.8% and 74.8%), but also increased the nicking (7.7%). Of note, a yield of 74.8% is 90.3% of the limit yield (MW ratio between circular RNA and its precursor). For PIE, 8min of circularization gives an efficiency of 86.7% with yield and nicking rates of 65.5% and 9.3%. Consequently, TRIC-V2 provides higher efficiency, higher yield, and lower nicking compared to PIE, all without the issues of native exon sequences in the resulting circular RNAs.

Overall, the post transcriptional circularization protocol and highly efficient TRIC-V2 allow the production of circular RNAs in a clean, highly efficient manner which results in a high final yield.

Example 6 - Immunogenicity of TRIC-V2

It has been shown that bacterial sequences introduced to circular GOIs by the PIE method are immunogenic (Liu, C. X. et al. Mol Cell 82, 420-434 e426, (2022)). As described above, TRIC-V2 can produce circular genes of interest without any bacterial sequence. The inventors determined whether circular RNAs produced by TRIC-V2 were less immunogenic than those produced by PIE.

Constructs that would produce circular EGFP and circular Nano luciferase (Nluc) by PIE (SEQ ID NOs: 48 and 49, respectively) or TRIC-V2 (SEQ ID NOs: 50 and 51, respectively) were created.

Circular RNAs were purified by gel filtration (using an SRT-2000 SEC column (Sepax)) and RNase R digestion.

As shown in Figure 11B, circular RNA can be separated from the spliced group I intron but not the nicked RNA and full length precursor. To further remove the nicked RNA and full-length precursor, the circular RNA was digested by an exonuclease, RNase R, at 37 °C for 1h. The circular RNAs were then cleaned by the RNA Clean & Concentrator Kits (ZYMO RESEARCH). Purified circular RNAs were checked by 6M-1.5% urea-agarose gel. As shown in Figure 11C, the intron and full-length precursor were not visible in purified circular RNAs. However, there were still some nicked products which were likely produced during the gel running process.

A549 cells were then transfected with the circular RNAs, and the expression levels of immune factors (IL6, CCL5, and INF beta) was monitored by RT-qPCR. Various controls were included in the experiment, such as mocks (Mock and Lipo), a positive control (poly I:C), unmodified linear mRNAs (lin. EGFP and lin. Nluc) and modified linear mRNAs (lin. EGFP-modi and lin. Nluc-modi). GAPDH was used as the internal reference. Figures 11D-F depict the results. The results showed that poly I:C and

unmodified linear mRNAs elicited a strong expression of immune factors, whereas the negative controls (mock and lipo) did not induce significant expression. Moreover, as expected, modified linear mRNAs did not induce a robust immune response. When it comes to circular RNAs, it was observed that they generally exhibited lower immunogenicity compared to unmodified linear mRNAs.

Specifically, for circular Nluc, the PIE-circ. Nluc triggered a significantly stronger expression of immune factors compared to the TRIC-circ. Nluc. Notably, the TRIC-circ. Nluc behaved similarly to the nonimmunogenic mocks, suggesting a low immunogenicity. Regarding circular EGFP, circular RNAs produced by both methods induced a comparable number of immune responses. These findings demonstrate that the immunogenicity of circular RNA produced by TRIC-V2 is lower than that of PIE, particularly in the case of circular Nluc. Furthermore, unmodified TRIC circular RNA can exhibit a similar level of immunogenicity as nonimmunogenic mocks.

The expression of Nluc from circular and linear RNAs in HEK293 cells over 7 days was then tested (Figure 11G). The circular form of Nluc was had increased and more persistent protein expression compared to linear mRNA. In some cases, expression from circular RNAs may need to be discontinued. One way to achieve this is by using siRNAs. Multiple siRNA target sites (msITS) were introduced 5' of the IRES sequence in circular RNAs produced by TRIC-V2 (constructs shown in Figure 11A). msiTS-circ. Nluc was transfected into HEK293 cells and siRNAs were applied to the cells on day 3 following transfection. As shown in Figure 11H, expression of Nluc is reduced 5 hours later, demonstrating that circular RNAs produced by TRIC-V2 can be effectively targeted with siRNAs.

Example 7 - Rolling Circle Replication

Circular RNAs (circRNAs) are more stable than mRNAs and are thus promising alternatives to mRNAs for therapeutics. Natural circRNAs can serve as sponges for microRNAs and proteins but are also templates for translation. In the presence of an internal ribosome entry site (IRES) for initiation, circRNAs can be translated efficiently. CircRNAs that lack an in-frame termination codon allow ribosomes to translate completely around the circRNA multiple times, resulting in a polypeptide, a process known as rolling circle translation (RCT) (Figure 15a). The polypeptide can in principle be cleaved by a protease or a self-cleaving sequence, thus allowing multiple copies of the GOI to be made in a single round of initiation. In principle, RCT can be 100-fold more efficient than single-shot translation. However, RCT encounters two challenges: low initiation efficiency and accessory sequences introduced by currently used in vitro circularization methods.

Unlike single-shot translation, RCT primarily produces a polypeptide. This can be converted into protein monomers provided with 2A skipping sequences (Figure 15b). Instead of using circRNAs with only a CDS, which are inefficient in translation initiation, we selected a potent short IRES (OR4F17) for the RCT construct. As expected, the circOR4F17-Nluc-RCT produced a significant number of proteins. However, the translation was still orders of magnitude less efficient compared to N1Ψ-modified mRNA.

We then asked whether potent viral IRESs could be used for RCT. However, viral IRESs are typically long and highly structured, which could block RCT either by in-frame stop codons or by causing ribosome stalling. Nevertheless, we tested the 373 nt CSFV IRES. As expected, multiple stop codons were identified in each frame. We therefore engineered frame 1 to eliminate stop codons and align with the CDS (Figure 15c). Notably, the RCT using CSFV IRES showed over 10-fold increase in Nluc expression compared to its single-shot translation and was over 7,000-fold more efficient than the RCT using the OR4F17 IRES (Figure 15d, e).

Example 8 - Loop Optimisation

As has been shown, the eACA structure is essential for circularization to happen. This eACA is a stem-loop structure which contains a stem ≥ 1 bp and a 7 nt loop with a U at the third position (Figure 16A a-0). This minimal structure enables circularization of GOIs with little restriction without introducing unwanted sequences.

To further release the restriction of the eACA, we tested multiple versions (Figure 16A b1-m12: SEQ ID Nos: 79 to 90) of the loop in circularization of EGFP. In these loops, loop length varies from 3-11 nt and the U was placed in 1st-5th positions. Circularization of EGFP was abolished in the case of L5_U2 and L3_U1, but worked efficiently in all other tested cases, including L5_U3 (Figure 16B). This indicates that the Loop needs to be longer than 4 nt and the U cannot be at the 1st or 2nd position.

Example 9 - Wobble base pair

The G and U wobble base pair between IGS (internal guide sequence) and GOIs is known to be important for ribozyme to function. To explore whether other base pairs are functional, we mutated the G in the IGS to U or A and the U in GOI to either A or C. This way, both new Watson-Crick base pair (U-A) and wobble base pairs (A-C) are included.

We produced full length precursors and performed circularization for 3 min and 30 min. The TRICv2(G-U)+ circularized the EGFP efficiently (Figure 17a). Meanwhile, it's clearly shown that both TRICv2(C-A) (SEQ ID NO: 92) and TRICv2(A-U) (SEQ ID NO: 91) can convert Full length to circular RNA. Circular RNAs from TRICv2(C-A) and TRICv2(A-U) were further confirmed by a urea-agarose gel (Figure 17b).

Example 10 – Comparison of TRIC-V2

Recent studies have described Tetrahymena thermophila (Tetra) group I intron-derived constructs, the Tetra-STS (RZ construct, WO2022/191642) and the Tetra-Rzy, for circRNA synthesis. We compared these to V2 by cloning CVB3-EGFP into Tetra-STS (AU-rich no. 16) and Tetra-Rzy (CVB3IRES-GFP) constructs (SEQ ID Nos: 94 and 95). V2 outperforms both constructs (Figure 18). Additionally, the Tetra-V2 (SEQ ID NO: 93) also produced circCVB3-EGFP efficiently, demonstrating that optimizations from the Ana intron can be effectively applied to other group I introns.

SEQUENCES

Nucleotide sequences referred to herein are set out in the following table.

SEQ ID NO	Description	Sequence
1	Ana group I intron (DNA)	AAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAACAACAGATAAATTACAGCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATG
2	T4 phage group I intron (DNA)	TAATTGAGGCCTGAGTATAAGGTGACTTATACCTTGAATCTATCTAAACGGGGAACCTCTCTAGTAGACA ATCCCGTGCTAAATTGTAGGACTTGCCCTTTAATAAACTCTTATATTTAAAGAGGTATTTATGAAAAGC GGAATTTATCAGATTAATAAATACTTTAAACAATAAAGTATATGTAGGAAGTGCTAAAGATTTTGAAAAGAG ATGGAAGAGGCATTTTAAAGATTTAGAAAAAGGATGCCATTCTTCTATAAACTTCAGAGGCTTTTAAACA AACATGGTAATGTGTTTGAATGTTCTATTTTGAAGAAATTCATATGAGAAAGATTTGATTATTGAACGA GAAAATTTTGGATTAAAGAGCTTAATTTCTAAAATTAATGGATACAATATTGCTGATGCAACGTTTGGTGAT ACATGTTCTACGCATCCATTAAGAAGAAATTTAAGAAACGTTCTGAAACTGTTAAAGCTAAGATGCT TAAACTTGGACCTGATGGTCGGAAGCTCTTACAGTAAACCCGGAAGTAAACCGGGCGTTGGAATCC AGAAACCCATAAGTTTTGAAGTGGTGTTCCGCATACAACTTCTGCTTATACCTGTAGTAAATGCAGAA ATCGTTCCAGGTGAAAATAATTCATTCTTTAATCATAAGCATTTCAGACATAAATCTAAAATATCAGAAA AGATGAAAGGTAAAAGCCTAGTAATATTAAGAAGATTTTCATGTGATGGGGTATTTTGTATTGTGCAGCA GATGCAGCTAGACATTTTAAATTTCTGCTGGATTAGTTACTTATCGTGTAATCTGATAAATGGAATTG GTTCTACATAAATGCCTAACGACTATCCCTTTGGGGAGTAGGGTCAAGTGACTCGAAACGATAGACAAC TGTCTTAACAAGTTGGAGATATAGTCTGCTCTGCATGGTGACATGCAGCTGGATATAATTCCGGGGTAAG ATTAACGACCTTATCTGAACATAATG
3	Tetrahymena ribosomal intron (DNA)	AAATAGCAATATTTACCTTTGGAGGGAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATA GATTGCATCGGTTTAAAAGGCAAGACCGTCAAATTGCGGGAAAGGGTCAACAGCCGTTTCAGTACCAAG TCTCAGGGGAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACATGGTCCCTAACCA CGCAGCCAAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCACAGACTAAATGTCGGTCCG GGAAGATGTATTCTTCTCATAAGATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGATGAAGTGATGC AACACTGGAGCCGCTGGGAACATAATTTGTATGCGAAAGTATATTGATTAGTTTTGGAGTACTCG
4	Azoarcus sp. BH72 Ile tRNA intron (DNA)	ATTTGATGTGCCTTGCGCCGGGAAACACGCAAGGGATGGTGTCAAATTCGGCGAAACCTAAGCGCC CGCCCGGGCGTATGGCAACGCCGAGCCAAGCTTCGGCGCCTGCGCCGATGAAGGTGTAGAGACTAGA CGGCACCCACCTAAGGCAAACGCTATGGTGAAGGCATAGTCCAGGGAGTGGCGAAAGTCACACAAACC GG
5	Staphylococcus phage Twort ribonucleotide reductase intron (DNA)	AAACAATTCTGCCCCCTATATTAGTAATAGTGTAGGTAAAAAACTTCCTTAATTCATGGGAAATCTCCCT GACTTTTATTATAAATTTGGTATAGTAAAAATGAGAAGGAGATAATCATGAGCAAGATAACCAATAGCAAA GAAACCCAAAAAGTACCTAACGCTACTAAAAGTATTTACCACCATATAAAAAAGTAAAAAGGATGGAAG TCATTAAATCACTTAATGAATTGGTAATTATCTTGTGCAACGACTAGAGAAAAGATAGTTTATTGTTACAG GCAGTAAATGAAGACTGAGTATCGTACACACAAGTGAGTGGAAACAGGAAGTATCCTAGAGTAACGACT AGGATAATGATATAGTCTGAACATTGTAGGTGACTACAAGAAGGTAAGGAGTAACGAACCTTATCGTAAC ATAATTG
6	Primer: qGAPDH-F	AGCTCATTTCTGGTATGACA
7	Primer: qGAPDH-R	AGGGGAGATTCAGTGTGGTG
8	Primer: qIL6-F	AGCCACTCACCTCTTCAGAAC
9	Primer: qIL6-R	GCCTCTTTGCTGCTTTCACAC
10	Primer: qCCL5-F	CACAGGTACCATGAAGGTCTC
11	Primer: qCCL5-R	GGTGTCCGAGGAATATGGG

12	Primer: qlFNβ-F	TCTAGCACTGGCTGGAATGAG
13	Primer: qlFNβ-F	GTTTCGGAGGTAACCTGTAAG
14	TRIC-V0 (3*Flag, construct)	AAATAATTGAGCCTTAGAGAAGAAATTCCTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGCT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAACACAGATAACTTACAGCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGTTGACCTTAAACGGTCGTGTG GTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAGATCACGAC ATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAAAAAGATATC
15	TRIC-V1.0 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTAAGTGGATGCTCTCAAACCT CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGT TGACCTTAAACGGTCGTGTGGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATG GTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAAACACCC GGAAGCCAACCGATTGACCGATATC
16	TRIC-V1.1 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTAAGTGGATGCTCTCAAACCT CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGT TGACCTTAAACGGTCGTGTGGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATG GTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAAACTTTC GGAAGCCAACCGATTGACCGATATC
17	TRIC-V1.2 (3*Flag)	GGTCAATCGGTTGGCTTCCGTAATTGAGCCTTAGAGAAGAAATTCCTTAAGTGGATGCTCTCAAACCTCAG GGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAACAA CAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGG TAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGTTG ACCTTAAACGGTCGTGTGGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGT GATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAAACCGGAAG CCAACCGATTGACCGATATC
18	3*Flag GOI	AAAATCCGTTGACCTTAAACGGTCGTGTGGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAG ATCACGATGGTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACT T
19	TRIC-V1.0 for EGFP	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTAAGTGGATGC TCTCAAACCTCAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCAACGCGGAAACGCAATAG CCGAAAAACAAAAACAAAAAAACAAAAAAACAAAAAAACAAAAACACATTAACAGCCTGTGGGT TGATCCCAACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCACTGTAAGTGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCAC GTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGCGGTTGAAGGAGAAA GCGTTCGTTATCCGGCCAACCTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCA GCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCAACGGGCGACCGTGGCGGTGGCTG CGTTGGCGGCTGCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATT GAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACCCCTCAAGC CAGAGGGCAGTGTGTCGTAAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCGCTGTTTCATT TTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCC GGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACAT TACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTACAGCGTGTCCGGCGAGGGCGAG GGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTCCCTG GCCACCCCTCGTGACCAACCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCATGAAGC AGCACGACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCT GAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACAGCC

		ACAACGTCTATATCATGGCCGACAAGCAGAAGACGGCATCAAGATTCGAAGTCCGCCACAACA TCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGT GCTGCTGCCGACAACCACTACCTGAGCACCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAAGCGCG ATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG TAAAAAAAACAAAAACAAAACGGCTATTATGCGTTACCGGCGAGACGCTACGGACTTAAACACCCGGAA GCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGATATC
20	TRIC-V1.0 for Fluc	GGATAATACGGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTCTTTAAGTGGATGC TCTCAAACCTAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCAACGCCGAAACGCAATAG CCGAAAAACAAAAACAAAAAACAAAAAACCAAAAAACAAAAACACATTAACAGCCTGTGGGT TGATCCCACCCACAGGCCCATTTGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCCCTGTTTTAT ACCCCTTCCCACTGTAAGTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACACAGCCAC GTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGCGGTTGAAGGAGAAA GCGTTCGTTATCCGGCCAACCTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCA GCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCAACGGGCGACCGTGGCGGTGGCTG CGTTGGCGGCTGCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCAAGAGTCTATT GAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGC CAGAGGGCAGTGTGTCGAACGGGCACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATT TTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCC GGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCCTTAGCTTGAAAGAGGTTAAACAT TACAATTCATTGTTAAGTTGAATACAGCAAAATGGAAGATGCCAAAAACATTAAGAAGGGCCAGCGCCA TTCTACCACTCGAAGACGGTACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGT GCCCGGCACCATCGCTTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGAT GAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCA GCGAGAATAGCTTGCACTTCTTCATGCCCGTGTGGGTGCCCTGTTTCATCGGTGTGGGTGTGGCCCGAG CTAACGACATCTAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCACCGCTCGTATTC GTGAGCAAGAAAGGGCTGCAAAAGATCCTAACGTGCAAAAGAAGCTACCGATCATACAAAGATCATC ATCATGGATAGCAAGACCGACTACCAGGGCTTCAAAGCATGTACACCTTCGTGACTTCCATTTGCCAC CCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGACCGGGACAAAACATCGCCCTGATCATG AACAGTAGTGGCAGTACCGGATTGCCAAGGGCGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATT CAGTCATGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGACACCGCTATCCTCAGCGTGGTGC CATTTACACAGGCTTCGGCATGTTACACACGCTGGGCTACTTGATCTGCGGCTTTCGGGTGTGCTCA TGTACCGCTTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTGG TGCCCACTATTAGCTTCTTCGCTAAGAGCACCTCATTGATAAGTACGACCTAAGCAACTTGCACGA GATCGCCAGCGCGGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTGGCTAAACGCTTCCACCTA CCAGGCATCCGCCAGGGCTACGGCCTGACAGAAACAACAGCGCCATTCTGATACCCCCGAAGGGGA CGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTGGACTTGGACACCG GCAAGACACTGGGTGTGAACAGCGCGCGAGCTGTGCGTCCGTGGCCCATGATCATGAGCGGCTA CGTTAACAACCCCGAGGCTACAAACGCACTCATGACAAGGACGGCTGGCTGCACAGCGGCGACATCG CCTACTGGGACGAGGACGAGCACTTCTTCATCGTGACCGGCTGAAGAGCCTGATCAAATACAAGGGC TACCAGGTAGCCCCAGCCGAACCTGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCCGGGT CGCCGGCCTGCCGACGACGATGCCGGCGAGCTGCCCGCCGAGTCTGCTGCTGGAACACGGTAAA ACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAAGAAGCTGCGCGG TGGTGTGTTGTTCTGTGGACGAGGTGCCTAAGGACTGACCGGTAAGTTGGACGCCCGAAGATCCGCG AGATTCTCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGTGAGCAGTGATTACAAGGATGATGATGATA AGTAAAAAAAACAAAAACAAAACGGCTATTATGCGTTACCGGCGAGACGCTACGGACTTAAACACCCG GAAGCCAACCGATTGACCAAAAAAAAAAAGCGGCCGCGATATC
21	TRIC-V1.0 for Spike	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTCTTTAAGTGGATGC TCTCAAACCTAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCAACGCCGAAACGCAATAG CCGAAAAACAAAAACAAAAAACAAAAAACCAAAAAACAAAAACACATTAACAGCCTGTGGGT TGATCCCACCCACAGGCCCATTTGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCCCTGTTTTAT ACCCCTTCCCACTGTAAGTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACACAGCCAC GTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGCGGTTGAAGGAGAAA

	GCGTTCGTTATCCGGCCAACACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCA GCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGCGGCTGGCTG CGTTGGCGGCTGCCATGGGAAACCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATT GAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACCCCTCAAGC CAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCGCTGTTTCATT TTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCC GGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAAACAT TACAATTCATTGTTAAGTTGAATACAGCAAAATGGGGATTCTGCCTAGCCCCGGCATGCCGCCCTCCTG AGCCTGGTGAGCCTGCTGAGCGTCTGTTGATGGGGTGCTGGCCGAGACCGGCACACAGTGCCTGA ACCTGACAAC TAGAACACAGCTGCCCCCAGCCTACACCAACAGCTTTACTAGGGGAGTGTACTACCCTG ATAAGTGCTTCAGGAGCAGTGTGCTGCACAGCACCCAGGACCTCTTTTTGCCCTTTTTCAGTAACGTGAC CTGGTTCCACGCTATCCACGTGAGCGGCACAAACGGCACAAAACGGTTCGATAACCCAGTGTGCCCTT CAACGACGGCGTCTACTTCGCCAGCACTGAGAAGTCCAACATTATCAGGGGGTGGATCTTCGGGACAAC CCTGGACAGCAAGACACAGAGCCTCCTGATTGTGAACAACGCCACCAACGTGGTCATTAAAGTCTGCGA GTTCCAGTTCTGCAACGACCCCTTTCTGGGGGTGACTACCACAAGAACAACAGTCTGGATGGAGTC CGAGTTTCGGGTGACTCCAGCGCCAACACTGCACATTTAGTACGTGAGCCAGCCCTTCTTGATGGA TCTGGAGGGCAAGCAGGGGAACCTTAAGAACCTTCGGGAGTTCGCTTTAAGAACATCGACGGGTACTT TAAGATCTACAGCAAGCACACCCCTATTAACCTGGTCCGCGATCTGCCCCAGGGCTTCAGCGCCCTGGA GCCACTGGTGGACCTCCCCATCGGGATCAACATTACTAGGTTCCAGACACTCTGGCCCTGCACCGGA GCTACCTGACTCCCGCGATAGCTCCAGCGGGTGGACTGCCGGGGCCGCCCTACTACGTCGGCTA CCTGCAGCCCAGGACATTCTGCTCAAGTACAACGAGAACGGCACCATCACTGATGCCGTGGACTGCG CTCTCGATCCCTGAGCGAGACAAAGTGACCCCTGAAGTCTTCACCGTGGAGAAGGGGATCTACCAG ACATCTAACTTCAGGGTGACGCCCACTGAGTCCATTGTGAGATTTCTAACATACCAACCTGTGCCCT TCGGCGAGGTCTTTAACGCCACACGGTTCGCCAGCGTCTACGCTTGAACAGGAAGCGGATCTCAAC TCGGTGGCCGACTACAGCGTGTGTACAACAGCGCCTCCTTCAGCACCTTTAAGTGCTACGGCGTGAG CCCCCTAAGCTGAACGACCTGTGCTTCACTAACGTGTACGCTGATAGCTTCGTGATTAGGGGCGATGA GGTCCGGCAGATTGCCCCCGCCAGACAGGGAAGATCGCTGACTACAATAAGTGCCCGATGACT TCACAGGGTGTGTGATCGCTGGAACAGCAACAACCTGGATAGCAAGGTGCGGGGGAACATAACTAC CTGTACAGGCTGTTAGAAAGTCCAACCTGAAGCCCTTCGAGCGGGATATCAGCACTGAGATCTACCAG GCGGGGAGCACCCCTGCAACGGCGTGCAGGGCTTTAACTGCTACTTTCCCTCCAGAGCTACGGCTT CCAACCCACCAACGGCGTGGGGTACCAGCCCTACCGGGTGTGGTCTTAGCTTCGAGCTCCTGCACG CCCCTGCAACCGTGTGCGGCCCAAGAAGAGCACAAACCTCGTGAAGAACAAGTGCCTGAACCTCAACT TCAACGGGCTGACCGGCACCGCGTGTGCTGACTGAGAGTAACAAGAAGTTTCTCCCTCCAGCAGTGC GGGAGGGATATTGCCGACACTACTGATGCCGTGCGCGACCCTCAGACCCCTGAGATTCTGGACATCACT CCCTGCTCCTTCGGCGGAGTGTCCGTGATCACACCAGGGACCAACACTAGCAACCAGGTGCGCGTCTT GTACCAGGACGTGAACGTGACCGAGGTGCCGTGGCCATCCACGCCGATCAGCTGACCCCACTTGGC GGGTGACTCCACCGGGAGCAACGTGTTTCAGACACGGGCGGGTGCTGATCGGGGCTGAGCATGT CAACAACAGCTACGAGTGCGATATTCTATCGGCGCCGGGATTTGCGCCAGCTACCAGACACAGACAAA CAGCCCTGGCTCGGCCAGCTCCGTGCGCCAGCCAGAGCATCATCGCTACACAATGTCTTGGGGGCCG AGAACAGCGTCGCTACAGCAACAACAGCATTGCCATCCCCACCAACTTCACAATCAGCGTCACAACCG AGATCCTGCCGTGAGCATGACTAAGACATCCGTGGATTGCACAATGTACATCTGCGGCGACTCCACAG AGTGACGTAACCTGCTGCTCCAGTACGGGTCTTTTGTACCCAGCTGAACAGAGCCTTGACCGGCATTG CCGTGGAGCAGGACAAGAACCCAGGAGGTGTTTCCCAGGTGAAGCAGATCTACAAGACACCCCCC ATAAAGGACTTCGGGGGCTTTAACTTCTCCAGATCTTGCCCGACCCAGCAAGCCCTCAAGCGGTCA TTCATTGAGGACCTGCTGTTTAACAAGGTGACCCTCGCTGATGCAGGGTTTATCAAGCAGTACGGCGAT TGCTGGGGGACATTGCCGCTAGGGACTTGATCTGCGCCAGAAGTTCAACGGCCTGACCGTGTCTCC CCCCCTGTTGACAGACGAGATGATCGCCAGTACACTAGCGCTCTGCTGGCCGGCACTATACCAGCG GCTGGACCTTTGGGGCCGGCGCCGCCCTCCAGATTCCATTGCCATGCAGATGGCTACCGCTTCAAC GGCATCGGCGTGAACAGCTGCTGTACGAGAACCAGAAGCTGATTGCCAACCAAGTTCAACTCCGC CATCGGCAAGATCCAGGACAGCCTGTCCAGCACAGCATCCGCCCTGGGGAAGTTGCAGGATGTGGTGA ACCAGAACGCCAGGCCCTGAACACACTGGTGAAGCAGCTGAGCAGCAACTTTGGCGCCATCAGTAGC GTCCTGAACGACATCCTGTCACGGCTTGATCCACCAGAGGCCGAGGTCCAGATCGATCGGCTGATTACA GGCCGGCTGCAGAGTCTCCAGACCTACGTACCCAGCAGCTGATCCGGGCTGCCGAGATTAGAGCCTC CGCCAATCTGCGCGCCACAAGATGAGTGAATGCGTCTGGGCCAGAGCAAAAGGGTCACTTTTGGC GGAAGGGGTACCACCTGATGTCTTTCCCCAGTCCGCCCTCACGGAGTGGTGTCTCTCATGTACAT ACGTGCCTGCCAGGAGAGAAGAACTTCACTACAGCCCCCGCCATTTGCCACGATGGGAAGGCCATTTT CCTAGAGAGGGCGTGTCTGTCCAACGGGACCCACTGGTTCGTGACTACGCGGAACCTTTATGAACCC CAGATTATCAACCCGATAACACATTTGTCTCCGGCAACTGTGATGTGGTCATCGGCATCGTGAACAACA
--	---

		CCGTGTACGACCCACTGCAGCCTGAGCTTGACAGCTTTAAGGAGGAGCTGGACAAGTACTTTAAGAACC ACACCAGCCCTGATGTGGACCTTGGCGACATCAGCGGCATTAACGCCAGCGTGGTGAACATTCAGAAG GAGATCGATAGACTGAACGAGGTGGCCAAGAACTTGAACGAGTCACTGATTGACCTGCAGGAGCTGGG GAAGTACGAGCAGTACATTAAGGGGTCCGGGAGGGAGAACCCTGACTTTCAGGGGGGGGGGGGGTCC GGCTACATTCCCAGGCCCAAGGGATGGCCAGGCCACGTGAGGAAGGATGGGGAGTGGGTGCTGC TGAGCACCTTCTTGGGCAGCGGCAGCAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAG GGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTG GCCCACCCTCGTGACCACCCTGACCTACGGCGTGACGTGCTTCAGCCGCTACCCCCGACCACATGAAGC AGCACGACTTCTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTCAAGGACG ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCT GAAGGGCATCGACTTCAAGGAGGACGGCAACATCTGGGGCACAAGCTGGAGTACAACACAAGCC ACAACGTCTATATCATGGCCGACAAGCAGAAGAAGCGCATCAAGGTGAACCTCAAGATCCGCCACAACA TCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAGACCCCCAACGAGAAGCGCG ATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG TAAAAAAAACAAAAACAAAACGGCTATTATGCGTTACCGCGGAGACGCTACGGACTTAAACACCCGGAA GCCAAGTATTGACCAAAAAAAAAAAGCGGCCGCGATATC
22	TRIC-V1.0 for Cas9	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGGAACGCAATAG CCGAAAAACAAAAACAAAAAACCACAAAAACAAACACATTAACAGCCTGTGGGT TGATCCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCAACTGTAAGTGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCAC GTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAA GCGTTCGTTATCCGGCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCA GCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCACCGGGCGACCGTGGCGGTGGCTG CGTTGGCGGCTGCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATT GAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAAGTGCAGGACACACCCCTCAAGC CAGAGGGCAGTGTGTCGTAAACGGCACTCTGCAGCGGAACCGACTACTTTGGGTGTCGTGTTTCATT TTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGCCATCC GGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACAT TACAATTCATTGTTAAGTTGAATACAGCAAAATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAA ATAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAA ATACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTATTTGACAGTGGAGAGACAGCGGA AGCGACTCGTCTCAACGGACAGCTCGTAGAAGGTATACAGTCCGGAAGAATCGTATTTGTTATCTACAG GAGATTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTCTTTTCATCGACTTGAAGAGTCTTTTTGGT GGAAGAAGACAAGAAGCATGAACGTCATCTATTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAG AAATATCCAATCTATCATCTGCGAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAAT CTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCAATTTTATTGATTGAGGGAGATTTAAATCCTGATA ATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACAATCAATATTTGAAGAAAACCTATT AACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATC TCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTATTTTGGGAATCTCATTGCTTTGTCATTGGGTTT GACCCCTAATTTAAATCAAATTTTGAATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACG ATGATGATTTAGATAATTTATTGGCGCAAATTTGGAGATCAATATGCTGATTTGTTTTGGCAGCTAAGAA TTATCAGATGCTATTTTACTTTTTCAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCTCCCTATCAGC TTCAATGATTAACGCTACGATGAACATCATCAAGACTTGACTCTTTTAAAGCTTTAGTTGACACAAC TTCCAGAAAAGTATAAAGAAATCTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGA GCTAGCCAAGAAGAATTTTATAAATTTATCAAAACCAATTTAGAAAAATGGATGGTACTGAGGAATTTAT GGTGAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTGACAACGGCTCTATTTCCCATCAA ATTCACCTGGGTGAGCTGCATGCTATTTTGAAGACAAGAAGACTTTTATCCATTTTAAAGACAATCG TGAGAAGATTGAAAAATCTTGACTTTTGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTC GTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAA AGGTGCTTCAGCTCAATCATTTATTGAACGCATGACAACTTTGATAAAAAATCTTCAAAATGAAAAAGTAC TACCAAAACATAGTTTGCTTTATGAGTATTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG AAGGAATGCGAAAACAGCATTCTTTTCAAGTGAACAGAAGAAGCCATTGTTGATTACTCTTCAAAAC

		AAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAAATAGAATGTTTTGATAGTGTGA AATTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAGA TAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGA AGATAGGGAGATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAG CTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGC AATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTGGCAATCGCAATTTATGCAGCTGATC CATGATGATAGTTTGACATTTAAGAAGACATTCAAAAAGCACAAAGTGTCTGGACAAGGCGATAGTTTAC ATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAATAAAGGTATTTACAGACTGTAAAGTTGTT GATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC AGACAACCTAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAAT TAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTAT TATCTCCAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAAGTGATTATGATGT CGATCACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATA AAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACCTATTGGAGACA ACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTAACGAAAGCTGAACGTGGAGGTTTGA GTGAAC TTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGC ACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGA TTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTCCAATTCTATAAAGTACGTGAGATTAACA ATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAAGTCTTTGATTAAGAAATATCCAAAA CTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCTGTAATGATTGCTAAGTCTGAGCA AGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACCTCTTCAAAACAGAAATTAC ACTTGCAAATGGAGAGATTGCAAAACGCCCTCTAATCGAACTAATGGGGAAACTGGAGAAATTGTCTG GGATAAAGGGCGAGATTTTCCACAGTGCGCAAAGTATTGTCCATGCCCAAGTCAATATTGTCAAGAAA ACAGAAGTACAGACAGGCGGATTCTCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTG CTCGTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCT AGTGGTTGCTAAGGTGAAAAAGGAAATCGAAGAAGTAAAAATCCGTTAAAGAGTTACTAGGGATCACA ATTATGGAAGAAGTTCCCTTTGAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAA AAAAGACTTAATCATTAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTAACGGATGCTGG CTAGTGCCGAGAAATTACAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTATATTTA GCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGC ATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTCTAAGCGTGTTATTTTAGCAGATGCC AATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATAT TATTCATTTATTTACGTTGACGAATCTTGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCG TAAACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATG AAACACGCATTGATTGAGTCAGCTAGGAGGTGACGGCAGCGGCAGCAGTATGGTGAGCAAGGGCGAG GAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACCGGCCACAAGTTTACG CGTGTCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTTATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCTCGTACCACCCTGACCTACGGCGTGACGTGCTTCAGCCG CTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGC GCACCATCTTCTCAAGGACGACGGCACTACAAGACCCGCGCCGAGGTGAAGTTGAGGGCGACACC CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCT GGAGTACAAC TACAAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGACGGCATCAAGGTGAA CTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCC CCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAA GACCCCAACGAGAAGCGCGATCAGATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGG CATGGACGAGCTGTACAAGTAAAAAAAACAAAAAACAAACGGCTATTATGCGTTACCGGCGAGACGCT ACGGACTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGATATC
23	TRIC-V1.0 for Factor 8	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCCACCCACGCCGGAACGCAATAG CCGAAAAACAAAAACAAAAAACAAAAAACAAAAAACAAAAACACATTAACAGCCTGTGGGT TGATCCCAACACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCAACGTGAAC TTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCAC GTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAA GCGTTCTGTTATCCGGCCAACCTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCA GCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCAACGGCGACCGTGGCGGTGGCTG

	CGTTGGCGGCCTGCCCATGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATT GAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACCCCTCAAGC CAGAGGGCAGTGTGTCGTAAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCGTGTTTCATT TTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCC GGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCCTTAGCTTGAAGAGGTTAAACAT TACAATTCATTGTTAAGTTGAATACAGCAAAATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAA ATAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAA ATACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTATTTGACAGTGGAGAGACAGCGGA AGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAG GAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTCATCGACTTGAAGAGTCTTTTTTGGT GGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGAAATATAGTAGATGAAGTTGCTTATCATGAG AAATATCCAATCTATCATCTCGCAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAAT CTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTGATTGAGGGAGATTTAAATCCTGATA ATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCTATT AACGCAAGTGGAGTAGATGCTAAAGCGATTCTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATC TCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTT GACCCCTAATTTTAAATCAAATTTTGATTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACG ATGATGATTTAGATAATTTATTGGCGCAAATGGAGATCAATATGCTGATTTGTTTTGGCAGCTAAGAAT TTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTAAATACTGAAATACTAAGGCTCCCTATCAGC TTCAATGATTAACGCTACGATGAACATCATCAAGACTTGACTCTTTTAAAGCTTTAGTTCGACAACAAC TTCCAGAAAAGTATAAAGAAATCTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGA GCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAATGGATGGTACTGAGGAATTATT GGTGAACTAAATCGTGAAGATTGCTGCGCAAGCAACGGACCTTGACAACGGCTCTATTTCCCATCAA ATTCATTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAGACAATCG TGAGAAGATTGAAAAATCTTGACTTTTGAATTCCTTATTATGTTGGTCCATTGGCGGTGGCAATAGTC GTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTGAGAAGTTGTCGATAA AGGTGCTTCAGCTCAATCATTATTGAACGCATGACAACTTTGATAAAAAATCTTCAAATGAAAAAGTAC TACCAAAACATAGTTTGCTTTATGAGTATTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG AAGGAATCGAAAAACCAGCATTCTTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTACTCTTCAAAC AAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAAATAGAATGTTTTGATAGTGTGA AATTTCAAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAGA TAAAGATTTTTTGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTATTTGA AGATAGGGAGATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAG CTTAAACGTCGCCGTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGC AATCTGGCAAAACAATATTAGATTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATC CATGATGATAGTTTGACATTTAAGAAGACATTCAAAAAGCACAAAGTGCTGGACAAGGCGATAGTTTAC ATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAATAAAGGTATTTTACAGACTGTAAAGTTGTT GATGAATTGGTCAAAGTAATGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC AGACAACCTAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAAT TAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTAT TATCTCAAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATTAATCGTTTAAAGTGATTATGATGT CGATCACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTAAACGCTTCTGATA AAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACTATTGGAGACA ACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAAGCTGAACGTGGAGTTTGA GTGAAC TTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGC ACAAATTTTGATAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGA TTACCTTAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTCCAATTCTATAAAGTACGTGAGATTAACA ATTACCATCATGCCCATGATGCGTATCTAAATGCCGTGTTGGAAGTCTTTGATTAAGAAATATCCAAAA CTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTGCTAAAATGATTGCTAAGTCTGAGCA AGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACCTCTTCAAAACAGAAATTAC ACTTGCAAATGGAGAGATTGCAAAACGCCCTCTAATCGAACTAATGGGGAAACTGGAGAAATTGTCTG GGATAAAGGGCGAGATTTTGCACAGTGCGCAAGTATTGTCCATGCCCAAGTCAATATTGTCAAGAAA ACAGAAGTACAGACAGCGGATTCTCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTG CTCGTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCT AGTGGTTGCTAAGGTGGAAGGGAAGTGAAGAAGTAAAAATCCGTTAAGAGTTACTAGGGATCACA ATTATGGAAGAAGTTCCCTTTGAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAA AAAAGACTTAATCATTAACCTAACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTAACGGATGCTGG CTAGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTATATTTA
--	--

		GCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGC ATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGATGCC AATTTAGATAAAGTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATAT TATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCG TAAACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATG AAACACGCATTGATTTGAGTCAGCTAGGAGGTGACGGCAGCGGCAGCAGTATGGTGAGCAAGGGCGAG GAGCTGTTACCGGGGTGGTGCCCATCTGGTCGAGCTGGACGGCAGCTAAACGGCCACAAGTTCAG CGTGTCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGAGTGCTTCAGCCG CTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGC GCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC CTGGTGAAACGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCT GGAGTAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAA CTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCC CCATCGGGCAGCGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAA GACCCCAACGAGAAGCGCGATCAGTGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGG CATGGACGAGCTGTACAAGTAAAAAACAACAAACGGCTATTATGCGTTACCGGCGAGACGCT ACGGACTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGATATC
24	EGFP GOI	AAAATCCGTTGACCTTAAACGGTCTGTGGGTTCAAGTCCCTCCACCCACGCCGGAACGCAATAGC CGAAAAACAAAAACAAAAAACAACAAAAAACCAAAAAACAAAACACATTAAACAGCCTGTGGGTT GATCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACG TTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAAGC GTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGC ACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGCGACCGTGGCGGTGGCTGCG TTGGCGGCCTGCCCATGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGA GCTAGTTGGTAGTCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCA GAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTT ATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCG GTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAACATT ACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTGG CCCACCTCGTGACCACCTGACCTACGGCGTGAGTGCTTCAGCCGCTACCCCGACCACATGAAGCA GCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTGAGGGCGACACCCTGGTGAACCGCATCGAGCTG AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACAAGCCA CAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAATTCAGATCCGCCACAACAT CGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTG CTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGA TCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGT AAAAAACAACAAACAGGCTATTATGCGTTACCGGCGAGACGCTACGGACTT
25	Fluc GOI	AAAATCCGTTGACCTTAAACGGTCTGTGGGTTCAAGTCCCTCCACCCACGCCGGAACGCAATAGC CGAAAAACAAAAACAAAAAACAACAAAAAACCAAAAAACAAAACACATTAAACAGCCTGTGGGTT GATCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACG TTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAAGC GTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGC ACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGCGACCGTGGCGGTGGCTGCG TTGGCGGCCTGCCCATGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGA GCTAGTTGGTAGTCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCA GAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTT ATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCG GTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAACATT ACAATTCATTGTTAAGTTGAATACAGCAAAATGGAAGATGCCAAAAACATTAAGAAGGGCCAGCGCCAT TCTACCCACTCGAAGACGGTACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTG CCCGGCACCATCGCCTTACCAGCGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGATG AGCGTTGCGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAACCATCGGATCGTGGTGTGACG

		CGAGAATAGCTTGCAGTTCTTCATGCCCGTGTGGGTGCCCTGTTTCATCGGTGTGGCTGTGGCCCCAGC TAACGACATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCTG GAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAGATCATCAT CATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATTGCCACC CGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGACCGGGACAAAACCATCGCCCTGATCATGAA CAGTAGTGGCAGTACCGGATTGCCCAAGGGCGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATTCA GTCATGCCCGCGACCCCATCTTCGGCAACCAGATCATCCCGACACCGCTATCCTCAGCGTGGTGCCAT TTCACCACGGCTTCGGCATGTTCAACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCTGTCTCATGT ACCGCTTCGAGGAGGAGCTATTCTTGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTGGTGC CCACACTATTTAGCTTCTTCGCTAAGAGCACCCCTCATTGATAAGTACGACCTAAGCAACTTGCACGAGAT CGCCAGCGGGCGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTGGCTAAACGCTTCCACCTACCA GGCATCCGCCAGGGCTACGGCCTGACAGAAACAACGAGCGCCATTCTGATACCCCCGAAGGGGACGA CAAGCCTGGCGCAGTAGGCAAGGTGGTCCCTTCTTCGAGGCTAAGGTGGTGGACTTGGACACCGGCA AGACACTGGGTGTGAACGAGCGCGGCGAGCTGTGCGTCCGTGGCCCCATGATCATGAGCGGTACGTT AACAACCCCGAGGCTACAAACGCACTCATCGACAAGGACGGCTGGCTGCACAGCGGCGACATCGCCTA CTGGGACGAGGACGAGCACTTCTTCATCGTGGACCGGCTGAAGAGCCTGATCAAATAAAGGGCTACC AGGTAGCCCCAGCCGAACCTGGAGAGCATCCTGTGCAACACCCCAACATCTTCGACGCCGGGGTCGCC GGCCTGCCCGACGACGATGCCGGCGAGCTGCCCGCCGAGTCGTGCTGGAACACGGTAAACCA TGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACGCCAAGAAGCTGCGCGGTGGT GTTGTGTTCTGTGGACGAGGTGCCTAAAGGACTGACCGGTAAGTTGGACGCCGCAAGATCCGCGAGAT TCTCATTAAAGCCAAGAAGGGCGCAAGATCGCCGTGAGCAGTGATTACAAGGATGATGATGATAAGTA AAAAAACAAAAACAAAACGGCTATTATGCGTTACCGGCGAGACGCTACGGACTT
26	Spike GOI	AAAATCCGTTGACCTTAACGGTCTGTGGGTTCAGTCCCTCCACCCCCACGCCGAAACGCAATAGC CGAAAAACAAAAACAAAAAACAAAAAACCAAAAAACAAAACACATTAAACAGCCTGTGGGT GATCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATA CCCCCTCCCCAACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACG TTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAAGC GTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGC ACTACCCAGTGTAGATCAGGTGATGAGTCACCGCATTCCCCACGGCGACCGTGGCGGTGGCTGCG TTGGCGGCTGCCATGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCAAGAGTCTATTGA GCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCTAAGTGGGAGCACACACCCCTCAAGCCA GAGGGCAGTGTGCTGAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTT ATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCG GTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACATT ACAATTCAATTGTTAAGTTGAATACAGCAAAATGGGGATTCTGCCTAGCCCCGGCATGCCCGCCCTCTG AGCCTGGTGAGCCTGCTGAGCGTCTCTGTTGATGGGTGCGTGCCGAGACCGGCACACAGTGCCTGA ACCTGACAACCTAGAACACAGCTGCCCCAGCCTACACCAACAGCTTTACTAGGGGAGTGTACTACCCTG ATAAGTGCTTCAGGAGCAGTGTGCTGCACAGCACCCAGGACCTCTTTTGGCCTTTTTCAGTAACGTGAC CTGTTTCCACGCTATCCACGTGAGCGGCACAAACGGCACAAAACGGTTCGATAACCCAGTGTGCCCTT CAACGACGGCGTCTACTTCGCCAGCACTGAGAAGTCCAACATTATCAGGGGGTGGATCTTCGGGACAAC CCTGGACAGCAAGACACAGAGCCTCTGATTGTGAACAACGCCACCAACGTGGTCTAAGGTCTGCGA GTTCCAGTTCTGCAACGACCCCTTTCTGGGGGTGTACTACCACAAGAACAACAAGTCTGGATGGAGTC CGAGTTTCCGGGTGTACTCCAGCGCCAACAACCTGCACATTTGAGTACGTGAGCCAGCCCTTCTTGATGGA TCTGGAGGGCAAGCAGGGGAACCTTAAGAACCTTCGGGAGTTCGTCTTAAGAACATCGACGGGTACTT TAAGATCTACAGCAAGCACACCCCTATTAACCTGGTCCGCGATCTGCCCCAGGGCTTCAGCGCCCTGGA GCCACTGGTGGACCTCCCCATCGGGATCAACATTACTAGGTTCCAGACACTCCTGGCCCTGCACCGGA GCTACCTGACTCCCGCGGATAGCTCCAGCGGGTGGACTGCCGGGGCCGCCCTACTACGTGCGCTA CCTGCAGCCCAGGACATTCTGCTCAAGTACAACGAGAACGGCACCATCACTGATGCCGTGGACTGCG CTCTCGATCCCTGAGCGAGACAAAGTGACCCCTGAAGTCTTACCCTGGAGAAAGGGGATCTACCAG ACATCTAACTTCAGGGTGCAGCCCACTGAGTCCATTGTGAGATTTCTAACATACCAACCTGTGCCCT TCGGCGAGGTCTTTAACGCCACACGGTTCGCCAGCGTCTACGCTTGAACAGGAAGCGGATCTCCAAC TGCGTGGCCGACTACAGCGTGTGTACAACAGCGCCTCCTTACGACCTTTAAGTGCTACGGCGTGAG CCCCACTAAGCTGAACGACCTGTGCTTCACTAACGTGTACGCTGATAGCTTCGTGATTAGGGGGGATGA GGTCCGGCAGATTGCCCCCGGCCAGACAGGGAAGATCGCTGACTACAATAAGCTGCCGATGACT TCACAGGGTGTGTATCGCTGGAACAGCAACAACCTGGATAGCAAGGTGGGGGGAACTACAATACTAC CTGTACAGGCTGTTCAAGAAAGTCCAACCTGAAGCCCTTCGAGCGGGATATCAGCACTGAGATCTACCAG GCCGGGAGCACCCCTGCAACGGCGTGCAGGGCTTTAACTGCTACTTTCCCTCCAGAGCTACGGCTT CCAACCCACCAACGGCGTGGGGTACCAGCCCTACCGGGTCTGGTGCTTAGCTTCAGCTCCTGCACG

		CCCCTGCAACCGTGTGCGGCCCAAGAAGAGCACAAACCTCGTGAAGAACAAGTGCCTGAACCTCAACT TCAACGGGCTGACCGGCACCGGCGTGCTGACTGAGAGTAACAAGAAGTTTCTCCCCTTCCAGCAGTGC GGGAGGGATATTGCCGACACTACTGATGCCGTGCGCGACCCTCAGACCCTTGAGATTCTGGACATCACT CCCTGCTCCTTCGGCGGAGTGTCCTGATCACACCAGGGACCAACACTAGCAACCAGGTGCGCGTCTT GTACCAGGACGTGAACGTGACCGAGGTGCCGTGGCCATCCACGCCGATCAGCTGACCCCAACTTGGC GGGTGTACTCCACCGGGAGCAACGTGTTTCAGACACGGGCGGGGTGCCTGATCGGGGCTGAGCATGT CAACAACAGCTACGAGTGCGATATTCTATCGCGCGCGGATTTCGCCAGCTACCAGACACAGACAAA CAGCCCTGGCTCGGCCAGTCCGTGCCAGCCAGAGCATCATCGCTACACAATGTCTTGGGGGCCG AGAACAGCGTCGCTACAGCAACAACAGCATTGCCATCCCCACCAACTTCACAATCAGCGTCACAACCG AGATCCTGCCCGTGAGCATGACTAAGACATCCGTGGATTGCACAATGTACATCTGCGGCGACTCCACAG AGTGCAGTAACCTGCTGCTCCAGTACGGGTCTTTTGTACCCAGCTGAACAGAGCCTTGACCGGCATTG CCGTGGAGCAGGACAAGAACACCCAGGAGGTGTTTCCCAGGTGAAGCAGATCTACAAGACACCCCCC ATAAAGGACTTCGGGGGCTTTAACTTCTCCAGATCTTGCCCGACCCAGCAAGCCCTCCAAGCGGTCA TTCATTGAGGACCTGCTGTTTAAACAGGTGACCCTCGCTGATGCAGGGTTTATCAAGCAGTACGGCGAT TGCCTGGGGGACATTGCCGCTAGGGACTTGATCTGCGCCAGAAAGTTCAACGGCCTGACCGTGCTCCC CCCCCTGTTGACAGACGAGATGATCGCCAGTACACTAGCGCTCTGCTGGCCGGCACTATCACCAGCG GCTGGACCTTGGGGCCGGCGCCGCCCTCCAGATTCCATTGCCATGCAGATGGCCTACCGCTTCAAC GGCATCGGCGTGACTCAGAACGTGCTGTACGAGAACAGAAAGCTGATTGCCAACAGTTCAACTCCGC CATCGGAAGATCCAGGACAGCCTGTCCAGCACAGCATCCGCCCTGGGGAAGTTGCAGGATGTGGTGA ACCAGAACGCCAGGCCCTGAACACACTGGTGAAGCAGCTGAGCAGCAACTTTGGCGCCATCAGTAGC GTCCTGAACGACATCCTGTCACGGCTTGATCCACCAGAGGCCGAGGTCCAGATCGATCGGCTGATTACA GGCCGGCTGCAGAGTCTCCAGACCTACGTCACCCAGCAGCTGATCCGGGCTGCCGAGATTAGAGCCTC CGCCAATCTGCGCGCCACAAGATGAGTGAATGCGTCTGGGCCAGAGCAAAAGGGTCGACTTTTTCG GGAAGGGGTACCACCTGATGTCTTTCCCAGTCCGCCCTCACGGAGTGGTGTCTCTCCATGTCACAT ACGTGCCTGCCAGGAGAAGAATTCCTACAGCCCCGCCATTTGCCACGATGGGAAGGCCCATTTTC CCTAGAGAGGGCGTGTTCTGTCCAACGGGACCCACTGGTTCGTGACTACGCGGAACTTTTATGAACCC CAGATTATCACAACCGATAACACATTTGTCTCCGGCAACTGTGATGTGGTCATCGGCATCGTGAACAACA CCGTGTACGACCCACTGCAGCCTGAGCTTGACAGCTTTAAGGAGGAGCTGGACAAGTACTTTAAGAACC ACACCAGCCCTGATGTGGACCTTGCGGACATCAGCGGCATTAACGCCAGCGTGGTGAACATTCAGAAG GAGATCGATAGACTGAACGAGGTGGCCAAGAACTTGAACGAGTCACTGATTGACCTGCAGGAGCTGGG GAAGTACGAGCAGTACATTAAGGGGTCCGGGAGGGAGAACCTGTACTTTCAGGGGGGGGGGGGGTCC GGCTACATTCCCAGGCCCAAGGGATGGCCAGGCCACGTGAGGAAGGATGGGGAGTGGGTGCTGC TGAGCACCTTCTTGGGACGCGGCAGCAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAG GCGGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTG GCCACCCCTCGTGACCACCTGACCTACGGCGTGAGTGCTTCAGCCGCTACCCCGACCACATGAAGC AGCACGACTTCTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCT GAAGGGCATCGACTTCAAGGAGGACGGCAACATCTGGGGCACAAGCTGGAGTACAATAACAACAGCC ACAACGTCTATATCATGGCCGACAAGCAGAAGAAGCGCATCAAGGTGAACCTCAAGATCCGCCACAACA TCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCGCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAAGCGCG ATCAGTGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG TAAAAAAAACAAAAACAAAACGGCTATTATCGGTTACCGGCGAGACGCTACGGACTT
27	Cas9 GOI	AAAATCCGTTGACCTTAAACGGTCTGTGGGTTCAAGTCCCTCCACCCACGCCGGAACGCAATAGC CGAAAAACAAAAACAAAAAAACAAAAAAACCAAAAAACAAAACACATTAAACAGCCTGTGGGTT GATCCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACG TTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAAGC GTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGC ACTACCCAGTGTAGATCAGGTGATGAGTACCGCATTCCCCACGGCGACCGTGGCGGTGGCTGCG TTGGCGGCCTGCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGA GCTAGTTGGTAGTCTCCGGCCCCCTGAATGCGGCTAATCTAACTGCGGAGCACACACCCCTCAAGCCA GAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTT ATTCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCG GTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACATT ACAATTCAATTGTTAAGTTGAATACAGCAAAATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAA TAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAA

	TACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTATTTGACAGTGGAGAGACAGCGGAA GCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAG GAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTCATCGACTTGAAGAGTCTTTTTGGT GGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGAAATATAGTAGATGAAGTTGCTTATCATGAG AAATATCCAATCTATCATCTGCGAAAAAATGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAAT CTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTGATTGAGGGAGATTTAAATCCTGATA ATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCTATT AACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATC TCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTT GACCCCTAATTTTAAATCAAATTTGATTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACG ATGATGATTTAGATAATTTATTGGCGCAAATGGAGATCAATATGCTGATTTGTTTTGGCAGCTAAGAAT TTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTAAATACTGAAATACTAAGGCTCCCCTATCAGC TTCAATGATTAACGCTACGATGAACATCATCAAGACTTGACTCTTTTAAAGCTTTAGTCGACAACAAC TTCCAGAAAAGTATAAAGAAATCTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGA GCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTAGAAAAATGGATGGTACTGAGGAATTATT GGTGAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTGACAACGGCTCTATTTCCCATCAA ATTCATTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAGACAATCG TGAGAAGATTGAAAAATCTTGACTTTTGAATTCCTATTATGTTGGTCCATTGGCGCGTGGCAATAGTC GTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTGAGAAGTTGTCGATAA AGGTGCTTCAGCTCAATCATTATTGAACGCATGACAACTTTGATAAAAAATCTTCCAAATGAAAAAGTAC TACCAAAACATAGTTTGCTTTATGAGTATTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG AAGGAATGCGAAAACAGCATTCTTTAGGTGAACAGAAGAAAGCCATTGTTGATTACTCTTCAAAAC AAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAAATAGAATGTTTTGATAGTGTGA AATTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAGA TAAAGATTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAATTTGACCTATTTGA AGATAGGGAGATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAG CTTAAACGTCGCCGTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGC AATCTGGCAAAACAATATTAGATTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTATGCAGCTGATC CATGATGATAGTTTGACATTTAAGAAGACATTCAAAAAGCACAAAGTGCTGGACAAGGCGATAGTTTAC ATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAATAAAGGTATTTTACAGACTGTAAAGTTGTT GATGAATTTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC AGACAACCTAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAAT TAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTAT TATCTCCAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATTAATCGTTTAAAGTGATTATGATGT CGATCACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATA AAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACCTATTGGAGACA ACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTAACGAAAGCTGAACGTGGAGGTTTGA GTGAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGC ACAAATTTTGATAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGA TTACCTTAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTCCAATTCTATAAAGTACGTGAGATTAACA ATTACCATCATGCCCATGATGCGTATCTAAATGCCGTGTTGGAAGTCTTTGATTAAGAAATATCCAAAA CTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTGCTAAAATGATTGCTAAGTCTGAGCA AGAAATAGGCAAAGCAACCGCAAAATATTTCTTTACTCTAATATCATGAACCTCTTCAAAACAGAAATTAC ACTTGCAAATGGAGAGATTGCGAAACGCCCTCTAATCGAAACTAATGGGGAACTGGAGAAATTGTCTG GGATAAAGGGCGAGATTTTGCCACAGTGCAGAAAGTATTGTCCATGCCCAAGTCAATATTGTCAAGAAA ACAGAAGTACAGACAGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTG CTCGTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCT AGTGGTTGCTAAGGTGAAAAAGGGAATCGAAGAAGTTAAATCCGTTAAAGAGTTACTAGGGATCACA ATTATGGAAGAAGTTCCCTTTGAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAA AAAAGACTTAATCATTAACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTAACGGATGCTGG CTAGTGCCGGAAGATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTATATTTA GCTAGTCATTATGAAAAGTTGAAGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGC ATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTATTTTAGCAGATGCC AATTTAGATAAAGTTCTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATAT TATTCATTTATTTACGTTGACGAATCTTGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCG TAAACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATG AAACACGCATTGATTGAGTCAGCTAGGAGGTGACGGCAGCGGCAGCAGTATGGTGAGCAAGGGCGAG GAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACCGGCCACAAGTTTCAG
--	--

		CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGAGTGCTTCAGCCG CTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGC GCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCT GGAGTAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAA CTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCC CCATCGGGCAGCGCCCCGTGCTGCTGCCCACAACCACTACCTGAGCACCAGTCCGCCCTGAGCAA GACCCCAACGAGAAGCGCGATCAGATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGG CATGGACGAGCTGTACAAGTAAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCGAGACGCT ACGGACTT
28	Factor 8 GOI	AAAATCCGTTGACCTTAAACGGTCTGTGTGGGTTCAAGTCCCTCCACCCCCACGCCGAAACGCAATAGC CGAAAAACAAAAACAAAAAAACAAAAAAACAAAAAAACAAAAACACATTAAACAGCCTGTGGGTT GATCCCAACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACG TTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGTTGAAGGAGAAAGC GTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGC ACTACCCAGTGTAGATCAGGTGATGAGTACCGCATTCCCCACGGCGACCGTGGCGGTGGCTGCG TTGGCGGCCCTGCCATGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGA GCTAGTTGGTAGTCTCCGGCCCCCTGAATGCGGCTAATCTAAGTGCAGGACACACACCCCTCAAGCCA GAGGGCAGTGTGTCTGAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCGGTGTTCATTTT ATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCG GTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACATT ACAATTCATTGTTAAGTTGAATACAGCAAAATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCTTTT GCGATTCTGCTTTAGTGCCACCAGAAGATACTACCTGGGTGCAGTGGAAGTGCATGGGACTATATGCA AAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTTCTCTAGAGTGCCAAAATCTTTCCATTCAAC ACCTCAGTCGTGTACAAAAGACTCTGTTTGTAGAATTACGGATCACCTTTTCAACATCGCTAAGCCAA GGCCACCCTGGATGGGTCTGCTAGGTCTACCATCCAGGCTGAGGTTTATGATACAGTGGTCATTACAC TTAAGAACATGGCTTCCATCCTGTCACTCTTATGCTGTTGGTGTATCTACTGGAAAGCTTCTGAGGG AGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGATGATAAGTCTTCCCTGGTGGAAGCCA TACATATGTCTGCCAGGTCTGAAAGAGAATGGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCA TATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGTAGAG AAGGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAAATTTATACTACTTTTTGCTGTATTTGATGA AGGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTCGGGC CTGGCCTAAAATGCACACAGTCAATGGTTATGTAACAGGTCTCTGCCAGGTCTGATTGGATGCCACAG GAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCTGAAGTGCATCAATATTCCTCGAAGGT CACACATTTCTGTGAGGAACCATCGCCAGGCGTCTTGAAAATCTGCCAATAACTTTCTTACTGCTC AAACACTCTTGATGGACCTTGACAGTTTCTACTGTTTTGTATATCTCTTCCACCAACATGATGGCATG GAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAATATGAAGAAGCG GAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGGTTTATGATGACAACCTCTCCT CCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTAAAACCTGGGTACATTACATTGCTGCTGAAGA GGAGGACTGGGACTATGCTCCCTTAGTCCTCGCCCCGATGACAGAAGTTATAAAAGTCAATATTTGAAC AATGGCCCTCAGCGGATTGGTAGGAAGTACAAAAAGTCCGATTTATGGCATAACAGATGAAACCTTTA AGACTCGTGAAGCTATTGAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACAC ACTGTTGATTATATTTAAGAATCAAGCAAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCC GTCCTTTGTATTCAAGGAGATTACAAAAGGTGTAACATTTGAAGATTTTCAATCTGCCAGGAGAA ATATTCAAATATAAATGGACAGTACTGTAGAAGATGGGCCAACTAAATCAGATCCTCGGTGCCTGACCC GCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCCTCATCTG CTACAAAGAATCTGTAGATCAAAGAGGAAACAGATAATGTCAGACAAGAGGAATGTCATCCTGTTTTCT GTATTTGATGAGAACCGAAGCTGGTACCTCACAGAGAATATACAACGCTTTCTCCCCAATCCAGCTGGAG TGCAGCTTGAGGATCCAGAGTTCCAAGCCTCAACATCATGCACAGCATCAATGGCTATGTTTTGATAG TTTGCAGTTGTAGTTTGTTCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCACAGACTGAC TTCTTTCTGTCTTCTCTGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTACCCCTATT CCCATTCTCAGGAGAACTGTCTTCATGTGATGGAACCCAGGTCTATGGATTCTGGGGTGGCCACAA CTCAGACTTTCGGAACAGAGGCATGACCGCCTTACTGAAGTTTCTAGTTGTGACAAGAACACTGGTGA TTATTACGAGGACAGTTATGAAGATATTTAGCATACTTGTGAGTAAAAACAATGCCATTGAACCAAGAA GCTTCTCCAGAATTCAAGACACCCCTAGCACTAGGCAAAAGCAATTTAATGCCACCACAATTCCAGAAAA TGACATAGAGAAGACTGACCCTTGGTTTGCACACAGAACACCTATGCCTAAAATACAAAATGTCTCCTCT

	AGTGATTGTTGATGCTCTTGCGACAGAGTCCTACTCCACATGGGCTATCCTTATCTGATCTCCAAGAAG CCAAATATGAGACTTTTTCTGATGATCCATCACCTGGAGCAATAGACAGTAATAACAGCCTGTCTGAAAT GACACACTTCAGGCCACAGCTCCATCACAGTGGGGACATGGTATTTACCCCTGAGTCAGGCCTCCAATT AAGATTAAATGAGAACTGGGGACAACGCAGCAACAGAGTTGAAGAACTTGATTTCAAAGTTTCTAGT ACATCAAATAATCTGATTCAACAATTCATCAGACAATTTGGCAGCAGGTAAGTATAACAAGTTCCTT AGGACCCCCAAGTATGCCAGTTCATTATGATAGTCAATTAGATACCACTCTATTTGGCAAAAAGTCATCTC CCCTTACTGAGTCTGGTGGACCTCTGAGCTTGAGTGAAGAAAATAATGATTCAAAGTTGTTAGAATCAGG TTTAATGAATAGCCAAGAAAGTTTCATGGGGAAAAATGTATCGTCAACAGAGAGTGGTAGGTTATTTAAA GGGAAAAGAGCTCATGGACCTGCTTTGTTGACTAAAGATAATGCCTTATTTCAAAGTTAGCATCTCTTTGTT AAAGACAAAACAACTTCCAATAATTCAGCAACTAATAGAAAGACTCACATTGATGGCCCATCATTATTAA TTGAGAATAGTCCATCAGTCTGGCAAAATATATTAGAAAGTGACACTGAGTTTAAAAAAGTGACACCTTTG ATTCATGACAGAATGCTTATGGACAAAATGCTACAGCTTTGAGGCTAAATCATATGTCAAATAAACTAC TTCATCAAAAACATGGAATGGTCCAACAGAAAAAGAGGGCCCCATTCCACCAGATGCACAAAATCCA GATATGTCGTTCTTAAGATGCTATTCTTGCCAGAATCAGCAAGGTGGATACAAAGGACTCATGGAAAGA ACTCTCTGAACCTCTGGGCAAGGCCCCAGTCCAAAGCAATTAGTATCCTTAGGACCAGAAAAATCTGTGG AAGGTGAGAATTTCTGTCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTTACAAGGACGTAGG ACTCAAAGAGATGGTTTTCCAAGCAGCAGAAACCTATTTCTTACTAATTTGGATAATTTACATGAAAATA ATACACACAATCAAGAAAAAAAATTCAGGAAGAAATAGAAAAGAGGAAACATTAATCCAAGAGAATGTA GTTTTGCCTCAGATACATACAGTACTGGCACTAAGAATTTTCATGAAGAACCTTTTCTTACTGAGCACTAG GCAAAATGTAGAAGTTTCATATGACGGGGCATATGCTCCAGTACTTCAAGATTTTAGGTCAATTAATGATT CAACAAATAGAACAAAGAAACACACAGCTCATTTCTCAAAAAAGGGGAGGAAGAAAACCTTGAAGGCTT GGGAAATCAAACCAAGCAAATTTAGAGAAATATGCATGCACCACAAGGATATCTCCTAATACAAGCCAG CAGAATTTGTACGCAACGTAGTAAGAGAGCTTTGAAACAATTCAGACTCCCACTAGAAGAAACAGAAC TTGAAAAAGGATAATTGTGGATGACACCTCAACCCAGTGGTCCAAAAACATGAACATTTGACCCCGAG CACCTCACACAGATAGACTACAATGAGAAGGAGAAAGGGGCCATTACTCAGTCTCCCTTATCAGATTG CCTTACGAGGAGTCATAGCATCCCTCAAGCAAATAGATCTCCATTACCCATTGCAAAGGTATCATCATTT CCATCTATTAGACCTATATATCTGACCAGGCTCTATTTCCAAGACAACCTTCTCATCTTCCAGCAGCATC TTATAGAAAGAAAGATTCTGGGGTCCAAGAAAGCAGTCATTTCTTACAAGGAGCCAAAAAAAATAACCTTT CTTTAGCCATTCTAACCTTGAGATGACTGGTGATCAAAGAGAGGTTGGCTCCCTGGGGACAAGTGCCA CAAATTCAGTCACATACAAGAAAGTTGAGAACACTGTTCTCCCGAAACCAGACTTGCCCAAAACATCTGG CAAAGTTGAATTGCTTCCAAAAGTTCATTTATCAGAAGGACCTATTCCTACGGAAACTAGCAATGGG TCTCCTGGCCATCTGGATCTCGTGAAGGGAGCCTTCTCAGGGAACAGAGGGAGCGATTAAGTGAAT GAAGCAAACAGACCTGGAAAAGTCCCTTTCTGAGAGTAGCAACAGAAAGCTCTGCAAGACTCCCTCC AAGCTATTGGATCCTCTTGCTTGGGATAACCACTATGGTACTCAGATACCAAAAGAAGAGTGGAATCCC AAGAGAAGTCACCAGAAAAACAGCTTTTAAAGAAAAGGATACCATTTTGCCCTGAACGCTTGTAAG CAATCATGCAATAGCAGCAATAAATGAGGGACAAAATAAGCCCGAAATAGAAGTCACCTGGGCAAGCA AGGTAGGACTGAAAGGCTGTGCTCTCAAAACCCACCAGTCTTGAACGCCATCAACGGGAAATAACTCG TACTACTCTTCAGTCAGATCAAGAGGAAATTGACTATGATGATACCATATCAGTTGAAATGAAGAAGGAA GATTTTGACATTTATGATGAGGATGAAAATCAGAGCCCCCGCAGCTTTCAAAGAAAACACGACACTATT TTATTGCTGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGTCCCCACATGTTCTAAGAAACAGGG CTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACTGATGGCTCCTTTACTCA GCCCTTATACCGTGGAGAATAAATGAACATTTGGGACTCCTGGGGCCATATATAAGAGCAGAAGTTGA AGATAATATCATGGTAACTTTAGAAATCAGGCCTCTCGTCCCTATTCTTCTATTCTAGCCTTATTTCTTA TGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGAAACCAAACTTAC TTTTGGAAGTGCAACATCATATGGCACCCACTAAAGATGAGTTTGACTGCAAGCCTGGGCTTATTTCT CTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGGTCTGCCACACTAACA CACTGAACCCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTACCATCTTTGATGA GACCAAAAGCTGGTACTTCACTGAAAATATGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGA AGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGCT TAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAACATCCATT TATTCATTTAGTGGACATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCT ATCCAGGTGTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAATTTGGCGGGTGAATGCCTTAT TGCGGAGCATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAGTGTGAGACTCCCT GGGAATGGCTTCTGGACACATTAGAGATTTTCAATTACAGCTTCAGGACAATATGGACAGTGGGCCCC AAAGCTGGCCAGACTTCATTATCCGGATCAATCAATGCCTGGAGCACCAAGGAGCCCTTTTCTGGATC AAGGTGGATCTGTTGGACCAATGATTATTCACGGCATCAAGACCCAGGGTGCCGTCAGAAGTTCTCC AGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAA ATTCCACTGGAACCTTAATGGTCTTCTTTGGCAATGTGGATTTCATCTGGGATAAAACACAATATTTTTAAC
--	--

		CCTCCAATTATTGCTCGATACATCCGTTTGCACCCAACCTCATTATAGCATTGCGAGCACTCTTCGCATGG AGTTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGC ACAGATTACTGCTTCATCCTACTTTACCAATATGTTTGCCACCTGGTCTCCTTCAAAGCTCGACTTCACC TCCAAGGGAGGAGTAATGCCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGTGCAAGTGGACTTCC AGAAGACAATGAAAGTCACAGGAGTAACACTCAGGGAGTAAAATCTCTGCTTACCAGCATGTATGTGAA GGAGTTCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTTCAAGATGGCAAAGTAAAG GTTTTTCAGGGAAATCAAGACTCCTTACACCTGTGGTGAACCTCTAGACCCACCGTTACTGACTCGCT ACCTTCGAATTCACCCCAAGAGTTGGGTGCACCAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAG GCACAGGACCTCTACGGCAGCGGCAGCAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGG TGCCCATCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGA GGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTCCCT GGCCACCTCGTGACCACCTGACCTACGGCGTGAGTGTTCAGCCGCTACCCCGACCACATGAAG CAGCAGCACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGACCATCTTCTTCAAGGAC GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCTGGTGAACCGCATCGAGC TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACAACAGC CACAACGTCTATATCATGGCCGACAAGCAGAGAAGACGGCATCAAGGTGAACCTCAAGATCCGCCACAAC ATCGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCG TGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGC GATCAGATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGCATGGACGAGCTGTACAA GTAAAAAACAACAAACACGCTATTATGCGTTACCGGCGAGACGCTACGGACTT
29	TRIC-V1.30 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAATACACTC GATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTATTATAAAGATCACGACATCGATTA CAAGGATGATGATGATAAGCTTAACTTTCCGAAGCCAACCGATTGACCGATATC
30	TRIC-V1.31 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAGCCGG CCGGAAGCCAACCGATACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGAT TATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGATGGTCTTCACTTAACTTTCCGAAG CCAACCGATTGACCGATATC
31	TRIC-V1.32 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAATACACTC GATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTATTATAAAGATCACGACATCGATTA CAAGGATGATGATGATAAGAGACGCTACGGACTTAACTTTCCGAAGCCAACCGATTGACCGATATC
32	TRIC-V1.33 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGT TGACCTTAAACGGTCGTGTGGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATG GTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGCTTAACTTTCCGAAGCCAACCG ATTGACCGATATC
33	TRIC-V1.34 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGT TGACCTTAAACGGTCGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGAT TATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGCTTAACTTTCCGAAGCCAACCGATTGA CCGATATC
34	TRIC-V1.35 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGT TGACCTTAAACGGTCGTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGAT TATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGCTTAACTTTCCGAAGCCAACCGATTGA CCGATATC

		TGACCTTTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGAT CACGACATCGATTACAAGGATGATGATGATAAGCTTAACTTTTCGGAAGCCAACCGATTGACCGATATC
35	TRIC-V1.36 (3*Flag)	GGGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC TCAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTA ACAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGA GGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCG TTTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGATCACGA CATCGATTACAAGGATGATGATGATAAGCTTAACTTTTCGGAAGCCAACCGATTGACCGATATC
36	TRIC-V1.37 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGT TGACCTTTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGAT CACGACATCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAACTTTTCGGAAGCCAACCGA TTGACCGATATC
37	TRIC-V1.38 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGT TACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGATCACGACA TCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAACTTTTCGGAAGCCAACCGATTGACCG ATATC
38	TRIC-V1.39 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGT TACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGATCACGACA TCGATTACAAGGATGATGATGATAAGAGACGCTACGGACTTAACTTTTCGGAAGCCAACCGATTGACCG ATATC
39	TRIC-V2.0 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAAACGGA TTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAAGATCACGAC ATCGATTACAAGGATGATGATGATAAGTCCGTCTTAAACACCCGGAAGCCAACCGATTGACCGATATC
40	TRIC-V2.1 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGT CTTATCATATTACACTCGATCTGGACTAAAGCTGCTCATGGATTACAAAGATCACGATGGTGATTATAAA GATCACGACATCGATTACAAGGATGATGATGATAAGACGGACTTAAACACCCGGAAGCCAACCGATTGA CCGATATC
41	TRIC-V2.2 (3*Flag)	GGTCAATCGGTTGGCTTCCGAAATAAAGTCCCTTAGAGAAGAAATTTCTTAAGTGGATGCTCTCAAAC CAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAA CAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGTTGATAAGT ACACTCGATCTGGACTAAAGCTGCTCGCCACCATGGATTACAAGGATGATGATGATAAGAGCAGCCATC ACCATCATCACCACCATCACAGTGTACTTAGATTACACCCGGAAGCCAACCGATTGACCGATATC
42	TRIC-V2 (EGFP)	GGATAATACGGTCAATCGGTTGGCTTCCGAAATAATTGACCCTTAGAGAAGAAATTTCTTAAGTGGATGC TCTCAAACCTCAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAACGGTCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTG AAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCTGGCCACCTCGTGACCACCTGACCTACGG CGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCG AAGGCTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCACTACAAGACCCGCGCCGAGGTG AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA CATCCTGGGGCACAAGCTGGAGTACAACACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA

		GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACC ACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACC CAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGC CGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGCAGCGGCGAGGGCAGGGGAAGTCTTCTA ACATGCGGGGACGTGGAGGAAAATCCCGGCCAGGTACTCCATCATGGTGAGCAAGGGCGAGGAGC TGTTACCGGGGTGGTGGCCATCCTGGTCGAGCTGGACGGTGACGTTAAACACCCGGAAGCCAACCTGA TTGACCAAAAAAAAAAAGCGCCGCGATATC
43	TRIC-V2 (NanoLuc)	TCGCGCGTTTCGGTGATGACGGTGAACCTCTGACACATGCAGTCCCGGAGACGGTCACAGCTTGT CTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGG GCTGGCTTAAGTATGCGGCATCAGAGCAGATTGTAAGTGAAGTGCACCAATGCGGTGTGAAATACCGC ACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTGCCATTAGGCTGCGCAACTGTTGGGAA GGGCGATCGGTGCGGGCCTCATCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGAT TAAGTTGGGTAAAGCCAGGGTTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGCAACGCGATGAC GATGGATAGCGATTATCGATGAGCTGACCCGATCGCCGCCGCGGAGGGTTCGCTTTGAGACGGGCG ACAGATTAATACGACTCACTATAGGGATAATACGGTCAATCGGTTGGCTTCCGAAATAAAGGTCTTAG AGAAGAAATCTTTAAGTGGATGCTCTCAACTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCC TGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAACAACAGATAAATTACAGCTAATCGGAAGGTGCAGAGA CTCGACGGGAGCTACCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCA GTAGCGAAAGCTGCGGGAGAATGTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTAC GCCGAACATGATCGACTATTTGCGACGGCCGTATGAAGGCATCGCCGTTCGACGGCAAAAAGATCAC TGTAACAGGGACCCTGTGGAACGGCAACAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCC TGCTGTTCCGAGTAACCATCAACGGAGTGACCGGTGGCGGTGTGCGAACGCATTCTGGCGGGCAGC GGCGAGGGCAGGGGAAGTCTTCAACATGCGGGGACGTGGAGGAAAATCCCGGCCAGGTACTCCCA TCATGGTCTTCACACTCGAAGATTTGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAG TCCTTGAAAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGGGGTGTCCGTAATCCGATCCAAAGGA TTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCGATGAAGGTCTGAGCG GCGACCAATGGGCCAGATCGAAAAATTTTAAAGGTGGTGTACCCTGTGGATGATCACCCTTTACACC CGGAAGCCAATGATTGACCAAAAAAAAAAAGCGGCCGCGATATCATCAGTTCTGGACCAGCGAGCTG TGCTGCGACTCGTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCC ACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATT AATTGCGTTGCGCTCACTGCCCCGTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG CCAACGCGCGGGGAGAGGCGGTTTGCATTGGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGC GCTCGGTGTTGCGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAA TCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCA GAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCGCT CTCCTGTTCCGACCTGTGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTT CTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGATGTCGTTCCGCTCAAGCTGGGCTGTGTGCACG AACCCCGCTTCAGCCGACCGCTGCGCCTTATCCGGTAATATCGTCTTGAAGTCAACCCGGTAAGAC ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT ACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGC TGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCG GTGGTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTT TCTACGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA GGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAGTATATATGAGTAACTT GGTCTGACAGTTAGAAAACTCATCGAGCATCAAGTGAAGTGAATTTATTCATATCAGGATTATCAATA CCATATTTTGAAGGCGGTTTCTGTAATGAAGGAGAAAACTACCGAGGCGAGTTCATAGGATGGCAA GATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCTCGTCAA AATAAGGTTATCAAGTGAGAAATACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAGAGTTTATGC ATTTCTTTCCAGACTTGTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCACCAACCAAAC CGTTATTCATTCTGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCCGTTAAAGGACAATTACAAAC AGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATA TTCTTCTAATACCTGGAATGCTGTTTTCCCTGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTA CGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATCCGTGACCCAGTTTGGCTGACCATCTCATCTG TAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTGGCGCATCGGGCTTCCCATACAA TCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCGAGCCCAATTTATACCATATAAATCAGCATCC ATGTTGGAATTAATCGCGGCTTCAAGCAAGACGTTTCCGTTGAATATGGCTCATTTTGTCTTCTTATGC TCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTTATTCATTATGGTGAAAGTTGGAAC

		CTCTTACGTGCCGATCACACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGA CATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCATC
44	TRIC-V2 (Znf609)	GGATAATACGGTCAATCGGTTGGCTTCCGAAATAACCGCCCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GGGAAGAGGGGGACAGTATGATGGAAGTGCAGGGGTGGATACAGGAGCTGTGGAGCCACTTGGGAGT ATAGCTATTGAGCCTGGGGCAGCGCTCAATCCTTTGGGAACTAAACCGGAGCCAGAGGAAGGGGAGAA TGAGTGTGCGCTGCTAAAGAAAGTCAAGTCTGAAAAGCAATGATGTTGCCACTGGGCATGTACTGACC AATGTGGCAGGTCTGAGAACATAGCTGAAGCTGAAAATAGGAAAGCTGGGGGCAAGGAAGAGCCTTGA ATCTTGAGGTGGGACGTTGACTCTAAGATGTCCTTGAGCAGTGGAGCCTCCGGAGGGAAAGGAGTGGGA TGCAAACCCGTTGAGACATACGACAGTGGGGATGAATGGGACATTGGAGTAGGGAATCTCATCATTGA CCTGGACGCCGATCTGGAAGAGGACCAGCAGAACTGGAAATGTCAGGCTCAAAGGAGGTGGGGATAC CGGCTCCCAATGCTGTGCCACACTACCAGACAACATCAAGTTTGTACCCCACTGCCAGGTCTCAAG GGAAGGAAGGCAATCAAAATCCAAAAGGAGTAAGAGTGGCAAAGACACTAGCAAACCCACTCCAGGGA CTTCCCTGTTCACTCCAAGTGAGGGGGCAGCTAGCAAGAAAGAGGTGCAGGGGCGCTCAGGAGATGGT GCCAATGCTGGAGGCCTGGTTGCTGCTATTGCTCCCAAGGGCTCAGAGAAGGCGGCTAAGGCATCCCG CAGTGTAGCCGGTTCAAAAAGGAGAAGGAGAACAGCTCATCTAAGAGCAAGAAGGAGAGAAGCGAAG GAGTGGGGACTTGTTCAGAAAAGGATCCTGGGGTCTCCAGCCTGTCCCCCTCGGTGGACACCCGGAA GCCAAGTATTGACCAAAAAAAAAAAGCGGCCGCGATATC
45	TRIC-V2 (EGFP), 15nt stem	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAATCCGTAGCGTCTCGCGAAAAACAAAAACAAAAAACAAAAAACAAAAAACAAAAACAAAAAC ATTAACACAGCCTGTGGGTTGATCCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTAC CTTTGTGCGCCTGTTTTATACCCCTCCCCCACTGTAAGTGAAGTAACACACACCCGATCAACAGTCA GCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTC ACGCGGTTGAAGGAGAAAGCGTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAAG TTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTACCCGATTCCCCACGGG CGACCGTGGCGGTGGCTGCGTTGGCGGCCGTGCCATGGGGAAACCCATGGGACGCTCTAATACAGAC ATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCCGTAATGCGGCTAATCTAAGTGC GAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTT TGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATA GCTATTGGATTGGCCATCCGGTACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGC TTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGA GCTGTTACCCGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCG TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACACCCGG CAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGCAGTGTTCAGCCGCT ACCCCGACCACATGAAGCAGCAGCACTTCTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGC ACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCAGGGCGACACCCCT GGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG AGTACAACCTACAACAGCCACAACGCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAC TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCT ATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGA CCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCA TGGACGAGCTGTACAAGTAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCGAGAGCGCTAC GGACTTAAACACCCGGAAGCCAAGTATTGACCAAAAAAAAAAAGCGGCCGCGATATC
46	TRIC-V2 (EGFP), 25nt stem	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAATCCGTAGCGTCTCGCCGTAACGCAGAAAAACAAAAACAAAAAACAAAAAACAAAAAACAAAA AACAAAAACACATTAACAGCCTGTGGGTTGATCCCACCCACAGGCCATTGGGCGCTAGCACTCTGGT ATCACGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCCACTGTAAGTGAAGTAACACACACCGA TCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAAT AGACTGCTCACGCGTTGAAGGAGAAAGCGTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACA CCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTACCCGATT

		CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAACCCATGGGACGCTCT AATACAGACATGGTGCAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATC CTAACTGCGGAGCACACACCTCAAGCCAGAGGGCAGTGTGTCGTAAACGGCAACTCTGCAGCGGAAC CGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGT TACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATA CCACTTAGCTTGAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAG GGCGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA AGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGC ACCACCGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGACGTGCTT CAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCC AGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGC GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCA CAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAA GGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGA ACACCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGGATCAC TCTCGGCATGGACGAGCTGTACAAGTAAAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCGA GACGCTACGACTTAAACACCCGGAAGCCAACCTGATTACCAAAAAAAAAAAGCGGCCGATATC
47	TRIC-V1 L15/R51 (EGFP), 40nt EGS	GGATAATACCCGTCGATTGTCCACTGGTCGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAA GAAATCTTTAAGTGATGCTCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAG CCAAGCCGAAGTAGTAATTAGTAAGTTAAACAAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCG ACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAGCCAATAGGCAGTAG CGAAAGCTGCGGGAGAATGAAATCCGTTGACCTTAAACGGTCTGTGGGTTCAAGTCCCTCCACCCCC ACGCCGGAACGCAATAGCCGAAAAACAAAAACAAAAACAAAAACAAAAACAAAAACAAAAACAC ATTAACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTAC CTTTGTGCGCCTGTTTTATACCCCTCCCCCACTGTAAGTAAAGTAACACACACCCGATCAACAGTCA GCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTC ACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACCTACTTCGAAAAACCTAGTAACACCGTGGAAG TTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTGATGAGTACCCGATTCCCCACGGG CGACCGTGCGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAACCCATGGGACGCTCTAATACAGAC ATGGTGGCAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCCCTGAATGCGGCTAATCTAAGTGC GAGCACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAAACGGGCAACTCTGCAGCGGAACCGACTACTT TGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATA GCTATTGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGC TTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGA GCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGG CAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGACGTCTCAGCCGCT ACCCCGACCACATGAAGCAGCAGCACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGC ACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCCT GGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG AGTACAACATAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAC TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCT ATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGA CCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGGATCACTCTCGGCA TGGACGAGCTGTACAAGTAAAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCGAGACGCTAC GGACTTAAACACCCGGAAGCCAACCTGATTGACCGACCAAGTGACAATCGACGGAACAAAAAAGCGG CCGCGATATC
48	PIE (EGFP) for immunogen icity	GGGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACCTAATCGGAAGGTGCAG AGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAGCCAATAG GCAGTAGCGAAAGCTGCAAGAGAATGAAATCCGTTGACCTTAAACGGTCTGTGGGTTCAAGTCCCTC CACCCACGCGGGAACGCAATAGCCGAAAAACAAAAACAAAAACAAAAACAAAAACAAAAAC AAAACACATTAACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATC ACGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCCACTGTAAGTAAAGTAACACACACCGATCA ACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAG ACTGCTCAGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACCTACTTCGAAAAACCTAGTAACACC GTGGAAGTTCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTGATGAGTACCCGATTCC CCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAACCCATGGGACGCTCTAA

		TACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCTGAATGCGGCTAATCCT AACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCG ACTACTTTGGGTGTCGGTGTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTA CCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACC ACTTAGCTTGAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGG GCGAGGAGCTGTTACCCGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAA GTTTCAGCGTGTCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTTCATCTGCA CCACCGGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGAGTGCTTC AGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCA GGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTGAGGGCG ACACCCTGGTGAACGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC AAGCTGGAGTACAATAACAAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAG GTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAA CACCCCATCGGCGACGGCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGA GCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACT CTCGGCATGGACGAGCTGTACAAGTAAAAAACAACAAAAACAAACGGCTATTATGCGTTACCGCGGAG ACGCTACGGACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAGTGGATGCTCTCAAACCTCAGGGAA ACCTAAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAGTGAGCA ATCGACGGATAACAGCATATCTAGCGGCCGCGATATC
49	PIE (Nluc) for immunogen icity	GGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAATAATCGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAAACCGTTGACCTTAAACGGTCGTGTGGGTTCAAGTCCCTCC ACCCACGCGCGAAACGCAATAGCCGAAAAACAAAAACAAAAAACAAAAAACAAAAAACAAAAACAA AAACACATTAACACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCGTAGCACTCTGGTATCA CGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCCAACTGTAACCTAGAAGTAACACACACCGATCAA CAGTCAGCGTGGCACCACGACGCTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGA CTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACCTACTTCGAAAAACCTAGTAACCCG TGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGTCAGGTCGATGAGTCACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGGAACCCATGGGACGCTCTAAT ACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCTGAATGCGGCTAATCCTA ACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGA CTACTTTGGGTGTCGGTGTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTAC CATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCA CTTAGCTTGAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTCTTCACACTC GAAGATTTGTTGGGACTGGCGACAGACAGCCGCTACAACCTGGACCAAGTCTTGAACAGGGAGG TGTGTCCAGTTTGTTCAGAATCTCGGGGTGTCGTAACCTCCGATCCAAAGGATTGCTCTGAGCGGTGA AAATGGGCTGAAGATCGACATCCATGTCATCATCCGATGAAGGTCTGAGCGGCGACCAATGGGCCA GATCGAAAAAATTTTAAAGGTGGTGTACCCTGTGGATGATCATCACTTAAAGGTGATCTGCACTATGGC ACACTGGTAATCGACGGGTTACGCCGAACATGATCGACTATTCGGACGGCGTATGAAGGCATCGCC GTGTTGACGGCAAAAAAGATCACTGTAAACAGGGACCCTGTGGAACGGCAACAAATTATCGACGAGCGC CTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGAGTGACCGGCTGGCGGCTGTG CGAACGCATTCTGGCGTGAAAAAACAACAAAAACAAACGGCTATTATGCGTTACCGGCGAGACGCTACG GACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAGTGGAACATCGACGG ATAACAGCATATCTAGCGGCCGCGATATC
50	TRIC-V2 (EGFP) for immunogen icity	GGATAATACGGTCAATCGGTTGGCTTCCGAAATAATTGACCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAACGGTCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTG AAGTTCATCTGCACCACCGGAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGG CGTGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCG AAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA CATCCTGGGGCACAAGCTGGAGTACAATAACAAGCCACAACGTCTATATCATGGCCGACAAGCAGAA GAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACC ACTACCAGCAGAACACCCCATCGGCGACGGCCCGTGCTGCTGCCGACAACCACTACCTGAGCACC CAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGC

		CGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTGAAAAACAAAAACAAAAACAAAAACAAAAAC AAACAAAAACAAAAACACATTAACACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCGCTA GCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCAACTGTAACCTAGAAGTAA CACACACCGATCAACAGTCAGCGTGGCACACCAGCCAGTTTTGATCAAGCACTTCTGTTACCCCGGAC TGAGTATCAATAGACTGCTCAGCGGTTGAAGGAGAAAGCGTTTCGTTATCCGGCCAACCTACTTCGAAAA ACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATGATCAGGTGATGA GTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGGAAACCCA TGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCTGAA TGGCGCTAATCCTAACTGCGGAGCACACACCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTC TGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTGGCTGCTTATGGTGACAA TTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTT GTTGGGTTTATACCACTTAGCTTGAAGAGGTTAAACATTACAATTTCATTGTTAAGTTGAATACAGCAAA ATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTCCCATCTGGTCGAGCTGGACGGTGACG TTAAACACCCGGAAGCCAACCTGATTGACCAAAAAAAAAAAGCGGCCGCGATATC
51	TRIC-V2 (Nluc) for immunogen icity	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATTTTAAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGAAGGTGCAGAGACTCGACGGGAGCTACCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAAGGAGGGCGCGCGGGTCCGGGCGCGGAAAAACAAAAACAAAAACAAAAACAAAAACAAAA AAAACAAAACACATTAACACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCGCTAGCACTCTG GTATCACGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCAACTGTAACCTAGAAGTAACACACACC GATCAACAGTCAGCGTGGCACACCAGCCAGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATC AATAGACTGCTCACGCGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACCTACTTCGAAAAACCTAGTA ACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATGATCAGGTGATGAGTACCCGC ATTCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGGAAACCATGGGACGC TCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCTGAATGCGGCTA ATCCTAACTGCGGAGCACACACCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGG AACCAGTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTGGCTGCTTATGGTGACAATTGAGAGAT CGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTT ATACCACTTAGCTTGAAGAGGTTAAACATTACAATTTCATTGTTAAGTTGAATACAGCAAAATGGTCTTC ACACTCGAAGATTTCTGTTGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCTTGAACA GGGAGGTGTGTCAGTTTGTTCAGAATCTCGGGGTGTCCGTAACCTCGATCCAAGGATTGTCTGAG CGGTGAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAT GGGCCAGATCGAAAAATTTTAAAGGTGGTGATCCCTGTGGATGATCATCACTTAAAGGTGATCCTGCAC TATGGCACACTGGTAATCGACGGGTTACGCCGAACATGATCGACTATTTCCGACGGCCGTATGAAGGC ATCGCCGTGTTGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGAC GAGCGCTGATCAACCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCG GCTGTGCGAACGCATTCTGGCGTAAAAACAAAAACAAAAACGCCCCCGGCCCGCGTCCTCCC TTAAACACCCGGAAGCCAACCTGATTGACCAAAAAAAAAAAGCGGCCGCGATATC
52	PIE-Ana	GGGAGACCCUCGACCGUCGAUUGUCCACUGGUCAACAUAUGAUGACUUAACUAAUCGGAAGGUGC AGAGACUCGACGGGAGCUACCCUAAACGUCAAGACGAGGUAAAGAGAGAGUCCAAUUCUCAAAGCCA AUAGGCAGUAGCGAAAGCUGCAAGAGAAUGAAAAUCCGUUGACCUUAAACGGUCUGUGGGUUAAG UCCCUCCACCCCCACGCCGGAACGCAUAGCCGAAAAACAAAAACAAAAACAAAAACAAAAACCA AAAAACAAAACACAUUAAACAGCCUGUGGGUUGAUCCACCCACAGGCCAUUGGGCGCUAGCAC UCUGGUUACACGGUACCUUUGUGCGCCUGUUUUAUACCCCUCCCCAACUGUAACUUAAGUAAC ACACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGUUAACCCGGA CUGAGUAUCAAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAACUACUUGCA AAAACCUAGUAACACCGUGGAAGUUGCAGAGUGUUCGUCACGACUACCCAGUGUAGAUACAGGUC GAUGAGUACCGCAUUCGCCACGGGCGACCGUGGCGUGGCGUGGCGUGGCGGCCUGCCCAUGGGG AAACCAUGGGACGCUCUAAUACAGACAUUGUGCGAAGAGUCUAAUUGAGCUAGUUGGUAGUCCUCCG GCCCCUGAUGCGGCUAAUCCUAAUCUGCGGAGCACACACCCUCAAGCCAGAGGGCAGUGUGUGUAA CGGGCAACUCUGCAGCGGAACGACUACUUUGGGUGUCCGUGUUUCAUUUUAUCCUUAUACUGGCU GCUUAUGGUGACAAUUGAGAGAUUGUACCAUUAUAGCUAAUUGGAUUGGCCAUCCGGUGACUAAUAGA GCUAAUUAUUAUCCCUUUGUUGGGUUUAUACCACUUAAGCUUUAAGAGGUUAAACAUUAACAAUUA UUGUUAAGUUAUACAGCAAAUUGGUGAGCAAGGGCGAGGAGCUUACCGGGGUGGUGCCCAU CCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGCGAGGGCGA UGCCACCUACGGCAAGCUGACCCUGAAGUUAUCUGACCAACGGCAAGCUGCCCGUGCCUGGCC CACCCUCGUGACCAACCCUGACCUACGGCGUGCAGUGCUUACGCCGUACCCGACCAUGAAGCAG

		CACGACUUCUUAAGUCCGCCAUGCCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCGCAUCGAGC UGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCCUGGGGACAAGCUGGAGUACAACUACAACAG CCACAACGUCUUAUUAUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGAUCGCCAC AACAUCCGAGGACGGCAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGACGGCC CCGUGCUGCUGCCCACAACCACUACCUGAGCACCAGUCCGCCUGAGCAAAGACCCCAACGAGAA GCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCU GUACAAGUAAAAAACAACAAACAAACGGCUAUUAUGCGUUAACGGCGAGACGCUACGGACUAAAA UAAUUGAGCCUUAAGAAGAAUUCUUAAGUGGAUGCUCUAAACUCAGGGAACCUAAAUUAGUU AUAGACAAGGCAUCCUGAGCCAAGCCGAAGUAGUAAUUAAGAACAGUGGACAAUCGACGGAUA ACAGCAUAUCUAGCGGCCGCGAU
53	TRIC-V1	GGGAUAAUACGGUCAAUACAGUUGGCUUCCGAAUUAUUGAGCCUUAAGAGAAGAAUUCUUAAGUGG AUGCUCUCAAACUCAGGGAACCUAAAUUCUAGCUUAUAGACAAGGCAUCCUGAGCCAAGCCGAAGUA GUAAUUAAGUUAAGUUAACAACAGAUAAUUAACGCUAAUCCGGAAGGUGCAGAGACUCGACGGGAGCUA CCCUAACGUCAAGACGAGGGUAAAGAGAGAGUCCAUAUUCUCAAAGCCAAUAGGCAUAGCGAAAGCU GCGGGAGAAUGAAAAUCCGUUGACCUUAAACGGUCGUGUGGGUUAAGUCCUCCACCCCCACGCC GGAAACGCAUAGCCGAAAAACAACAAAAACAAAAAACAACAAAAAACAACAAAAACAACACAUUAA AACAGCCUGUGGGUUGAUCCACCCACAGGCCAUUGGGCGCUAGCACUCUGGUUACACGGUACCU UUGUGCGCCUUGUUUAUACCCCUCCCCCAACUGUAACUUAAGAUAACACACACCGAUAACAGUC AGCGUGGCACACACAGCCACGUUUUAUCAAGCACUUCUGUUAACCCGGACUGAGUAUCAAUAGACUG CUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAAUACUUCGAAAAACCUAGUAACACCGU GGAAGUUGCAGAGUGUUCGCUACGACUACCCAGUGUAGAUACAGGUCGAUGAGUCACCGCAUUC CCCACGGGCGACCGUGGCUGGCGUGGCGUUGGCGGCCUGCCCAUGGGGAAACCUAGGGACGCUC UAAUACAGACAUGGUGCGAAGAGUCUUAUAGAGCUAGUUGGUAGUCCUCCGCCCCUGAAUGCGGCU AAUCCUAAUCUGCGGAGCACACACCCUCAAGCCAGAGGGCAGUGUGUCGUAAACGGGCAACUCUGCAGC GGAACCGACUACUUAUGGUGUCCGUGUUAUUAUCCUUAUACUGGCGCUUAUGGUGACAUAU GAGAGAUUGUUAACAUUAAGCUUAUUGGAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUAUCCCU UUGUUGGUGUUAUACCACUUAAGCUUGAAAGAGGUUAAACAUAUACAAUUAUUGUUAAGUUGAAUAC AGCAAAUUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGAC GGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAG CUGACCCUGAAGUUAUCUGCACCACCGCAAGCUGCCCGUGCCUUGGCCACCCUCGUGACCACC CUGACCUACGGCGUGCAGUGCUUACGCCGUACCCCGACCAUGAAGCAGCACGACUUCUUAAGU CCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACGACGGCAACUACAAGAC CCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCGCAUCGAGCUGAAGGGCAUCGACU CAAGGAGGACGGCAACAUCCUGGGGCAACAGCUGGAGUACAACUACAACAGCCACAACGUCUUAUAC AUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCCGAGGACGGCA GCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCAUCGGCGACGGCCCCGUGCUGCUGCCCG ACAACCACUACCUGAGCACCAGUCCGCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUUGGU CCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUGGCAUGGACGAGCUGUACAAGUAAAAA CAAAAAACAACGGCUAUUAUGCGUUAACGGCGAGACGCUACGGACUUAACACCCGGAAGCCAAC UGAUUGACCAAAAAAAAAAAGCGGCCGCGAU
54	Exemplary first loop	AAATAA
55	Exemplary second loop	ACACC
56	Exemplary first EGS	GGUCAUCCGGUUGGCUUCCG
57	Exemplary second EGS	CGGAAGCCAACCGAUUGACC
58	ACA sequence used for circularization of NLuc	GATCACCACCTTTAAGGTGATC

59	Nluc first (5') eACA portion	TTAAGGTGATC
60	Nluc second (3') eACA portion	GATCACCCT
61	TIRC V2 Nluc: eACA placed in UTR	GGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAAAAGGAGGGCGCGCGGGTCGGGGCGGCGGAAAAACAAATGGGCCAGATCGAAAAAACGAACATGA TCGACTATTTACACATTAACACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCGCTAGCACT CTGGTATCACGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCAACTGTAAGTTAGAAGTAACACAC ACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGT ATCAATAGACTGCTCAGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACCTTCCGAAAAACCTA GTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTGATGAGTCAC CGCATTCACCGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGGAAACCCATGGGA CGCTCTAATACAGACATGGTGCAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGG CTAATCCTAACTCGGAGCACACACCTCAAGCCAGAGGGCAGTGTGCTGAACGGGCAACTCTGCAG CGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATCTATACTGGCTGCTTATGGTGACAATTGAG AGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGG GTTTATACCACTAGCTTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGT CTTCACACTCGAAGATTTGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGA ACAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGGGGTGTCCGTAACCTCCGATCCAAGGATTGCTCT GAGCGGTGAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACC AAATGGGCCAGATCGAAAAATTTTAAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCT GCACTATGGCACACTGGTAATCGACGGGGTTACGCCAACATGATCGACTATTTGAGACGGCCGATGA AGGCATCGCCGTGTTGACGGCAAAAAGATCACTGTAAACAGGGACCCGTGGAACGGCAACAAAATTAT CGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCT GGCGGCTGTGCGAACGCATTCTGGCGTAAAAAAACAAAAACAAACCGCCCCCGGCCCGCGCTCC TCCCTTAACACCCGGAAGCCAACCTGATTGACCAAAAAAAGCGGCCGCGATATC
62	TRIC V2 EGFP: eACA placed in CDS	GGATAATACGGTCAATCGGTTGGCTTCCGAAATAATTGACCCTTAGAGAAGAAATCTTTAAGTGGATGC TCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGT CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAAT GAACGGTCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTG AAGTTCATCTGCACCACCGGAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGG CGTGCACTGCTTCAGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCG AAGGCTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA CATCCTGGGGCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA GAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACC ACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACC CAGTCCGCCCTGAGCAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTGCTGACCGC CGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAACAAAAACAAAAAACAAACAAA TGGGCCAGATCGAAAAAACGAACATGATCGACTATTTACACATTAACACAGCCTGTGGGTTGATCCCA CCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCTGTTTTATACCCCTC CCCCAACTGTAAGTTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGAT CAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCAGCGGTTGAAGGAGAAAGCGTTTCGT TATCCGGCCAACCTTCCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACC CCAGTGTAGATCAGGTGATGAGTACCGCATTCACCGGCGGCGACCGTGGCGGTGGCTGCGTTGGC GGCTGCCATGGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCAGAGTCTATTGAGCTAGT TGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCTCAAGCCAGAGGGC AGTGTGCTGAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATCTA TACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTA ATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTAGCTTGAAAGAGGTTAAACATTACAATTC TTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGTGGTGCCCATCCT

		GGTCGAGCTGGACGGTGACGTTAAACACCCGGAAGCCAAGTATTGACCAAAAAAAAAAAGCGGCCGCGATATC
63	TRIC-V1: 1.0-2, Short GOI	GGGUCAAUCGGUUGGCUUCCGAAUAAUUGAGCCUUAGAGAAGAAUUCUUUAAGUGGAUGCUCUCA AACUCAGGGAAACCUAAAUUCUAGCUAUAGACAAGGCCAUCCUGAGCCAAGCCGAAGUAGUAAUUAAGUA AGUUAACAACAGAUAAAUACAGCUAAUUCGGAAGGUGCAGAGACUCGACGGGAGCUACCCUAACGUC AAGACGAGGGUAAAGAGAGAGUCCAAUUCUCAAAGCCAAUAGGCAGUAGCGAAAGCUGCGGGAGAAU GAAAAUCCGUUGACCUUAAACGGUCGUGUGGUACACUCGAUCUGGACUAAAGCUGCUCAUGGAUUAC AAAGAUCACGAUGGUGAUUUAAGAUACGACAUCAUUAACAAGGAUGAUGAUGAUAGAGACGCUA CGGACUUAACACCCGGAAGCCAACCGAUUGACCGAU
64	PIE-Fluc	GGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACATAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAGTCCCTCC ACCCCACGCCGGAACGCAATAGCCGAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAA AAACACATTAAACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCCTAGCACTCTGGTATCA CGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACGTAACTTAGAAGTAACACACACCGATCAA CAGTCAGCGTGGCACACCGACCGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGA CTGCTCACGCGGTTGAAGGAGAAAGCGTTTCGTTATCCGCCAACCTACTTCGAAAAACCTAGTAACACCG TGAAGATTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGAAACCCATGGGACGCTCTAAT ACAGACATGGTGCAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGCCCCGTAATGCGGCTAATCCTA ACTGCGGAGCACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAAACGGGCAACTCTGCAGCGGAACCGA CTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTAC CATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTGGGTTTATACCA CTTAGCTTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGAAGATGCCAAA AACATTAAGAAGGGCCAGCGCCATTCTACCCACTCGAAGACGGTACCGCCGGCGAGCAGCTGCACAA AGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCGCCTTACCGACGCACATATCGAGGTGGACA TTACCTACGCCGAGTACTTCGAGATGAGCGTTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATA CAAACCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCCCGTGTGGGTGCCCTGT TCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGC ATCAGCCAGCCACCGTCGTATTCTGTGAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAA GCTACCGATCATACAAAGATCATCATCATGGATAGCAAGACCGACTACCAGGGCTTCCAAGCATGTACA CCTTCGTGACTTCCCATTTGCCACCCGGCTTCAACGAGTACGACTTCGTGCCGAGAGCTTCGACCGGG ACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCAAGGGCGTAGCCCTACCG CACCGCACCGCTTGTGTCCGATTAGTCATGCCCGGACCCCATCTTCGGCAACCGAGATCATCCCGAC ACCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCTTCGGCATGTTACCACGCTGGGCTACTTGATC TCGGCTTTTCGGGTCGTGCTCATGTACCGCTTCGAGGAGGAGCTATTCTTCGCGAGCTTGCAAGACTAT AAGATTCAATCTGCCCTGCTGGTGCCACACTATTTAGCTTCTTCGCTAAGAGCACCCCTATTGATAAGT ACGACCTAAGCAACTTGCACGAGATCGCCAGCGCGGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGC CGTGGCTAAACGCTTCCACCTACCAGGCATCCGCCAGGGCTACGGCCTGACAGAAACAACAGCGCCA TTCTGATCACCCCGAAGGGGACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCT AAGGTGGTGGACTTGGACACCGGCAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTG GCCCATGATCATGAGCGGCTACGTTAAACACCCCGAGGCTACAAACGCACTCATCGACAAGGACGGC TGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGTGGACCGGCTGAA GAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAAGTGGAGAGCATCCTGCTGCAACACC CCAACATCTTCGACGCCGGGGTGCCTGGCCTGCCGACGAGATGCCGGCGAGCTGCCCGCCGAGT CGTCGTGCTGGAACACGGTAAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTA CAACCGCCAAGAAGCTGCGCGGTGGTGTGTTCTGTTGACGAGGTGCCTAAAGGACTGACCGGTAAG TTGGACGCCCCGAAGATCCGCGAGATTCTCATTAAAGGCCAAGAAGGGCGGCAAGATCGCCGTGAGCAG TGATTACAAGGATGATGATGATAAGTAAAAAACAAAAAACAAACCGGCTATTATGCGTTACCGGCGAG ACGCTACGGACTTAATAATTGAGCCTTAAAGAGAAATTCTTTAAGTGGATGCTCTCAAACCTCAGGGAA ACCTAAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACAGTGGACA ATCGACGGATAACAGCATATCTAGCGGCCGCGATATC
65	PIE-Spike	GGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACATAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTTGACCTTAAACGGTCGTGTGGGTTCAGTCCCTCC ACCCCACGCCGGAACGCAATAGCCGAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAA AAACACATTAAACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCCTAGCACTCTGGTATCA CGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACGTAACTTAGAAGTAACACACACCGATCAA CGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCAACGTAACTTAGAAGTAACACACACCGATCAA

	CAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGA CTGCTCACGCGGTTGAAGGAGAAAGCGTTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCG TGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCCACTGATAGATCAGGTCGATGAGTCACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGAAACCCATGGGACGCTCTAAT ACAGACATGGTGCAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTA ACTGCGGAGCACACACCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGA CTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTAC CATATAGCTATTGGATTGGCCATCCGGTACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCA CTTAGCTTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGGGATTCTGCCT AGCCCCGGCATGCCCGCCCTCCTGAGCCTGGTGAAGCTGCTGAGCGTCTGTTGATGGGGTGGCTGG CCGAGACCGGCACACAGTGCCTGAACCTGACAAGTACAGCTGCCCCAGCCTACACCAACAGC TTTACTAGGGGAGTGTACTACCCTGATAAGTGTTCAGGAGCAGTGTGCTGCACAGCACCCAGGACCTC TTTTTGCCTTTTTTCAGTAACGTGACCTGTTCCACGCTATCCAGTGAGCGGCACAAACGGGCACAAAAC GGTTCGATAACCCAGTGTGCCCTTCAACGACGGCGTCTACTTCGCCAGCACTGAGAAGTCCAACATTA TCAGGGGGTGGATCTTCGGGACAACCTGGACAGCAAGACACAGAGCCTCCTGATTGTGAACAACGCC ACCAACGTGGTCATTAAGGTCTGCGAGTTCAGTTCTGCAACGACCCCTTTCTGGGGGTGTAATACCAC AAGAACAACAAGTCTGGATGGAGTCCGAGTTTCGGGTGTAATCCAGCGCCAACAACCTGCACATTTGAG TACGTGAGCCAGCCCTTCTTGATGGATCTGGAGGGCAAGCAGGGGAACTTAAGAACCCTTCGGGAGTTC GTCTTTAAGAATCGACGGGTACTTTAAGATCTACAGCAAGCACACCCCTATTAACCTGGTCCGCGATC TGCCCCAGGGCTTCAGCGCCCTGGAGCCACTGGTGGACCTCCCCATCGGGATCAACATTACTAGGTTT CAGACACTCCTGGCCCTGCACCGGAGCTACCTGACTCCCGGCGATAGCTCCAGCGGGTGGACTGCCG GGGCGCGCCCTACTACGTGGCTACCTGCAGCCAGGACATTCTGCTCAAGTACAACGAGAACGGC ACCATCACTGATGCCGTGGACTGCGCTCTCGATCCCTGAGCGAGACAAAGTGCACCCTGAAGTCTTC ACCGTGGAGAAGGGGATCTACCAGACATCTAACTTCAGGGTGCAGCCCACTGAGTCCATTGTGAGATT CCTAACATCACCAACCTGTGCCCTTCGGCGAGGTCTTAACGCCACACGGTTCGCCAGCGTCTACGCT TGGAACAGGAAGCGGATCTCCAAGTGCCTGGCCGACTACAGCGTGTGTACAACAGCGCTCCTTCAG CACCTTTAAGTGTACGGCGTGAAGCCCACTAAGCTGAACGACCTGTGCTTCACTAACGTGACGCTGA TAGCTTCGTGATTAGGGGCGATGAGGTCCGGCAGATTGCCCCGGCCAGACAGGGAAGATCGCTGACT ACAACATAAGCTGCCCCGATGACTTCACAGGGTGTGTGATCGCCTGGAACAGCAACAACCTGGATAGCA AGGTCCGGGGGAACTACAACCTACCTGTACAGGCTGTTTCAAGTCCAACCTGAAGCCCTTCGAGCGG GATATCAGCACTGAGATCTACCAGGCCGGGAGCACCCCTGCAACGGCGTGCAGGGCTTTAACTGCTA CTTTCCCTCCAGAGCTACGGCTTCAACCCACCAACGGCGTGGGGTACCAGCCCTACCGGGTCTGTGG TGCTTAGCTTCGAGCTCCTGCACGCCCTGCAACCGTGTGCGGCCCAAGAAGAGCACAAACCTCGTG AAGAACAAGTGCCTGAACCTTCAACTTCAACGGGCTGACCGGCACCGGCGTGTGACTGAGAGTAACAA GAAGTTTCTCCCTTCCAGAGTGCCTGGGAGGGATATTGCCGACACTACTGATGCCGTGCGCGACCCCTC AGACCTTGAGATTCTGGACATCACTCCCTGCTCCTTCGGCGGAGTGTCCGTGATCACACCAGGGACCA ACACTAGCAACCAAGTGCCTGCTTGTACCAGGACGTGAACCTGCACCGAGGTGCCCGTGGCCATCCAC GCCGATCAGCTGACCCCAACTTGGCGGGTGTACTCCACCGGGAGCAACGTGTTTCAGACACGGGCGCG GTGCTGATCGGGGCTGAGCATGTCAACAACAGCTACGAGTGCATATTCCTATCGGCGCCGGGATTT GCGCCAGCTACCAGACACAGACAACAGCCCTGGCTCGGCCAGCTCCGTGCCAGCCAGAGCATCATC GCCTACACAATGTCTTGGGGGCCGAGAACAGCGTGCCTACAGCAACAACAGCATTGCCATCCCCAC CAACTTACAATCAGCGTCAACAACGAGATCCTGCCGTGAGCATGACTAAGACATCCGTGGATTGCAC AATGTACATCTGCGGCGACTCCACAGAGTGCAGTAACCTGCTGCTCCAGTACGGGTCTTTGTACCCA GCTGAACAGAGCCTTGACCGGCATTGCCGTGGAGCAGGACAAGAACACCCAGGAGGTGTTTGCCAGG TGAAGCAGATCTACAAGACACCCCCATAAAGGACTTCGGGGGCTTTAACTTCTCCAGATCTTGCCCG ACCCAGCAAGCCCTCAAGCGGTCAATTCATTGAGGACCTGCTGTTTAAAGGTGACCCTCGCTGATG CAGGGTTTATCAAGCAGTACGGCGATTGCCGTGGGGGACATTGCCGTAGGGACTTGATCTGCGCCAG AAGTTCAACGGCCTGACCGTGTCCCCCCCCCTGTTGACAGACGAGATGATCGCCAGTACACTAGCGCT CTGCTGGCCGGCACTATCACCAGCGGTGACCTTTGGGGCCGGCGCCGCCCTCCAGATTCCATTGCG CATGCAGATGGCCTACCGCTTCAACGGCATCGCGTGAACGACGTGCTGTACGAGAACCAGAAGC TGATTGCCAACAGTTCAACTCCGCCATCGGCAAGATCCAGGACAGCCTGTCCAGCACAGCATCGCCCC TGGGGAAGTTGCAGGATGTGGTGAACAGAACGCCAGGCCCTGAACACACTGGTGAAGCAGCTGAGC AGCAACTTTGGCGCCATCAGTAGCGTCTGAACGACATCCTGTCACGGCTTGATCCACCAGAGGCCGA GGTCCAGATCGATCGGCTGATTACAGGCCGGTGCAGAGTCTCCAGACCTACGTACCCAGCAGCTGA TCCGGGCTGCCGAGATTAGAGCCTCCGCCAATCTCGCCGCCACAAAGATGAGTGAATGCGTCTGGGC CAGAGCAAAAGGGTCGACTTTTGGGGAAGGGGTACCACCTGATGTCCTTTCCCAAGTCCGCCCTCA CGGAGTGGTGTCTCCATGTCACATACGTGCTGCCAGGAGAGAAGAACTTCACTACAGCCCCCGCCAT TTGCCACGATGGGAAGGCCCATTTCCCTAGAGAGGGCGTGTCTGTGCAACGGGACCCACTGGTTGCG
--	---

		TGACTCAGCGGAAC TTTTATGAACCCAGATTATCACAACCGATAACACATTTGTCTCCGGCAACTGTGA TGTGGTCATCGGCATCGTGAACAACACCGTGTACGACCCACTGCAGCCTGAGCTTGACAGCTTTAAGGA GGAGCTGGACAAGTACTTTAAGAACCACACCAGCCCTGATGTGGACCTTGGCGACATCAGCGGCATTAA CGCCAGCGTGGTGAACATTCAGAAGGAGATCGATAGACTGAACGAGGTGGCCAAGAACTTGAACGAGT CACTGATTGACCTGCAGGAGCTGGGAAGTACGAGCAGTACATTAAGGGGTCCGGGAGGGAGAACCTG TACTTTAGGGGGGGGGGGGGTCCGGCTACATTCCCAGGCCCAAGGGATGGCCAGGCCTACGTGA GGAAGGATGGGAGTGGGTGCTGCTGAGCACCTTCTTGGGCAGCGGCAGCAGTATGGTGAGCAAGGG CGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAACCGGCCACAAG TTCAGCGTGTCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCAC CACCGGCAAGCTGCCCGTCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCA GCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAG GAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCAGGTGAAGTTCGAGGGCGA CACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA AGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAGCGCATCAAGG TGAAC TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAAC ACCCCATCGGCGACGGCCCGTGTCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAG CAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGTGGAGTTCGTACCGGCCCGGGATCACTC TCGGCATGGACGAGCTGTACAAGTAAAAAACAACCAACGGCTATTATCGTTACCGCGGAGA CGCTACGGACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAGTGGATGCTCTCAAACCTCAGGGAAA CCTAAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAGTGGACAA TCGACGGATAACAGCATATCTAGCGGCCGCGATATC
66	PIE-Cas9	GGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAAC TAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTTGACCTTAAACGGTCTGTGGGTTCAGTCCCTCC ACCCACGCGCGAAACGCAATAGCCGAAAAACAAAAACAAAAACAAAAACAAAAACAAAAACAA AAACACATTAACAGCCTGTGGGTGATCCACCCACAGGCCATTGGGCGTAGCACTCTGGTATCA CGGTACCTTTGTGCGCTGTTTATACCCCTCCCCAAGTGAAC TTAGAAGTAACACACACCGATCAA CAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGA CTGCTCACGCGGTTGAAGGAGAAAGCGTTCTGTTATCCGGCCAAC TACTTCGAAAAACCTAGTAACCCG TGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCTGATGAGTCACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCTGCCATGGGGAACCCATGGGACGCTCTAAT ACAGACATGGTGCAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTA ACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAAACGGGCAACTCTGCAGCGGAACCGA CTACTTTGGGTGTCCGTGTTTCATTTTATCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTAC CATATAGCTATTGGATTGGCCATCCGGTACTAATAGAGCTATTATATATCCCTTTGTGGGTTTATACCA CTTAGCTTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGATAAGAAATAC TCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGCGGTGATCACTGATGAATATAAGGTTCCG TCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTT ATTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTC GGAAGAATCGTATTTGTTATCTACAGGAGATTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTT CATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCACTCTATTTTTGAAATAT AGTAGATGAAGTTGCTTATCATGAGAAATATCCAATCTATCATCTGCGAAAAAATTTGGTAGATTCTA CTGATAAAGCGGATTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTCTGGTCATTTTTTG ATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACAA TCAATTATTTGAAGAAAACCTATTAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCAGATTGA GTAAATCAAGACGATTAGAAAATCTATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAA TCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTTTGATTGGCAGAAGATGCTAAAT ACAGCTTTCAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTTGAGATCAATATGCTG ATTTGTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTAGATATCCTAAGAGTAAATACTGAAA TAACTAAGGCTCCCTATCAGCTTCAATGATTAACGCTACGATGAACATCATCAAGACTTGACTCTTTTA AAAGCTTAGTTGACACAACACTTCCAGAAAAGTATAAAGAAATCTTTTTGATCAATCAAAAAACGGATA TGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTATAAATTTATCAAACCAATTTTAGAAAAA TGGATGGTACTGAGGAATTATTGGTGAACATAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGA CAACGGCTCTATTTCCCATCAAATTCACTTGGGTGAGCTGCATGCTATTTGAGAAGACAAGAAGACTTT TATCCATTTTTAAAGACAATCGTGAGAAGATTGAAAAATCTTGACTTTTGAATTCCTTATTATGTTGGT CCATTGGCGCGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGCTGAAGAAACAATTACCCCATGG AATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAACGCATGACAACTTTGATAA AAATCTTCAAATGAAAAAGTACTACCAAAACATAGTTTGTCTTATGAGTATTTACGGTTTATAACGAATT

		GACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAGAAGAAAGCC ATTGTTGATTTACTCTTCAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAA ATAGAATGTTTTGATAGTGTTGAAATTTGAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCA TGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATAT TGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAACATATGCTCACCTCT TTGATGATAAGGTGATGAACAGCTTAAACGTCGCCGTTACTGTTGGGGACGTTTGTCTCGAAAATT GATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGCCA ATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAGAAGACATTCAAAAAGCACAAAGTG TCTGGACAAGGCGATAGTTTACATGAACATATTGCAAAATTAAGCTGGTAGCCCTGCTATTA AAAAAGGTAT TTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTT ATTGAAATGGCACGTGAAAATCAGACAACCTCAAAAGGGCCAGAAAATTCGCGAGAGCGTATGAAACGA ATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGATTCTTAAGAGCATCCTGTTGAAAATACTCAATTGCA AAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATATTA ATCGTTTAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTCCTTAAGACGATTCAATAGACAAT AAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAA AGATGAAAACCTATTGGAGACAACCTTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACG AAAGCTGAACGTGGAGGTTTGAGTGAACCTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTC GCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAA CTTATTCGAGAGGTTAAAGTGATTACCTTAAATCTAAATTAGTTTCTGACTCCGAAAAGATTTCCAATTC TATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAAGT CTTTGATTAAGAAATATCCAAAACCTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCTGA AAATGATTGCTAAGTCTGAGCAAGAAATAGGCAAGCAACCGCAAAATATTTCTTTACTCTAATATCATG AACTTCTTCAAACAGAAATTACACTTGCAATGGAGAGATTGCAAAACGCCCTCTAATCGAAACTAATG GGGAAACTGGAGAAATTGTCTGGGATAAAGGCGAGATTTGCCACAGTGCGCAAGTATTGTCCATGC CCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAA AAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGT CCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAGGGAAGTGAAGAAGTTAAATCC GTTAAAGAGTTACTAGGGATACAATTATGGAAGAAGTTCCTTTGAAAAAATCCGATTGACTTTTTAGA AGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAACTACCTAAATATAGTCTTTTTGAGTTAGA AAACGGTCGTAAACGGATGCTGGCTAGTGCCGAGAAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAG CAAATATGTGAATTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAAC AAAAACAATTGTTGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAATCAGTGAATTTTCT AAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCCTTAGTGCATATAACAAACATAGAGACAAACC AATACGTGAACAAGCAGAAAATATTATTCATTATTTACGTTGACGAATCTTGAGCTCCCGCTGCTTTTA AATATTTTGATACAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTATAGATGCCACTCTTATCC ATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCAGCTAGGAGGTGACGGCAGCGGCA GCAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGG CGACGTAACGGCCACAAGTTCAGCGTGTCCGGCAGGGCGAGGGCGATGCCACCTACGGCAAGCTG ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGAC CTACGGCGTGACGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTCAAGTCCGCCAT GCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCG AGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGAC GGCAACATCCTGGGGCACAAAGCTGGAGTACAACACAACGCCACAACGTCTATATCATGGCCGACAAG CAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGC CGACCACTACCAGCAGAACACCCCATCGGGCAGCGCCCCGTGCTGCTGCCGACAACCACTACCTGA GCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTG ACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAACAACAAAAACAAACCGGCTA TTATGCGTTACCGGCGAGACGCTACGGACTTAATAATTGAGCCTTAAGAAGAAATCTTTAAGTGGAT GCTCTCAAACCTCAGGAAACCTAAATCTAGTTATAGACAAGCAATCCTGAGCCAAGCCGAAGTAGTAAT TAGTAAGACCAGTGGACAATCGACGGATAACAGCATATCTAGCGGCCGCGATATC
67	PIE-Factor 8	GGAGACCCTCGACCGTCGATTGTCCACTGGTCAACAATAGATGACTTACAACATAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTTGACCTTAAACGGTCGTGTGGTTCAAGTCCCTCC ACCCACACGCCGGAACGCAATAGCCGAAAAACAAAAAACAACAAAAAACAACAAAAAACA AAACACATTAACACAGCCTGTGGTTGATCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCA CGGTACCTTTGTGCGCTGTTTTATACCCCTCCCCAAGTGAACCTAGAAGTAACACACACCGATCAA CAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGA CTGCTCACGCGGTTGAAGGAGAAAGCGTTGTTATCCGGCCAACCTACTTCGAAAACCTAGTAACACCG

	TGAAGTTGCAGAGTGTTCGCTCAGCACTACCCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGGAAACCCATGGGACGCTCTAAT ACAGACATGGTGCAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCTGAATGCGGCTAATCCTA ACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAAACGGGCAACTCTGCAGCGGAACCGA CTACTTTGGGTGTCCGTGTTTCATTTTATCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTAC CATATAGCTATTGGATTGGCCATCCGGTACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCA CTTAGCTTGAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGCAATAGAGCTC TCCACCTGCTTCTTTCTGTGCCTTTTGCATTCTGCTTATAGTCCACCAGAAGATACTACCTGGGTGCAG TGGAAGTGTCTGGGACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTTCCTCCTA GAGTGCCAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAATTCACGGAT CACCTTTTCAACATCGCTAAGCCAAGGCCACCCTGGATGGGTCTGCTAGGTCTACCATCCAGGCTGAG GTTTATGATACAGTGGTCATTACACTTAAGAACATGGCTTCCATCCTGTCACTTTCATGCTGTTGGTGT ATCCTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGAAGATGA TAAAGTCTTCCCTGGTGGAAGCCATACATATGTCTGGCAGGTCTGAAAGAGAATGGTCCAATGGCCTC TGACCCACTGTGCCTTACCTACTCATATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTC ATTGGAGCCCTACTAGTATGTAGAGAAGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAAATTT ATACTACTTTTTGCTGTATTTGATGAAGGGAAAAGTTGGCACTCAGAAACAAAGAATCCTTGATGCAGG ATAGGGATGCTGCATCTGCTCGGGCTGGCCTAAAATGCACACAGTCAATGGTTATGTAACAGGTCTC TGCCAGGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGACCACTCCTG AAGTGCACCTCAATATTCCTCGAAGGTCACACATTTCTTGTGAGGAACCATCGCCAGGCGTCTTGGAAAT CTCGCCAATAACTTTCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTTGTCATA TCTCTTCCCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCAAC TACGAATGAAAAATAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGT CAGGTTTGATGATGACAACCTCTCCTTCTTTATCAAATTCGCTCAGTTGCCAAGAAGCATCCTAAAACCT GGGTACATTACATTGCTGCTGAAGAGGAGGACTGGGACTATGCTCCCTTAGTCTCGCCCCGATGACA GAAGTTATAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAGGAAGTACAAAAAGTCCGATT TATGGCATAACAGATGAAACCTTTAAGACTCGTGAAGCTATTCAGCATGAATCAGGAATCTTGGGACCT TTACTTTATGGGAAGTTGGAGACACACTGTTGATTATATTTAAGAATCAAGCAAGCAGACCATATAACAT CTACCCTCACGGAATCACTGATGTCCGTCTTTGATTCAAGGAGATTACCAAAAGGTGTAAACATTTG AAGGATTTTCCAATTCTGCCAGGAGAAATATTCAAATATAAATGGACAGTACTGTAGAAGATGGGCCAA CTAAATCAGATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCTTC AGGACTCATTGGCCCTCTCCTCATCTGTACAAAGAATCTGTAGATCAAAGAGGAAACCAGATAATGTCA GACAAGAGGAATGTCATCCTGTTTTCTGATTTTGTAGAGAACCGAAGCTGGTACCTCACAGAGAATATAC AACGCTTTCTCCCAATCCAGCTGGAGTGCAGCTTGAGGATCCAGAGTTCCAAGCCTCCAACATCATGC ACAGCATCAATGGCTATGTTTTGTAGTTTGCAGTTGTCAGTTTGTTCATGAGGTGGCATACTGGTA CATTCTAAGCATTGGAGCACAGACTGACTTCTTTCTGTCTTCTCTGGATATACCTTCAAACACAAAA TGGTCTATGAAGACACACTACCCCTATTCCCATTCTCAGGAGAACTGTCTTCATGTCGATGGAAAACCC AGGTCTATGGATTCTGGGTGCCACAACCTCAGACTTTCGGAACAGAGGCATGACCGCTTACTGAAGGT TTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGATATTTACGATACTTGCTGA GTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCAGAATTCAAGACACCCTAGCACTAGGCAAAAGCA ATTTAATGCCACCACAATTCCAGAAAATGACATAGAGAAGACTGACCTTGGTTTGCACACAGAACACCT ATGCCTAAAATACAAAATGTCTCCTCTAGTGATTTGTTGATGCTCTTGCACAGAGTCTTACTCCACATG GGCTATCCTTATCTGATCTCCAAGAAGCCAAATATGAGACTTTTTCTGATGATCCATCACCTGGAGCAATA GACAGTAATAACAGCCTGTCTGAAATGACACACTCAGGCCACAGCTCCATCACAGTGGGACATGGTA TTTACCCTGAGTCAGGCCTCCAATTAAGATTAAATGAGAACTGGGGACAACCTGCAGCAACAGAGTTGA AGAACTTGATTTCAAAGTTTCTAGTACATCAAATAATCTGATTTCAACAATTCATCAGACAATTTGGCAG CAGGTACTGATAATAACAAGTTCTTAGGACCCCCAAGTATGCCAGTTCAATTATGATAGTCAATTAGATACC ACTCTATTTGGCAAAAAGTCATCTCCCTTACTGAGTCTGGTGGACCTCTGAGCTTGAGTGAAGAAAATA ATGATTCAAAGTTGTTAGAATCAGGTTAATGAATAGCCAAGAAAGTTCATGGGGAAAAATGTATCGTCA ACAGAGAGTGGTAGGTTATTTAAAGGGAAAAGAGCTCATGGACCTGCTTTGTTGACTAAAGATAATGCCT TATTCAAAGTTAGCATCTTTTGTAAAGACAAAACAACTTCCAATAATTCAGCAACTAATAGAAAGACTC ACATTGATGGCCCATCATTATTAATTGAGAATAGTCCATCAGTCTGGCAAAATATATTAGAAAGTGACACT GAGTTTAAAAAAGTGACACCTTTGATTCATGACAGAATGCTTATGGACAAAATGCTACAGCTTTGAGGC TAAATCATATGTCAAATAAACTACTTCATCAAAAACATGGAAATGGTCCAACAGAAAAAGAGGGCCC CATTCCACCAGATGCACAAAATCCAGATATGTCGTTCTTAAGATGCTATTCTTGCCAGAATCAGCAAGG TGGATACAAAGGACTCATGGAAGAAGTCTCTGAACTCTGGGCAAGGCCCCAGTCCAAAGCAATTAGTA TCCTTAGGACCAGAAAAATCTGTGGAAGGTGAGAATTTCTGTCTGAGAAAAACAAAGTGGTAGTAGGAA AGGGTGAATTTACAAAGGACGTAGGACTCAAAGAGATGGTTTTTCCAAGCAGCAGAAACCTATTTCTTAC
--	---

	TAACTTGGATAATTTACATGAAAATAATACACACAATCAAGAAAAAAATTCAGGAAGAAATAGAAAAGA AGGAAACATTAATCCAAGAGAATGTAGTTTTGCCTCAGATACATACAGTGACTGGCACTAAGAATTCAT GAAGAACCTTTTCTTACTGAGCACTAGGCAAAATGTAGAAGGTTTCATATGACGGGGCATATGCTCCAGTA CTTCAAGATTTTAGGTCATTAATGATTCAACAAATAGAACAAAGAAACACACAGCTCATTCTCAAAAAA AGGGGAGGAAGAAAACCTTGAAGGCTTGGGAAATCAAACCAAGCAAATTGTAGAGAAATATGCATGCAC CACAAGGATATCTCCTAATACAAGCCAGCAGAATTTTGTACGCAACGTAGTAAGAGAGCTTTGAAACAA TTCAGACTCCCACTAGAAGAAACAGAACTTGAAAAAGGATAATTGTGGATGACACCTCAACCCAGTGGT CCAAAAACATGAAACATTTGACCCCGAGCACCTCAGACAGATAGACTACAATGAGAAGGAGAAAGGGG CCATTACTCAGTCTCCCTTATCAGATTGCCTTACGAGGAGTCATAGCATCCCTCAAGCAAATAGATCTCC ATTACCCATTGCAAAGGTATCATCATTTCCATCTATTAGACCTATATATCTGACCAGGGTCCATTCCAAG ACAACTCTTCTCATCTTCCAGCAGCATCTTATAGAAAGAAAGATTCTGGGGTCCAAGAAAGCAGTCATTT CTTACAAGGAGCCAAAAAAATAACCTTTCTTTAGCCATTCTAACCTTGGAGATGACTGGTGATCAAAGA GAGGTTGGCTCCCTGGGGACAAGTGCCACAAATTCAGTCACATACAAGAAAGTTGAGAACACTGTTCTC CCGAAACCAGACTTGCCCAAAACATCTGGCAAAGTTGAATTGCTTCCAAAAGTTCACATTTATCAGAAGG ACCTATTCCCTACGGAACTAGCAATGGGTCTCCTGGCCATCTGGATCTCGTGGAAGGGAGCCTTCTTC AGGGAACAGAGGGAGCGATTAAGTGAATGAAGCAAACAGACCTGGAAGTTCCCTTTCTGAGAGTAG CAACAGAAAGCTCTGCAAAGACTCCCTCCAAGCTATTGGATCCTCTTGCTTGGGATAACCACTATGGTAC TCAGATACCAAAAGAAGAGTGGAAATCCCAAGAGAAGTCACCAGAAAAACAGCTTTTAAGAAAAAGGAT ACCATTTTGTCCCTGAACGCTTGTGAAAGCAATCATGCAATAGCAGCAATAAATGAGGGACAAAATAAGC CCGAAATAGAAGTCACCTGGGCAAGCAAGGTAGGACTGAAAGGCTGTGCTCTCAAAACCCACCAGTCT TGAAACGCCATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTATGATGA TACCATATCAGTTGAAATGAAGAAGGAAGATTTTACATTTATGATGAGGATGAAAATCAGAGCCCCCGC AGCTTTCAAAAGAAAACACGACACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTA GCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCTCAGTTCAAGAAAGTTGTTTCCA GGAATTTACTGATGGCTCCTTTACTCAGCCCTTATACCGTGGAGAATAAATGAACATTTGGGACTCCTG GGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACCTTCAGAAATCAGGCCTCTCGTCCCT ATTCCTTCTATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAATTT GTCAAGCCTAATGAAACCAAACTTACTTTTTGAAAGTGCAACATCATATGGCACCCACTAAAGATGAGT TTGACTGCAAAGCCTGGGCTTATTTCTGATGTTGACCTGGAAGAGATGTGCACTCAGGCCTGATTGG ACCCCTTCTGGTCTGCCACACTAACACACTGAACCTGCTCATGGGAGACAAAGTACAGTACAGGAATT TGCTCTGTTTTTACCATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAATATGAAAGAAACTGCA GGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGTTCATGCAATCAATGG CTACATAATGGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGC ATGGGCAGCAATGAAAACATCCATTCTATTCTTCACTGGACATGTGTTCACTGTACGAAAAAAGAGG AGTATAAAATGGCACTGTACAATCTCTATCCAGGTGTTTTTGTGACAGTGGAAATGTTACCATCCAAAGCT GGAATTTGGCGGGTGAATGCCTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTTCTGGTG TACAGCAATAAGTGCAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAAGATTACAGCTT CAGGACAATATGGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATCCGGATCAATCAATGCCTGGA GCACCAAGGAGCCCTTTTCTTGATCAAGGTGGATCTGTTGGCACCATGATTATTCACGGCATCAAGA CCCAGGGTGGCGTCAGAAAGTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGG GAAGAAGTGGCAGACTTATCGAGGAAATTCAGTGAACCTTAATGGTCTCTTTGGCAATGTGGATTCA TCTGGGATAAAACACAATATTTTAAACCTCCAATTATTGCTCGATACATCCGTTTGCACCAACTCATT TAGCATTGCGAGCACTCTTCGATGGAGTTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGGA ATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCTACTTTACCAATATGTTGCCACCTG GTCTCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATGCCTGGAGACCTCAGGTGAATAATCC AAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACAGGAGTAACTACTCAGGGAGTAAA ATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACT CTCTTTTTTCAAGATGGCAAAGTAAAGGTTTTTCAGGGAAATCAAGACTCCTTCACACCTGTGGTGAAC CTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTACCCCCAGAGTTGGGTGCACAGATTGCC TGAGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTCTACGGCAGCGGCAGCAGTATGGTGAGCAA GGGCGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTGCAGCTGGACGGCGACGTAACGGGCCAC AAGTTACGCGTGTCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTG CACCACCGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCTGACCTACGGCGTGACGTGCT TCAGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTCAAGTCCGCCATGCCGAAGGCTACGTC CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGG CGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC ACAAGCTGGAGTACAACACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCA AGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGCCGACCACTACCAGCAG
--	---

		AACACCCCATCGGCGACGGCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCT GAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCA CTCTCGGCATGGACGAGCTGTACAAGTAAAAAACAAAAACAAACGGCTATTATGCGTTACCGGCG AGACGCTACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAGTGGATGCTCTCAAACCTAGGG AAACCTAAATCTAGTTATAGACAAGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGACCAAGTGA CAATCGACGGATAACAGCATATCTAGCGGCCGCGATATC
68	Lin. EGFP	GAGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGTGAGCAAGGGC GAGGAGCTGTTACCGGGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTT CAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCA CCGGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAAGTCTCAGC CGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGA GCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG CTGGAGTACAACATAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTG AACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACAC CCCCATCGGCGACGGCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCA AAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTC GGCATGGACGAGCTGTACAAGCATGGCTTCCCGCCGAGGTGGAGGAGCAGGATGATGGCAGCTGC CCATGTCTTGCCCGAGGAGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTAGGATCAAT GTGGGCAGCGGCGAGGGCAGGGGAAGTCTTCAACATCGGGGACGTGGAGGAAAATCCCGGCCCAA GCGGCAGCTAGTGACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCTTTCCCGTCTGGGT ACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCCCACTCACCACCTC TGCTAGTTCAGACACCTCCCAAGCAGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCA CGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTAACTAAGCTATACTAACCCAGGGTTG GTCAATTCGTGCCAGCCACACCCTGGAGCTAGCGCGGCCGCGATATC
69	Lin. Nluc	GAGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGTCTTCACACTC GAAGATTTGTTGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCTTGAACAGGGAGG TGTGTCCAGTTTGTTCAGAATCTCGGGGTGTCCGTAACCTCCGATCCAAAGGATTGCTGAGCGGTGA AAATGGGTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAATGGGCCA GATCGAAAAATTTTAAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGC ACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTCGGACGGCCGTATGAAGGCATCGCC GTGTTGACGGCAAAAAGATCACTGTAAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC CTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGTGTG CGAACGCATTCTGGCGGGCAGCGGCGAGGGCAGGGGAAGTCTTCAACATCGGGGACGTGGAGGAA AATCCCGGCCAAGCGGCAGCTAGTGACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCTT TCCCGTCTGGGTACCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCC CACTCACCACCTCTGCTAGTTCAGACACCTCCCAAGCAGCAGCAATGCAGCTCAAAACGCTTAGCCT AGCCACACCCCGACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTAACTAAGCTATACT AACCCAGGGTTGGTCAATTCGTGCCAGCCACACCCTGGAGCTAGCGCGGCCGCGATATC
70	siRNA	CAAATGGGCCAGATCGAAA
71	CVB3-IRES	TTAAACAGCCTGTGGGTGATCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTATCACGGTACC TTTGTGCGCTGTTTTATACCCCTCCCCAAGTGAACCTAGAAAGTAACACACACCGATCAACAGTCAG CGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTACCCCGGACTGAGTATCAATAGACTGCTCA CGCGGTGAAGGAGAAAGCGTTTCTGATCCGGCCAAGTCTCGAAAAACCTAGTAACACCGTGGAAGT TGCAGAGTGTTCGCTCAGCACTACCCAGTGATGATCAGGTGATGAGTACCGCATTCACCGGCGG GACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGGAAACCCATGGGACGCTCTAATACAGACAT GGTGCAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCTGAATGCGGCTAATCCTAACTGCGGA GCACACACCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTG GGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGC TATTGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTT GAAAGAGGTTAAACATTACAATTCATTGTTAAGTTGAATACAGCAA
72	CroV-IRES	GTATAAGAGACAGGTGTTTGCCTTGCTTCGGACTGGCATCTTGGGACCAACCCCTTTTCCCCAGCC ATGGGTAAATGGCAATAAAGGACGTAAACACTTTGTAACCATTAAGCTTTGTAATTTGTAACCACTAAG CTTTGTGCACATAATGTAACCATCAAGCTTGTAGTCCCAGCAGGAGGTTTGCATGCTTGTAGCCGAAAT GGGGCTCGACCCCCCATAGTAGGATACTTGATTTTGCATTCCATTGTGGACCTGCAAACTCTACACATAG AGGCTTTGCTTGCATCTAAACACCTGAGTACAGTGTGTACCTAGACCCTATAGTACGGGAGGACCGTTT GTTTCTCAATAACCCATACATAATAGGCTAGGTGGGCATGCCCAATTTGCAAGATCCAGACTGGGGGT

		CGGTCTGGGCAGGGTTAGATCCCTGTTAGCTACTGCCTGATAGGGTGGTGCTCAACCATGTGTAGTTA AATTGAGCTGTTTATATACC
73	CSFV- IRES	GTATACGAGGTTAGTTCATTCTCGTATACACGATTGGACAAATCAAATTATAATTTGGTTCAGGGCCTCC CTCCAGCGACGGCCGAACCTGGGCTAGCCATGCCATAGTAGGACTAGCAAACGGAGGGACTAGCCGTA GTGGCGAGCTCCCTGGGTGGTCTAAGTCTGAGTACAGGACAGTCGTCACTAGTTCGACGTGAGCAGA AGCCACCTCGAGATGCTACGTGGACGAGGGCATGCCAAGACACACCTTAACCCTAGCGGGGGTCCG TAGGGTGAATCACACCACGTGATGGGAGTACGACCTGATAGGGCGCTGCAGAGGCCCACTATTAGGC TAGTATAAAAATCTCTGCTGTACATGGCAC
74	CircCSFV- Nluc	TTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGAC TATTTTCGACGGCCGTATGAAGGCATCGCCGTGTTTCGACGGCAAAAAGATCACTGTAACAGG GACCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTG CTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGG GCAGCGGCGAGGGCAGGGGAAGTCTTCTAACATGCGGGGACGTGGAGGAAAATCCCGGCC CAAGCGGCAGCTAGTGAGTATACGAGGTTAGTTCATTCTCGTATACACGATTGGACAAATCA AAATTATAATTTGGTTCAGGGCCTCCCTCCAGCGACGGCCGAACCTGGGCTAGCCATGCCCCA TAGTAGGACTAGCAAACGGAGGGACTAGCCGTAGTGGCGAGCTCCCTGGGTGGTCTAAGT CCTGAGTACAGGACAGTCGTCACTAGTTCGACGTGAGCAGAAGCCACCTCGAGATGCTA CGTGGACGAGGGCATGCCAAGACACACCTTAACCCTAGCGGGGGTCTGCTAGGGTGAAT CACACCACGTGATGGGAGTACGACCTGATAGGGCGCTGCAGAGGCCCACTATTAGGCTAG TATAAAAATCTCTGCTGTACATGGCACAGTGGTGATCATCCACAGGGTACACCACCTTAAAA ATTTTTTCGATCTGGCCCATTTGGTCGCCGCTCAGACCTTCATACGGGATGATGACATGGATG TCGATCTTCAGCCCATTTTACCCTGCTCAGGACAATCCTTTGGATCGGAGTTACGGACACCCC GAGATTCTGAAACAACTGGACACACCTCCCTGTTCAAGGACTTGGTCCAGGTTGTAGCCGG CTGTCTGTGCCAGTCCCCAACGAAATCTTCGAGTGTGAAGACCAT (CSFV IRES in bold)
75	CircCSFV- Nluc_RCT	TTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGAC TATTTTCGACGGCCGTATGAAGGCATCGCCGTGTTTCGACGGCAAAAAGATCACTGTAACAGG GACCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTG CTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGG GCAGCGGCGAGGGCAGGGGAAGTCTTCTAACATGCGGGGACGTGGAGGAAAATCCCGGCC CAAGCGGCAGCTAGTGAGTATACGAGGTTAGTTCATTCTCGTATACACGATTGGACAAATCA AAATTACAAATTTGGTTCAGGGCCTCCCTCCAGCGACGGCCGAACCTGGGCTAGCCATGCCCCA TAGTAGGACTAGCAAACGGAGGGACTAGCCGTAGTGGCGAGCTCCCTGGGTGGTCTAAGT CCTGAGTACAGGACAGTCGTCACTAGTTCGACGTGAGCAGAAGCCACCTCGAGATGCTA CGTGGACGAGGGCATGCCAAGACACACCTTAACCCTAGCGGGGGTCTGCTAGGGTGAAT CACACCACGTGATGGGAGTACGACCTGATAGGGCGCTGCAGAGGCCCACTATTAGGCTAG TATAAAAATCTCTGCTGTACATGGCACAGTGGTGATCATCCACAGGGTACACCACCTTAAAA ATTTTTTCGATCTGGCCCATTTGGTCGCCGCTCAGACCTTCATACGGGATGATGACATGGATG TCGATCTTCAGCCCATTTTACCCTGCTCAGGACAATCCTTTGGATCGGAGTTACGGACACCCC GAGATTCTGAAACAACTGGACACACCTCCCTGTTCAAGGACTTGGTCCAGGTTGTAGCCGG CTGTCTGTGCCAGTCCCCAACGAAATCTTCGAGTGTGAAGACCAT (CSFV IRES in bold; mutations underlined)
76	CircCSFV- Nluc-3*Flag	GGATAATACGGTCAATCGGTTGGCTTCCGAAATAAAAGGTCCTTAGAGAAGAAATTCTTTAAG TGGATGCTCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAG CCGAAGTAGTAATTAGTAAGTTAAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAAT AGGCAGTAGCGAAAGCTGCGGGAGAATGTTAAGGTGATCCTGCACTATGGCACACTGGTAAT CGACGGGGTTACGCCGAACATGATCGACTATTTTCGACGGCCGTATGAAGGCATCGCCGTG TTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGA GCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGC TGGCGGCTGTGCGAACGCATTCTGGCGGGCAGCGGCGAGGGCAGGGGAAGTCTTCTAACAT GCGGGGACGTGGAGGAAAATCCCGGCCAACGCGCAGCTAGTGAGTATACGAGGTTAGTTC ATTCTCGTATACACGATTGGACAAATCAAATTATAATTTGGTTCAGGGCCTCCCTCCAGCG ACGGCCGAACCTGGGCTAGCCATGCCATAGTAGGACTAGCAAACGGAGGGACTAGCCGTA GTGGCGAGCTCCCTGGGTGGTCTAAGTCTGAGTACAGGACAGTCGTCACTAGTTCGACGT GAGCAGAAGCCACCTCGAGATGCTACGTGGACGAGGGCATGCCAAGACACACCTTAAC CCTAGCGGGGGTCTGCTAGGGTGAATCACACCACGTGATGGGAGTACGACCTGATAGGGC GCTGCAGAGGCCCACTATTAGGCTAGTATAAAAATCTCTGCTGTACATGGCACATGGATTAC

		AAAGATCACGATGGTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGGGC GGCTCAGTCTTCACACTCGAAGATTCGTTGGGACTGGCGACAGACAGCCGGCTACAACCT GGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGGGGTGCCGTAA CTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATC ATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTAAGGTGGT GTACCCTGTGGATGATCACCACCTTACACCCGGAAGCCAATGATTGACCAAAAAAAAAAAGC CGGCCGAT (CSFV IRES in bold)
77	CircCSFV- Nluc- 3*Flag_RC T	GGATAATACGGTCAATCGGTTGGCTCCGAAATAAAAGGTCCTTAGAGAAGAAATTCTTTAAG TGGATGCTCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAAG CCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAAT AGGCAGTAGCGAAAGCTGCGGGAGAATGTTAAGGTGATCTGCACTATGGCACACTGGTAAT CGACGGGGTTACGCCGAACATGATCGACTATTCGGACGGCCGTATGAAGGCATCGCCGTG TTCGACGGCAAAAAGATCACTGTAAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGA GCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGC TGGCGGCTGTGCGAACGCATTCTGGCGGGCAGCGGCGAGGGCAGGGGAAGTCTTCTAACAT GCGGGGACGTGGAGGAAAATCCCGGCCAAGCGGCAGCGTATACGAGGTTAGTTTCATTCTC GTATACACGATTGGACAAATCAAAATTACAAATTTGGTTACGGGCTCCCTCCAGCGACGGC CGAACTGGGCTAGCCATGCCATAGTAGGACTAGCAAACGGAGGGACTAGCCGTAGTGGC GAGCTCCCTGGGTGGTCTAAGTCTGAGTACAGGACAGTCGTCAGTAGTTCGACGTGAGCA GAAGCCACCTCGAGATGCTACGTGGACGAGGGCATGCCAAGACACACCTTAACCTAG CGGGGTGCTAGGGTGAAATCACACCACGTGATGGGAGTACGACCTGATAGGGCGCTGC AGAGGCCACTATTAGGCTAGTATAAAAATCTCTGCTGTACATGGCACATGGATTACAAAGA TCACGATGGTGATTATAAAGATCACGACATCGATTACAAGGATGATGATGATAAGGGCGGCTC AATGGTCTTCACACTCGAAGATTCGTTGGGACTGGCGACAGACAGCCGGCTACAACCTGG ACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGGGGTGCCGTAACCT CCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCAT CCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTAAGGTGGTGT ACCCTGTGGATGATCACCACCTTACACCCGGAAGCCAATGATTGACCAAAAAAAAAAAGCG GCCGCGAT (CSFV IRES in bold; mutations underlined)
78	CircOR4F1 7_RCT	TTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGAC TATTCGGACGGCCGTATGAAGGCATCGCCGTGTTGACGGCAAAAAGATCACTGTAACAGG GACCTGTGGAACGGCAACAAAATTATCGACGAGCGCTGATCAACCCCGACGGCTCCCTG CTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGG GCAGCGGCGAGGGCAGGGGAAGTCTTCAACATCGGGGACGTGGAGGAAAATCCCGGCC CAGGT ACTCCC ATCATGGTCTTCACACTCGAAGATTCGTTGGGACTGGCGACAGACAGCC GGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGG GGTGTCCGTAACCTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACA TCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATT TTTAAGGTGGTGTACCCTGTGGATGATCACCACCT (IRES in bold underlined)
79	L7-1bp	GGGATAATACGGTCAATCAGTTGGCTCCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTCAGGGAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATG AAC GTCCGTAGCGTCTCGCAAAAAACAAAAA CAAAAAAACAAAAAAACCAAAAAACAAAACACATTAACACAGCCTGTGGGTTGATCCC ACCCACAGGCCCATTTGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACAC CAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGG GTTGAAGGAGAAAGCGTTCTGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGATCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAA

		CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCACTGCT TCAGCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGACTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
80	L7-2bp	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAGTCCGTAGCGTCTCGCAAAAAACAAAAA CAAAAAACAAAAAACCACAAAAACAAACACATTAACACAGCCTGTGGGTTGATCCC ACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACAC CAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGATCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCTGTA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCACTGCT TCAGCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGACTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined

81	L7+1bp	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAACCGTAGCGTCTCGCGAAAAACAAAAAA CAAAAAAACAAAAAAACCAAAAAACAAAACACATTAACACAGCCTGTGGGTTGATCCC ACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACAC CAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTTCGCTCAGCACTACCCAGTGATAGATCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCTGTA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTG CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCAAGTGCT TCAGCCGCTACCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAAACAAAAACAAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGACTTAACACCCGGAAGCCAAGTATTGACCAAAAAAAAAAAGCGGCCGCGAT Bold underlined - eACA_loop_Left (ending with U (T)) and eACA_loop_Right
82	L7+2bp	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAACACGTAGCGTCTCGCGAAAAACAAAAAA CAAAAAAACAAAAAAACCAAAAAACAAAACACATTAACACAGCCTGTGGGTTGATCCC ACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACAC CAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTTCGCTCAGCACTACCCAGTGATAGATCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCTGTA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTG CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCAAGTGCT TCAGCCGCTACCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC

		ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCAGCGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGGACTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
83	p1+1nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGTAGCGTCTCGCAAAAAACAAAAA CAAAAAAACAAAAAACCACAAAAACAAACACATTAAACAGCCTGTGGGTTGATCCC ACCCACAGGCCCATTTGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACAC CAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGTCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCGTTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACCGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCAGTGCT TCAGCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCAGCGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGACCTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
84	p1+1nt_p1 ex-1nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGTAGCGTCTCGCAAAAAACAAAAA AAAAAACAACAAAAAACCACAAAAACAAACACATTAAACAGCCTGTGGGTTGATCCCA CCACAGGCCCATTTGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACC AGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGTCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCGTTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA

		TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCAAGTGCT TCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACAAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGATCCTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
85	p1+2nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGTAGCGTCTCGCAAAAAACAAAAA CAAAAAACAAAAAACCACAAAAACAAACACATTAACACAGCCTGTGGGTTGATCCC ACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTAT ACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACAC CAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGATCAGGTGATGAGTACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCTGTA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCAAGTGCT TCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACAAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGATCCCTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
86	p1+2nt_p1 ex-2nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGTAGCGTCTCGCAAAAAACAAAAA AAAAAACAACAAAAAACCACAAAAACAAACACATTAACACAGCCTGTGGGTTGATCCCA CCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACC AGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCG

		GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGATCAGGTCGATGAGTCACCGCATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCTGTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTG CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCAACCTCGTGACCACCCTGACCTACGGCGTGCACTGCT TCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAAAACAAAAACAAAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGATTCCCTTAAACACCCGGAAGCCAACCTGATTGACCAAAAAAAAAAAGCGGCCGCG AT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
87	p1ex-1nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGTAGCGTCTCGCGAAAAACAAAAACA AAAAAACAAAAAAAAAACCAAAAAACAAAACACATTAACACAGCCTGTGGGTTGATCCAC CCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTATAC CCCCTCCCCCAACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCA GCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGT TGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAG TTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGATCAGGTCGATGAGTCACCGCATTCC CCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGGAC GCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCTGA ATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCTGTAAACG GGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTGGC TGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTA ATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACATT ACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCGG GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCC GGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG GCAAGCTGCCCGTGCCCTGGCCCAACCTCGTGACCACCCTGACCTACGGCGTGCACTGCTT CAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGC TACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGG TGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGA GGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCA TGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAAAACAAAAACAAAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGACTTAAACACCCGGAAGCCAACCTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined

88	p1ex-2nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGTAGCGTCTCGCGAAAAACAAAAACAA AAAAACAAAAAAACCAAAAAACAAAAACACATTAACAGCCTGTGGGTTGATCCCACC CACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATACC CCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCCAG CCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGGTT GAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGT TGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGGACG CTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCCCTGAA TGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTGCTAACGG GCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCT GCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAA TAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAAACATTA CAATTCATTGTAAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGG GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCG GCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGG CAAGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCACTGCTTC AGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTA CGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATG GCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTCAAGATCCGCCACAACATCGAGGACG GCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCT GCTGCCCCACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAG CGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCATGGACG AGCTGTACAAGTAAAAAAACAAAAACAAACCGCTATTATGCGTTACCGGCGAGACGCTAC GGACTTAAACACCCGGAAGCCAAGTATTGACCAAAAAAAAGAGCGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
89	p1ex+1nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATTCCTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAAACCCGTAGCGTCTCGCGAAAAACAAAAA ACAAAAAAACAAAAAAACCAAAAAACAAAAACACATTAACAGCCTGTGGGTTGATCC CACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTA TACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACA CCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGC GGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGG AAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCAT TCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGG GACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCC TGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTGCTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAAACA TTACAATTCATTGTAAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTG CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCACTGCT TCAGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATC

		ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCAGCGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACGCGTATTATGCGTTACCGGCGAGACG CTACGGAGCTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
90	p1ex+2nt	GGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTGAGCCTTAGAGAAGAAATCTTTAA GTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATAGACAAGGCAATCCTGAGCCAA GCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAGCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAA TAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAACCTCCGTAGCGTCTCGCAAAAAACAAAA AACAAAAAACAACAAAAAACCACAAAAACAAACACATTAAACAGCCTGTGGGTTGATCC CACCCACAGGCCCATTTGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTA TACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACA CCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGC GGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGG AAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTCGATGAGTCACCGCAT TCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGGAAACCCATGG GACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCTCCGGCCCC TGAATGCGGCTAATCCTAACTGCGGAGCACACACCCCTCAAGCCAGAGGGCAGTGTGTCGTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAACA TTACAATTCAATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCG GGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTG CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCT TCAGCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCAGCGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAACAACAAACGCGTATTATGCGTTACCGGCGAGACG CTACGGAGCTTAAACACCCGGAAGCCAACTGATTGACCAAAAAAAAAAAGCGGCCGCGAT eACA_loop_Left (ending with U (T)) and eACA_loop_Right Bold underlined
91	TRICv2- UA	CAGATTAATACGACTCACTATAGGGATAATACGGTCAATCAGTTGGCTTCGAAATAATTTAG CCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAACCTAGGGAAACCTAAATCTAGCTATA GACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTTAACAACAGATAACTTACAG CTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAAATCCGTAG CGTCTCGCAAAAAACAAAAAACAACAAAAAACAACAAAAAACAACACATTAAA ACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTA CCTTTGTGCGCCTGTTTTATACCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATC AACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGT ATCAATAGACTGCTCACGCGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAA AAACCTAGTAACACCGTGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCA GGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTG CCCATGGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAG TTGGTAGTCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCC AGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTG TTTCATTTTATTCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTAT TGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTA

		GCTTGAAAGAGGTTAAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAA GGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACC CTGACCTACGGCGTGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCAGCTTCTT CAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAATA CAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTTC AGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACAC CCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCC CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCG CCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAAAACAAAAACAAACGGCTA TTATGCGTTACCGGCGAGACGCTACGGACTAAAACACCCGGAAGCCAAGTATTGACCAAAA AAAAAAAAGCGGCCGCGATATCATCAGTTCTGGACC
92	TRICv2- AC	CAGATTAATACGACTCACTATAGGGATAATACGGTCAATCAGTTGGCTTCCGAAATAATTAAG CCTTAGAGAAGAAATCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGCTATA GACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAAGTAAACAACAGATAACTTACAG CTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCGGGAGAATGAAATCCGTAG CGTCTCGCGAAAAACAAAAACAAAAAAAACAAAAAAAACCAAAAAACAAACACATTAAA ACAGCCTGTGGGTTGATCCACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTA CCTTTGTGCGCCTGTTTTATACCCCTCCCCCACTGTAAGTAAAGTAACACACACCGATC AACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGT ATCAATAGACTGCTCACGCGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACCTACTTCGAA AAACCTAGTAACACCGTGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCA GGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTG CCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAG TTGGTAGTCCTCCGGCCCCGTAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCC AGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTG TTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTAT TGGATTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTA GCTTGAAAGAGGTTAAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAA GGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACC CTGACCTACGGCGTGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCAGCTTCTT CAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAATA CAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTTC AGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACAC CCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCC CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCG CCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAAAACAAAAACAAACGGCTA TTATGCGTTACCGGCGAGACGCTACGGACTAAAACACCCGGAAGCCAAGTATTGACCAAAA AAAAAAAAGCGGCCGCGATATCATCAGTTCTGGACC
93	TRICv2- Tetra	GGGATAATACGGTCAATCAGTTGGCTTCCG<u>AAATAGCAATATTTACCTTTGGAGGGAAAAAGT</u> <u>TATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTTAAAAGGCAA</u> <u>GACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCAGGGGAAACT</u> <u>TTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACATGGTCTTAACCACGCAGC</u> <u>CAAGTCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTTACAGACTAAATGTCGGTC</u> <u>GGGGAAGATGTATTCTTCTCATAAGATATAGTCGGACCTCTCCTTAATGGGAGCTAGCGGA</u> <u>TGAAGTATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGCGAAAGTATATTGATTAG</u> <u>TTTTGGAGTACTCG</u>TAAGGTACCTCCGTAGCGTCTCGCCGTAACGCAGAAAAACAAAAAC AAAAAAAACAAAAAAAACCAAAAAACAAACACATTAAACAGCCTGTGGGTTGATCCCA

		CCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATA CCCCCTCCCCCACTGTAACCTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACC AGCCACGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATCAATAGACTGCTCACGCG GTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGA AGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGTAGATCAGGTGATGAGTACCCGATT CCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCATGGGAAACCCATGGG ACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCTGTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCTATACTG GCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACA TTACAATTCAATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCG GGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTG CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCT TCAGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGA CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGTAAAAAAAACAAAAACAAAAACGGCTATTATGCGTTACCGGCGAGACG CTACGGACTCTCTAAACCAACACACCCGGAAGCCAACCTGATTGACCAAAAAAAAAAAGCGG CCGCGAT Group I intron bold underlined
94	Tetra- Rzy	TGCGCTCTGCCCACTGACGGGACCCGGAGCGATCGCAGATCCTTCGAATTC<u>TCCTTGAGGC</u> <u>TGAAAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCATCGGTTT</u> <u>AAAAGGCAAGACCGTCAAATTGCGGGAAGGGGTCAACAGCCGTTCAGTACCAAGTCTCA</u> <u>GGGGAACCTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACATGGTCTTAA</u> <u>CCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCACAGACTA</u> <u>AATGTCGGTCCGGGAAGATGTATTCTTCTCATAAGATATAGTCGGACCTCTCCTTAATGGGA</u> <u>GCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATCGGAAAGTAT</u> <u>ATTGATTAGTTTTGGAGTACTCG</u>TCAAGGACGACGGCAAATACAAGACCCGCGCCGTGGTGA AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCACCGACTTCAAGGAGGA CGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTTAACAGCCACAACGTCTATATCACCG CCGACAAGCAGAAGAACGGCATCAAGGCGAACTTACCGTGCGCCACAACGTGGAGGACGG CAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG CTGCCCGACAACCACTACCTGAGCACCCAGACCGTGCTGAGCAAAGACCCCAACGAGAAGC GCGATCACATGGTCTGTCATGAGTATGTGAACGCCGCGGGATCACTTAGAAAAACAAAAAC ACATTAACAGCCTGTGGGTTGATCCACCCACAGGCCATTGGGCGCTAGCACTCTGGTA TCACGGTACCTTTGTGCGCCTGTTTTATACCCCTCCCCCAACTGTAACCTAGAAAGTAACACA CACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTACCCCG GACTGAGTATCAATAGACTGCTCAGCGGTTGAAGGAGAAAGCGTTCTGTTATCCGGCCAACT ACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGT GTAGATCAGGTGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGG CGGCCTGCCCATGGGGAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATT GAGCTAGTTGGTAGTCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACC TCAAGCCAGAGGGCAGTGTGTGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGT GTCCGTGTTTCATTTTATTCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATA TAGCTATTGGATTGGCCATCCGGTACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATA CCACTTAGCTTGAAAGAGGTTAAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGT GAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGA CGTAAACGGCCACAAGTTCAGCGTGCAGCGGCGAGGGCGAGGGCGATGCCACCATTGGCAA GCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTG ACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCATGAAGCAGCAGC

		ACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATCAGCTTCAGGCG GCCGCGAAGGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCAGAT Group I intron bold underlined
95	Tetra- STS	GGGGTAAAT<u>AAAAGTTATCAGGCATGCACCTGGTAGCTAGTCTTTAAACCAATAGATTGCAT</u> <u>CGGTTTAAAAGGCAAGACCGTCAAATTGCGGGAAAGGGGTCAACAGCCGTTCAGTACCAA</u> <u>GTCTCAGGGGAAACTTTGAGATGGCCTTGCAAAGGGTATGGTAATAAGCTGACGGACATG</u> <u>GTCTTAACCACGCAGCCAAGTCCTAAGTCAACAGATCTTCTGTTGATATGGATGCAGTTCA</u> <u>CAGACTAAATGTCGGTCGGGGAAGATGTATTCTTCTCATAAGATATAGTCGGACCTCTCCTT</u> <u>AATGGGAGCTAGCGGATGAAGTGATGCAACACTGGAGCCGCTGGGAACTAATTTGTATGC</u> <u>GAAAGTATATTGATTAGTTTGGAGTACTCG</u>CGAGTACGCTTAAGAAAAAAAAAATTTAAACA GCCTGTGGGTTGATCCACCCACAGGCCCATTTGGGCGCTAGCACTCTGGTATCACGGTACCT TTGTGCGCCTGTTTTATACCCCTCCCCCAACTGTAACCTAGAAGTAACACACACCGATCAAC AGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGACTGAGTATC AATAGACTGCTCACGCGTTGAAGGAGAAAAGCGTTCTGTTATCCGGCCAACCTACTTCGAAAAA CCTAGTAACACCGTGGAAGTTGCAGAGTGTTCGCTCAGCACTACCCAGTGATAGATCAGGT CGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCC ATGGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTG GTAGTCCTCCGGCCCCCTGAATGCGGCTAATCTAACTGCGGAGCACACACCCTCAAGCCAGA GGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTC ATTTTATTCCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGA TTGGCCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTT GAAAGAGGTTAAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGC GAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGA AGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCACCCTGAC CTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGT CCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA CAAGACCCGCGCCGAGGTGAAGTTGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTACAACA GCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATC CGCCACAACATCGAGGACGGCAGCGTGCACTCGCCGACCACTACCAGCAGAACACCCCCA TCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCAGTCCGCCCTGAG CAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCTGTACCGCCGCGGGG ATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAAAAAAAAACCGGTATTTAT Group I intron bold underlined
96	Modified CSFV IRES	GTATACGAGGTTAGTTTATTCTCGTATACACGATTGGACAAATCAAAATTCAATTTGGTTCAG GGCCTCCCTCCAGCGACGGCCGAAGTGGGCTAGCCATGCCATAGTAGGACTAGCAAACGG AGGGACTAGCCGTAGTGCGGAGCTCCCTGGGTGGTCTAAGTCCTGAGTACAGGACAGTCGT CAGTGGTTTCGACGTGAGCAGAAGCCACCTCGAGATGCTACGTGGACGAGGGCATGCCAA GACACACCTTAACCCTAGCGGGGGTCTAGGGTGAATCACACCACGTGATGGGAGTACG ACCTGATAGGGCGCTGCAGAGGCCCACTATTAGGCTAGTATAAAATCTCTGCTGTACATGG CAC

CLAIMS

1. A recombinant nucleic acid molecule for making a circular RNA, comprising in the 5' to 3' direction:
 - 5 a) an internal guide sequence (IGS),
 - b) a ribozyme,
 - c) a first portion of an extended anticodon arm (eACA) sequence,
 - d) a gene of interest, and
 - e) a second portion of the eACA sequence,
- 10 wherein a nucleotide in the second portion of the eACA sequence forms a wobble base pair with a nucleotide in the IGS.
2. The recombinant nucleic acid molecule of claim 1, wherein the first portion of the eACA sequence comprises a first eACA stem portion and a first eACA loop portion.
- 15 3. The recombinant nucleic acid molecule of claim 1 or 2, wherein the second portion of the eACA sequence comprises a second eACA stem portion and a second eACA loop portion.
4. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second eACA stem portions are each at least 5 nucleotides in length.
- 20 5. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second eACA stem portions are each at least 10 nucleotides in length.
- 25 6. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second eACA stem portions are each at least 15 nucleotides in length.
7. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second eACA stem portions are each at least 20 nucleotides in length.
- 30 8. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second eACA stem portions are each at least 25 nucleotides in length.
9. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second eACA stem portions are between 1 and 50 nucleotides in length.
- 35 10. The recombinant nucleic acid molecule of any preceding claim, wherein the first eACA loop portion is 1-20 nucleotides in length.

11. The recombinant nucleic acid molecule of any preceding claim, wherein the second eACA loop portion is 1-20 nucleotides in length.
12. The recombinant nucleic acid molecule of any preceding claim, wherein the first eACA loop portion is 4 nucleotides in length and wherein the second eACA loop portion is 3 nucleotides in length.
13. The recombinant nucleic acid molecule of any preceding claim, further comprising a P1 extension 3' to the second portion of the eACA sequence.
14. The recombinant nucleic acid molecule of claim 13, wherein the P1 extension is 2 nucleotides in length.
15. The recombinant nucleic acid molecule of claim 13 or 14, wherein the second portion of the eACA sequence and the P1 extension together are capable of forming a P1 region.
16. The recombinant nucleic acid molecule of claim 15, wherein the P1 region is complementary to the IGS.
17. The recombinant nucleic acid molecule of any preceding claim, wherein a 5' portion of the first portion of the eACA sequence is capable of forming a P10 region.
18. The recombinant nucleic acid molecule of claim 17, wherein the 5' portion of the first portion of the eACA sequence that forms a P10 region is complementary to the IGS.
19. The recombinant nucleic acid molecule of claim 17 or 18, wherein the 5' portion of the first portion of the eACA sequence that forms a P10 region is 2 nucleotides in length.
20. The recombinant nucleic acid molecule of any preceding claim, wherein the nucleotide in the second portion of the eACA sequence that forms a wobble base pair is uracil or cytosine.
21. The recombinant nucleic acid molecule of any preceding claim, wherein the nucleotide in the second portion of the eACA sequence that forms a wobble base pair is the last nucleotide in the second portion of the eACA sequence.
22. The recombinant nucleic acid molecule of any preceding claim, wherein the second eACA loop portion comprises uracil or cytosine as a nucleotide other than the first or second nucleotide in the 5' to 3' direction.
23. The recombinant nucleic acid molecule of any preceding claim, wherein the second eACA loop portion comprises uracil or cytosine as the third nucleotide in the 5' to 3' direction.

24. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second portions of the eACA sequence together are capable of forming a stem loop structure.

5 25. The recombinant nucleic acid molecule of claim 24, wherein the loop of the stem loop structure is 3-40 nucleotides in length.

26. The recombinant nucleic acid molecule of claim 24 or 25, wherein the loop of the stem loop structure is greater than 4 nucleotides in length.

10

27. The recombinant nucleic acid molecule of any preceding claim, wherein the ribozyme is a group I intron.

15

28. The recombinant nucleic acid molecule of any preceding claim, wherein the wobble base pair comprises a GU or AC wobble base pair.

29. The recombinant nucleic acid molecule of any preceding claim, wherein the second portion of the eACA sequence is complementary to the IGS except for the wobble base pair.

20

30. The recombinant nucleic acid molecule of any preceding claim, further comprising a first extended guide sequence (EGS) 5' to the IGS, and a second EGS 3' to the second portion of the eACA sequence.

25

31. The recombinant nucleic acid molecule of claim 30, wherein the first EGS is at least 50% complementary to the second EGS.

32. The recombinant nucleic acid molecule of claim 31, wherein the first EGS is 100% complementary to the second EGS.

30

33. The recombinant nucleic acid molecule of any of claims 30-32, wherein the first and second EGS are each between 1 and 500 nucleotides in length.

34. The recombinant nucleic acid molecule of any preceding claim, wherein the first and second portions of the eACA sequence naturally occur in the gene of interest.

35

35. The recombinant nucleic acid molecule of any of claims 1-34, wherein the first and second portions of the eACA sequence are derived from human ribosomal RNA (rRNA).

36. The recombinant nucleic acid molecule of any of claims 1-33 or 35, wherein the nucleotide sequence of the gene of interest is modified to include the first and second portions of the eACA sequence.

5 37. The recombinant nucleic acid molecule of any of claims 15-36, further comprising a first loop sequence 5' to the IGS, and a second loop sequence 3' to the P1 extension.

38. The recombinant nucleic acid molecule of claim 37, wherein the first and second loop sequences are each between 1 and 10 nucleotides in length.

10

39. The recombinant nucleic acid molecule of claim 37 or 38, wherein the first loop sequence is 6 nucleotides in length.

15

40. The recombinant nucleic acid molecule of any of claims 37 to 39, wherein the second loop sequence is 5 nucleotides in length.

41. The recombinant nucleic acid molecule of any of claims 37 to 40, wherein the first loop sequence is 6 nucleotides in length and the second loop sequence is 5 nucleotides in length.

20

42. The recombinant nucleic acid molecule of any of claims 37 to 41, wherein the first loop sequence is non-complementary to the second loop sequence.

43. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest is at least 500 nucleotides in length.

25

44. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest is at least 2000 nucleotides in length.

30

45. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest is at least 4000 nucleotides in length.

46. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest is at least 6000 nucleotides in length.

35

47. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest is at least 8000 nucleotides in length.

48. The recombinant nucleic acid molecule of any preceding claim, comprising in a 5' to 3' direction:

40

a) a first EGS,

- b) a first loop sequence,
- c) an internal guide sequence (IGS),
- d) a ribozyme,
- e) a first portion of an extended anticodon arm (eACA) sequence,
- 5 f) a gene of interest,
- g) a second portion of the eACA sequence,
- h) a P1 extension,
- i) a second loop sequence, and
- 10 j) a second EGS.

49. The recombinant nucleic acid molecule of any preceding claim, further comprising a T7 high efficiency sequence.

50. The recombinant nucleic acid molecule of any preceding claim, further comprising at least one
15 restriction enzyme cleavage site.

51. The recombinant nucleic acid molecule of any preceding claim, further comprising a poly(A) tail.

20 52. The recombinant nucleic acid molecule of any preceding claim, wherein the recombinant nucleic acid molecule is a DNA template molecule.

53. The recombinant nucleic acid molecule of claim 52, further comprising a T7 promoter
25 sequence.

54. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest is a noncoding RNA.

55. The recombinant nucleic acid molecule of claims 1-53, wherein the gene of interest encodes a
30 polypeptide.

56. The recombinant nucleic acid molecule of any preceding claim, wherein the recombinant nucleic acid molecule is devoid of in-frame stop codons.

35 57. The recombinant nucleic acid molecule of any preceding claim, wherein the gene of interest comprises a translation initiation element.

58. The recombinant nucleic acid molecule of claim 57, wherein the translation initiation element is a modified viral IRES that lacks in-frame stop codons.

59. Use of a recombinant nucleic acid molecule according to any preceding claim in a method of making a circular RNA.
60. A method for producing a circular RNA, comprising:
- 5 a) providing a recombinant nucleic acid molecule according to any of claims 1-58, and
b) circularizing the recombinant nucleic acid molecule.
61. A method for producing a circular RNA, comprising:
- 10 a) providing a recombinant nucleic acid molecule according to any of claims 1-58,
b) transcribing the recombinant nucleic acid molecule to produce an RNA precursor, and
c) circularizing the RNA precursor.
62. The method of claim 61, wherein splicing is suppressed during step (b).
- 15 63. The method of claim 61 or 62, wherein step (b) is performed in the presence of nucleoside triphosphates (NTPs) at a concentration of approximately 24 mM and Mg²⁺ at a concentration of 16 mM or less.
64. A method for producing a circular RNA, comprising:
- 20 a) identifying a gene of interest comprising a sequence capable of forming an eACA stem-loop,
b) preparing a recombinant nucleic acid molecule comprising, in a 5' to 3' direction: an internal guide sequence (IGS) – a ribozyme – a sequence encoding the gene of interest, and
c) circularizing the recombinant nucleic acid molecule.
- 25 65. A method for producing a circular RNA, comprising:
- a) providing a recombinant nucleic acid molecule according to any of claims 1-58, and
b) transcribing and circularizing the recombinant nucleic acid molecule,
wherein circularization occurs co-transcriptionally.
- 30 66. A circular RNA comprising a sequence encoding a gene of interest, wherein the circular RNA does not comprise exogenous splicing sequences.
67. A circular RNA comprising a sequence encoding a gene of interest, wherein the circular RNA does not comprise RNA from a circularization agent.
- 35 68. The circular RNA of claim 66 or 67, wherein the circularization agent is a ribozyme.
69. The circular RNA of any of claims 66-68, wherein the circularization agent is a group 1 intron.

70. The circular RNA of any of claims 66-69, wherein the sequence encoding a gene of interest comprises a sequence capable of forming an eACA stem loop structure.

71. The circular RNA of any of claims 66-70, wherein the eACA stem loop structure comprises a stem of at least 5 base pairs.

72. The circular RNA of any of claims 66 to 71, wherein the eACA stem loop structure comprises a loop at least 7 nucleotides in length.

73. The circular RNA of any of claims 66 to 72, wherein the circular RNA has reduced immunogenicity compared to a circular RNA which comprises exogenous splicing sequences.

74. The circular RNA of any of claims 66 to 73, wherein the circular RNA has reduced immunogenicity compared to a circular RNA produced using the permuted intron-exon (PIE) approach.

75. The circular RNA of any of claims 66 to 74, wherein the sequence encoding a gene of interest is at least 2000, at least 3000, at least 4000, at least 5000, or at least 6000 nucleotides in length.

76. A circular RNA obtained or obtainable by the method of any of claims 60 to 65, wherein the circular RNA is less immunogenic compared to a circular RNA which comprises exogenous exon sequences.

77. A circular RNA obtained or obtainable by the method of any of claims 60 to 65, wherein the circular RNA does not comprise exogenous exon sequences.

78. Use of a circular RNA according to any of claims 66-77 in an in vitro method of expressing a protein in a cell.

79. A method for expressing a gene of interest in a cell, the method comprising (a) circularizing a recombinant nucleic acid molecule according to any of claims 1-58 to provide a circular RNA comprising the gene of interest, and (b) administering the circular RNA to the cell.

80. A method of treating a disease in a subject, the method comprising (a) circularizing a recombinant nucleic acid molecule according to any of claims 1-58 to provide a circular RNA, and (b) administering the circular RNA to the subject.

81. A recombinant nucleic acid molecule according to any of claims 1-58, for use as a medicament.

82. A recombinant nucleic acid molecule according to any of claims 1-58, for use in a method of treating a disease in a subject.

83. A circular RNA obtainable by the method of any of claims 60 to 65.

5

84. A circular RNA comprising a sequence encoding a gene of interest, wherein the circular RNA comprises an eACA sequence.

1/40

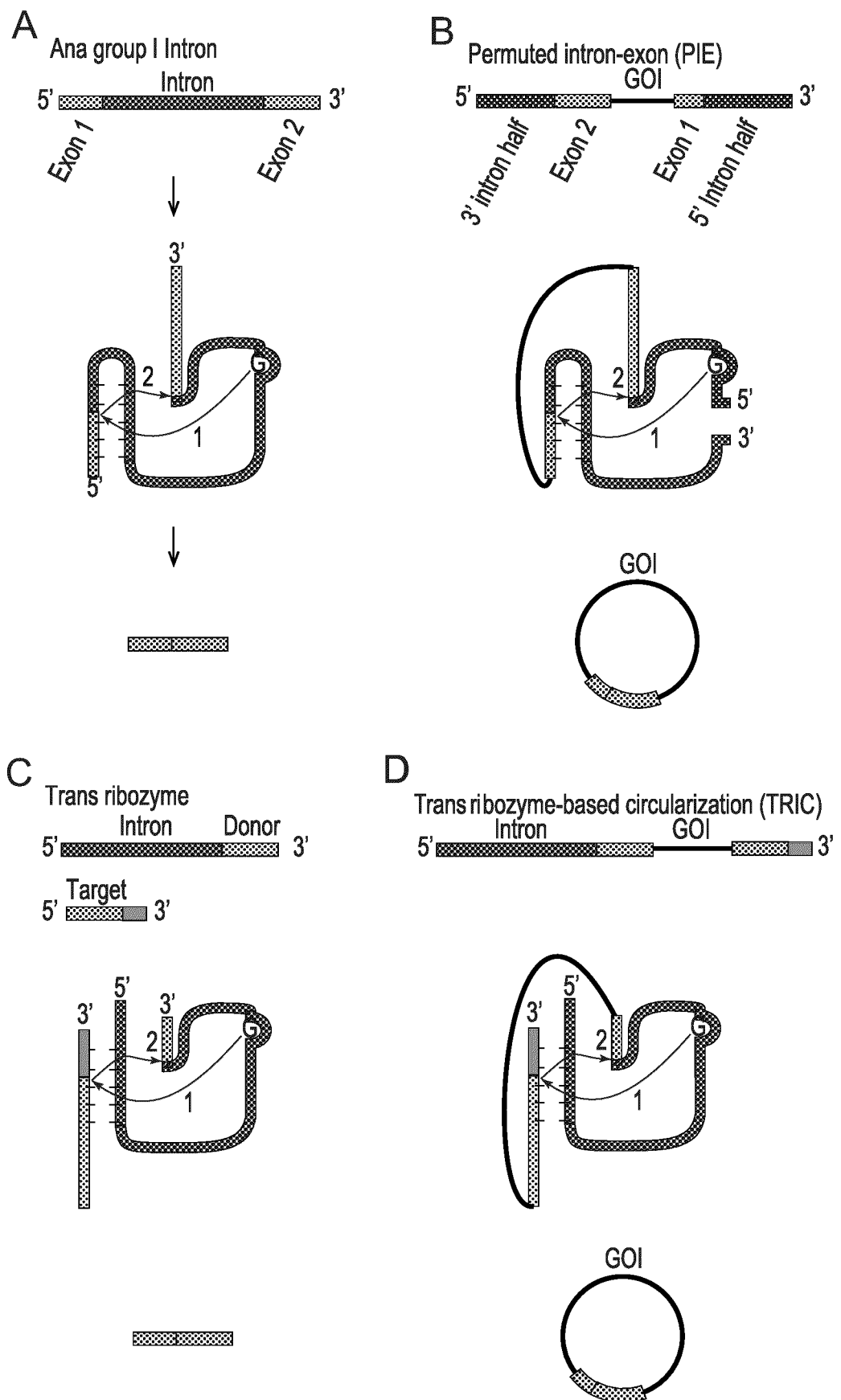


FIG. 1

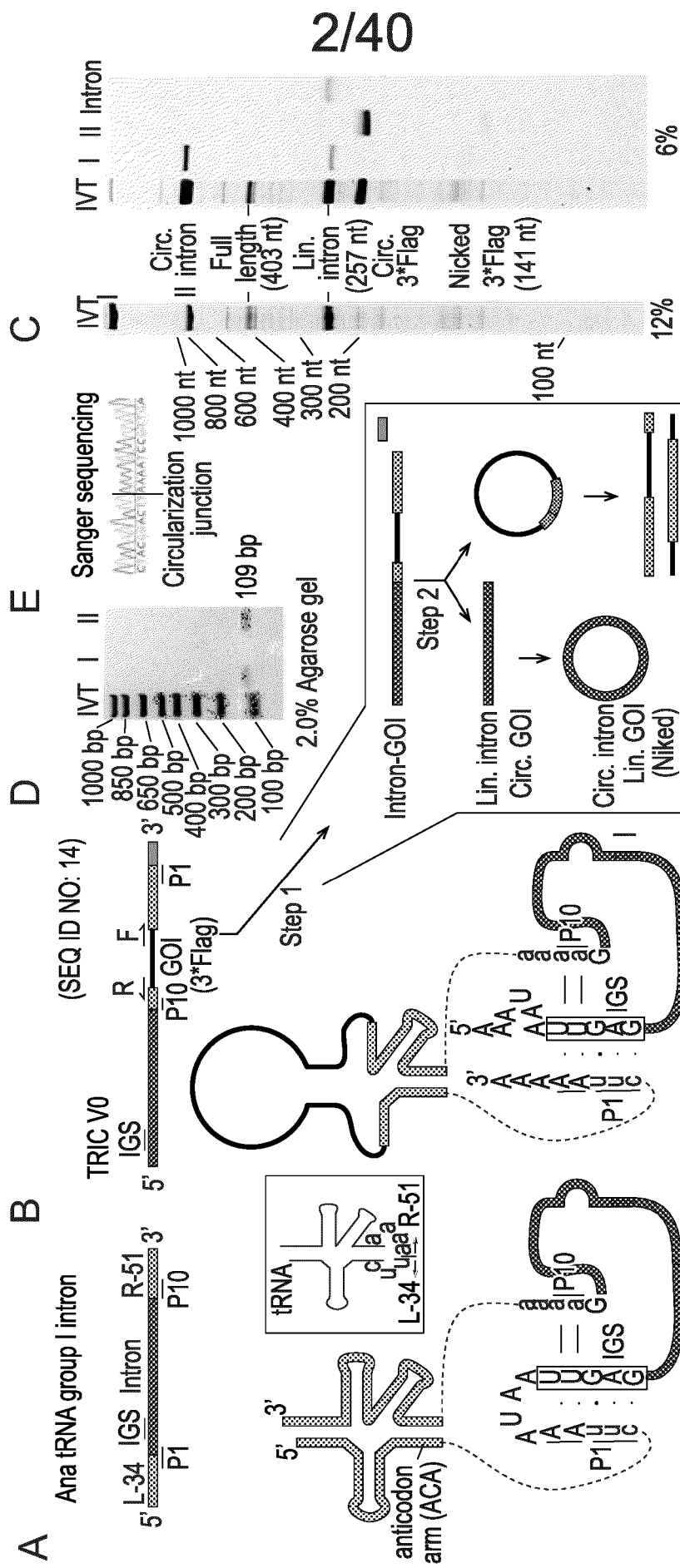
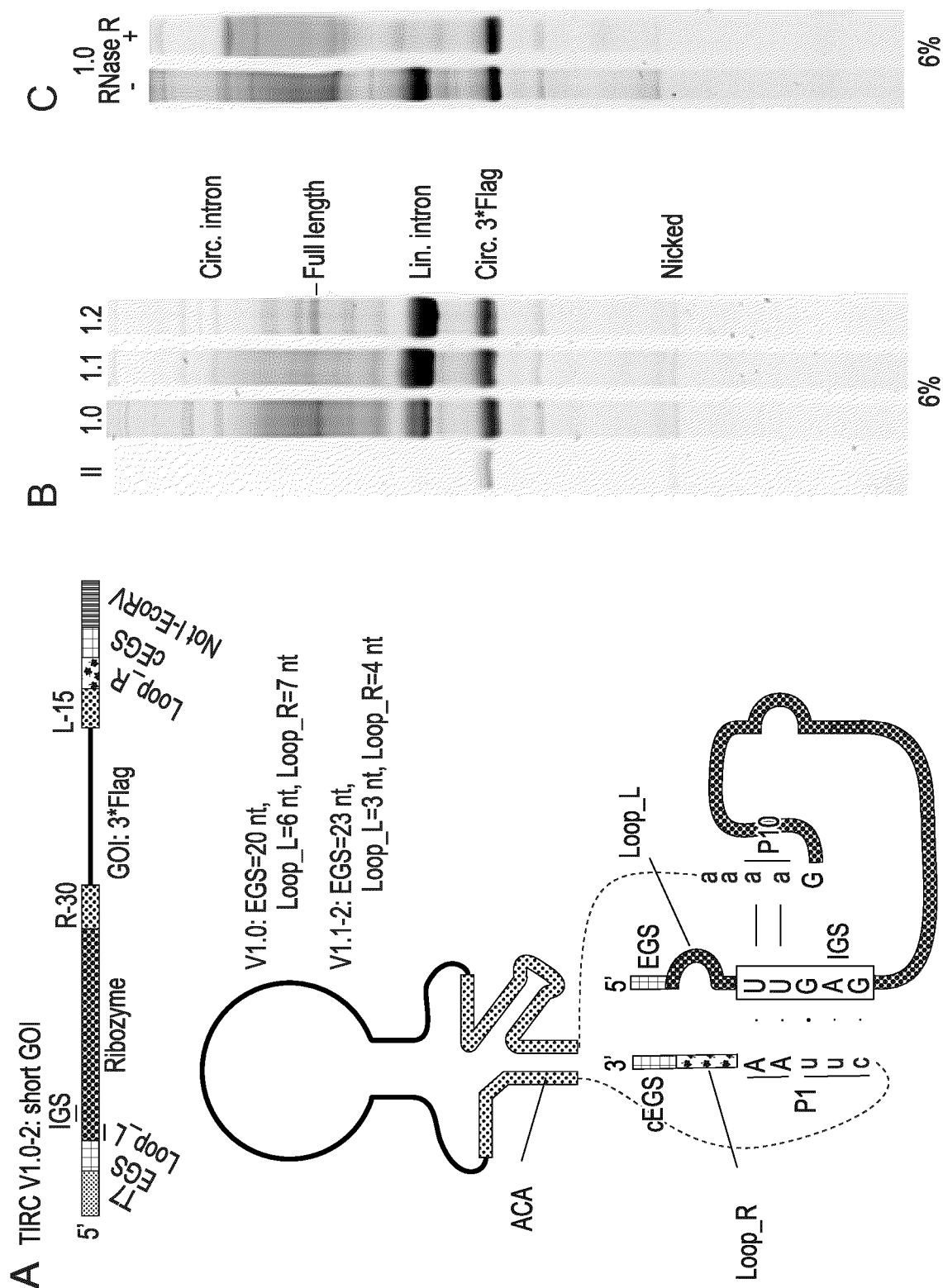
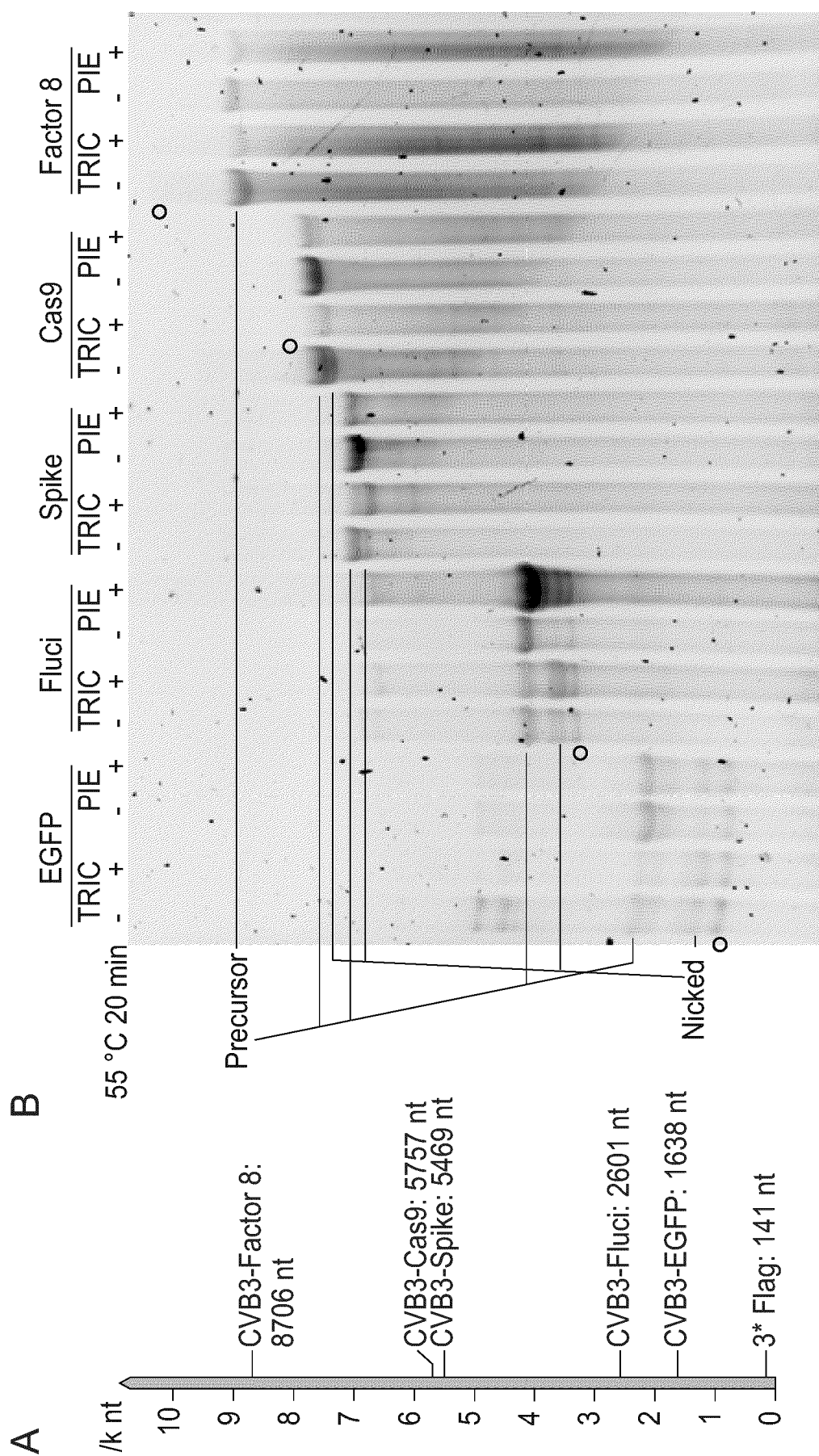


FIG. 2

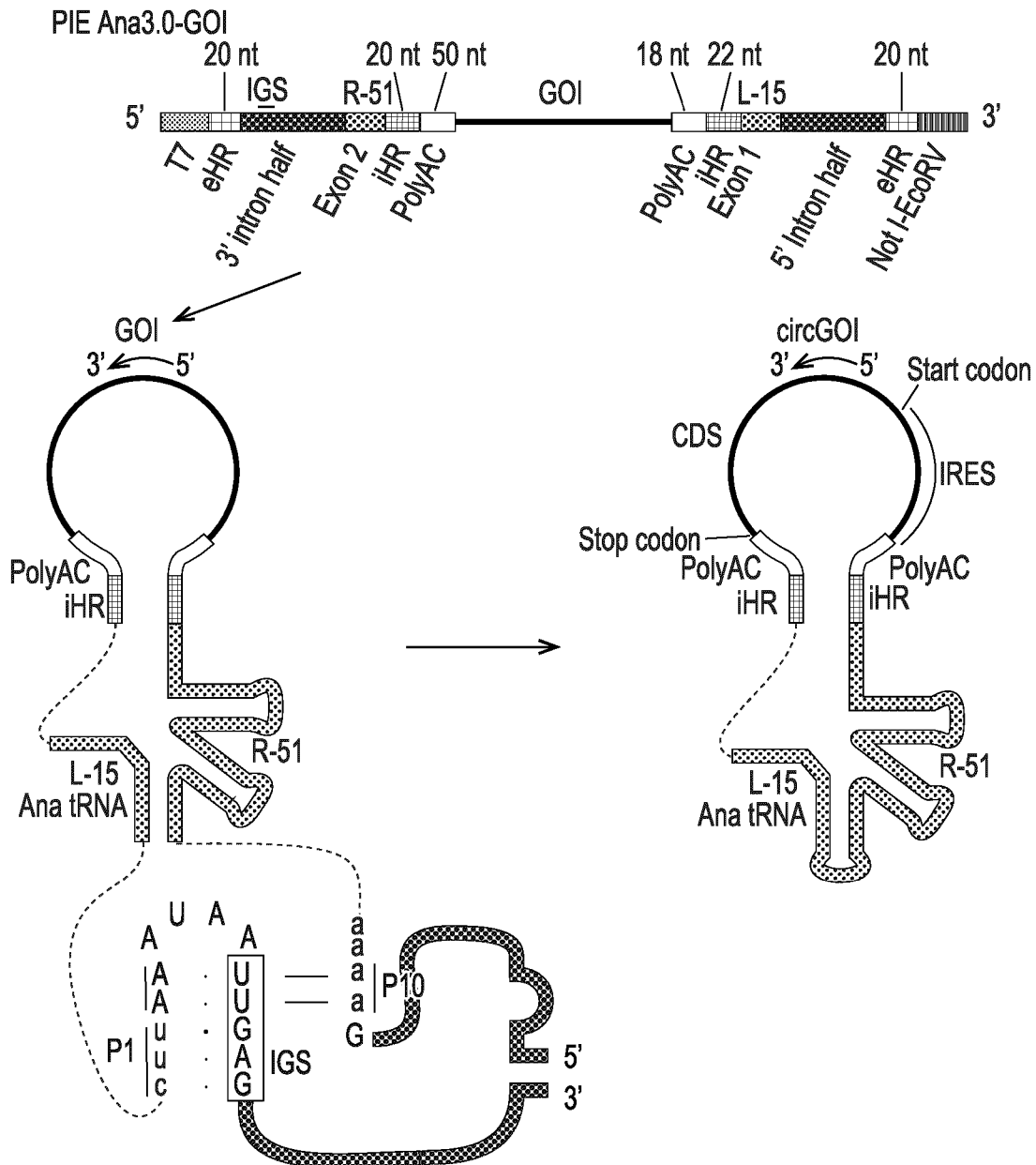


3
6
—
L

4. G

5/40

A



T7: T7 promotor
e/iHR: external/internal homology arm
IGS: internal guide sequence
R-51/L-15: Ana tRNA right arm-51 nt, left arm-15 nt
GOI: R51-iHR-PolyAC-CVB3-start codon-CDS-stop codon-PolyAC-iHR-L15 CDS=EGFP/FlucI/Spike-EGFP/Cas9-EGFP/Factor 8-EGFP
Not I-EcoR V: restriction enzyme cleavage sites

FIG. 5A

A

Not I-EcoR V

(SEQ ID NO: 52)

CDS=EGFP/Fluci/Spike-EGFP/Cas9-EGFP/Factor 8-EGFP

SUBSTITUTE SHEET (RULE 26)

B

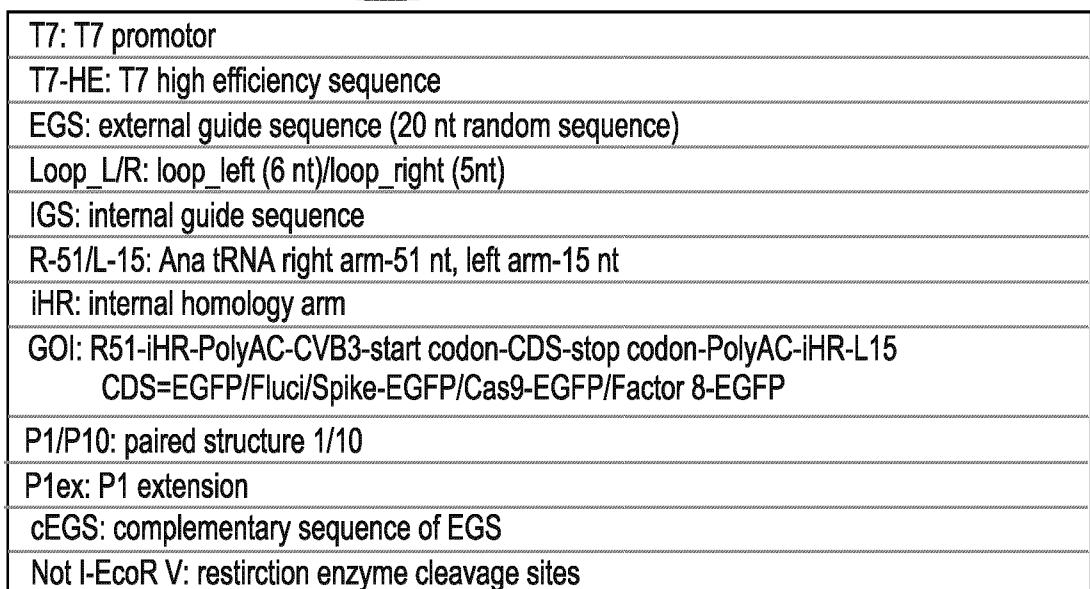


FIG. 5B

8/40

TRIC-V1: long GOIs

T7-HE [EGS] [Ribozyme] [Loop_L] [IGS] circGOI [Ana tRNA] [P1ex] [Loop_R] cEGS PolyA Not I-EcoR V

GGAUAAUAC GGUCAAUCAGUUGGCUUCCG AAAUAAUUGAGCCUUAAGAGAAGAAAUU
 CUUUAAGUGGAUGCUCUCAACUCAGGGAAACCUAAAUCUAGCUAUAGACAAGGCAAUC
 CUGAGCCAAGCCGAAGUAGUAAUUAAGUUAACAACAGAUAAACUUAACAGCUAAUCGG
 AAGGUGCAGAGACUCGACGGGAGCUACCCUAACGUAAGACGAGGGUAAAGAGAGAGU
 CCAAUUCUCAAGCCAAUAGGCAGUAGCGAAAGCUGCGGGAGAAUG AAAAUCCGUUGA
 CCUUAACGGUCGUGUGGGUUAAGUCCCUCCACCCCAACGCGCGAAACGCAAUAGCC
 GAAAAACAAAAACAAAAAAACAAAAAAACCAAAAAACAAAACACAUUAAAAACAGCC
 UGUGGGUUGAUUCCACCCACAGGCCCAUUGGGCGCUAGCACUCUGGUUACACGGUACC
 UUUGUGCGCCUGUUUUAUACCCCUCCCAACUGUAACUUAAGAAGUAACACACACCGA
 UCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACUUCUGUUACCCCGGA
 CUGAGUAUCAUAGACUGCUCACGCGGUUGAAGGAGAAAGCGUUCGUUAUCCGGCCAA
 CUACUUCGAAAAACCUAGUAACACCCGUGGAAGUUGCAGAGUGUUUCGCUCAGCACUACC
 CCAGUGUAGAUAGGUCGAUGAGUCACCGCAUUCCCCACGGGCGACCGUGGCGGUGGC
 UGCGUUGGCGGCCUGCCAUUGGGAAACCAUUGGGACGCUCUAAUACAGACAUGGUGC
 GAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAUCCUAAC
 UGCGGAGCACACACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGC
 GGAACCGACUACUUGGGUGUCCGUGUUUCAUUUUAUCCUAUACUGGCUGCUUAUGG
 UGACAAUUGAGAGAUUCGUUACCAUUAUAGCUAUUGGAUUGGCCAUCCGGUGACUAAUAGA
 GCUAUUAUUAUCCCUUUGUUGGGUUUAUACCACUUAAGCUUGAAAGAGGUUAAAACAUU
 ACAAUUCAUUGUUAAGUUGAAUACAGCAAAUUGGUGAGCAAGGGCGAGGAGCUGUUCAC
 CGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGC
 GUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUAUC
 UGCACCACCGGCAAGCUGCCCGUGCCCGGCCACCCUCGUGACCACCCUGACCUACG
 GCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUAAGUC
 CGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACGACGGCAAC
 UACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCGCAUCGAGC
 UGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGCACAAGCUGGAGUACAA
 CUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGA
 ACUUCAAGAUCCGCCACAACAUUCGAGGACGGCAGCGUGCAGCUCGCCGACCACUACCA
 GCAGAACACCCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGC
 ACCCAGUCCGCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGG
 AGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAAA
 AAAAAACAAAACGGCUAUUAUGCGUUAACGGCGAGACGCUACGGACUU AA ACACC
 GGUCAAUCAGUUGGCUUCCG AAAAAAAAAA GCGGCCGCGAU (SEQ ID NO: 53)

circGOI: R51-iHR-PolyAC-CVB3-start codon-CDS-stop codon-PolyAC-iHR-L15

CDS=EGFP/Fluci/Spike-EGFP/Cas9-EGFP/Factor 8-EGFP

FIG. 5B (continued)

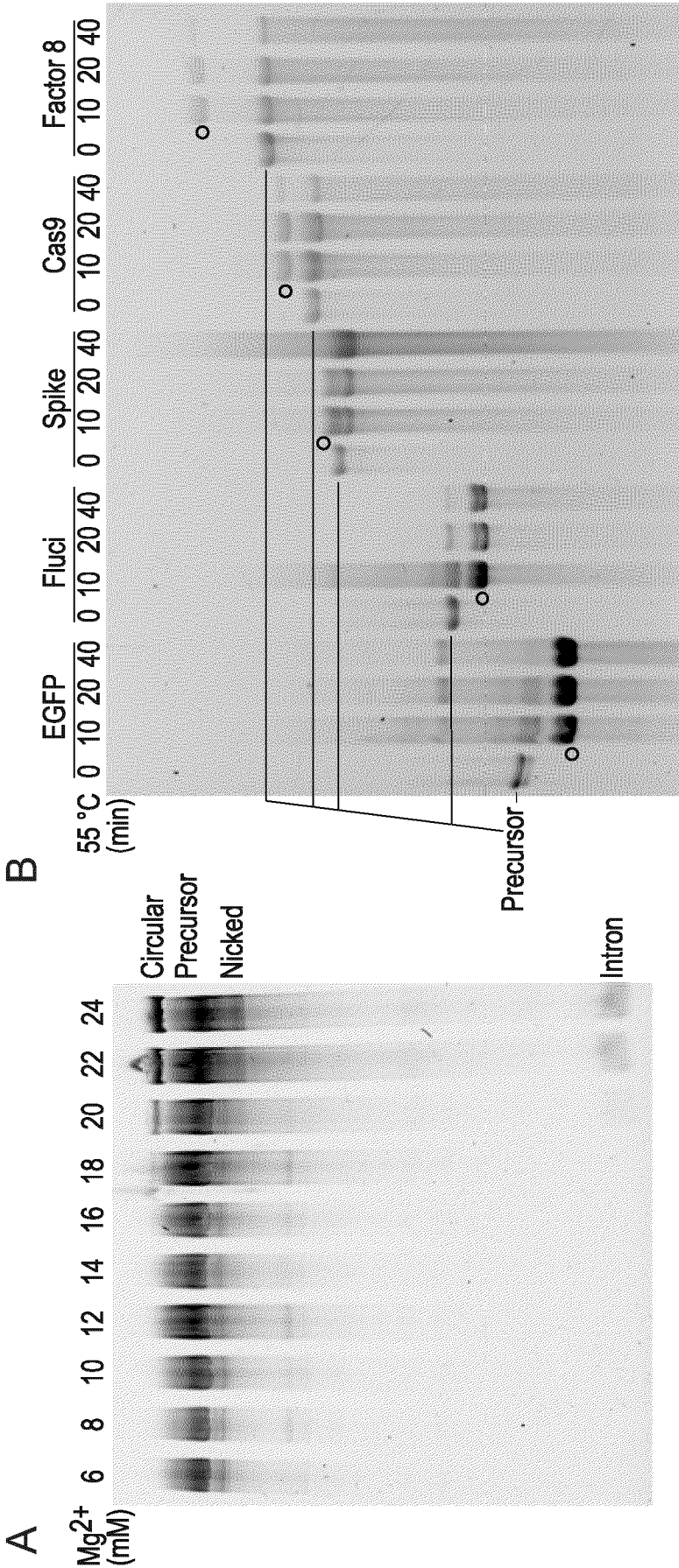


FIG. 6

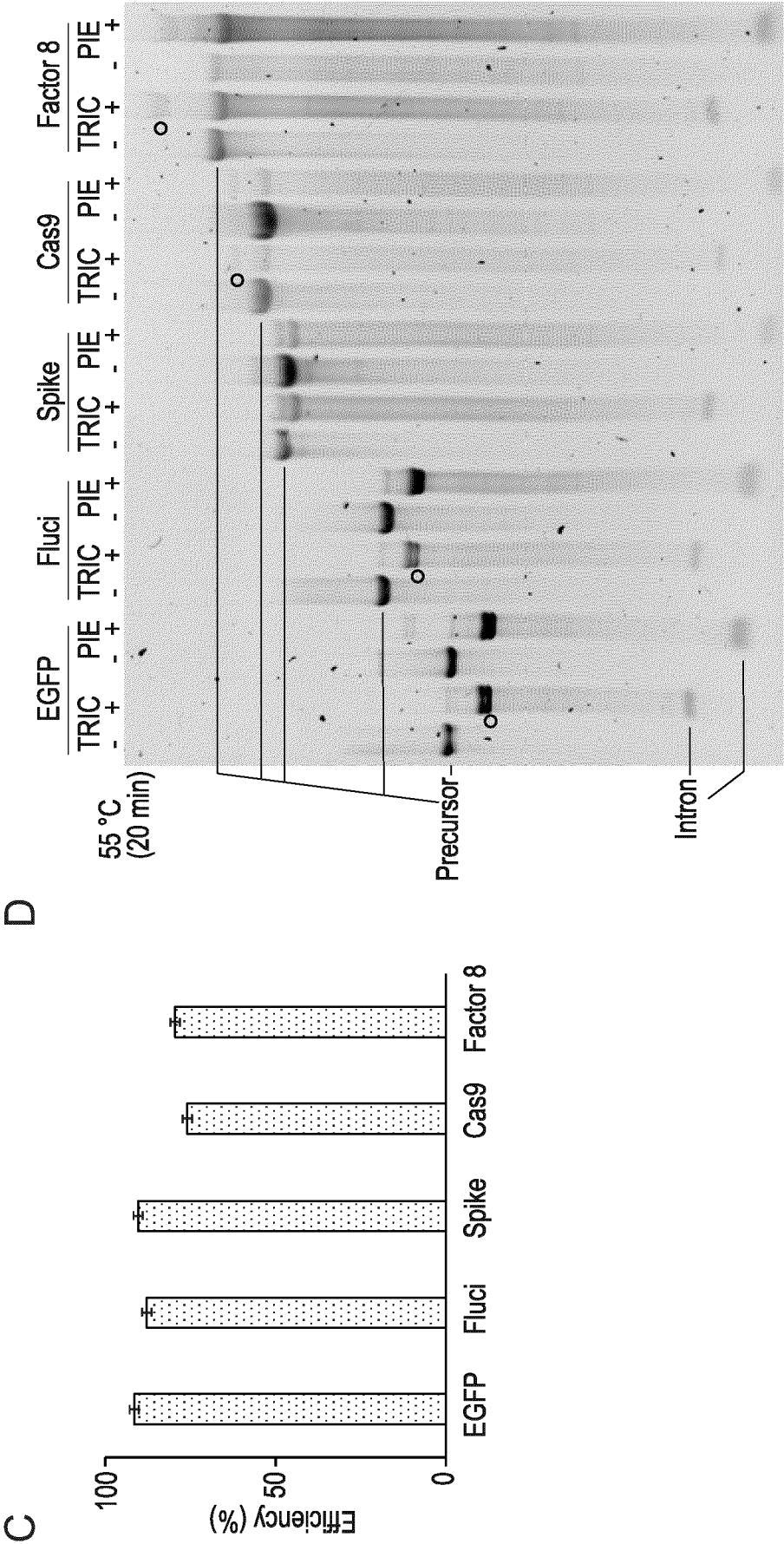


FIG. 6B (continued)

12/40

TRIC-V1: 1.0-2, short GOI

EGS

Ribozyme

Loop_L

IGS

circGOI

Ana tRNA

P1ex

Loop_R

cEGS

EcoR V

G

GGUCAAU

CGGUUG

GCCU

UCCG

AAAUA

UUGAG

CCU

UAGAGA

AAG

AAAUUC

UUUAAG

UGGAUG

CUCU

CAACUC

CAGGG

AAACCU

AAAUCU

AGCUAU

AGACAAG

GCAAUCC

UGAGCC

AAGCCG

AAGUAG

UAAUUAG

UAAGUU

AACAAC

AGAUAA

CUUAC

AGCUAA

UCGGA

AAGGUG

CAGAGA

CUCGAC

GGGAG

CUACCC

UAACGU

CAAGAC

GAGGG

UAAAG

AGAGA

GUCCA

AUUCU

CAAAG

CCAAU

AGGC

AGUAG

CGAA

AGCUG

CGGG

GAG

AAUG

AAAAU

CCGU

UGACCU

UAAAC

GGUCG

UGUGG

UACAC

UCGA

UCUGG

ACUAA

AGCUG

CUCAU

GGAUU

ACAA

GAUCAC

GAUGG

UGA

UUAUAA

GAUCAC

GACAUC

GAUUAC

AAGGA

UGAUG

AUGAU

AAG

AGACG

CUACG

GACUU

AA

ACACC

GGUCA

AUCG

GUUG

GCCU

UCCG

GAU

(SEQ ID N0: 63)

circGOI: R30-3*Flag-L15

T7: T7 promotor
EGS: external guide sequence (20 nt random sequence)
Loop_L/R: loop_left (6 nt)/loop_right (5nt)
IGS: internal guide sequence
R-51/L-15: Ana tRNA right arm-30 nt, left arm-15 nt (1.0-2)
circGOI: R30-3*Flag-L15
P1/P10: paired structure 1/10
P1ex: P1 extension
cEGS: complementary sequence of EGS
EcoR V: restriction enzyme cleavage sites

V1.3	0	1	2	3	4	5	6	7	8	9
R	2	2	2	30	25	16	9	16	9	9
L	3	3	15	3	3	3	3	15	15	8

FIG. 7B

13/40

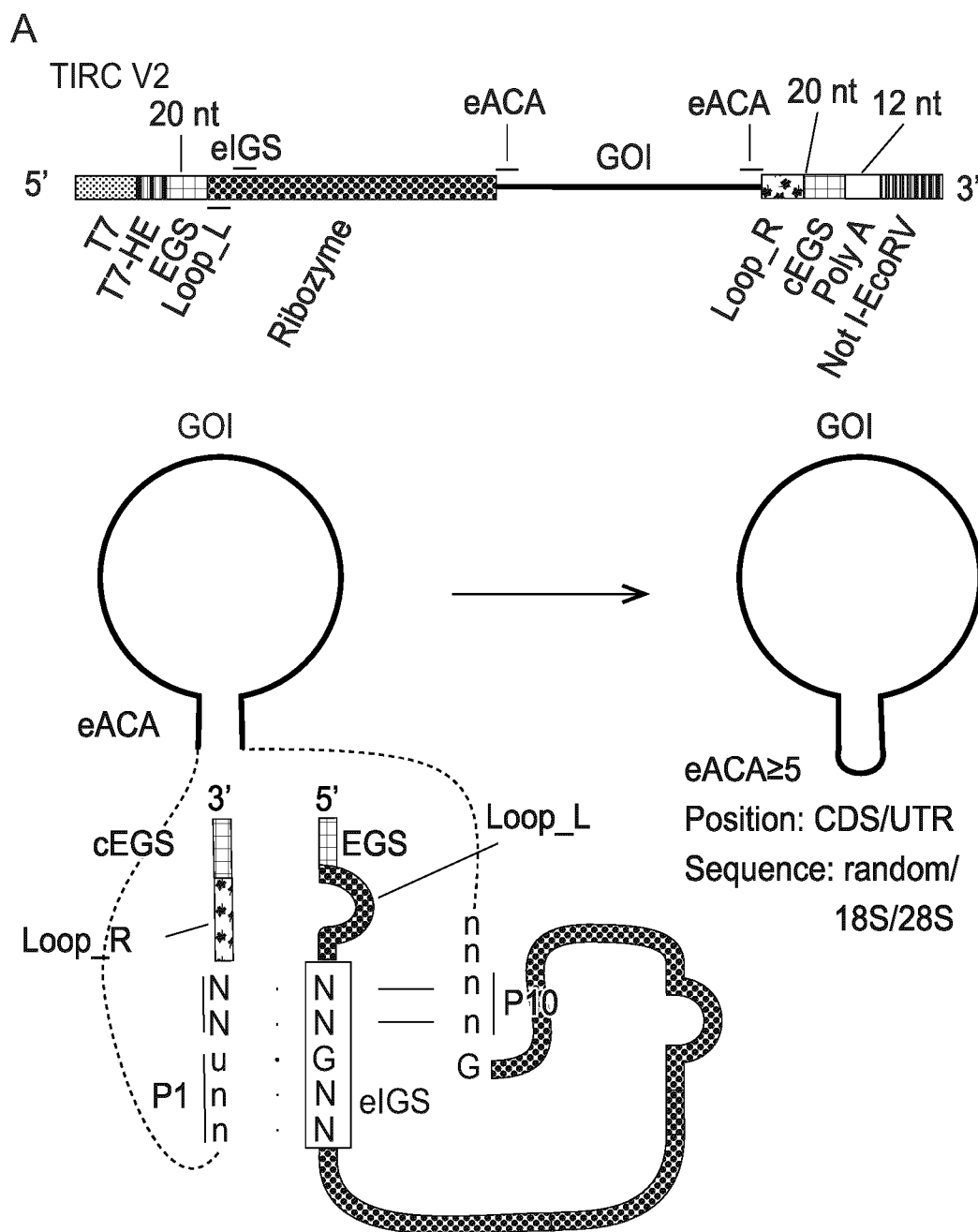


FIG. 8

14/40

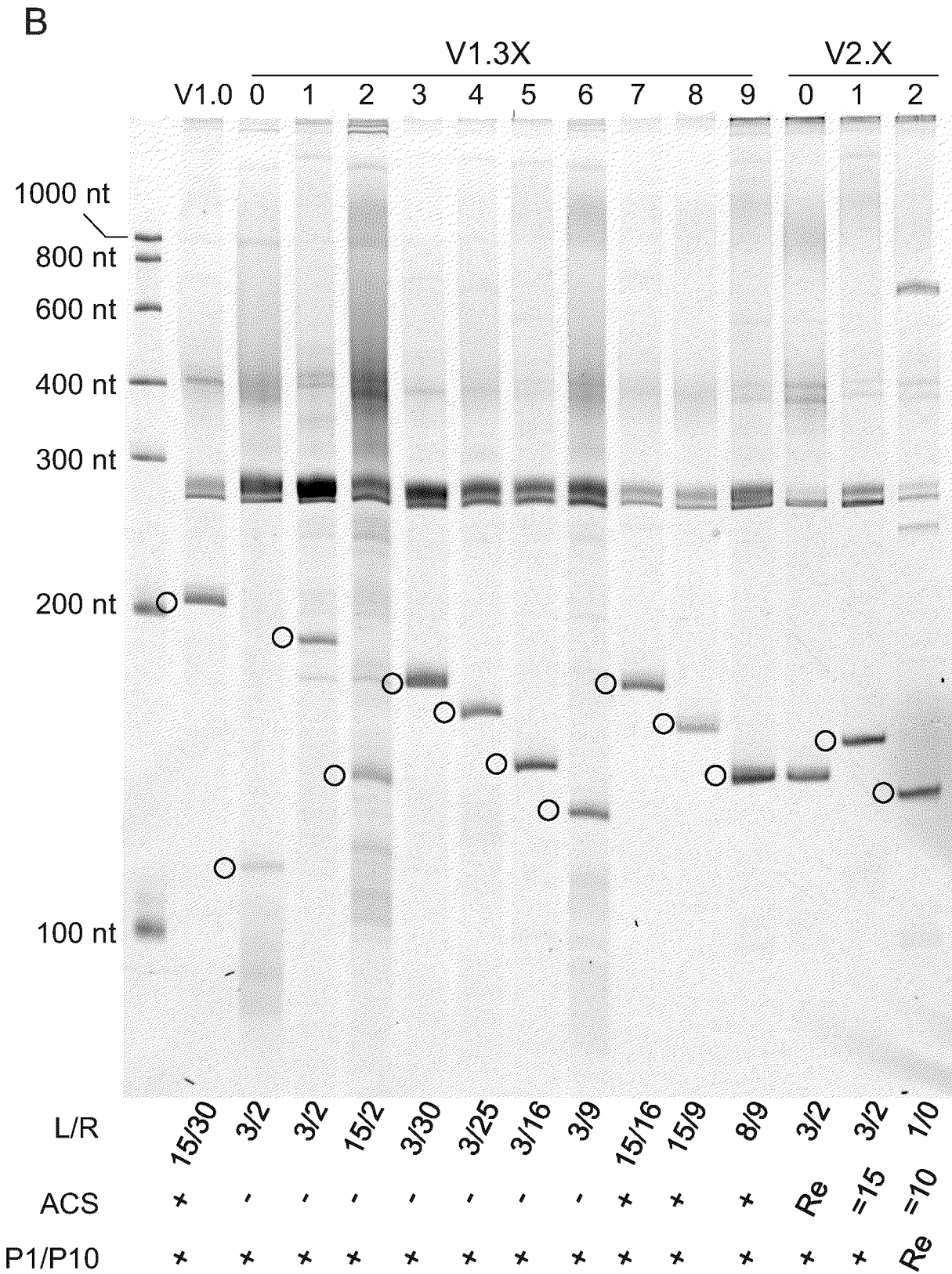


FIG. 8 (continued)

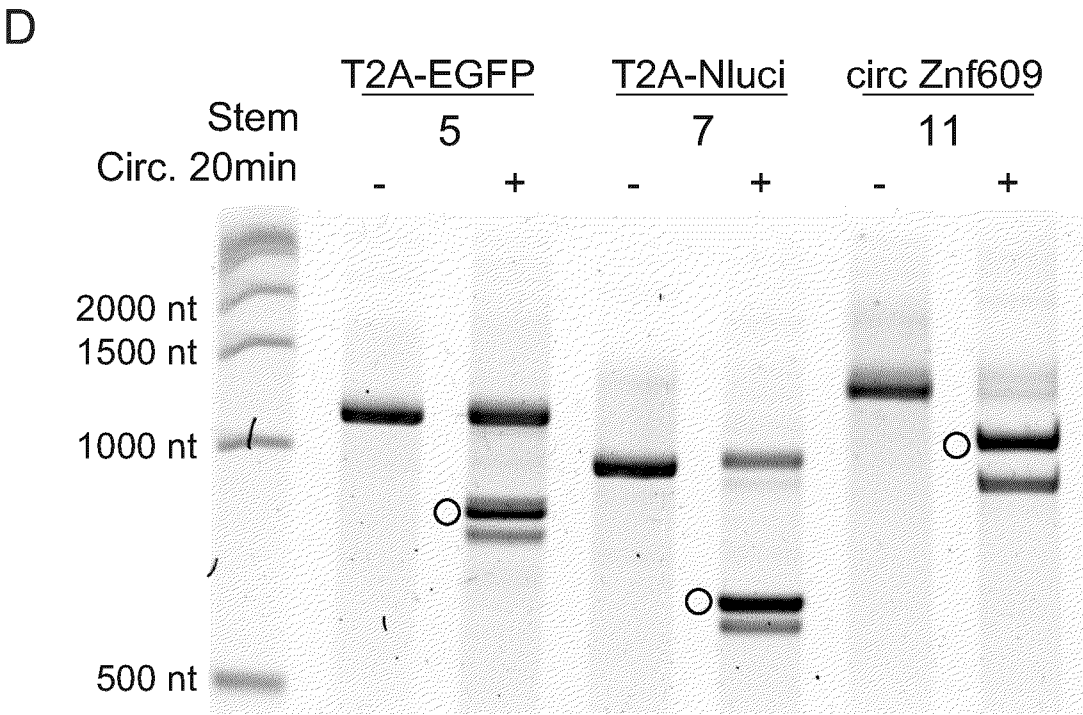
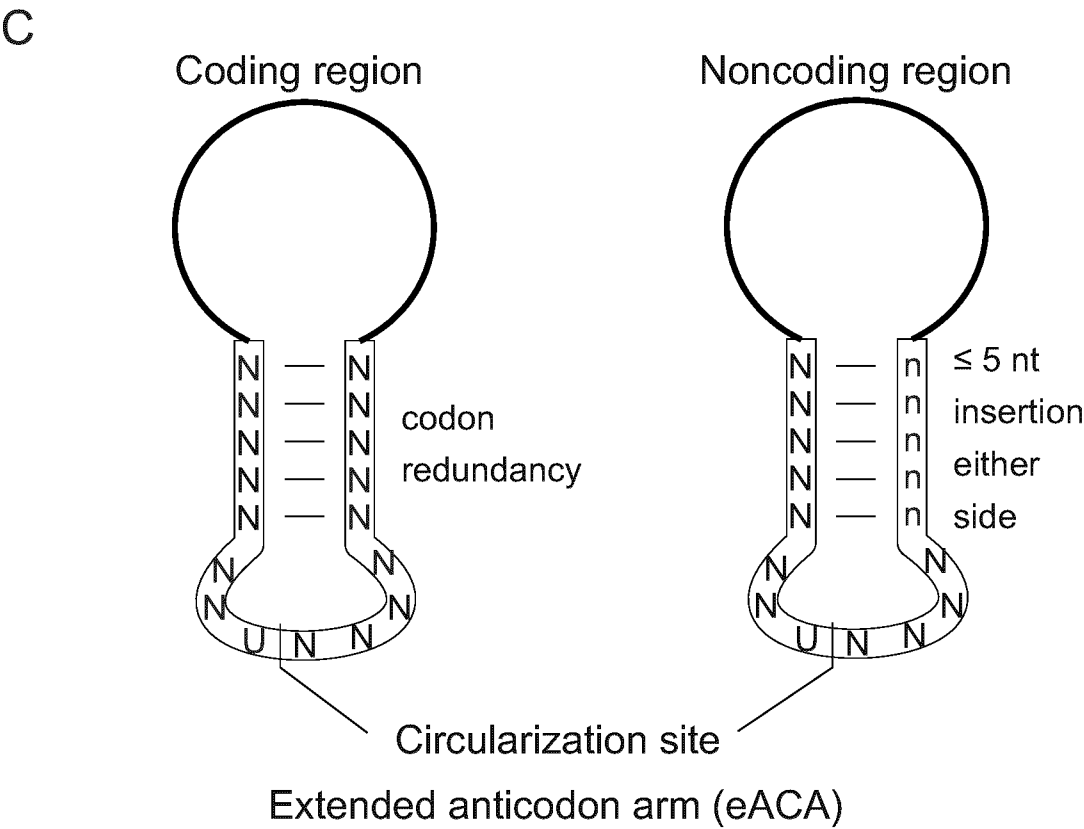


FIG. 8 (continued)

16/40

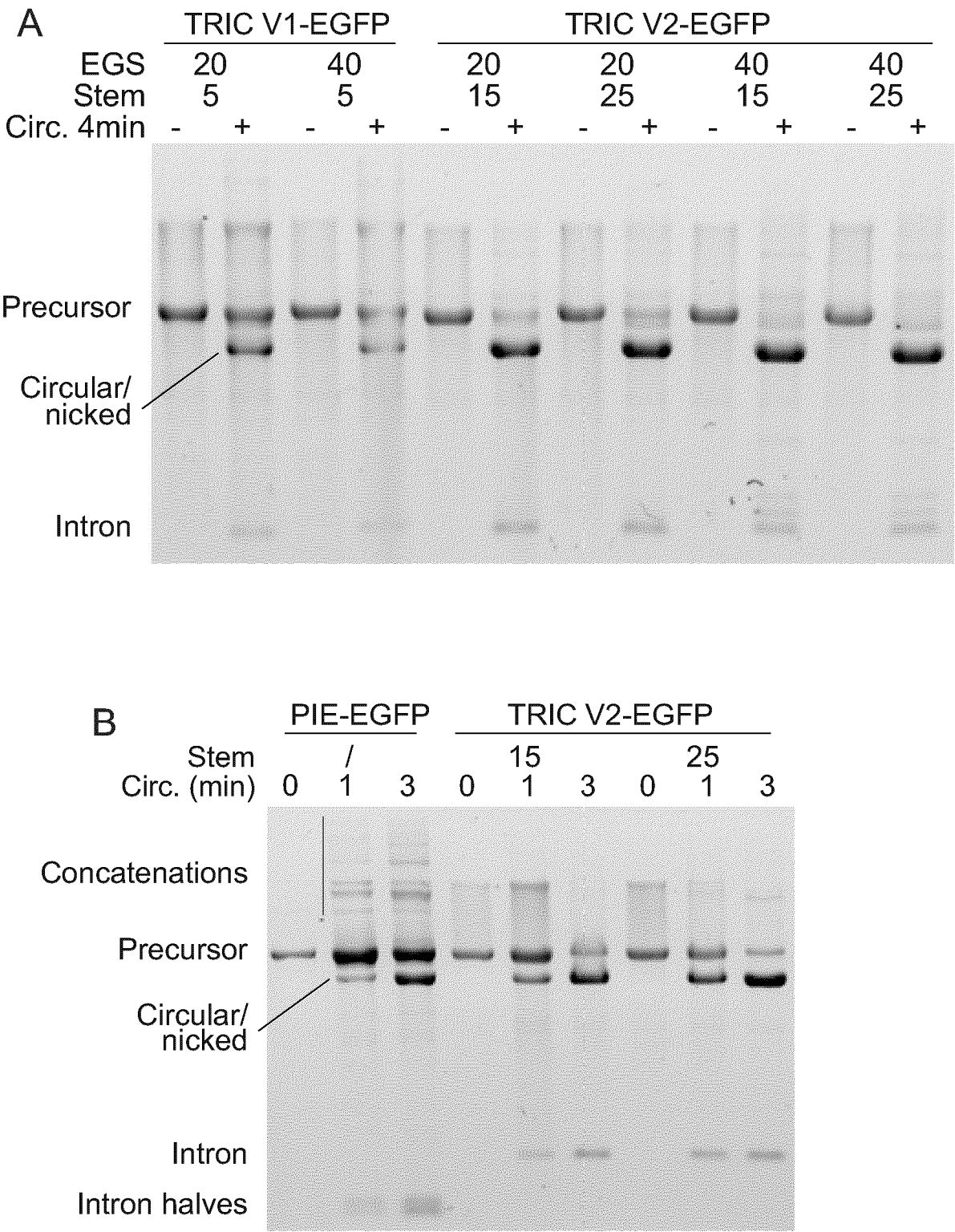
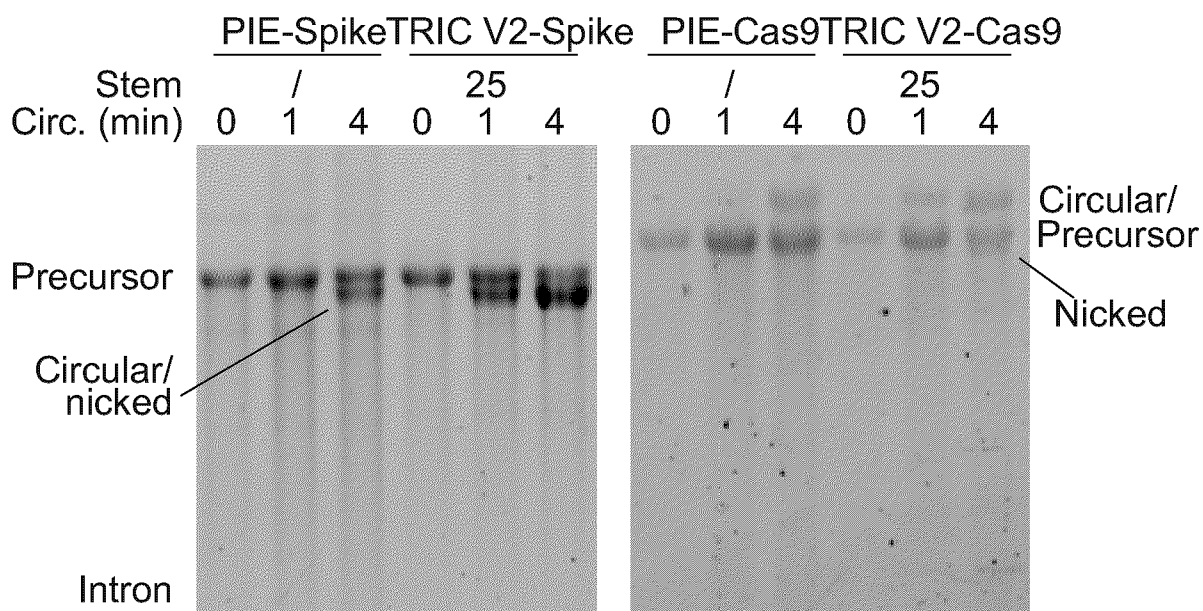


FIG. 9

17/40

C



D

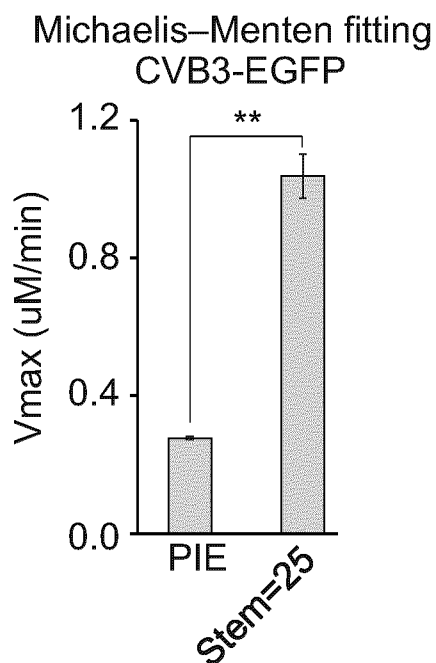


FIG. 9 (continued)

18/40

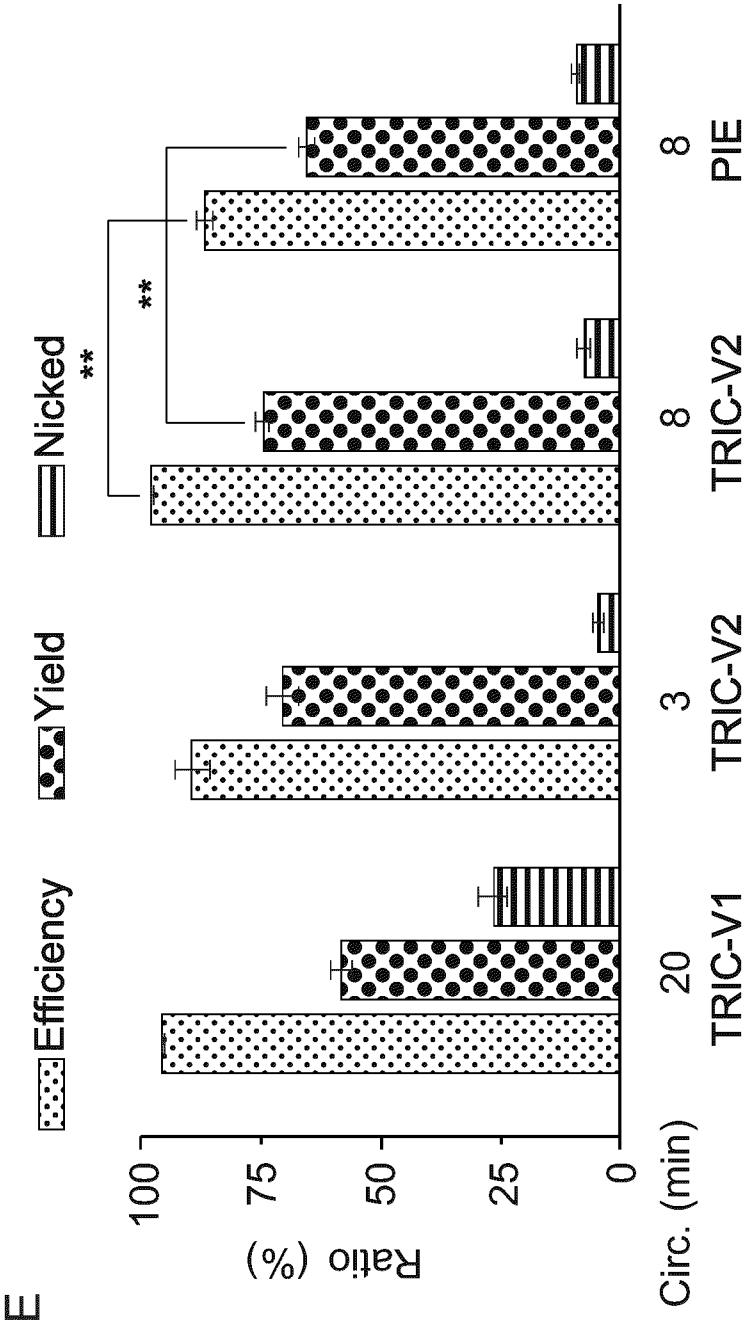


FIG. 9 (continued)

19/40

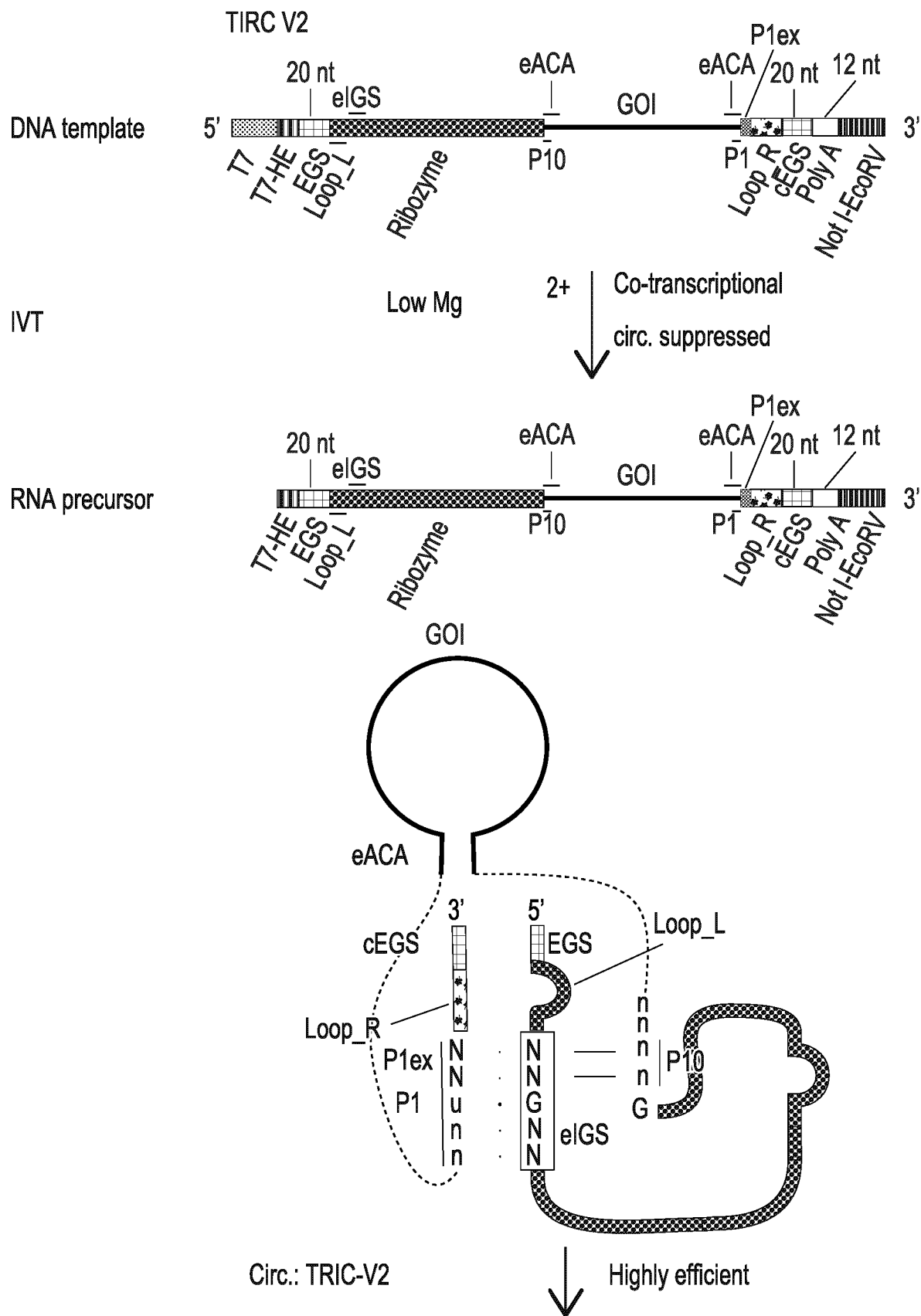


FIG. 10A

20/40

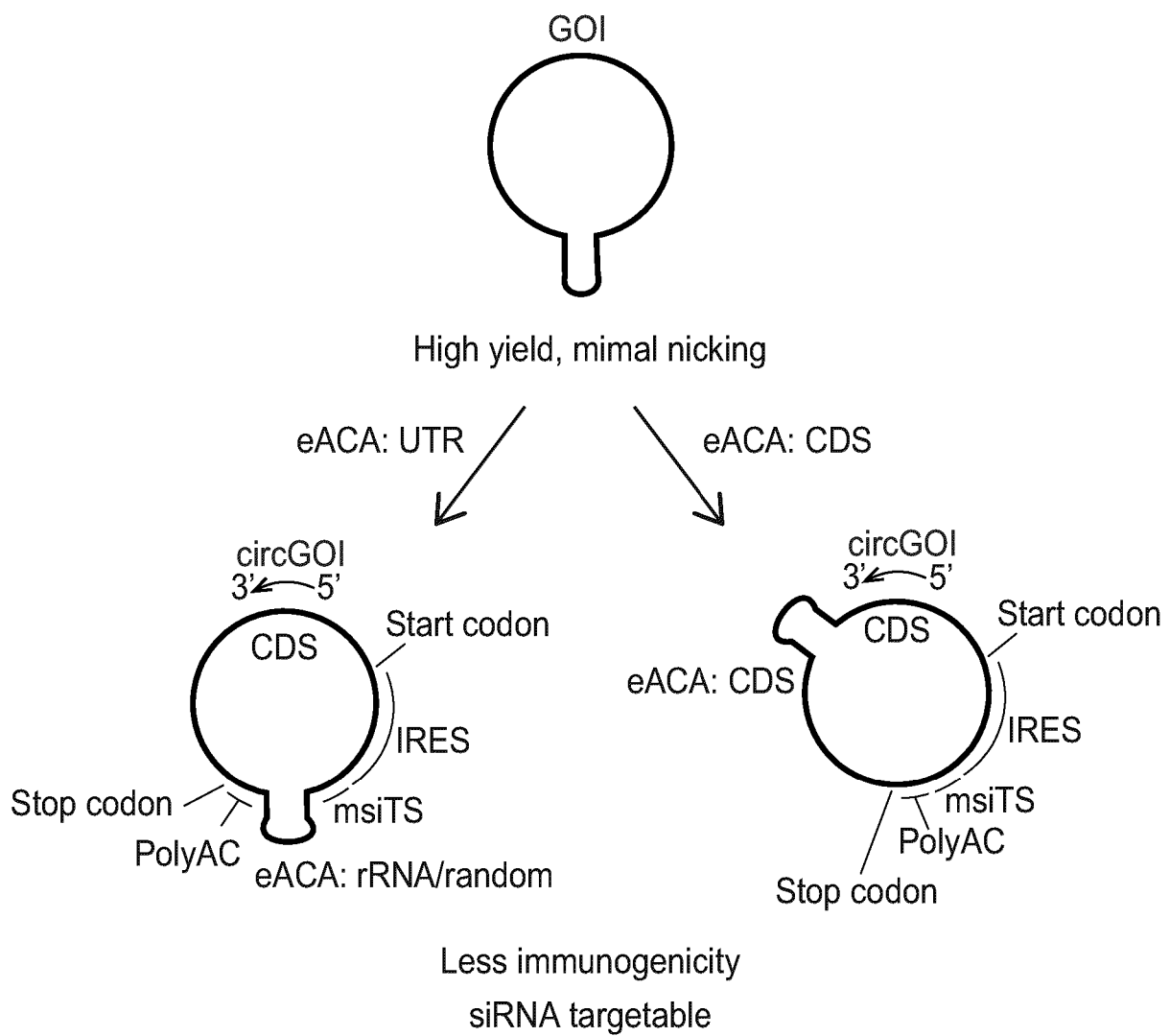


FIG. 10A (continued)

21/40

T7: T7 promotor
T7-HE: T7 high efficiency sequence
EGS: external guide sequence (20/40 nt random sequence)
Loop_L/R: loop_left (6 nt)/loop_right (5nt)
eIGS: extended internal guide sequence (random with G 3 rd)
eACA: extended anticodon arm sequence: random with an U-3 rd 7 nt loop/18S/28S; position; UTR/CDS stem length: 5-25 nt (tested: 5, 7, 10, 11, 15, 17, 24, 25)
circ.GOI-short: 3*Flag circ.GOI-long: 1: eACA (UTR)-msiTS-IRES-start codon-CDS-stop codon-PolyAC-eACA (UTR) 2: eACA (CDS)-stop codon-PolyAC-msiTS-IRES-start codon-eACA (CDS) IRES= CVB3, CSFV, CroV, CDS= EGFP, Spike-EGFP, Cas9-EGFP, Nluc, circZNF609, T2A-EGFP
P1/P10: paired structure 1/10
P1ex: P1 extension
msiTS: multiple siRNA target sequence
cEGS: complementary sequence of EGS
Not I-EcoR V: restriction enzyme cleavage sites

FIG. 10B

22/40

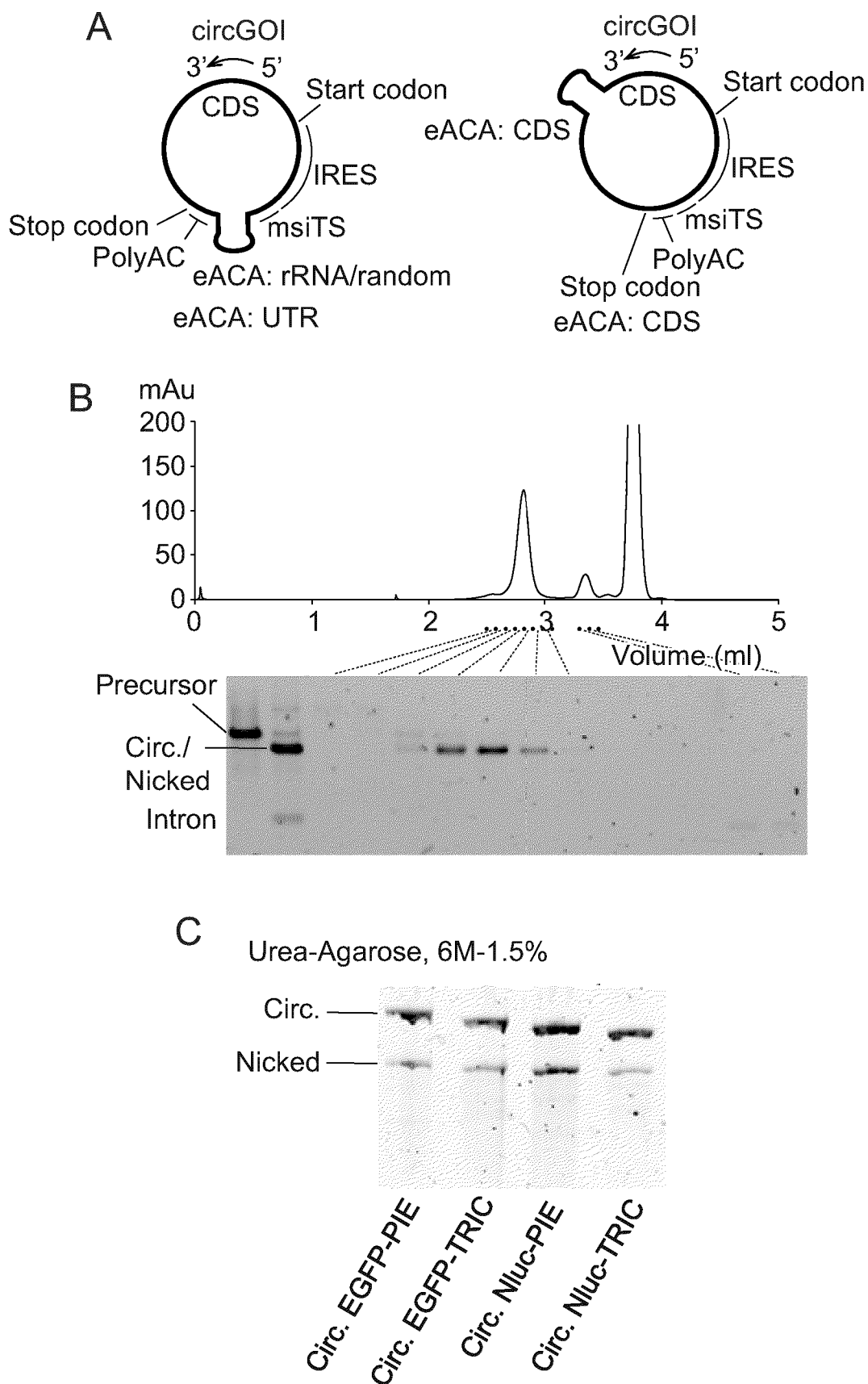


FIG. 11

23/40

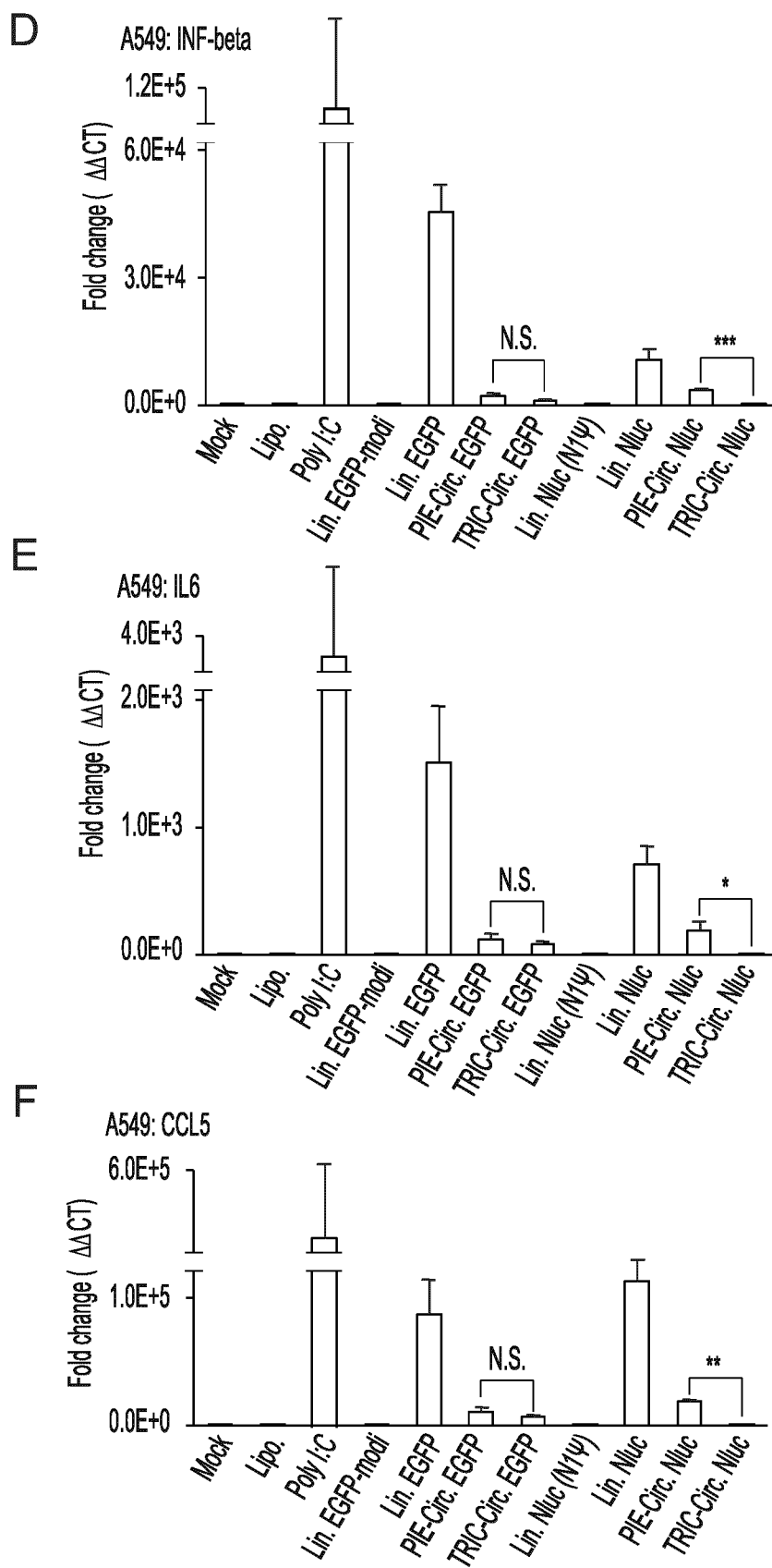


FIG. 11 (continued)

24/40

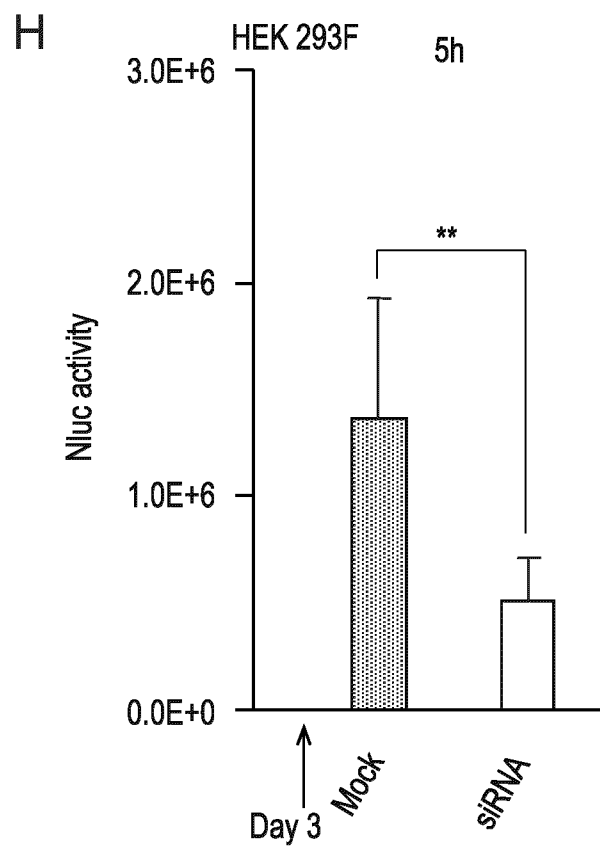
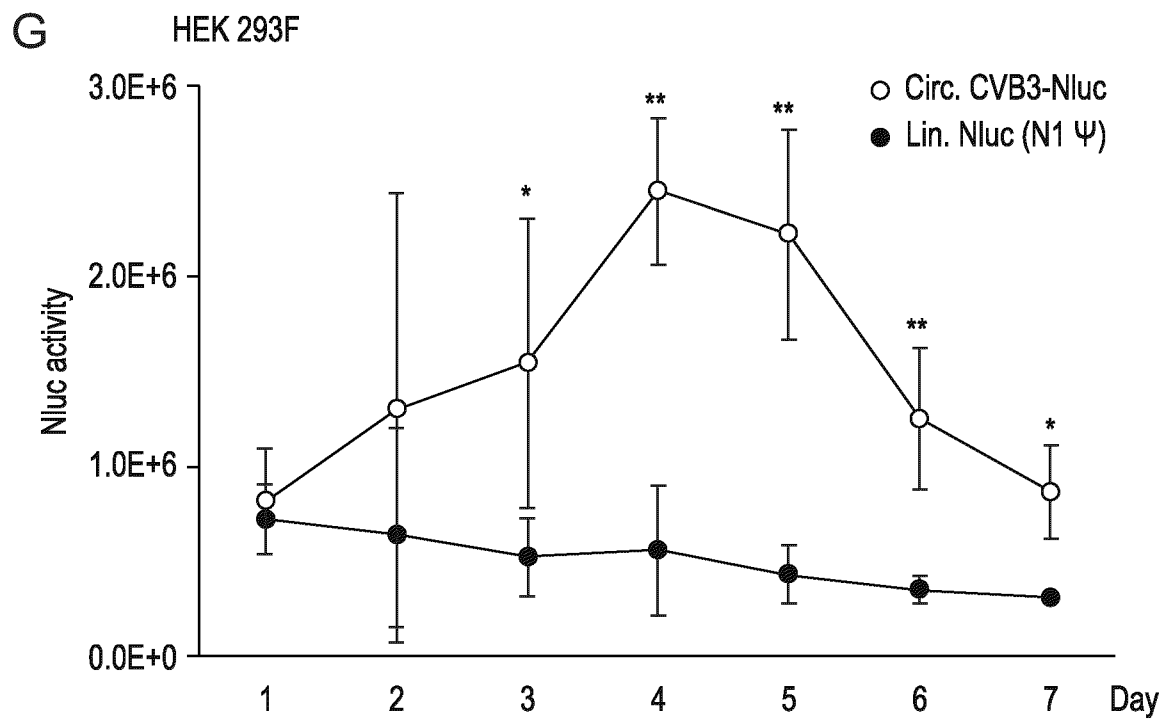
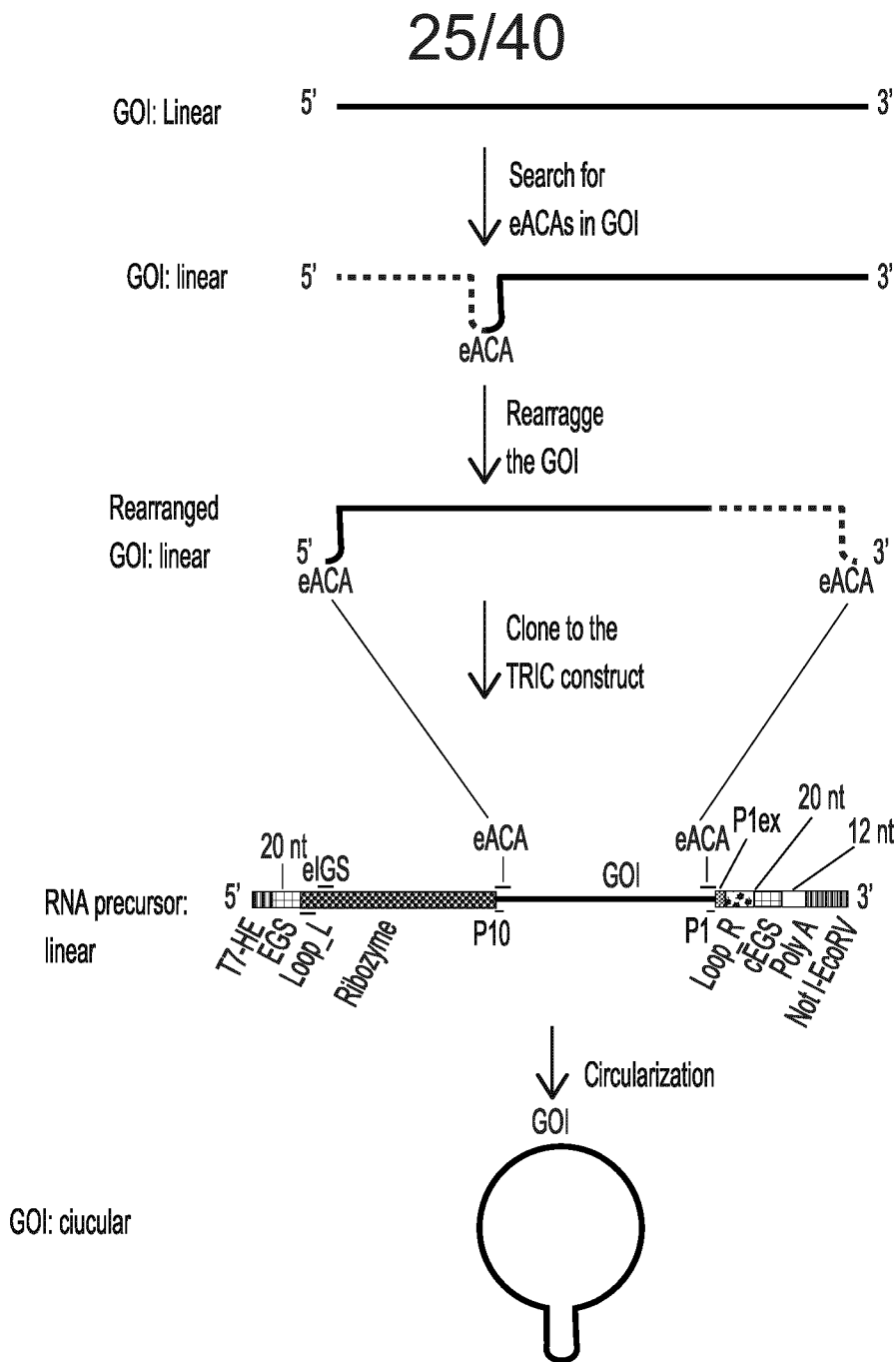


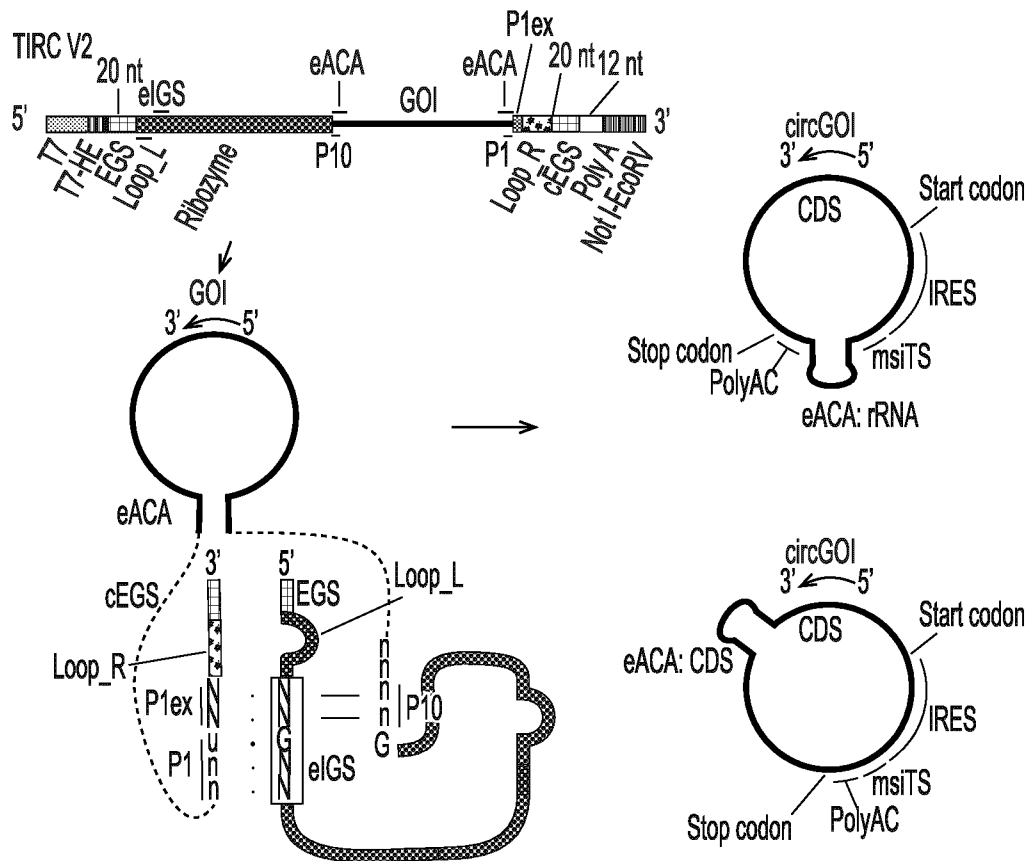
FIG. 11 (continued)



GOI: gene of interest
T7-HE: T7 high efficiency sequence
EGS: external guide sequence (20/40 nt random sequence)
Loop_L/R: loop_left (6 nt)/loop_right (7nt)
eIGS: extended internal guide sequence (random with G 3rd)
eACA: extended anticodon arm
P1/P10: paired structure 1/10
P1ex: P1 extension
cEGS: complementary sequence of EGS
Not I-EcoR V: restirction enzyme cleavage sites

FIG. 12

26/40



T7: T7 promotor
T7-HE: T7 high efficiency sequence
EGS: external guide sequence (20/40 nt random sequence)
Loop_L/R: loop_left (6 nt)/loop_right (5nt)
eIGS: extended internal guide sequence (random with G 3 rd)
eACA: extended anticodon arm sequence: random with an U-3 rd 7 nt loop/18S/28S; position; UTR/CDS stem length: 5-25 nt (tested: 5, 7, 10, 11, 15, 17, 24, 25)
circ.GOI-short: 3*Flag circ.GOI-long: 1: eACA (UTR)-msiTS-IRES-start codon-CDS-stop codon-PolyAC-eACA (UTR) 2: eACA (CDS)-stop codon-PolyAC-msiTS-IRES-start codon-eACA (CDS) IRES= CVB3, CSFV, CroV, CDS= EGFP, Spike-EGFP, Cas9-EGFP, Nluc, circZNF609, T2A-EGFP
msiTS: multiple siRNA target sequence
P1/P10: paired structure 1/10
P1ex: P1 extension
cEGS: complementary sequence of EGS
Not I-EcoR V: restriction enzyme cleavage sites

FIG. 13A

27/40

TIRC V2-Nluc: eACA placed in UTR

T7-HE[EGS] [Ribozyme] [Loop_L] [IGS] circGOI [eACA] [P1ex] [Loop_R] cEGS PolyA Not I-EcoR V

GGGAUAAUAC [GGUCAAUACAGUUGGCUUCCG] [AAAUAAUUGAGCCUUAGAGAAGAAAUU]
 CUUUAAAGUGGAUGCUCUCAAAACUCAGGGAAACCUAAAUCUAGCUAUAGACAAGGCAAUC
 CUGAGCCAAGCCGAAGUAGUAAUAGUAAGUUAACAACAGAUACUACAGCUAAUCGG
 AAGGUGCAGAGACUCGACGGGAGCUACCCUAACGUAAGACGAGGGUAAAGAGAGAGU
 CCAAUUCUCAAGCCAAUAGGCAGUAGCGAAAGCUGCGGGAGAAUG [AAAAGGAGGGCG]
 CGCGGGUCGGGGCGGCGG [AAAACAAUUGGGCCAGAUCGAAAAAACGAACAUGAUCGA
 CUAUUUCACACAUUAAAACAGCCUGUGGGUUGAUCCCACCCACAGGCCCAUUGGGCGC
 UAGCACUCUGGUUACACGGUACCUUUGUGCGCCUGUUUUUUAUACCCCUCCCCCAACUG
 UAACUUAGAAGUAACACACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGA
 UCAAGCACUUCUGUUACCCCGGACUGAGUAUCAUAGACUGCUCACGCGGUUGAAGGA
 GAAAGCGUUCGUUAUCCGGCCAACUACUUCGAAAAACCUAGUAACACCGUGGAAGUUG
 CAGAGUGUUUCGCUCAGCACUACCCAGUGUAGAUAGGUCGAUGAGUCACCGCAUUC
 CCCACGGGCGACCGUGGGCGGUGGCUGCGUUGGCGGCCUGCCCAUGGGGAAACCCAUG
 GGACGCUCUAAUACAGACAUGGUGCGAAGAGUCUAUUGAGCUAGUUGGUAGUCCUCCG
 GCCCCUGAAUGCGGCUAUCCUAACUGCGGAGCACACACCCUCAAGCCAGAGGGCAGU
 GUGUCGUAAACGGGCAACUCUGCAGCGGAACCGACUACUUUGGGUGUCCGUGUUUCAU
 UUAUUCUUAUACUGGCUGCUUAUGGUGACAAUUGAGAGAUCGUUACCAUUAAGCUAU
 GGAUUGGCCAUCCGGUGACUAAUAGAGCUAUUAUUAUACCCUUUGUUGGGUUUAUACC
 ACUUAAGCUUGAAAGAGGUUAAAACAUUACAUAUUAUUAAGUUGAAUACAGCAAAU
 GGUCUUCACACUCGAAGAUUUCGUUGGGGACUGGCGACAGACAGCCGGCUACAACCUG
 GACCAAGUCCUUGAACAGGGAGGUGUGUCCAGUUUGUUUCAGAAUCUCGGGGUGUCC
 GUAACUCCGAUCCAAAGGAUUGUCCUGAGCGGUGAAAAUGGGCUGAAGAUCGACAUC
 AUGUCAUCAUCCCGUAUGAAGGUCUGAGCGGCGACCAAUUGGGCCAGAUCGAAAAAU
 UUUUAAGGUGGUGUACCCUGUGGAUGAUCAUCACUUUAAGGUGAUCCUGCACUAUGGC
 ACACUGGUAAUCGACGGGGUACGCCGAACAUGAUCGACUAAUUCGGACGGCCGUAUG
 AAGGCAUCGCCGUGUUCGACGGCAAAAAGAUACUGUAACAGGGACCCUGUGGAACGG
 CAACAAAUAUUCGACGAGCGCCUGAUCAACCCCGACGGCUCCUGUGUUCGAGUA
 ACCAUCAACGGAGUGACCGGCUGGCGGCUGUGCGAACGCAUUCUGGCGUGAAAAAAC
 AAAAAACAAAACCGCCCCCGGCCCGCGUCCUCCCU [AA] ACACC CGGAAGCCAACU
 GAUUGACC AAAAAAAAAA GCGGCCGCGAU (SEQ ID NO: 61)

circ.GOI-long:

eACA (UTR)-msiTS-IRES-start codon-CDS-stop codon-PolyAC-eACA (UTR)

IRES= CVB3, CDS= EGFP, Spike-EGFP, Cas9-EGFP, Nluc

FIG. 13B

28/40

TIRC V2-EGFP: eACA placed in CDS

T7-HE [EGS] [Ribozyme] [Loop_L] [IGS] circGOI [eACA] [P1ex] [Loop_R] cEGS PolyA Not I-EcoR V

GGGAAUAAUAC GGUCAAUCAGUUGGCUUCCG AAAUAAUUGAGCCUUAGAGAAGAAAUU
 CUUUAAAGUGGAUGCUCUCAACUCAGGGAAACCUAAAUCUAGCUAUAGACAAGGCAAUC
 CUGAGCCAAGCCGAAGUAGUAAUAGUAAGUUAACAACAGAUAAACUUAACAGCUAAUCGG
 AAGGUGCAGAGACUCGACGGGAGCUACCCUAACGUCUAGACGAGGGUAAAGAGAGAGU
 CCAAUUCUCAAGCCAAUAGGCAGUAGCGAAAGCUGCGGGAGAAUG AACGGUCACAAG
 UUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGU
 UCAUCUGCACCACCGGCAAGCUGCCCGUGCCCGUGGCCACCCUCGUGACCACCCUGAC
 CUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
 AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACGACG
 GCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCGCAU
 CGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCCUGGGGCACAAGCUGGAG
 UACAACUACAACAGCCACAACGUCUUAUAUCAUGGCCGACAAGCAGAAGAACGGCAUCAA
 GGUGAACUUAAGAUCGCCACAACAUCGAGGACGGCAGCGUGCAGCUCGCCGACCAC
 UACCAGCAGAACACCCCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACC
 UGAGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAGAAGCGCGAUACAUUGGUCCU
 GCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCUGUACAAGUG
 AAAAAACAAAAACAAAAAAACAAACAAUUGGGCCAGAUCGAAAAAACGAACAUGAUCG
 ACUAAUUCACACAUAUAAACAGCCUGUGGGUUGAUCCACCCACAGGCCCAUUGGGCG
 CUAGCACUCUGGUUAUCACGGUACCUUUGUGCGCCUGUUUUUAUACCCCCUCCCCAACU
 GUAACUUAAGAAGUAACACACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUG
 AUCAAGCACUUCUGUUACCCCGGACUGAGUAUCAAUAGACUGCUCACGCGGUUGAAGG
 AGAAAGCGUUCGUUAUCCGGCCAACUACUUCGAAAAACCUAGUAACACCGUGGAAGUUG
 CAGAGUGUUUCGCUCAGCACUACCCAGUGUAGAUCAGGUCGAUGAGUCACCGCAUUC
 CCCACGGGCGACCGUGGCGGUGGCUGCGUUGGCGGCCUGCCCAUGGGGAAACCCAUG
 GGACGCUCUAAUACAGACAUGGUGCGAAGAGUCUUAUUGAGCUAGUUGGUAGUCCUCCG
 GCCCCUGAAUGCGGCUAAUCCUAACUGCGGAGCACACCCUCAAGCCAGAGGGCAGU
 GUGUCGUAACGGGCAACUCUGCAGCGGAACCGACUACUUGGGUGUCCGUGUUUCAUU
 UUAUUCUUAUACUGGCUGCUUAUUGGUGACAAUUGAGAGAUCGUUAACCAUAUAGCUAUUG
 GAUUGGCCAUCCGGUGACUAAUAGAGCUAAUUAUAUAUCCCUUUGUUGGGUUUAUACCA
 CUUAGCUUGAAAGAGGUUAAACAUUACAUAUUAUUGUUAAGUUGAAUACAGCAAAUUG
 GUGAGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGAC
 GGUGACGUU AA ACACC CGGAAGCCAACUGAUUGACC AAAAAAAAAA GCGGCCGC
 GAU

(SEQ ID NO: 62)

circ.GOI-long:

eACA (CDS)-stop codon-PolyA-msiTS-IRES-start codon-eACA (CDS)

IRES= CVB3, CSFV, CroV, CDS= EGFP, Nluc, circZNF609, T2A-EGFP

FIG. 13B (continued)

30/40

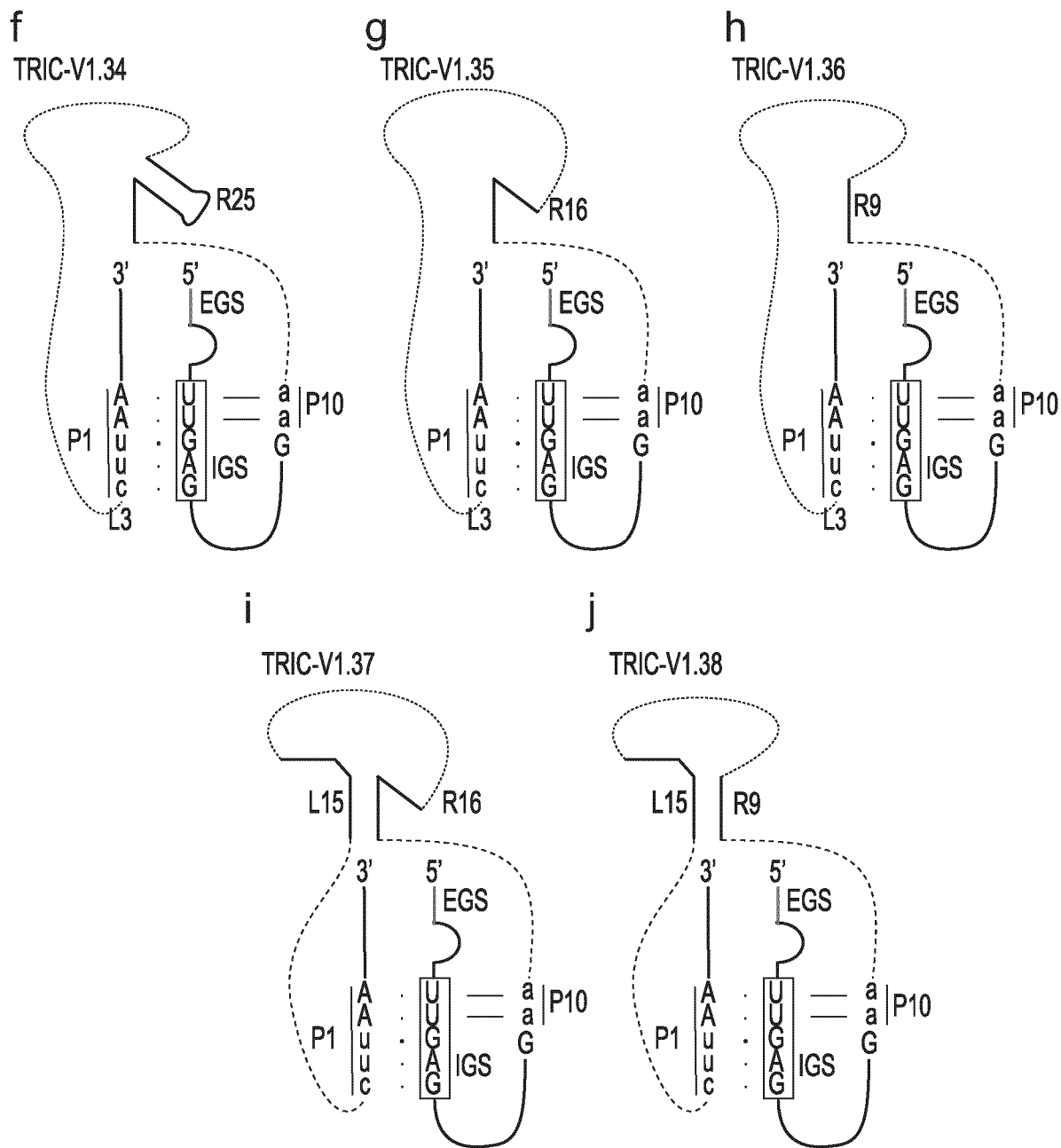


FIG. 14B (continued)

31/40

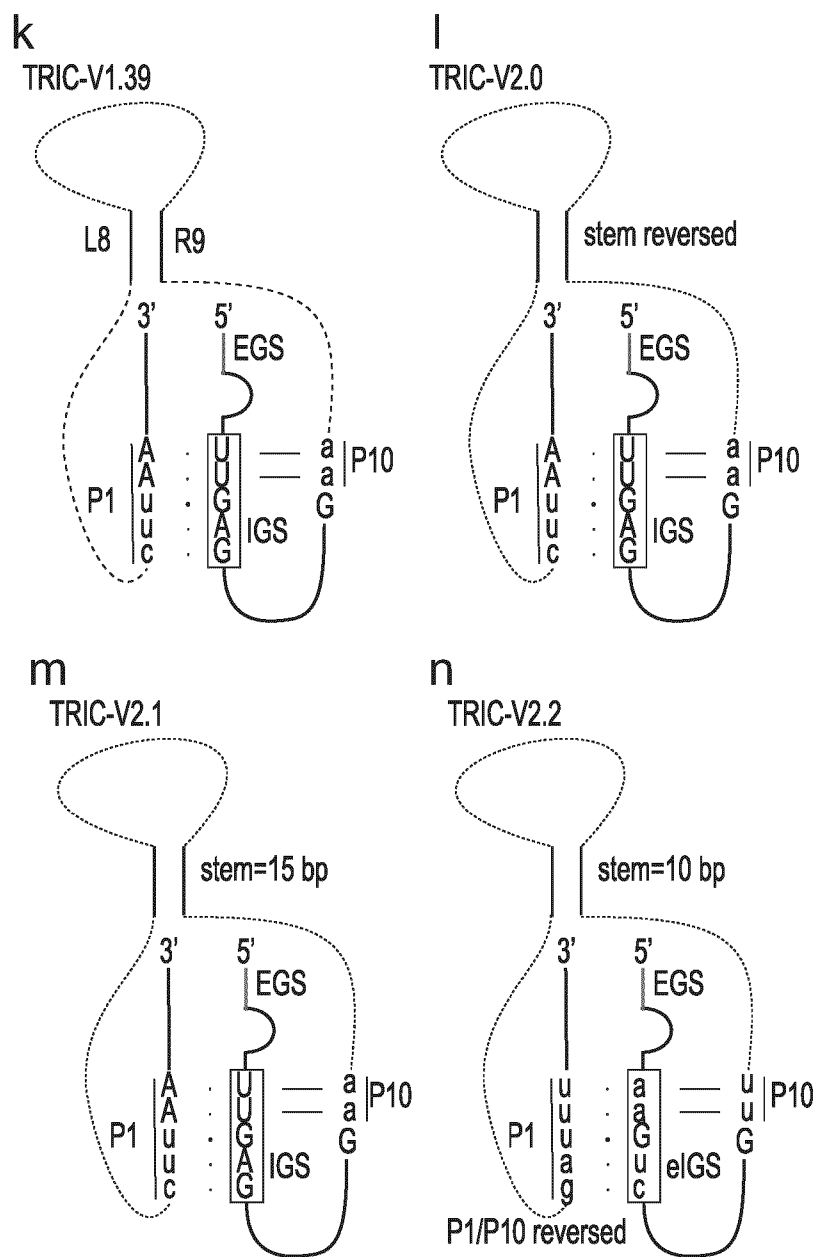


FIG. 14B (continued)

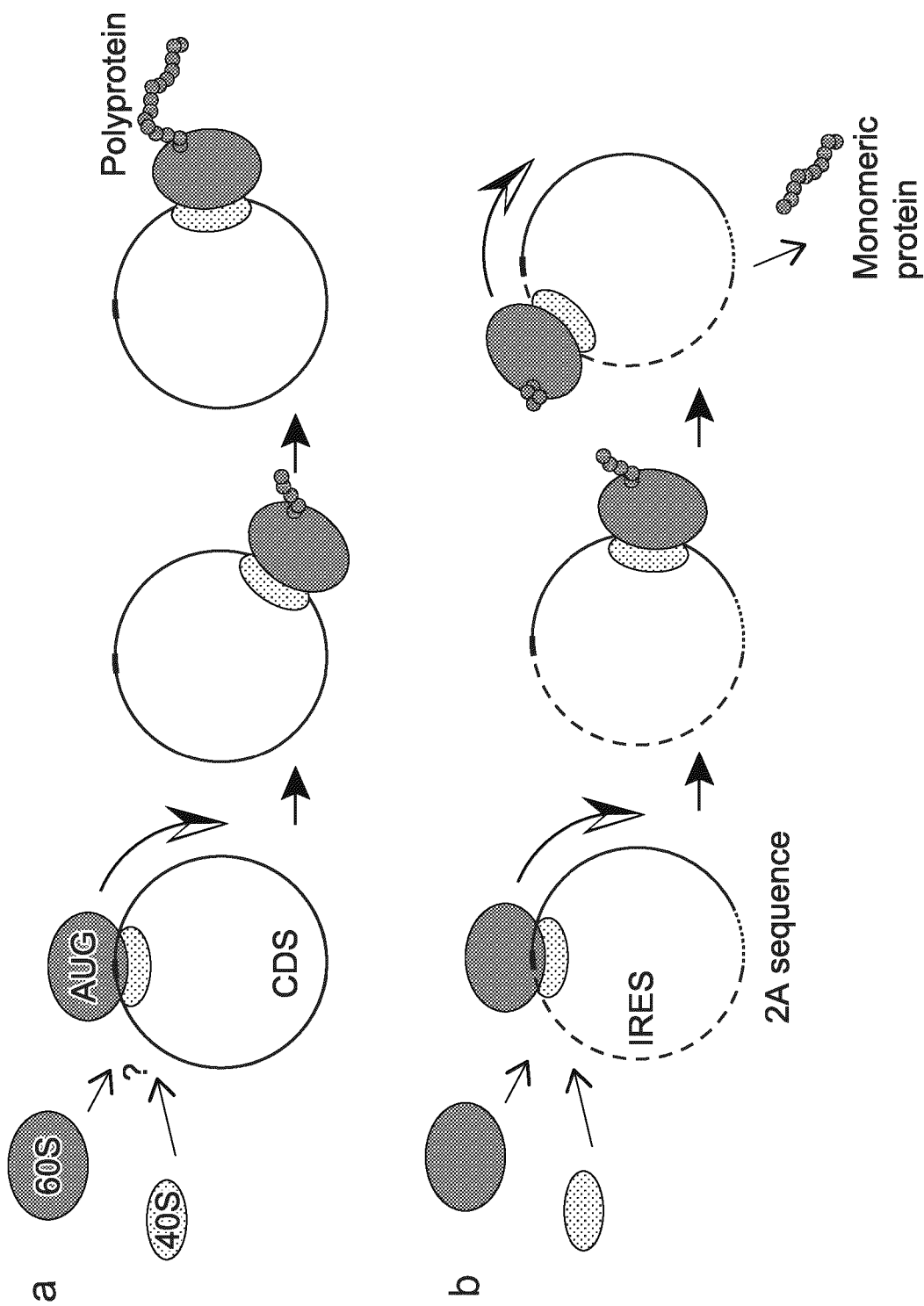


FIG. 15



34/40

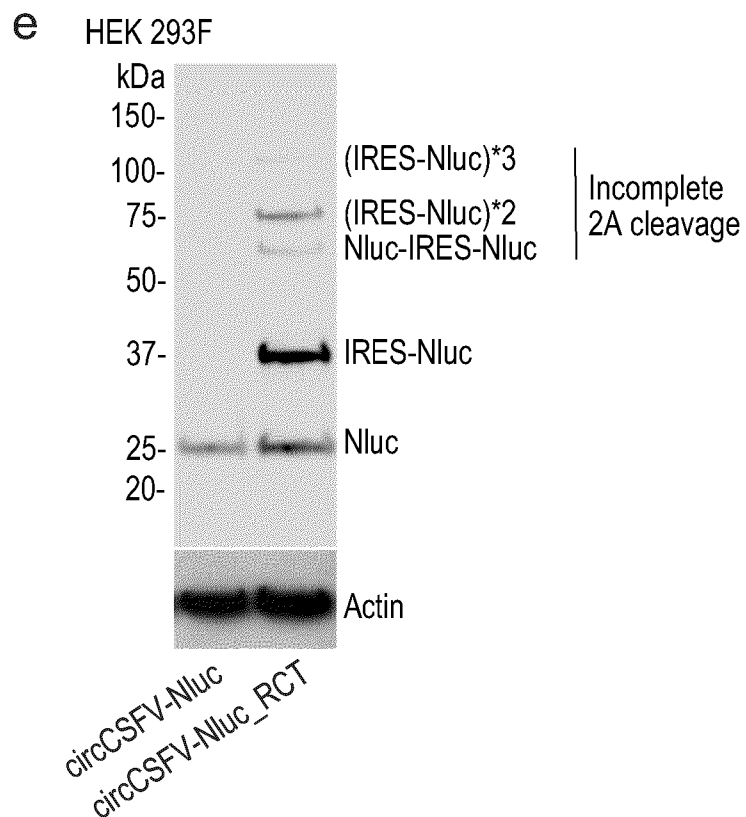
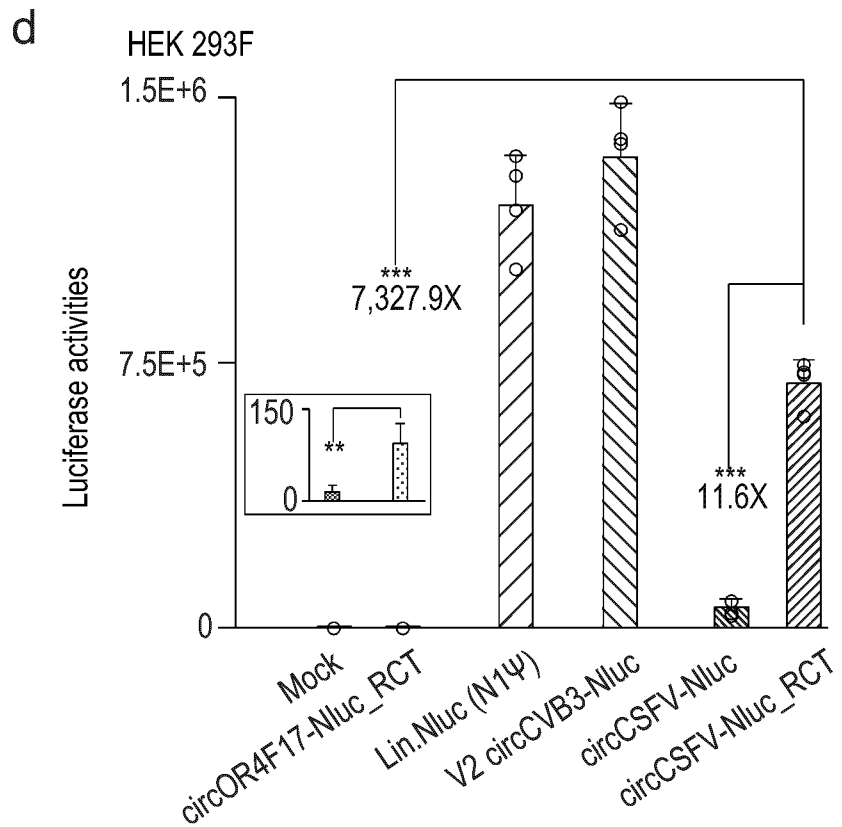


FIG. 15 (continued)

35/40

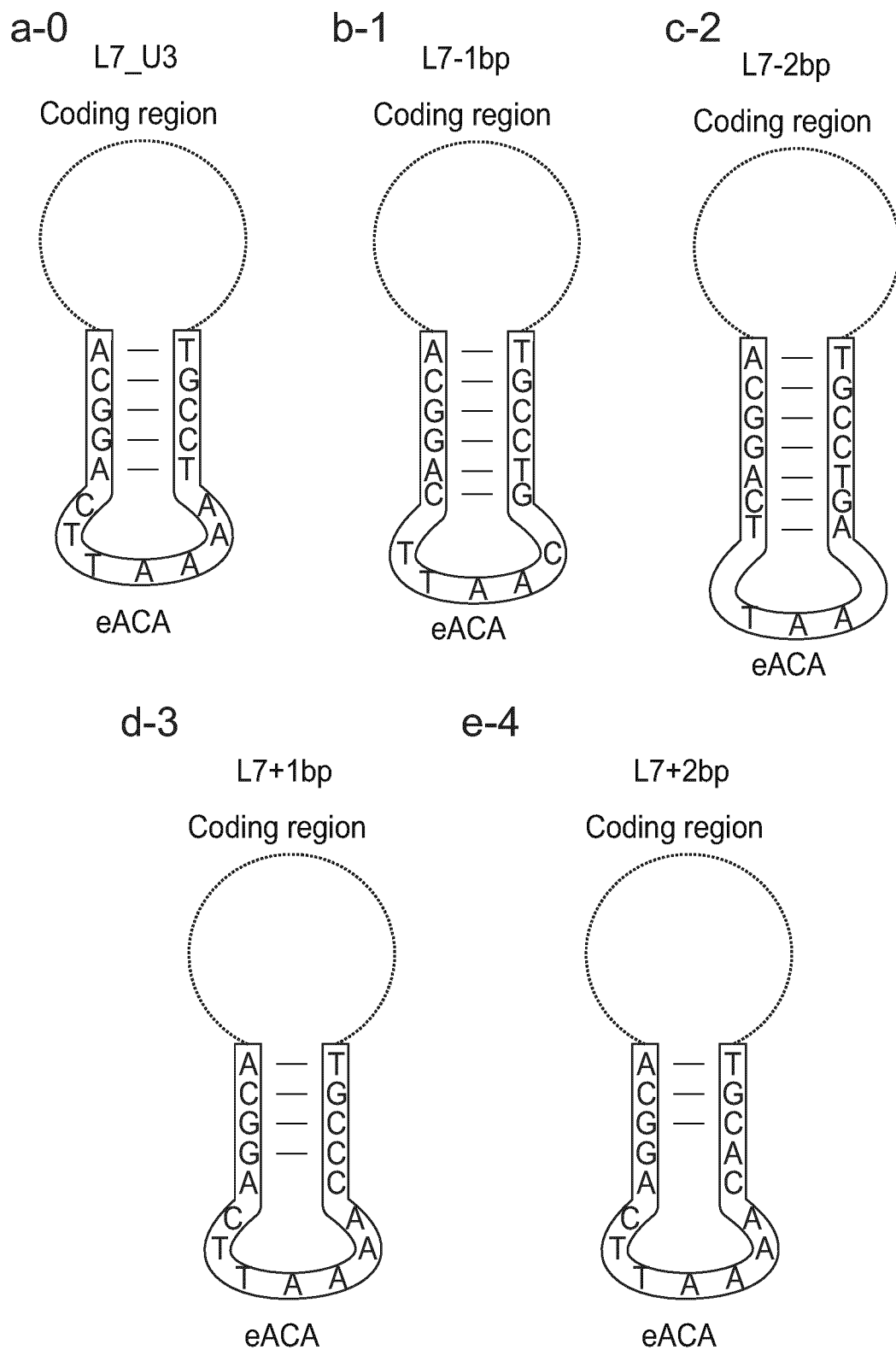


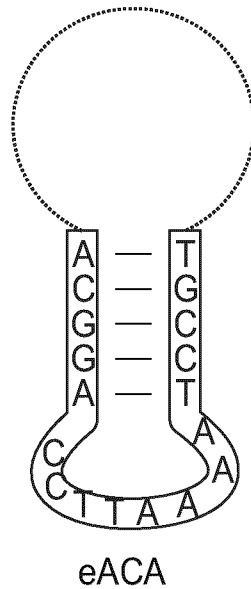
FIG. 16A

36/40

f-5

L7_p1+1nt

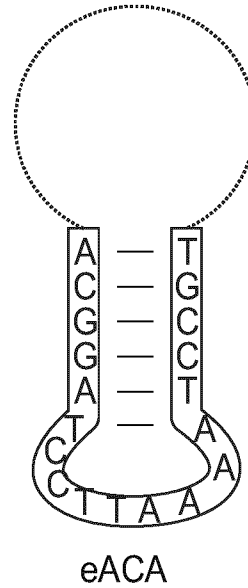
Coding region



g-6

L7_p1+1nt_p1ex-1nt

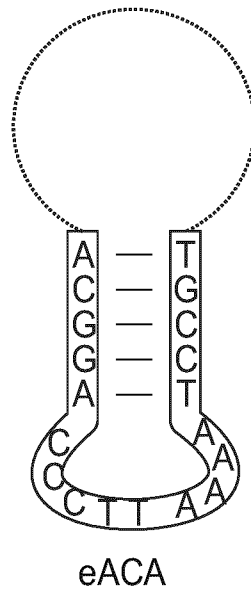
Coding region



h-7

L7_p1+2nt

Coding region



i-8

L7_p1+2nt_p1ex-2nt

Coding region

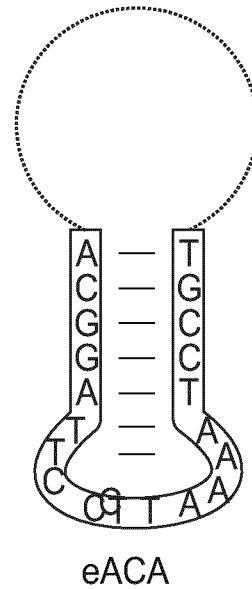
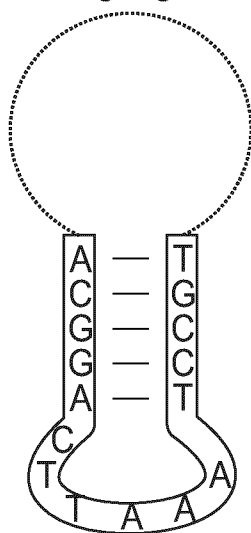


FIG. 16A (continued)

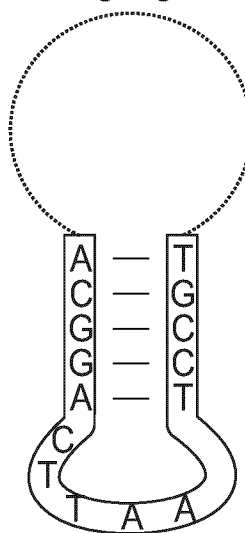
37/40

j-9

L7_p1ex-1nt
Coding region

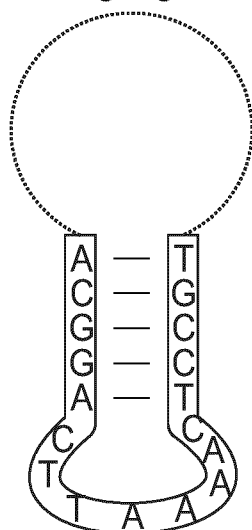
eACA

k-10

L7_p1ex-2nt
Coding region

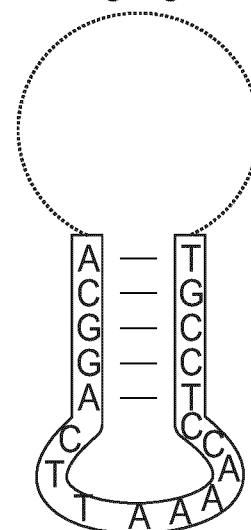
eACA

l-11

L7_p1ex+1nt
Coding region

eACA

m-12

L7_p1ex+2nt
Coding region

eACA

FIG. 16A (continued)

38/40

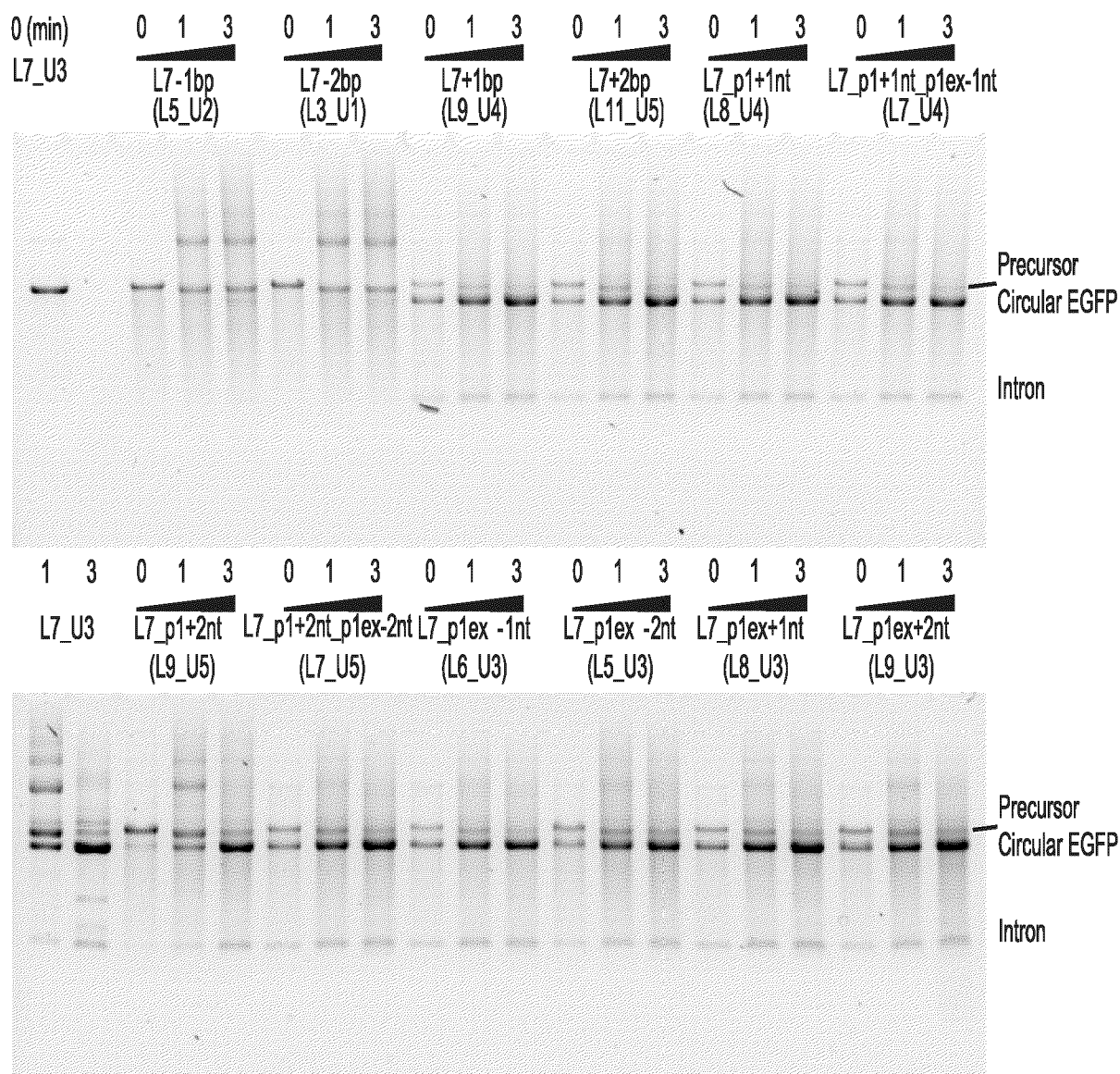


FIG. 16B

39/40

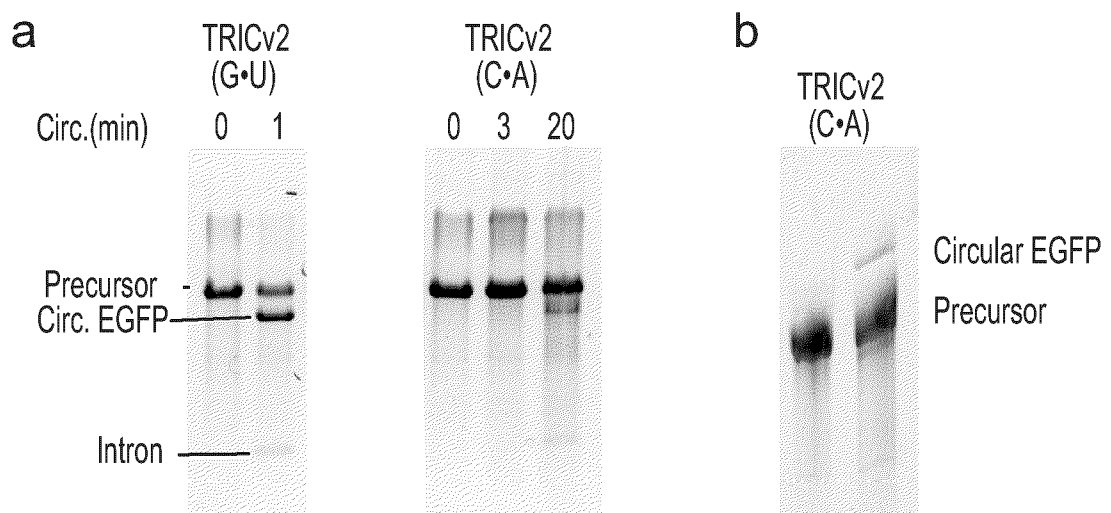
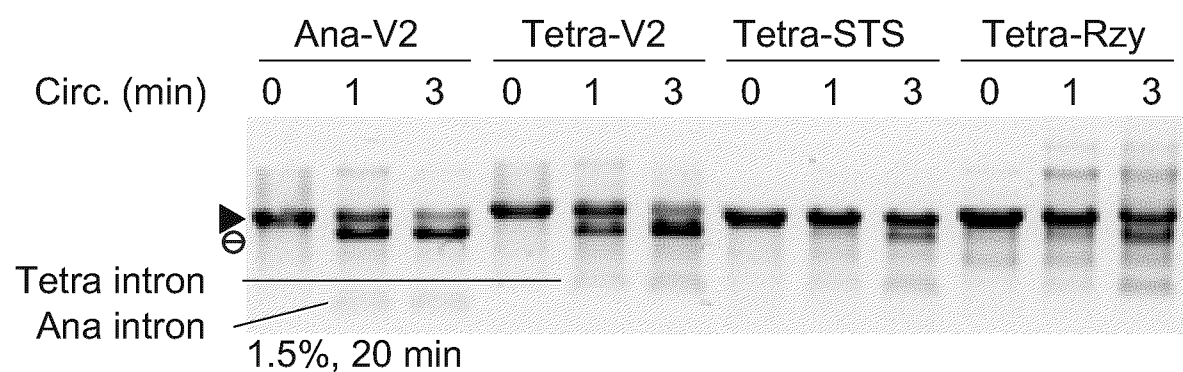


FIG. 17

40/40

a



b

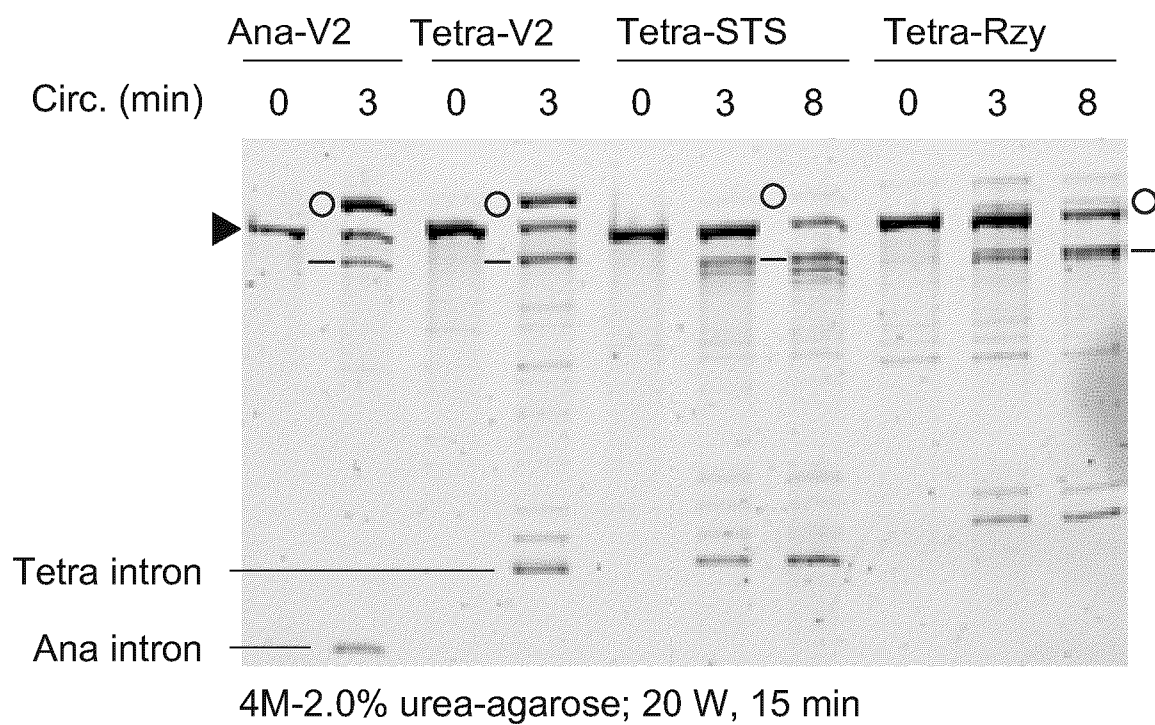


FIG. 18