



- (51) International Patent Classification: Not classified
- (21) International Application Number: PCT/IB2016/001902
- (22) International Filing Date: 14 October 2016 (14.10.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 62/242,884 16 October 2015 (16.10.2015) US
- (71) Applicant: ASTRAZENECA AB [SE/SE]; SE-151 85 Södertälje (SE).
- (72) Inventors: MARESCA, Marcello; SE-151 85 Södertälje (SE). BOHLOOLY-YEGANEH, Mohammad; SE-151 85 Södertälje (SE). JAISWAL, Himjyot; SE-151 85 Södertälje (SE). MAYR, Lorenz, M.; SE-151 85 Södertälje (SE). XU, Xiufeng; SE-151 85 Södertälje (SE).
- (74) Agents: WINTER, Christopher, Spencer et al.; MedImmune Limited, Milstein Building, Granta Park, Cambridge CB21 6GH (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

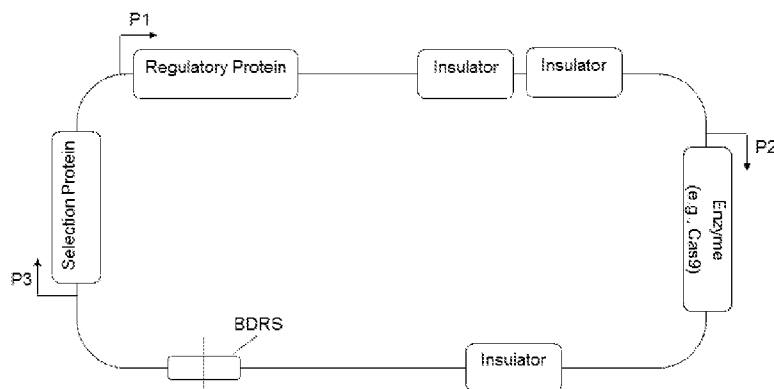
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: INDUCIBLE MODIFICATION OF A CELL GENOME

Fig. 1



(57) Abstract: The present disclosure is directed, in some embodiments, to compositions and methods for inducible modification of a cell genome.

INDUCIBLE MODIFICATION OF A CELL GENOME

RELATED APPLICATION

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional
5 application number 62/242,884, filed October 16, 2015, which is incorporated by reference
herein in its entirety.

BACKGROUND

Genome editing is a type of genetic engineering in which a genome is modified (*e.g.*,
10 DNA is introduced, removed or replaced) using engineered nucleases. Typically, the
nucleases create specific double-stranded break (DSBs) at desired locations in the genome
and harness the cell's endogenous mechanisms to repair the induced break by natural
processes of homologous recombination (HR) and nonhomologous end-joining (NHEJ).
There are currently four families of engineered nucleases being used: zinc finger nucleases
15 (ZFNs), transcription activator-like effector nucleases (TALENs), the CRISPR/Cas system,
and engineered meganuclease re-engineered homing endonucleases.

SUMMARY

Provided herein, in some embodiments, are engineered nucleic acid constructs
20 ("engineered constructs") useful for cross-species integration and introducing into a genome
an inducible genome editing system. In some embodiments, the engineered constructs are
introduced into a genome without also introducing vector (plasmid) material. The engineered
constructs of the present disclosure, in some embodiments, permit spatially-controlled and
temporally-controlled activation of target gene expression following site-specific integration
25 into a genome via the non-homologous end-joining (NHEJ) pathway (*see, e.g.*, Maresca *et al.*
Genome Res. 2013 Mar;23(3):539-46, incorporated herein by reference). Unlike presently-
available systems, which typically coordinate targeted modification of a cell genome through
the use of at least two independent constructs, the genome editing systems of the present
disclosure rely on activation of a single construct that comprises the genetic elements used to
30 express a regulatory protein as well as the inducible genetic elements used to express a target
gene. This single-construct configuration results in tightly regulated and substantially non-
leaky target gene expression, thereby providing more precise and efficient genome editing
capability relative to presently-available systems.

Thus, some embodiments of the present disclosure provide engineered nucleic acid
35 constructs that comprise (a) a promoter operably linked to a nucleic acid encoding a

regulatory protein (*e.g.*, an inducer protein or a repressor protein); (b) an inducible promoter operably linked to a nucleic acid encoding an enzyme that cleaves nucleic acid (*e.g.*, Cas9 nuclease, Cpf1 nuclease, or a functional equivalent thereof), a nucleic acid encoding an enzyme that nicks nucleic acid (*e.g.*, Cas9 nickase), or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid (*e.g.*, Cre recombinase), wherein activity of the inducible promoter is modulated by the regulatory protein; (c) at least two insulators (*e.g.*, mammalian insulators) located downstream from (a) and upstream from (b); at least one insulator located downstream from (b) and upstream from (a); and (e) at least one deoxyribonucleic acid (DNA)-binding domain recognition sequence located downstream from (b) and upstream from (a). In some embodiments, at least two DNA-binding domain recognition sequence located downstream from (b) and upstream from (a). In some embodiments, the inducible promoter is operably linked to a nucleic acid encoding an enzyme that regulates gene expression (*e.g.*, Cas9 fused to KRAB, Cas9 fused to VP64, Cas9 fused to p300), or a nucleic acid encoding an enzyme that modifies a nucleotide base (*e.g.*, Cas9 dead or nickase fused to AID/ApoBEC domains and to an inhibitor of uracil glycosylase).

The present disclosure further provides, in some embodiments, vectors comprising an engineered nucleic acid construct, cells comprising an engineered nucleic acid construct, or cells comprising vectors containing an engineered nucleic acid construct.

Also provided herein, in some embodiments, are methods of modifying a cell genome. For example, methods of the present disclosure may be used to delete (knockout) a gene of interest, introduce (knockin) a gene of interest, or modify a gene of interest.

In some embodiments, an enzyme is a nuclease, a nickase or a recombinase.

In some embodiments, an inducible promoter is a tissue-specific inducible promoter or a developmental-specific inducible promoter.

In some embodiments, the regulatory protein is a tetracycline-controlled transactivator (tTA) protein, a reverse tetracycline-controlled transactivator (rtTA) protein, or a Lac repressor protein.

In some embodiments, the DNA-binding domain recognition sequence is cleaved by a nuclease having a FokI nuclease domain. Non-limiting examples of such “hybrid nucleases” include zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and hybrid Cas9-FokI nucleases.

In some embodiments, a DNA-binding domain recognition sequence is cleaved by a nuclease not having a FokI nuclease domain. For example, Cpf1 nuclease may be used to

cleave a DNA-binding recognition sequence. Other nucleases, similar in structure and function to those nucleases described herein, may be used in the present methods.

Also provided herein is a transgenic mouse comprising in the genome of the mouse an engineered nucleic acid construct (*e.g.*, a TOICas construct) as provided herein. In some
5 embodiments, the engineered nucleic acid construct is integrated in the Rosa26 locus of the mouse genome. It should be understood that while the Rosa26 locus is exemplified in some embodiments, the present disclosure is not limited to genomic integration at the Rosa26 locus. The engineered constructs of the present disclosure may be integrated into any locus in the mouse genome (or the human genome when applicable to the generation of TOIC cell
10 lines, such as TOIC human iPSC lines).

In some embodiments, the mouse is immunocompetent. In some embodiments, expression of the enzyme (*e.g.*, Cas9) is not detectable in the absence of induction of the inducible promoter (*e.g.*, in the absence of Dox administration).

In some embodiments, an engineered nucleic acid of the present disclosure comprises
15 the sequence of SEQ ID NO: 8.

A transgenic mouse (*e.g.*, an immunocompetent mouse) comprising in the genome of the mouse an engineered nucleic acid construct that comprises the sequence of SEQ ID NO: 8 (TOICas9) is also provided herein. In some embodiments, the engineered nucleic acid construct comprising the sequence of SEQ ID NO: 8 is integrated in the Rosa26 locus (or
20 another locus) of the mouse genome.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings are not intended to be drawn to scale. For purposes of clarity, not every component may be labeled in every drawing.

Fig. 1 is a schematic of an example of an engineered nucleic acid construct in
25 accordance with the present disclosure. Promoter P1 is operably linked to a nucleic acid encoding a regulatory protein (*e.g.*, an inducer protein or a repressor protein), which is upstream from two (*e.g.*, at least two) insulators. The insulators are upstream from promoter P2, which is operably linked to a nucleic acid encoding an enzyme that cleaves (*e.g.*, Cas9
30 nuclease, Cpf1 nuclease, or a functional equivalent thereof) or nicks (*e.g.*, Cas9 nickase) nucleic acid, or catalyzes exchange of nucleic acid (*e.g.*, Cre recombinase). The enzyme is upstream from an additional insulator, which is upstream from a (*e.g.*, at least one) deoxyribonucleic acid (DNA)-binding domain recognition sequence (BDRS). The size of the engineered construct may be, for example, at least 15 kilobases (kb).

Fig. 2 is a schematic of an example of an engineered nucleic acid construct of the present disclosure used for insertion in the adeno-associated virus integration site 1 (AAVS1) locus (human) or the ROSA β geo26 (ROSA26) locus (mouse). The size of the engineered construct is 17575 kb.

5 Figs. 3A-3B show images of induced pluripotent stem cells (iPSCs) containing a doxycycline-inducible system of the present disclosure, as depicted in Fig. 3, in the absence of doxycycline (-dox) (Fig. 4A) and in the presence of doxycycline (+dox). Green fluorescent protein (GFP) is an indicator of activation of nuclease (*e.g.*, Cas9 nuclease, Cpf1 nuclease, or a functional equivalent thereof) or recombinase (*e.g.*, Cre) expression. GFP is
10 expressed only in the presence of doxycycline, indicating no “leakage” of the system (Fig. 3B).

Fig. 4 shows an electrophoretic gel image representative of induction at the level of DNA. A Surveyor assay was used to cut a known human gene, Usp14 (note that this concept is applicable to any sequence, *e.g.*, of the human genome). The assay showed that the
15 inducible system is active in iPSCs only when doxycycline is added to the cell culture medium and a guide RNA (gRNA) is present. Different pools or single clones (C6, C25, C41) of cells were analyzed, each containing an inducible Cas9 system of the present disclosure. Cleavage is indicated by the release of additional lower bands in the electrophoresis gel. The same efficiency of cleavage was observed among all the cells
20 analyzed.

Fig. 5 is a schematic of an example of an engineered nucleic acid construct in accordance with the present disclosure. Promoter P1 is operably linked to a nucleic acid encoding an enzyme that cleaves, a nucleic acid encoding an enzyme that nicks nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid, wherein the
25 nucleic acid encoding the enzyme is flanked by a modified version of estrogen receptor (ERT2) sequences. The enzyme is upstream from a (*e.g.*, at least one) deoxyribonucleic acid (DNA)-binding domain recognition sequence (BDRS).

Fig. 6 is a schematic of an example of an engineered construct used to target the MTH1 gene. Promoter P1 is operably linked to a nucleic acid encoding a regulatory protein
30 (*e.g.*, an inducer protein or a repressor protein), which is upstream from two (*e.g.*, at least two) insulators. The insulators are upstream from promoter P2, which is operably linked to a nucleic acid encoding an enzyme that cleaves (*e.g.*, Cas9 nuclease, Cpf1 nuclease, or a functional equivalent thereof), a nucleic acid encoding an enzyme that nicks (*e.g.*, Cas9 nickase) nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of
35 nucleic acid (*e.g.*, Cre recombinase). The enzyme is upstream from a nucleic acid encoding a

guide RNA (gRNA), which is upstream from an additional insulator, which is upstream from a (e.g., at least one) deoxyribonucleic acid (DNA)-binding domain recognition sequence (BDRS). In this example, the gRNA is specific for MTH1 gene, is constitutively expressed and will bind Cas9 to form an active complex upon expression of Cas9. The construct in this
5 examples comprises DNA-binding domain recognition sequences that correspond to sequences located in the AAVS1 locus of the human genome. The construct was integrated in the AAVS1 locus by and recombinant clones were selected. The recombinant clones were assayed for cleavage of the endogenous MTH1 locus upon induction.

Fig. 7 is a graph showing cleavage efficiency as determined by TIDE software. The
10 analysis shows that 85% of the cells contain a deletion (knockout) in the MTH1 gene, while no deletion was observed in absence of induction. The data from this analysis shows that MTH1 is not lethal when combined with a particular mutation present in the cell line tested.

Fig. 8 shows schematics of different engineered nucleic acid constructs, one of which (top panel) expresses GFP upon induction with doxycycline.

15 Fig. 9 shows data confirming integration into the Rosa26 locus and germline transmission of the engineered construct depicted in Fig. 8, top panel.

Fig. 10 shows a schematic of an example of an engineered TOICas construct. Cas9 is fused to GFP by T2A peptide and is flanked by insulators for tight regulation.

20 Fig. 11 is an electrophoretic gel image showing efficient cleavage of DNA in the presence of Dox and guide RNA.

Fig. 12 shows a schematic of a gene target strategy using AAV-U6 gRNA (p53^{-/-}/KrasG12D) construct (top panel). A Kras target site is located in an intron to cause less damage possible to the gene in case of failure of HDR ssODN-mediated point-mutation “repair”. Trp53 CRISPR follows a “common” strategy (bottom panel).

25 Fig. 13 shows a GFP curve followed over time after Dox stimulation (48h) and removal (112h) of cells.

Fig. 14 is an electrophoretic gel image showing results from a surveyor assay on fibroblasts stimulated with Dox and infected with AAV-gRNA-Kras to produce both point mutation and indel.

30 Fig. 15 is an electrophoretic gel image showing results of a surveyor nuclease assay demonstrating the precise cut in Trp53 gene (red asterisks) in the lungs of mice administered intra-tracheal with AAV-gRNA-Trp53 (#37, #42, #52), validating the function of TOICas.

Fig. 16 shows data from a Sanger sequencing analysis demonstrating that TOICas9 (inducible Cas9) cleavage is not observed in the absence of Dox.

Fig. 17 is a graph showing GFP expression and Cas9 expression as a proportion of the parental population and the stable pools.

Fig. 18 is a graph showing the effect of Cas9 expression on acute DNA damage (gH2AX foci formation).

5 Fig. 19 is a graph showing the effect of Cas9 expression on persistent DNA damage (micronuclei formation).

Figs. 20-28 show plasmid maps of examples of TOICas constructs encompassed by the present disclosure. The sequences of Figs. 20-28 correspond respectively to SEQ ID NO: 12-20.

10

DETAILED DESCRIPTION

The present disclosure provides engineered nucleic acid constructs (*e.g.*, “TOICas constructs”) used for cross-species integration and introducing into a genome an inducible genome editing system, in some embodiments, without also introducing vector (plasmid) material. Engineered constructs of the present disclosure facilitate site-specific integration of a linearized form of the construct into a single locus of a genome, into multiple different loci of a genome, or into loci of different genomes (of different species). This direct integration depends on coordinated enzymatic cleavage of the construct and the targeted genomic locus, and ligation of the linearized construct into the genomic locus, for example, via the non-homologous end-joining (NHEJ) pathway (*see, e.g., Maresca et al. Genome Res. 2013 Mar;23(3):539-46, incorporated herein by reference*).

Custom-designed nucleases, such as Cas9, Cpf1, zinc finger nucleases, Tale nucleases, and functional equivalents thereof contain a DNA cleavage domain and a DNA binding domain assembled from optimized DNA binding modules. In cells, these nucleases generate a double-strand break in the genome at or near a sequence recognized by the DNA binding domain of the nuclease (a “DNA-binding domain recognition sequence”) and induce DNA damage repair. The engineered constructs of the present disclosure are based, in part, on results showing that an episomal construct can be ligated into a target genomic locus if the construct and the locus contain the same DNA-binding domain recognition sequence.

30 In some embodiments, multiple DNA-binding domain recognition sequences are positioned in a construct such that they flank unwanted vector (*e.g.*, bacterial plasmid) DNA. This configuration results in the removal of vector DNA upon integration of the construct into a genome.

The engineered constructs (*e.g.*, “TOICas constructs”) of the present disclosure are versatile in that they comprise, in some embodiments, the genetic elements used to induce

35

gene expression in a temporally-controlled and spatially-controlled manner, an array (*e.g.*, two or more) of DNA-binding domain recognition sequences that facilitate site-specific integration of the construct into multiple different loci, and are substantially non-leaky. Thus, the engineered constructs of the present disclosure provide more precise and efficient genome editing capability relative to presently-available genome editing systems.

TOICas Constructs and Transgenic Animals

Engineered constructs are herein referred to, in some embodiments, as TOIC, TOICas or TOICas9 constructs, which include a nucleic acid encoding a Cas9 enzyme. It should be understood that in any of the TOIC, TOICas or TOICas9 constructs, the nucleic acid encoding a Cas9 enzyme may be replaced with another enzyme that cleaves nucleic acid, nicks nucleic acid, catalyzes the exchange of nucleic acid, regulates gene expression, or modifies a nucleotide base. Non-limiting examples of TOIC constructs are depicted in Figs. 20-28, the nucleic acid sequences of which are represented by SEQ ID NO: 12-20, respectively.

TOIC constructs, in some embodiments, include any one of Figs. 20-28, for example, having a sequence of any one of SEQ ID NO: 12-20, respectively. The TOICas constructs of the present disclosure, as discussed in greater detail below, include a promoter (P1) (*e.g.*, CMV, CAGG, CBh, or EF1alpha, or a tissue-specific promoter) operably linked to a regulatory protein, which is upstream of two insulators positioned in tandem (insulator 1 and insulator 2), which are upstream of a promoter (P2) (*e.g.*, an arrayed sequence bound by the regulator protein) operably lined to an enzyme (*e.g.*, Cas9), which is upstream of an insulator (insulator 3), which is upstream of a DNA-binding domain recognition site (BDRS) (*e.g.*, nuclease, a recombinase, or an integrase), which is optionally upstream of a promoter (P3) operably linked to a nucleic acid encoding a selection protein (*e.g.*, a drug selection protein or a fluorescent marker) (*see, e.g.*, Fig. 1 and Fig. 10). In some embodiments, a TOIC construct comprises a nucleic acid encoding a guide RNA (gRNA), which may be located, for example, between the enzyme and insulator 3, or between insulator 3 and the BDRS.

Also provided herein are transgenic animals, such as transgenic mouse models, comprising a TOIC construct. While many embodiments described herein refer to transgenic mouse models, it should be understood that the disclosure covers a variety of transgenic animal models (invertebrates and vertebrates), including, but not limited to: *Amphimedon queenslandica*, *Arbacia punctulata*, *Aplysia*, *Branchiostoma floridae*, *Caenorhabditis elegans*, *Caledia captiva*, *Callosobruchus maculatus*, *Chorthippus parallelus*, *Ciona intestinalis*, *Daphnia spp.*, *Coelopidae*, *Diopsidae*, *Drosophila*, *Euprymna scolopes*, *Galleria*

mellonella, *Gryllus bimaculatus*, *Hydra*, *Loligo pealei*, *Macrostomum lignano*, *Mnemiopsis leidyi*, *Nematostella vectensis*, *Oikopleura dioica*, *Oscarella carmela*, *Parhyale hawaiiensis*, *Platynereis dumerilii*, *Podisma* spp., *Pristionchus pacificus*, *Scathophaga stercoraria*, *Schmidtea mediterranea*, *Stomatogastric*, *Strongylocentrotus purpuratus*, *Symsagittifera roscoffensis*, *Tribolium castaneum*, and *Trichoplax adhaerens*, *Tubifex tubifex* (invertebrates); and *Bombina*, Carolina anole (*Anolis carolinensis*), Cat (*Felis sylvestris catus*), Chicken (*Gallus gallus domesticus*) - Cotton rat (*Sigmodon hispidus*), Dog (*Canis lupus familiaris*), Golden hamster (*Mesocricetus auratus*), Guinea pig (*Cavia porcellus*), Little brown bat (*Myotis lucifugus*), Medaka (*Oryzias latipes*, or Japanese ricefish), Mouse (*Mus musculus*),

10 Naked mole-rat (*Heterocephalus glaber*), *Nothobranchius furzeri*, Pigeon (*Columba livia domestica*), *Poecilia reticulata*, Rat (*Rattus norvegicus*), Rhesus macaque (or rhesus monkey) (*Macaca mulatta*), Sea lamprey (*Petromyzon marinus*), Takifugu (*Takifugu rubripes*, a pufferfish), Three-spined stickleback (*Gasterosteus aculeatus*), *Xenopus tropicalis* and *Xenopus laevis* (African clawed frog), Zebra finch (*Taeniopygia guttata*), and Zebrafish

15 (*Danio rerio*, a freshwater fish).

In some embodiments, the transgenic animal model is selected from a fish, a frog, a bird, a mouse, a rat, a hamster, a cat, a dog, a pig, a sheep and a monkey. In some embodiments, the transgenic animal model is a mouse model. In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 8.

20 Also provided herein is a transgenic animal (e.g., mouse), for example, an immunocompetent animal (e.g., mouse), comprising in the genome of the animal (e.g., mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 8 integrated, for example, in the Rosa26 locus of the animal (e.g., mouse) genome.

In some embodiments, provided herein is an engineered nucleic acid comprising the

25 sequence of SEQ ID NO: 12. Also provided herein is a transgenic animal (e.g., mouse), for example, an immunocompetent animal (e.g., mouse), comprising in the genome of the animal (e.g., mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 12 integrated, for example, in the Rosa26 locus of the animal (e.g., mouse) genome.

In some embodiments, provided herein is an engineered nucleic acid comprising the

30 sequence of SEQ ID NO: 13. Also provided herein is a transgenic animal (e.g., mouse), for example, an immunocompetent animal (e.g., mouse), comprising in the genome of the animal (e.g., mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 13 integrated, for example, in the Rosa26 locus of the animal (e.g., mouse) genome.

In some embodiments, provided herein is an engineered nucleic acid comprising the

35 sequence of SEQ ID NO: 14. Also provided herein is a transgenic animal (e.g., mouse), for

example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 14 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

5 In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 15. Also provided herein is a transgenic animal (*e.g.*, mouse), for example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 15 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

10 In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 16. Also provided herein is a transgenic animal (*e.g.*, mouse), for example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 16 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

15 In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 17. Also provided herein is a transgenic animal (*e.g.*, mouse), for example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 17 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

20 In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 18. Also provided herein is a transgenic animal (*e.g.*, mouse), for example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 18 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

25 In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 19. Also provided herein is a transgenic animal (*e.g.*, mouse), for example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 19 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

30 In some embodiments, provided herein is an engineered nucleic acid comprising the sequence of SEQ ID NO: 20. Also provided herein is a transgenic animal (*e.g.*, mouse), for example, an immunocompetent animal (*e.g.*, mouse), comprising in the genome of the animal (*e.g.*, mouse) an engineered nucleic acid construct comprising the sequence of SEQ ID NO: 29 integrated, for example, in the Rosa26 locus of the animal (*e.g.*, mouse) genome.

35 Induction of enzyme (*e.g.*, Cas9) expression in an animal model may be achieved by administering doxycycline or other appropriate induction agent (depending on the particular

induction system used in the TOIC construct). In some embodiments, the induction agent (agent that directly or indirectly activates the inducible promoter of the TOIC construct) is administered to an animal via injection (*e.g.*, tail vein injection) or oral gavage.

Transgenic animals, as provided herein, may be used to generate knockout or knockdown alleles, or to overexpress a gene or knock a gene into a particular loci, by homologous recombination or by non-homologous end joining. This may be achieved, for example, by administering to the animal a template DNA (*e.g.*, containing a modification of interest) and a nucleic acid encoding a gRNA targeting a loci and/or gene of interest.

Examples of genes of interest include, but are not limited to, oncogenes such as *Pik3ca*, *Kras*, *Braf*, *Nras*, and tumor suppressor genes such as *Pten*, *p53*, *Rb*, *Apc*, *p16/p19*, *Brca1*, *Brca2*, *Lkb1*. Various disease models may be produced by combining TOIC transgenic mice with template DNA and gRNA targeting a gene or genes of interest. Such models include, but are not limited to, lung cancer (*e.g.*, *Kras*, *Lkb*, *p53* and/or *Rb*; *e.g.*, targeting *MAPK*, metabolism), pancreatic cancer (*e.g.*, *Kras*, *p53*, *p16/p19* and/or *Pdx1*; *e.g.*, targeting *MAPK*), prostate cancer (*e.g.*, *Pten*, *Brca1*, *Brca2* and/or *p53*; *e.g.*, targeting *PI3K*, *AR*, *ASO*, *DDR*), breast cancer (*e.g.*, *Pik2ca*, *p53* and/or *Pten*; *e.g.*, targeting *PI3K*, *SERD*), ovarian cancer (*e.g.*, *Brca1*, *Brca2*, *p53* and/or *Rb*; *e.g.*, targeting *Erk*, *MEK*, *Kras*, *ASO*, *modp53*), melanoma (*e.g.*, *Braf* and/or *Nras*; *e.g.*, targeting *MAPK*) and colorectal cancer (*e.g.*, *Pik3ca*, *Kras* and/or *Apc*; *e.g.*, targeting *MAPK*, *PI3K*).

A nucleic acid encoding a gRNA and associated template DNA may be administered to an animal via intratracheal, intravenous, or intraperitoneal transduction using a virus (*e.g.*, adeno-associated virus or adenovirus), for example.

In some embodiments, a nucleic acid encoding a gRNA is integrated into the genome of the transgenic animal model. For example, a tissue-specific gRNA or a constitutively-expressed gRNA may be integrated into the genome of the transgenic animal model.

Also provided herein are organoids (three-dimensional organ-bud grown *in vitro*) derived from (obtained from) transgenic animals of the present disclosure. Thus, an organoid may comprise any of the TOIC constructs.

Advantageously, TOIC animals (*e.g.*, mice), for example, those described in the Examples, may be immunocompetent (able to produce a normal immune response following exposure to an antigen). Also encompassed herein, however, are immunocompromised (have a weakened immune system) TOIC animals.

Enzymes for Genomic Integration

Engineered constructs (*e.g.*, “TOICas constructs”) of the present disclosure are used to facilitate direct, site-specific ligation of a linearized form of the construct into a single locus or multiple different loci of a single genome or multiple different genomes. This direct ligation occurs through the non-homologous end-joining (NHEJ) pathway (*see, e.g.*, Maresca *et al. Genome Res.* 2013 Mar;23(3):539-46, incorporated herein by reference). Site-specific integration depends on the presence of hybrid nucleases that contain a DNA binding domain and a DNA cleavage domain (typically a FokI domain) and the presence of nucleic acids that contain at least one DNA-binding domain recognition sequence. A “DNA-binding domain recognition sequence” is a nucleotide sequence to which a nuclease DNA-binding domain binds and a nuclease DNA cleavage domain cleaves. Engineered constructs contain at least one DNA-binding domain recognition sequence that is recognized and cleaved by a hybrid nuclease. Cleavage of the engineered construct results in a linearized form, which can then be “ligated” into a genome in a site-specific manner.

Engineered constructs (*e.g.*, “TOICas constructs”) of the present disclosure, in some embodiments, comprise a single DNA-binding domain recognition sequence (BDRS) or an array (*e.g.*, two or more) of DNA-binding domain recognition sequences, which facilitate site-specific genomic integration of the nucleic acid. Advantageously, an engineered construct of the present disclosure can be used to facilitate site-specific ligation of a linearized form of the construct into multiple different loci of several different genomes, which is useful for cross-species integration of the same construct.

In some embodiments, a DNA-binding domain recognition sequence of an engineered construct corresponds to a sequence located in the Rosa26 locus such that the nucleic acid may be integrated in a mouse genome. In some embodiments, a DNA-binding domain recognition sequence of an engineered construct corresponds to a sequence located in the AAVS1 locus such that the nucleic acid may be integrated in a human genome. Other DNA-binding domain recognition sequence located in other genomic loci are encompassed by the present disclosure.

Examples of hybrid nuclease for use in linearizing an engineered construct include, without limitation, zinc finger nucleases (ZFNs), Tale nucleases (TALENs), dCas9-FokI fusion proteins (catalytically inactive Cas9 fused to FokI), Cas9, Cas9 nickase fused to FokI, and Cas9 variants evolved to generate overhangs.

In some embodiments, the DNA-binding domain recognition sequence is a ZFN DNA binding domain recognition sequence, which is bound by one or more zinc finger(s). The DNA-binding domain of individual ZFNs may contain between three and six individual zinc

finger repeats and can each recognize between 9 and 18 base pairs. If the zinc finger domains are specific for their intended target site, then even a pair of 3-finger ZFNs that recognize a total of 18 base pairs can target a single locus in a mammalian genome.

In some embodiments, the DNA-binding domain recognition sequence is a TALEN
5 DNA binding domain recognition sequence, which is bound by one or more TAL effector unit(s). TAL effectors are proteins secreted by *Xanthomonas* bacteria. The DNA binding domain typically contains a repeated highly conserved 33–34 amino acid sequence with the exception of the 12th and 13th amino acids. These two locations are highly variable (Repeat Variable Di-residue, RVD) and show a strong correlation with specific nucleotide recognition
10 (Boch *et al. Science* 326 (5959): 1509–12, 2009; Moscou *et al. Science* 326 (5959): 1501, 2009, each of which is incorporated by reference herein). In some embodiments, specific DNA-binding domains are engineered by selecting a combination of repeat segments containing the appropriate RVDs (Boch *et al. Nature Biotechnology* 29 (2): 135–6, 2011).

In some embodiments, the DNA-binding domain recognition sequence is a sequence
15 complementary (*e.g.*, 100% complementary) to two co-expressed guide RNAs. In some embodiments, the DNA-binding domain recognition sequence is a sequence that is at least 80%, at least 85%, at least 90%, at least 95% or at least 98% complementary to two co-expressed guide RNAs. In such embodiments, a catalytically inactive Cas9 (dCas9) fused to FokI nuclease may be used to generate double strand breaks in an engineered nucleic acid.

It should be understood that the engineered constructs of the present disclosure may
20 comprise a nucleic acid encoding a nuclease (*e.g.*, Cas9 nuclease, Cpf1 nuclease, or a functional equivalent thereof) and contain a DNA-binding nuclease recognition sequence that is not necessarily recognized by the nuclease of the engineered nucleic acid. For example, an engineered construct for use in genomic editing may encode Cas9 (*e.g.*, wild-type or
25 otherwise catalytically active Cas9 for the purpose of editing the genome of a cell) or Cpf1 nuclease and may also contain a DNA-binding nuclease recognition sequence specific for a zinc finger nuclease or a catalytically inactive Cas9 (dCas) fused to FokI nuclease. If this is the case, it may be necessary to introduce into a cell another nucleic acid encoding the zinc finger nuclease or the dCas9 fused to FokI that specifically recognizes and cleaves the DNA-
30 binding nuclease recognition.

Enzymes for Genomic Editing

Also described herein are nucleic acids that encode enzymes that cleave nucleic acid, nick nucleic acid, or catalyze exchange of nucleic acid. Enzymes that cleave nucleic acids

are referred to as nucleases. Enzymes that nick nucleic acids are referred to as nickases. Enzymes that catalyze exchange of nucleic acid are referred to as recombinases.

“Cleavage” refers to the process by which a nuclease cuts (hydrolyzes) each nucleic acid backbone (e.g., sugar-phosphate backbone) of a double-stranded nucleic acid. Thus, the nuclease makes two incisions: one in the backbone between the nucleotide subunits of one strand of the double-stranded nucleic acid, and another in the backbone between the nucleotide subunits of the other strand of the double-stranded nucleic acid. Cleavage of a single nucleic acid molecule typically results in the production of two separate nucleic acid molecules. “Nicking,” by contrast, refers to the process by which a nickase cuts only one strand of a double-stranded nucleic acid. “Catalyzing exchange of nucleic acid” refers to the process by which genetic material is broken and joined to other genetic material and encompasses genetic recombination. Recombination is recombining or rearranging genetic material, for example, by crossing over in chromosomes or by joining segments of DNA.

In some embodiments, an engineered nucleic acid encodes a nuclease. Nucleases of the present disclose may be engineered to cut a pre-determined nucleotide sequence, permitting, for example, efficient engineering of genetic information and the creation of a variety of diverse nucleic acid modifications. Examples of engineered nucleases include, without limitation, DNA-guided endonucleases, RNA-guided endonucleases (RGENs) such as Cas9 or Cpf1, zinc finger nucleases (ZFNs) (Kim *et al. Proc Natl Acad Sci USA* 93 (3): 1156–60, 1996; Bitinaite *et al. Proc Natl Acad Sci USA* 95 (18): 10570–5, 1998; and Cathomen *et al. Mol. Ther.* 16 (7): 1200–7, 2008), TAL effector nucleases (TALENs, transcription activator-like effector nucleases) (Boch *et al. Science* 326 (5959): 1509–12, 2009; Christian *et al. Genetics* 186 (2): 757–61, 2010); and Miller *et al. Nature Biotechnology* 29 (2): 143–8, 2011) (Table 1), and functional equivalents thereof. Nucleases typically comprise a DNA binding domain, which recognizes and binds to a particular DNA sequence, and a DNA cleavage domain, which cleaves the DNA at or near (e.g., within 10 nucleotides of) the DNA binding domain. For example, ZFNs comprise zinc finger domains, which bind DNA, and a Fok I domain, which cleaves the DNA (Kim *et al. Natl Acad Sci USA* 93 (3): 1156–60, 1996). Similarly, TALENs comprise TAL effector units, which bind DNA, and a Fok I domain, which cleaves DNA. The RNA-guided Cas9 nuclease cleaves the DNA, but to do so, it must first be guided to the target cleavage site by a guide RNA, which is complementary to and binds to the DNA cleavage site, as described elsewhere herein.

Table 1. Examples of Engineered Nuclease Systems

Nuclease	DNA binding domain	DNA cleavage domain
RGEN	Guide RNA that hybridizes to the target DNA (1:1 nucleotide base pairing)	Cas9 protein (contains two nuclease domains); or or Cpf1 protein
ZFN	Zinc fingers (each module recognizes 3 bp of target sequence)	Fok I restriction enzyme nuclease domain (requires dimerization for cleavage)
TALEN	TAL effector units (each module recognizes 1 bp of target sequence)	Fok I restriction enzyme nuclease domain (requires dimerization for cleavage)

Cas9 (CRISPR associated protein 9) is an RNA-guided DNA nuclease associated with the CRISPR (Clustered Regularly Interspersed Palindromic Repeats) adaptive immunity system in *Streptococcus pyogenes*, among other bacteria. CRISPR systems for editing, regulating and targeting genomes may comprise at least two distinct components: (1) a guide RNA (gRNA) and (2) Cas9. A gRNA is a single chimeric transcript that combines the targeting specificity of endogenous bacterial CRISPR targeting RNA (crRNA) with the scaffolding properties of trans-activating crRNA (tracrRNA). Typically, a gRNA used for genome editing is transcribed from either a plasmid or a genomic locus within a cell. The gRNA transcript forms a complex with Cas9, and then the gRNA/Cas9 complex is recruited to a target sequence as a result of the base-pairing between the crRNA sequence and its complementary target sequence in genomic DNA, for example.

In a typical synthetic CRISPR/Cas9 genome editing system, a genomic sequence of interest (genomic target sequence) is modified by use of a gRNA complementary to the sequence of interest, which directs the gRNA/Cas9 complex to the target (Sander JD *et al.*, 2014 *Nature Biotechnology* 32, 247-355, incorporated by reference herein). The Cas9 endonuclease cuts the genomic target DNA upstream of a protospacer adjacent motif (PAM), resulting in double-strand breaks. Repair of the double-strand breaks often results in inserts or deletions at the double-strand break site. This CRISPR/Cas9 system is often used to edit the genome of a cell, each iteration requiring the design and introduction of a new gRNA sequence specific to a target sequence of interest.

In some embodiments, an engineered construct of the present disclosure comprises a promoter (*e.g.*, an inducible promoter) operably linked to a nucleic acid encoding a guide RNA (*e.g.*, downstream from a nucleic acid encoding a Cas9 nuclease), which guides the Cas9 nuclease to a genomic target (modification) site. Enzymes that are functionally similar to Cas9 may be used in accordance with the present disclosure.

Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system (Zetsche *et al.*, 2015, *Cell* 163: 1-13, incorporated by reference herein). Cpf1, like Cas9, is a two-component RNA programmable DNA nuclease. Targeted DNA is cleaved as a 5-nt staggered cut distal to a 5' T-rich protospacer adjacent motif (PAM). There are two Cpf1
5 orthologs that exhibit robust nuclease activity in human cells, either of which may be used as provided herein. Enzymes that are functionally similar to Cpf1 may be used in accordance with the present disclosure.

In other embodiments, a nucleic acid encoding a guide RNA is introduced into a host cell that is modified to express (*e.g.*, stably express) in the cell genome an engineered
10 construct of the present disclosure (*e.g.*, a construct encoding a Cas9 nuclease) and is used to replace any unwanted DNA introduced into the host cell during modification of the host cell genome. For example, a cell, such as a stem cell (*e.g.*, a pluripotent stem cells), may be modified to express (*e.g.*, stably express) in the cell genome an engineered construct comprising (a) a promoter operably linked to a nucleic acid encoding a regulatory protein, (b)
15 an inducible promoter operably linked to a nucleic acid encoding Cas9 or Cpf1, wherein activity of the inducible promoter is modulated by the regulatory protein, (c) at least two insulators located downstream from (a) and upstream from (b), and (d) at least one insulator located downstream from (b). The modification of this host cell may result in the introduction of vector (*e.g.*, bacterial plasmid) DNA in the genome of the cell as well as other
20 unwanted DNA (*e.g.*, sequence encoding a selection marker) following site-specific integration of the construct. This vector DNA and any other unwanted DNA can be removed, for example, by introducing (a) an engineered nucleic acid encoding a guide RNA flanked by DNA-binding domain recognition sequences, (b) an engineered nucleic acid encoding a hybrid nuclease that recognizes and cleaves DNA-binding domain recognition sequences
25 flanking the guide RNA as well as DNA-binding domain recognition sequences flanking unwanted sequence located in the genome of the cells, and (c) and an orthogonal Cas9 (Cas9 obtained from a species different than the species from which the host cell Cas9 was obtained). For example, the host cell may be engineered to express Cas9 obtained from *Streptococcus pyogenes* and the orthogonal Cas9 obtained from *Streptococcus aureus*,
30 *Streptococcus thermophilis* or *Neisseria meningitidis*. Other orthogonal Cas9 nucleases are encompassed by the present disclosure. The hybrid nuclease, the guide RNA and the orthogonal Cas9 may be included on the same construct (*e.g.*, vector) or each on a separate construct. The guide RNA replaces the vector DNA or any unwanted DNA in the genome of the host cell.

In some embodiments, an orthogonal Cas9 is used to replace bacterial plasmid sequence integrated in a host cell genome with a guide RNA of interest and, optionally, a selection marker. In some embodiments, the host cell is a pluripotent stem cell (*e.g.*, a human pluripotent stem cell, such as a human induced pluripotent stem cell) or an embryonic stem cell (*e.g.*, a mouse embryonic stem cell used for the generation of a mouse model system). In these cell types, bacterial plasmid sequence has the potential of inactivating a locus of interest, thus removal of the bacterial plasmid sequence is preferred.

In some embodiments, an orthogonal Cas9 or Cpf1 and a guide RNA are used to delete expression (knockout), reduce expression (knockdown) or increase expression (overexpress) a gene of interest in a differentiated state following a pluripotent state.

In some embodiments vector DNA or any unwanted DNA in the genome of the host cell is removed by using a site-specific recombinase (*e.g.*, Cre, FLP, Dre, Vike or a combination thereof).

The process of introducing an engineered nucleic acid construct into the genome of a cell and removing vector DNA or any unwanted DNA may be performed in a single step (*e.g.*, all constructs are delivered to the host cell(s) simultaneously) or in multiple steps (*e.g.*, each construct is delivered to the host cell(s) sequentially).

The present disclosure also includes the use of catalytically inactive forms of any of the nucleases described herein. For example, in some embodiments, a catalytically inactive form of Cas9 (dCas9) or a catalytically inactive form of Cpf1, which can knockdown gene expression by interfering with transcription, may be used as provided herein. In some embodiments, a dCas9 (or catalytically inactive form of Cpf1 or other nuclease) is fused to a repressor peptide (a peptide that represses transcription, *e.g.*, Cas9-KRAB (Urrutia 2003 *Genome Biol.* 4(10): 231)). In some embodiments, a dCas9 (or catalytically inactive form of Cpf1 or other nuclease) is fused to an activator peptide (a peptide that activates or increase transcription, *e.g.*, Cas9-VP64 (Beerli *et al.* 1998 *Proc Natl Acad Sci USA.* 95(25):14628-33)). In some embodiments, a dCas9 (or catalytically inactive form of Cpf1 or other nuclease) is fused to an epigenomic regulator (*e.g.*, Cas9-DNMT or Cas9-p500). In some embodiments, a dCas9 (or catalytically inactive form of Cpf1 or other nuclease) is fused to FokI nuclease to generate double strand breaks at sequences homologous to two co-expressed gRNAs.

In some embodiments, an engineered nucleic acid expressing dCas9 (or other catalytically inactive nuclease) is used to image specific sequences in the genome (*see, e.g.*, Chen B., *et al.* (2013) *Cell* 155(7): 1479-1491, incorporated herein by reference).

In some embodiments, wild-type or unmodified Cas9 or Cpf1 (or other catalytically inactive nuclease) fused to a repressor peptide (*e.g.*, Cas9-KRAB) or an activator peptide (*e.g.*, Cas9-VP64) is used in combination with a short gRNAs to regulate gene expression (*see, e.g.*, Kiani S, *et al.* (2015) "Cas9 gRNA engineering for genome editing, activation and repression." *Nature Methods* (2015) [epub ahead of print], incorporated herein by reference).

In some embodiments, an engineered nucleic acid encodes a nickase. A nickase is an enzyme that generates a single-strand break in a double-stranded nucleic acid. In some embodiments, the nickase is Cas9 nickase (Cong *et al.* 2013 *Science* 339(6121): 819-823; Shen *et al.* 2014 *Nature Methods* 11, 399–402). Cas9 nickase generates a single-strand DNA break (nick) at a specific location based on a co-expressed gRNA-defined target sequence, rather than a double-strand DNA break (cut) produced by the wild type enzyme. Nicks are preferentially repaired in a cell by homology directed repair (HDR), using the intact strand as the template. HDR has high fidelity and rarely results in errors. Two adjacent, opposite strand nicks can cause a double strand break (DSB) and trigger error-prone non-homologous end joining (NHEJ) repair; however, in the presence of a repair template, the double nicks can be repaired by HDR. Double nicking typically reduces unwanted off-target effects.

In some embodiments, an engineered nucleic acid encodes a recombinase. Recombinases, typically derived from bacteria and fungi, catalyze directionally sensitive DNA exchange reactions between short (*e.g.*, 30–40 nucleotides) target sequences that are specific to each recombinase. These reactions enable four basic functions—excision/insertion, inversion, translocation and cassette exchange—which may be used individually or in combination. Examples of recombinases for use as provided herein include, without limitation, Cre recombinase, FLP recombinase, Hin recombinase (Dhar *et al.* 2004 *Cell* 119 (1): 33–45; Sanders *et al.* 2004 *Mol Biol* 340 (4): 753–66; Kamtekar *et al.* 2006 *Proc Natl Acad Sci USA* 103 (28): 10642–7; Li *et al.* 2005 *Science* 309 (5738): 1210–5) and Tre recombinase (Sarkar *et al.* 2007 *Science* 316 (5833): 1912–15).

Cre recombinase is a tyrosine recombinase enzyme derived from the P1 bacteriophage. The enzyme uses a topoisomerase I-like mechanism to carry out site-specific recombination. The enzyme (*e.g.*, 38kDa) is a member of the integrase family of site-specific recombinase and catalyzes site-specific recombination between two DNA recognition sites (loxP sites). The ~34 base pair (bp) loxP recognition site contains two ~13 bp palindromic sequences that flank an ~8 bp spacer region. The products of Cre-mediated recombination at loxP sites are dependent upon the location and relative orientation of the loxP sites. Two separate DNA species both containing loxP sites can undergo fusion as the result of Cre mediated recombination. DNA sequences found between two loxP sites are said to be

"floxed". The products of Cre mediated recombination depends upon the orientation of the loxP sites. DNA found between two loxP sites oriented in the same direction are excised as a circular loop of DNA, while intervening DNA between two loxP sites that are oppositely orientated are inverted. Cre recombinase requires no additional cofactors (such as ATP) or accessory proteins for its function.

Flp-*FRT* recombination is a site-directed recombination technology analogous to Cre-*lox* recombination. Flp-*FRT* recombination involves the recombination of sequences between short flippase recognition target (FRT) sites by the recombinase (Flp) derived from the 2 μ m plasmid of baker's yeast *Saccharomyces cerevisiae* (Zhu et al. 1995 *Journal of Biological Chemistry* 270 (39): 23044–54; Schlake et al. 1994 *Biochemistry* 33 (43): 12746–12751; and Turan et al 2010 *J. Mol. Biol.* 402 (1): 52–69).

In some embodiments, an engineered nucleic acid encodes a base editing enzyme. A base editing enzyme is a fusion of a DNA binding protein (e.g., Cas9, TALE, ZF) to a specific effector that induces base exchange (e.g., C to T) in the proximity of a DNA binding site (see, e.g., Komor, A.C., et al. *Nature*, 2016; and Nishida, K., et al. *Science* 353: 6305, 2016).

Control of Gene Expression

Engineered constructs (e.g., "TOICas constructs") of the present disclosure permit spatial control of genomic editing, temporal control of genomic editing, of a combination of spatial and temporal control. Spatial control generally refers to the activation of transcription within specific tissues of an organism. Temporal control generally refers to the activation of transcription at specific times during development. Spatial control and/or temporal control may result from use of a cell-specific or tissue-specific promoter driving nucleic acid expression, from control over the time during which an effector substance is delivered to a cell or organism to induce or repress nucleic acid expression, or from a combination thereof, for example. In some embodiments, a cell-specific or tissue-specific promoter drives expression of a nucleic acid to which it is operably linked only during a particular phase of cell specification or cell differentiation. In some embodiments, an effector substance (e.g., Dox) is delivered to a cell or organism only during a particular phase of cell specification or cell differentiation.

Further, the engineered constructs (e.g., "TOICas constructs") of the present disclosure substantially reduce leaky gene expression. A gene expression system may be considered "leaky" if gene transcription is initiated in the absence of a regulatory protein or in an uncontrolled manner. As described elsewhere herein, regulatory proteins bind to

promoters to regulate transcriptional activity. The expression of a gene is considered “leaky” if expression occurs in the absence of the regulatory protein intended to bind the promoter that controls expression of the gene. Expression of a gene is considered “substantially non-leaky” if the level of gene expression in the absence of the regulatory protein is less than 15% (e.g., less than 10%, less than 5%, less than 2%, less than 1%, less than 0.5%) of the level of gene expression in the presence of the regulatory protein. As depicted in Figs. 1 and 2, for example, at least one insulator (e.g., one, two or more insulators) is positioned between the nucleic acid encoding the regulatory protein and the downstream promoter (P2) (to which the regulatory protein binds) controlling expression of the enzyme (e.g., nuclease or recombinase). An “insulator” is a nucleotide sequence that blocks the interaction between enhancers and promoters. It should be understood that insulator(s) may be positioned, in some embodiments, between any two promoters driving gene expression so as to prevent transcriptional activation of the downstream promoter upon transcriptional activation of the upstream promoter.

Thus, engineered constructs (e.g., “TOICas constructs”), in some embodiments, comprise at least one insulator. In some embodiments, an engineered construct comprises at least 2, at least 3 or at least 4 insulators. In some embodiments, an insulator comprises a mammalian insulator. For example, the insulator may comprise a (at least one) human insulator, such as 5’HS5, DMD/ICR, BEAD-1, apoB (-57 kb), apoB (+43 kb), or DM1 site 1 or DM1 site 2 (Table 2). In some embodiments, the insulator may comprise a (at least one) *Mus musculus* insulator, such as BEAD-1, HS2-6 or DMR/ICR. See, e.g., Bell *et al.*, Curr Opin Genet Dev. 1999 Apr;9(2):191-8; Science. 2001 Jan 19;291(5503):447-50; West *et al.*, Genes Dev. 2002 Feb 1;16(3):271-88; and Ziebarth *et al.*, Nucleic Acids Research. 2013; 41(D1):D188-D194, each of which is incorporated herein by reference.

In some embodiments, an insulator comprises a (at least one) non-mammalian insulator. For example, the insulator may comprise a (at least one) *Drosophila melanogaster* insulator, such as *scs/scs*’, *gypsy*, *Fab-7*, *Fab-8*, *fa^{swb}* or the *eve* promoter. In some embodiments, the insulator may comprise a *Saccharomyces cerevisiae* insulator, including *HMR tRNA^{Thr}*, *Chal* UAS, UAS_{tpg} or *STAR*. In some embodiments, the insulator may comprise a (at least one) *Gallus gallus* insulator, such as *Lys 5’A*, *HS4*, or *3’HS*. In some embodiments, the insulator may comprise *sns*, a *Paracentrotus lividus* insulator, *URI*, a *Hemicentrotus pulcherrimus* insulator, or *RO*, a *Xenopus laevis* insulator.

Table 2

5'HS5 (<i>Homo sapiens</i>)	CATCTTGGACCATTAGCTCCACAGGTATCTTCTTC CCTCTAGTGGTCATAACAGCAGCTTCAGCTACCTC TC (SEQ ID NO: 1)
apoB (-57 kb) (<i>Homo sapiens</i>)	CAAATTATCCTGCCCCCTAGACATAACCTCCC (SEQ ID NO: 2)
BEAD-1 (<i>Homo sapiens</i>)	TGCATTGGCTGCCCAGGCCTGCACTGCCGCTGCC GGCAGGGGTCCAGTCCACGAGACCCAGCTCCCTG C (SEQ ID NO: 3)
DM1 site 1 (<i>Homo sapiens</i>)	GCCGGCCGCGGACCCGGCCCCCTCCCTCCCCGGCC GCTAGGGGGCGGGCCCGGATCACAGGA (SEQ ID NO: 4)
DM1 site 2 (<i>Homo sapiens</i>)	CATGCACAAGAAAGCTTTGCACTTTGCGAACCAA CGATAGGTGGGGGTGCGTGGAGGATGG (SEQ ID NO: 5)

“Downstream” and “upstream” refer to the relative position of nucleic acid (*e.g.*, DNA or RNA). Each strand of DNA or RNA has a 5' end and a 3' end, so named for the carbon position on the deoxyribose (or ribose) ring. By convention, upstream and downstream relate to the 5' to 3' direction in which RNA transcription takes place. Upstream is toward the 5' end of the RNA molecule and downstream is toward the 3' end. When considering double-stranded DNA, upstream is toward the 5' end of the coding strand for the gene of interest and downstream is toward the 3' end.

Inducible Gene Editing Systems

Inducible gene editing systems are useful for temporally-controlled, spatially-controlled, and both temporally-controlled and spatially-controlled modification of genes. Some embodiments utilize a Tet-Off or Tet-On inducible system. Other embodiments utilize a tamoxifen-inducible system. Yet other embodiments utilize an isopropyl β -D-1-thiogalactopyranoside (IPTG)-inducible System.

Tetracycline-Inducible System

Tet technology comprises two complementary control circuits, initially described as the tTA dependent (Gossen *et al. Proc Natl Acad Sci USA*. 1992 Jun 15;89(12):5547-51) and rtTA dependent (Gossen *et al. Science*. 1995 Jun 23;268(5218):1766-9) expression systems. They are now commonly referred to as the Tet-Off system (tTA dependent) and the Tet-On system (rtTA dependent). In each system, a recombinant tetracycline controlled transcription factor (tTA or rtTA) interacts with a tTA/rtTA responsive promoter, Ptet, to drive expression of the gene of interest. Expression is regulated by the effector substance tetracycline (Tc) or one of its derivatives. Tet-On systems respond to doxycycline (Dox). Tetracyclines act at

the level of DNA binding of tetracycline-controlled transactivator (tTA) and reverse tetracycline-controlled transactivator (rtTA) transcription factors. rtTA requires a tetracycline ligand for DNA binding and transcription. By contrast, the interaction between tTA and DNA is prevented by tetracycline. Thus, the two versions of the Tet system respond to tetracyclines differently and may be used in a complementary manner.

tTA is a hybrid transcription factor resulting from the fusion of the prokaryotic Tet repressor, TetR, with a eukaryotic transcriptional transactivation domain (*e.g.*, HSV VP16). The TetR moiety confers sequence specific DNA binding, sensitivity to tetracyclines and dimerization to the tTA fusion protein. Accordingly, the response of both TetR and tTA to tetracyclines is similar: binding of the antibiotic lowers their affinity to their common cognate binding sites, the tet operators.

rtTA differs from tTA by a few point mutations within TetR. These, however, result in a complete reversal of tetracycline responsiveness of this transcription factor. rtTA requires tetracyclines for binding to tetO. Specific tetracycline derivatives such as doxycycline (Dox) or anhydrotetracycline (ATc) may be used to exploit the rtTA phenotype.

Ptet is a synthetic promoter responsive to both tTA and rtTA. It is comprised of a minimal RNA polymerase II promoter (transcriptionally silent in the absence of additional transcription factor binding sites) fused to multimerized tetO sequences. This arrangement makes the activity of Ptet dependent on the binding of tTA or rtTA. The design of such synthetic tTA/rtTA responsive promoters is flexible with respect to both the origin of the minimal promoter as well as the exact arrangement of the operators. The original version, for example, which consists of a CMV minimal promoter fused to an array of seven tetO sequences is designated Ptet-1. It is commercially distributed as part of the pTRE vector series (for tetracycline responsive element), somewhat in line with the prevailing eukaryotic nomenclature.

In some embodiments, doxycycline, a tetracycline derivative, is the effector substance used for a Tet-On or a Tet-Off system. Doxycycline binds with high affinity to tTA as well as to rtTA and, thus, is fully effective in a Tet-Off system at concentrations as low as 1-2 ng/ml in the case of tTA, for example. In a Tet-On system, concentrations as low as 80 ng/ml, in the case of rtTA2-syn1, for example, are effective.

In some embodiments of the present disclosure, an engineered construct comprises a promoter (*e.g.*, CAG) operably linked to a nucleic acid (*e.g.*, gene) encoding rtTA, which is located upstream from a Ptet promoter operably linked to a nucleic acid encoding an enzyme that cleaves (*e.g.*, Cas9 nuclease), a nucleic acid encoding an enzyme that nicks (*e.g.*, Cas9 nickase) nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of

nucleic acid (*e.g.*, Cre recombinase). Typically, at least one (*e.g.*, one, two or more) insulator is located between the nucleic acid encoding rtTA and the Ptet promoter such that activating transcription of the nucleic acid encoding rtTA does not also activate transcription of the enzyme in the absence of a suitable effector substance, such as doxycycline.

5 In some embodiments of the present disclosure, an engineered construct comprises a promoter (*e.g.*, CAG) operably linked to a nucleic acid (*e.g.*, gene) encoding tTA, which is located upstream from a Ptet promoter operably linked to a nucleic acid encoding an enzyme that cleaves (*e.g.*, Cas9 nuclease), a nucleic acid encoding an enzyme that nicks (*e.g.*, Cas9
10 nickase) nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid (*e.g.*, Cre recombinase). Typically, at least one (*e.g.*, one, two or more) insulator is located between the nucleic acid encoding rtTA and the Ptet promoter such that activating transcription of the nucleic acid encoding tTA does not also activate transcription of the enzyme in the absence of a suitable effector substance, such as doxycycline.

IPTG-Inducible System

15 Some embodiments of the present disclosure utilize a system that relies on the presence of a lactose (*lac*) repressor protein and a *lac* operon. The *lac* repressor is a DNA-binding protein that binds to the *lac* operon and inhibits expression of a nucleic acid operably linked to the *lac* operon. The presence of allolactose or an allolactose mimic, such as isopropyl β -D-1-thiogalactopyranoside (IPTG), inhibits the DNA binding ability of the *lac*
20 repressor protein. This loss of DNA binding by the *lac* repressor is used for transcriptional activation of the *lac* operon and expression of any nucleic acid linked to that operon.

The *lac* operon contains three structural genes, and a promoter, a terminator, regulator, and an operator. The three structural genes are *lacZ*, *lacY*, and *lacA*. *lacZ* encodes β -galactosidase (LacZ), an intracellular enzyme that cleaves the disaccharide lactose into
25 glucose and galactose; *lacY* encodes lactose permease (LacY), a transmembrane symporter that pumps β -galactosides into the cell using a proton gradient in the same direction; and *lacA* encodes galactoside O-acetyltransferase (LacA), an enzyme that transfers an acetyl group from acetyl-CoA to β -galactosides.

In some embodiments of the present disclosure, an engineered construct comprises a
30 promoter (*e.g.*, CAG) operably linked to a nucleic acid (*e.g.*, gene) encoding the *lac* repressor protein, which is located upstream from a *lac* operon operably linked to a nucleic acid encoding an enzyme that cleaves (*e.g.*, Cas9 nuclease), a nucleic acid encoding an enzyme that nicks (*e.g.*, Cas9 nickase) nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid (*e.g.*, Cre recombinase). Typically, at least one (*e.g.*, one,
35 two or more) insulator is located between the nucleic acid encoding the *lac* repressor protein

and the *lac* operon such that activating transcription of the nucleic acid encoding the lac repressor protein does not also activate transcription of the enzyme in the absence of a suitable effector substance, such as IPTG.

Tamoxifen-Inducible System

5 Also provided herein are engineered constructs comprising (a) a promoter operably linked to a nucleic acid encoding a Cas9 nuclease that does not comprise a nuclear localization signal, wherein the nucleic acid encoding the enzyme is flanked by estrogen receptor (ERT2) sequences, and (b) a deoxyribonucleic acid (DNA)-binding recognition sequence. A nuclear localization signal or sequence (NLS) is an amino acid sequence that
10 'tags' a protein for import into the cell nucleus by nuclear transport. Typically, this signal contains one or more short sequences of positively charged lysine residues or arginine residues exposed on the protein surface. Different nuclear localized proteins may share the same NLS. An NLS has the opposite function of a nuclear export signal, which targets proteins out of the nucleus. A Cas9 nuclease that does not comprise a nuclear localization
15 signal is a modified Cas9 protein that is not imported into the cell nucleus by nuclear transport.

Cas9 activity can be regulated by fusing the nuclease to a modified fragment of the estrogen receptor (ERT2). In some embodiments, the ERT2 is a modified version of the ER receptor that is highly selective for tamoxifen binding relative to endogenous estrogen. Cas9
20 fused to a modified fragment is sequestered outside of the nucleus where it cannot direct recombination. In the presence of estrogen receptor antagonists (*e.g.* tamoxifen), Cas9 can relocate into the nucleus where it is able to function. In some embodiments, an engineered nucleic acid is flanked by ERT2 fragments (comprises an ERT2 fragment on either end of the protein (ERT2Cas9ERT2; SEQ ID NO: 7)). If ERT2Cas9ERT2 expression is driven by
25 (controlled by) tissue specific promoters, genes of interest can be modified in a tissue of interest at any time in their development. Once this information is recorded, the cells can be monitored over the lifetime of the animal. In some embodiments, an engineered construct comprises a nucleic acid that encodes a Cas9 nuclease that does not comprise a nuclear localization signal, wherein the nucleic acid encoding the enzyme is flanked by estrogen
30 receptor (ERT2) sequences (*e.g.*, SEQ ID NO: 7).

Nucleic Acids

An "engineered construct" refers to an artificially constructed segment (linear or circular) of nucleic acid that is used for introduction into a cell. Engineered constructs (*e.g.*,
35 "TOICas constructs") typically contain at least one promoter operably linked to a nucleic acid

encoding a protein of interest. Embodiments of the present disclosure provide engineered nucleic acids encoding elements of an inducible nucleic acid expression system. An “engineered nucleic acid” is a nucleic acid (*e.g.*, at least two nucleotides covalently linked together, and in some instances, containing phosphodiester bonds, referred to as a phosphodiester “backbone”) that does not occur in nature. Engineered nucleic acids include recombinant nucleic acids and synthetic nucleic acids. A “recombinant nucleic acid” is a molecule that is constructed by joining nucleic acids (*e.g.*, isolated nucleic acids, synthetic nucleic acids or a combination thereof) and, in some embodiments, can replicate in a living cell. A “synthetic nucleic acid” is a molecule that is amplified or chemically, or by other means, synthesized. A synthetic nucleic acid includes those that are chemically modified, or otherwise modified, but can base pair with (also referred to as “binding to,” *e.g.*, transiently or stably) naturally-occurring nucleic acid molecules. Recombinant and synthetic nucleic acids also include those molecules that result from the replication of either of the foregoing.

While an engineered nucleic acid, as a whole, is not naturally-occurring, it may include wild-type nucleotide sequences. In some embodiments, an engineered nucleic acid comprises nucleotide sequences obtained from different organisms (*e.g.*, obtained from different species). For example, in some embodiments, an engineered nucleic acid includes a murine nucleotide sequence, a bacterial nucleotide sequence, a human nucleotide sequence, a viral nucleotide sequence, or a combination of any two or more of the foregoing sequences.

In some embodiments, an engineered nucleic acid of the present disclosure may comprise a backbone other than a phosphodiester backbone. For example, an engineered nucleic acid, in some embodiments, may comprise phosphoramidate, phosphorothioate, phosphorodithioate, O-methylphosphoroamidite linkages, peptide nucleic acids or a combination of any two or more of the foregoing linkages. An engineered nucleic acid may be single-stranded (ss) or double-stranded (ds), as specified, or an engineered nucleic acid may contain portions of both single-stranded and double-stranded sequence. In some embodiments, an engineered nucleic acid contains portions of triple-stranded sequence. An engineered nucleic acid may comprise DNA (*e.g.*, genomic DNA, cDNA or a combination of genomic DNA and cDNA), RNA or a hybrid molecule, for example, where the nucleic acid contains any combination of deoxyribonucleotides and ribonucleotides (*e.g.*, artificial or natural), and any combination of two or more bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine, hypoxanthine, isocytosine and isoguanine.

Engineered nucleic acids of the present disclosure may be produced using standard molecular biology methods (*see, e.g.*, *Green and Sambrook, Molecular Cloning, A Laboratory Manual*, 2012, Cold Spring Harbor Press). In some embodiments, nucleic acids

are produced using GIBSON ASSEMBLY® Cloning (see, *e.g.*, Gibson, D.G. *et al. Nature Methods*, 343–345, 2009; and Gibson, D.G. *et al. Nature Methods*, 901–903, 2010, each of which is incorporated by reference herein). GIBSON ASSEMBLY® typically uses three enzymatic activities in a single-tube reaction: 5′ exonuclease, the 3′ extension activity of a DNA polymerase and DNA ligase activity. The 5′ exonuclease activity chews back the 5′ end sequences and exposes the complementary sequence for annealing. The polymerase activity then fills in the gaps on the annealed regions. A DNA ligase then seals the nick and covalently links the DNA fragments together. The overlapping sequence of adjoining fragments is much longer than those used in Golden Gate Assembly, and therefore results in a higher percentage of correct assemblies. Other methods of producing engineered nucleic acids are known in the art and may be used in accordance with the present disclosure.

Genetic Elements

Engineered nucleic acids of the present disclosure may include one or more genetic elements. A “genetic element” refers to a sequence of nucleotides that has a role in nucleic acid expression (*e.g.*, promoters, insulators, enhancers, terminators and molecular (*e.g.*, DNA or protein) binding regions) or encodes a product of a nucleic acid (*e.g.*, a sequence of nucleotides encoding a regulatory protein or a sequence of nucleotides encoding an enzyme that cleaves nucleic acid, a nucleic acid encoding an enzyme that nicks nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid).

Expression of engineered nucleic acids is typically driven by a promoter operably linked to the engineered nucleic acid. A “promoter” refers to a control region of a nucleic acid at which initiation and rate of transcription of the remainder of a nucleic acid sequence are controlled. A promoter drives transcription of the nucleic acid sequence that it regulates, thus, it is typically located at or near the transcriptional start site of a gene. A promoter, in some embodiments, is 100 to 1000 nucleotides in length. A promoter may also contain sub-regions at which regulatory proteins and other molecules may bind, such as RNA polymerase and other transcription factors. Promoters may be constitutive (*e.g.*, CAG promoter, cytomegalovirus (CMV) promoter), inducible (also referred to as activatable), repressible, tissue-specific, developmental stage-specific or any combination of two or more of the foregoing.

A promoter is considered to be “operably linked” when it is in a correct functional location and orientation relative to a sequence of nucleic acid that it regulates (*e.g.*, to control (“drive”) transcriptional initiation and/or expression of that sequence).

A promoter, in some embodiments, is naturally associated with a nucleic acid and may be obtained by isolating the 5′ non-coding sequence(s) located upstream of the coding

region of the given nucleic acid. Such a promoter is referred to as an “endogenous” promoter.

A promoter, in some embodiments, is not naturally associated with a nucleic acid. Such a promoter is referred to as a “heterologous” promoter and includes, for example, promoters that regulate other nucleic acids and promoters obtained from other cells. A heterologous promoter may be synthetic or recombinant. Synthetic heterologous promoters, in some embodiments, contain various elements obtained from known transcriptional regulatory regions. Synthetic heterologous promoters, in some embodiments, contain mutations that alter expression through methods of genetic engineering that are known in the art. Recombinant heterologous promoters, in some embodiments, are produced by recombinant cloning, nucleic acid amplification (*e.g.*, polymerase chain reaction (PCR)), or a combination of recombinant cloning and nucleic acid amplification (see U.S. Pat. No. 4,683,202 and U.S. Pat. No. 5,928,906). Other methods of producing synthetic and recombinant heterologous promoters are contemplated herein.

A promoter, in some embodiments, is an inducible promoter. An “inducible promoter” regulates (*e.g.*, activates or inactivates) transcriptional activity of a nucleic acid to which it is operably linked when the promoter is influenced by or contacted by a corresponding regulatory protein.

Thus, a “regulatory protein,” as used herein, is a protein that modulates (*e.g.*, activates or inactivates) transcriptional activity from a promoter (*e.g.*, an inducible promoter). In some embodiments, a regulatory protein binds directly to an inducible promoter (*e.g.*, to a sequence of nucleotides within a promoter). In some embodiments, a regulatory protein binds to a region upstream from an inducible promoter (*e.g.*, within 50 to 100 nucleotides upstream from an inducible promoter). In some embodiments, a regulatory protein binds proximal to (*e.g.*, adjacent to) an inducible promoter. Examples of regulatory proteins include, without limitation, tetracycline-controlled transactivator (tTA) transcription factor, reverse tetracycline-controlled transactivator (rtTA) transcription factor, and Lac repressor protein.

A regulatory protein that modulates transcription may activate or inactivate transcription, depending on the system used. Activation of transcription may involve directly acting on a promoter to drive transcription or indirectly acting on a promoter by inactivation of a repressor element (*e.g.*, repressor protein) that is preventing the promoter from driving transcription. Conversely, inactivation of transcription may involve directly acting on a promoter to prevent transcription or indirectly acting on a promoter by activating a repressor element that then acts on the promoter.

Activity of a regulatory protein is often regulated by an effector substance. An “effector substance” is any substance that modulates (*e.g.*, activates or inactivates) activity of a regulatory protein. An effector substance may be an endogenous or exogenous condition (*e.g.*, light or heat), compound (*e.g.*, chemical or non-chemical compound) or other protein that regulates (*e.g.*, directly or indirectly) activity of a regulatory protein.

For example, in the Tet-Off System (also referred to as the tTA-dependent system) and the Tet-On System (also referred to as the rtTA-dependent system), a recombinant tetracycline controlled transcription factor (tTA or rtTA) (a “regulatory protein”) interacts with a tTA/rtTA inducible promoter, *P*_{tet}, to drive expression of the gene operably linked to the promoter. Gene expression is regulated by the effector substance tetracycline or one of its derivatives. Tetracyclines act at the level of DNA binding of tTA and rtTA transcription factors. rtTA requires a tetracycline ligand for DNA binding and hence, transcription. In contrast, the interaction between tTA and DNA is prevented by tetracycline. Thus, the Tet-Off System and the Tet-On System, two versions of the Tet System, respond to tetracyclines (and their derivatives, such as doxycycline) differently.

Typically, the administration or removal of an effector substance results in a switch between activation and inactivation of transcription of the operably linked nucleic acid sequence. Thus, the active state of a promoter operably linked to a nucleic acid sequence refers to the state when the promoter is actively regulating transcription of the nucleic acid sequence (*e.g.*, the linked nucleic acid sequence is expressed). Conversely, the inactive state of a promoter operably linked to a nucleic acid sequence refers to the state when the promoter is not actively regulating transcription of the nucleic acid sequence (*e.g.*, the linked nucleic acid sequence is not expressed).

Examples of effector substances that regulate inducible promoters (*e.g.*, via regulation of a regulatory protein) include, without limitation, physiological conditions, such as changes in light, pH, temperature, radiation, osmotic pressure, saline gradients and cell surface binding. Inducible promoters may also be regulated by varying the concentration of extrinsic or intrinsic effector substances. Examples of extrinsic effector substances include, without limitation, amino acids and amino acid analogs, saccharides and polysaccharides, nucleic acids, protein transcriptional activators and repressors, cytokines, toxins, petroleum-based compounds, metal (*e.g.*, copper) containing compounds, salts, ions, enzyme substrate analogs, hormones or combinations of any two or more of the foregoing. Other effector substances are known in the art and may be used in accordance with the present disclosure.

Examples of inducible promoters include, without limitation, chemically- or biochemically-regulated and physically-regulated promoters, such as alcohol-regulated

promoters, tetracycline-regulated promoters (*e.g.*, anhydrotetracycline (aTc)-responsive promoters and other tetracycline-responsive promoter systems, which include a tetracycline repressor protein (tetR), a tetracycline operator sequence (tetO) and a tetracycline transactivator fusion protein (tTA)), steroid-regulated promoters (*e.g.*, promoters based on the rat glucocorticoid receptor, human estrogen receptor, moth ecdysone receptors, and promoters from the steroid/retinoid/thyroid receptor superfamily), metal-regulated promoters (*e.g.*, promoters derived from metallothionein (proteins that bind and sequester metal ions) genes from yeast, mouse and human), pathogenesis-regulated promoters (*e.g.*, induced by salicylic acid, ethylene or benzothiadiazole (BTH)), temperature/heat-inducible promoters (*e.g.*, heat shock promoters), and light-regulated promoters (*e.g.*, light responsive promoters from plant cells). Other inducible promoters are known in the art and may be used in accordance with the present disclosure.

Enhancers

Engineered nucleic acids, in some embodiments, comprise enhancers. An “enhancer” is a cis-acting regulatory sequence of nucleotides involved in the transcriptional activation of a nucleic acid sequence operably linked to a promoter. The enhancer may be located at any functional location upstream or downstream from the promoter.

Terminators

Engineered nucleic acids, in some embodiments, comprise terminators. A “terminator” is a sequence of nucleotides that causes transcription to stop. A terminator may be unidirectional or bidirectional. A terminator comprises a DNA sequence involved in specific termination of an RNA transcript by an RNA polymerase and prevents transcriptional activation of downstream nucleic acid sequences by upstream promoters.

The most commonly used type of terminator is a forward terminator. When placed downstream of a nucleic acid sequence that is usually transcribed, a forward transcriptional terminator will cause transcription to abort. In some embodiments, bidirectional transcriptional terminators are used, which usually cause transcription to terminate on both the forward and reverse strand. In some embodiments, reverse transcriptional terminators are provided, which usually terminate transcription on the reverse strand only.

Examples of terminators for use in accordance with the present disclosure include, without limitation, termination sequences of genes such as, for example, the bovine growth hormone terminator, and viral termination sequences such as, for example, the T0 terminator, the TE terminator, Lambda T1 and the T1T2 terminator found in bacterial systems. In some embodiments, the termination signal may be a sequence that cannot be transcribed or translated, such as those resulting from a sequence truncation.

Selectable Markers

Engineered constructs (*e.g.*, “TOICas constructs”), in some embodiments, comprise a nucleic acid encoding a selectable marker protein. A selectable marker is a gene introduced into a cell that confers a trait suitable for artificial selection. A selectable marker may be, for example, an antibiotic resistance gene. Non-limiting examples of antibiotic resistance genes include gene encoding resistance to ampicillin, chloroamphenicol, tetracycline or kanamycin. For example, beta-lactamase confers ampicillin resistance to bacterial hosts, the *neo* gene obtained from Tn5, confers resistance to kanamycin in bacteria and geneticin in eukaryotic cells, the mutant *FabI* gene (mFabI) obtained from the *Escherichia coli* genome confers triclosan resistance to the host, and URA3, an orotidine-5' phosphate decarboxylase obtained from yeast is a positive and negative selectable marker.

Vectors

Embodiments of the present disclosure provide vectors comprising engineered nucleic acids encoding elements of an inducible nucleic acid expression system. A “vector” refers to a nucleic acid (*e.g.*, DNA) used as a vehicle to carry genetic material (*e.g.*, an engineered nucleic acid) into a cell where, for example, it can be replicated and/or expressed. It should be understood that the term “vector,” as used herein, does not encompass lentiviral vectors. Thus, in some embodiments, an engineered nucleic acid of the present disclosure is cloned into a vector, or delivered to a cell via a vector, that is not a lentiviral vector. Vectors for use as provided here are typically engineered and include episomal expression vectors. Examples of episomal expression vectors include, without limitation, plasmids, which are double-stranded generally circular DNA sequences that are capable of automatically replicating in a host cell, and viral vectors, which may be based on sequences from DNA viruses, such as BK virus, bovine papilloma virus 1 and Epstein-Barr virus.

Baculovirus vectors (Kost *et al.* 2005 *Nat Biotechnol.* 2005 May; 23(5): 567–575) are also provided herein. Baculovirus gene expression systems and gene delivery systems are known (*see, e.g.*, Makela AR *et al. Cold Spring Harb Protoc.* 2010 Mar;2010(3), incorporated herein by reference) and may be used in accordance with the present disclosure. In some embodiments, a baculovirus vector is used to deliver TOICas constructs, as provided herein, to cells, such as mammalian cells.

A vector may also be an engineered bacterial artificial chromosome (BAC) (O’Conner M. *et al.* 1989 *Science* 244 (4910): 1307–1312; Shizuya H. *et al.* 1992 *Proc Natl Acad Sci USA* 89 (18): 8794–8797; and Shizuya H *et al.* 2001 *Keio J Med.* 50 (1): 26–30, each of which is incorporated herein by reference) or a yeast artificial chromosome (YAC)

(Struhl K *et al.* 1979 *PNAS* 76(3): 1035-39, incorporated herein by reference) where, for example, the nucleic acid encoding the regulatory protein, the nucleic acid encoding the nuclease or the nucleic acid encoding the recombinase is placed under an inducible promoter (*e.g.*, a tissue-specific promoter) or a housekeeping-gene promoter. The promoter present in the BAC or YAC may, in some embodiments, regulate a downstream gene.

Vectors comprising engineered nucleic acids (or the engineered nucleic acids themselves), in some embodiments, are larger than typical expression constructs. In some embodiments, engineered nucleic acids (or vectors comprising an engineered nucleic acid) of the present disclosure are at least 12 kilobases (kb). For example, the engineered nucleic acids (or vectors comprising an engineered nucleic acid) may be at least 13 kb, at least 14 kb, at least 15 kb, at least 16 kb, at least 17 kb, at least 18 kb, at least 19 kb or at least 20 kb. In some embodiments, the engineered nucleic acids (or vectors comprising an engineered nucleic acid) are 15 kb to 20 kb, 15 kb to 30 kb, 15 kb to 40 kb, 20 kb to 30 kb, or 20 kb to 40 kb.

Cells

Engineered constructs (*e.g.*, “TOICas constructs”) of the present disclosure may be introduced into a variety of different cells. Examples of cells into which an engineered construct may be introduced include, without limitation, mammalian cells, insect cells, bacterial cells and yeast cells. Mammalian cells may be human cells, primate cells (*e.g.*, vero cells), rat cells (*e.g.*, GH3 cells, OC23 cells) or mouse cells (*e.g.*, MC3T3 cells), for example. There are a variety of human cell lines, including, without limitation, HEK cells (*e.g.*, HEK 293 or HEK 293T cells), HeLa cells, cancer cells from the National Cancer Institute's 60 cancer cell lines (NCI60), DU145 (prostate cancer) cells, Lncap (prostate cancer) cells, MCF-7 (breast cancer) cells, MDA-MB-438 (breast cancer) cells, PC3 (prostate cancer) cells, T47D (breast cancer) cells, THP-1 (acute myeloid leukemia) cells, U87 (glioblastoma) cells, SHSY5Y human neuroblastoma cells (cloned from a myeloma) and Saos-2 (bone cancer) cells.

In some embodiments, engineered constructs are expressed in stem cells (*e.g.*, human stem cells) such as, for example, pluripotent stem cells (*e.g.*, human pluripotent stem cells including human induced pluripotent stem cells (hiPSCs)). A “stem cell” refers to a cell with the ability to divide for indefinite periods in culture and to give rise to specialized cells. A “pluripotent stem cell” refers to a type of stem cell that is capable of differentiating into all tissues of an organism, but not alone capable of sustaining full organismal development. A “human induced pluripotent stem cell” refers to a somatic (*e.g.*, mature or adult) cell that has

been reprogrammed to an embryonic stem cell-like state by being forced to express genes and factors important for maintaining the defining properties of embryonic stem cells (*see, e.g.,* Takahashi and Yamanaka, 2006 *Cell* 126 (4): 663–76, incorporated by reference herein).

Human induced pluripotent stem cell express stem cell markers and are capable of generating cells characteristic of all three germ layers (ectoderm, endoderm, mesoderm).

Additional non-limiting examples of cell lines that may be used in accordance with the present disclosure include 293-T, 293-T, 3T3, 4T1, 721, 9L, A-549, A172, A20, A253, A2780, A2780ADR, A2780cis, A431, ALC, B16, B35, BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C2C12, C3H-10T1/2, C6, C6/36, Cal-27, CGR8, CHO, CML T1, CMT, COR-L23, COR-L23/5010, COR-L23/CPR, COR-L23/R23, COS-7, COV-434, CT26, D17, DH82, DU145, DuCaP, E14Tg2a, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, Hepa1c1c7, High Five cells, HL-60, HMEC, HT-29, HUVEC, J558L cells, Jurkat, JY cells, K562 cells, KCL22, KG1, Ku812, KYO1, LNCap, Ma-Mel 1, 2, 3....48, MC-38, MCF-10A, MCF-7, MDA-MB-231, MDA-MB-435, MDA-MB-468, MDCK II, MG63, MONO-MAC 6, MOR/0.2R, MRC5, MTD-1A, MyEnd, NALM-1, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NW-145, OPCN/OPCT Peer, PNT-1A/PNT 2, PTK2, Raji, RBL cells, RenCa, RIN-5F, RMA/RMAS, S2, Saos-2 cells, Sf21, Sf9, SiHa, SKBR3, SKOV-3, T-47D, T2, T84, THP1, U373, U87, U937, VCaP, WM39, WT-49, X63, YAC-1 and YAR cells.

In some embodiments, engineered constructs of the present disclosure are introduced into human cells, for example, lymphocytes, such as T cells (*e.g.,* CD8⁺ cells, CD4⁺ cells), B cells or natural killer cells (NK cells).

Cells of the present disclosure, in some embodiments, are modified. A modified cell is a cell that contains an exogenous nucleic acid or a nucleic acid that does not occur in nature. In some embodiments, a modified cell contains a mutation in a genomic nucleic acid. In some embodiments, a modified cell contains an exogenous independently replicating nucleic acid (*e.g.,* an engineered nucleic acid present on an episomal vector). In some embodiments, a modified cell is produced by introducing a foreign or exogenous nucleic acid into a cell.

An engineered construct may be introduced into a cell by methods, such as, for example, electroporation (*see, e.g.,* Heiser W.C. *Transcription Factor Protocols: Methods in Molecular Biology*TM 2000; 130: 117-134), chemical (*e.g.,* calcium phosphate or lipid), transfection (*see, e.g.,* Lewis W.H., *et al., Somatic Cell Genet.* 1980 May; 6(3): 333-47; Chen C., *et al., Mol Cell Biol.* 1987 August; 7(8): 2745–2752), fusion with bacterial protoplasts containing recombinant plasmids (*see, e.g.,* Schaffner W. *Proc Natl Acad Sci USA.* 1980 Apr;

77(4): 2163-7), or microinjection of purified DNA directly into the nucleus of the cell (*see, e.g.,* Capecchi M.R. *Cell*. 1980 Nov; 22(2 Pt 2): 479-88).

Mammalian cells (*e.g.,* human cells) modified to comprise an engineered construct of the present disclosure may be cultured (*e.g.,* maintained in cell culture) using conventional mammalian cell culture methods (*see, e.g.,* Phelan M.C. *Curr Protoc Cell Biol*. 2007 Sep; Chapter 1: Unit 1.1, incorporated by reference herein). For example, cells may be grown and maintained at an appropriate temperature and gas mixture (*e.g.,* 37 °C, 5% CO₂ for mammalian cells) in a cell incubator. Culture conditions may vary for each cell type. For example, cell growth media may vary in pH, glucose concentration, growth factors, and the presence of other nutrients. Growth factors used to supplement media are often derived from the serum of animal blood, such as fetal bovine serum (FBS), bovine calf serum, equine serum and/or porcine serum. In some embodiments, culture media used as provided herein may be commercially available and/or well-described (*see, e.g.,* Birch J. R., R.G. Spier (Ed.) *Encyclopedia of Cell Technology*, Wiley. 411-424, 2000; Keen M. J. *Cytotechnology* 17: 125-132, 1995; Zang, *et al. Bio/Technology*. 13: 389-392, 1995). In some embodiments, chemically defined media is used.

The inducible genome editing systems of the present disclosure permit temporally- and spatially-controlled modification of genome, which is useful for a variety of applications. For example, they may be used to generate animal models, cell lines and induced pluripotent stem cells (iPSCs), where regulation of a target gene during a stage of differentiation must be tightly controlled. Inducible genome editing systems may also be used to generate tissue specific, inducible knockouts or knockins. Other applications include gene therapy.

EXAMPLES

Example 1: Protocol for insertion of inducible cassette

1. Grow cells and split around 500000 cells in a fresh well of a 6-well plate the day before transfection.
2. Transfect cells with ZFNs vectors targeting the Nuclease binding site locus and the TOICas inducible construct using the appropriate transfection reagents (usually Lipofectamine LTX or Fugene HD (most often the ration ZFNs/TOICas is 3:1).
3. 3 days after transfection split the cell in 20 cm well and select with appropriate selection (G418 for dox inducible, puromycine for C9Ert2).
4. 3 weeks after selection, pick clones and check for integration in safe harbor locus.
5. gRNA can be inserted as Lentivirus infection or re-targeting the locus with the inducible construct or transient transfected.

6. Induction is performed by adding Doxycycline (10 $\mu\text{g/ml}$) or Tamoxifen (0.5 μM) to the medium.
7. Cleavage is assessed 3 days after induction.

5 **Example 2: Generation of Cell Lines Using TOICas Construct**

Human induced pluripotent stem cells (hiPSCs) are a valuable tool to study the multistep differentiation processes and for generating cell-based *in vitro* disease models. They can also be used for chemical screens and cellular therapies. Each of these applications can be significantly enhanced by genome engineering; in particular the inducible knock-in or
10 knock out of genes of interest facilitate the study of its function at different stages of differentiation or disease progression without the use of animal models.

A TOICas construct (Fig. 10) was targeted to the human safe harbor AAVS1 locus in hiPSCs using a tightly-regulated one vector system to achieve doxycycline (Dox) inducible Cas9. In the presence of Dox, the transactivator (3G) binds to responsive element and drives
15 the expression of the protein of interest, Cas9 T2A GFP.

The iPSC line generated shows no detectable GFP expression in the absence of Dox, while a clear signal of GFP was detected in the presence of Dox (data not shown). GFP could be detected with as low as 5 ng/ml of Dox, and the signal increased with increasing concentrations (0 ng/ml, 1 ng/ml, 5 ng/ml, 10 ng/ml, 50 ng/ml and 100 ng/ml) over
20 the course of 40 hours, showing that the system can be tuned as required (fluorescent data images not shown). There was no detectable genome cleavage (indicated by indels) on a gRNA target locus in the absence of either Dox or a targeting guide RNA, whereas efficient cleavage is observed in the presence of Dox and guide RNA, showing that the system is efficient with no leaky expression of Cas9 (Fig. 11).

25 Single or multiple copies of a construct could be inserted using the method of the present disclosure. Higher copy number corresponded with higher protein (*e.g.*, GFP) expression level.

Karyotype is particularly important in cells that are used for differentiation processes, such as induced pluripotent stem cells (iPSC) and mouse embryonic stem cells (mESC).
30 Chromosome analysis of fixed cell suspensions from the human induced pluripotent stem cell line showed an apparently normal male karyotype in 20 cells examined, indicating that the integration method, as provided herein, does not interfere with pluripotency in induced pluripotent stem cells containing the inducible Cas9 system (data not shown).

A time-course of Cas9 induction was also performed following Dox treatment of the cells, showing detectable GFP levels after only 6 hours and persisting for at least 40 hours. Maximum expression was observed at 30 hours (data not shown).

An induced pluripotent stem cell (iPSC) line expressing a variant of Cas9 that is enzymatically inactive (Cas9 dead) was also generated using the same TOICas system. When compared to Cas9 line in the same TOICas system, toxicity of Cas9 was observed only when the system was overexpressed (100 ng/mL for 2 weeks of expression). The toxicity was indicated by induction of differentiation of the iPSC line, as indicated by morphological changes and by reduction in the expression of pluripotency genes (*e.g.*, *Oct4*). This indicates that the nuclease activity of Cas9 is associated with toxicity in induced pluripotent stem cells.

Next, gene knockout efficiency was assessed in iPSCs or in cells derived from the iPSCs. Efficient gene knockout was achieved by transfecting synthetic gRNA, synthetic crRNA/tracrRNA and gRNA encoded in an expression plasmid. The efficiency and the precision of deletion increased by using 2 or more gRNA constructs targeting nearby sequences (~10-400 bp apart).

Example 3: Human Endocannabinoid Receptor Deletions

Endocannabinoids (eCBs) play an important role in a diverse range of physiological and pathological processes including neural development, immune function, pain, energy homeostasis, lipid and glucose metabolism. Although endocannabinoids requirement has been extensively studied in mouse little is known about the role of these small signaling lipids in human neuronal development and pathology.

In order to study the need of endocannabinoid signaling in the human system, human induced pluripotent stem cells (hiPSCs) pools were generated using TOICas, in which the human endocannabinoid receptor 1 and 2 genes (CNR1 and CNR2) were targeted, alone or in combination (single and double knock-out, respectively).

Precise deletion was achieved in the CNR1 and 2 genes using synthetic dual guide RNA in presence of Cas9 upon a single Dox induction in hiPSC (data not shown).

After one week in cell culture, cells were passaged and genomic DNA was extracted. Precise deletion in the CNR1 and CNR2 coding sequence were confirmed by junction PCR, cell assay, and Sanger sequencing (data not shown).

Example 4: TOICAS-KRAB for CRISPR KO and CRISPRi

An iPSC cell line containing a modified version of TOIC was also generated. The normal Cas9 was fused to KRAB domain. A gRNA with a normal/conventional protospacer length or a gRNA with short protospacer length was used to induce respectively CRISPR KO or CRISPRi (*see, e.g., methods in Kiani, S., et al. Nature methods* 12(11): 1051-1054, 2015). This system was also used to target the *Oct4* gene and achieved knock out or knock down of the *Oct4* gene by using the alternative gRNA indicated above. This system is generally applicable to other modifications, such CRISPRa, base editing and to other orthogonal Cas9 enzymes.

Example 5: Generation of Functional Mouse Model in Immunocompetent Mice

Tet-On inducible Cas9 (iCas9) mouse for ex vivo primary cell modelling. CRISPR-Cas9 is a modular and versatile tool for genetic perturbation. It is composed by the ribonucleoprotein endonuclease (Cas9) and a guide RNA (gRNA). Cultured cells are relatively amenable to CRISPR-Cas9 engineering. However, application in living animals cast some major challenge, given the complexity to deliver all the components, precisely and simultaneously, in the tissue of interest. Moreover, conventional knockout strategies affect every cells in an animal, so that its often impossible to distinguish primary and secondary changes in complex phenotype. Finally, any genetic change has potential consequences on mouse development that either preclude or complicate studies on adult animals (*e.g., embryonic lethality, abnormalities, metabolic defects, cancer*). Therefore, there is a growing necessity of a precise and temporal editing system to use *in vivo*.

Provided herein is a tight, one component system to achieve doxycycline inducible expression of Cas9 in immunocompetent mice without any other phenotypic consequence. An engineered TOICas construct of the present disclosure (Fig. 10) was inserted in the safe r26 locus in mouse embryonic stem cells (mES). Clones were successfully generated with either single or double copy integration in r26 locus. After doxycycline induction and further confirmation of Cas9 expression, different clones were selected to inject into the blastocyst to produce chimeric mice (data not shown). The expression of Cas9 after dox treatment correlates with the number of copies of the TOICas construct in the r26 locus. Introducing multiple copies into a specific locus is only achievable by using a NHEJ-based method for integration. This is the first animal model generated by this method.

In order to prove the functionality of TOICas in mouse cells, ear fibroblasts were isolated from the founder mice and cultivate *ex vivo*. The cells were then stimulated with

Dox. After imaging under fluorescence microscope, GFP expression was confirmed (data not shown).

Tet-On inducible Cas9 (iCas9) mouse for AAV-mediated in vivo genome editing and cancer modeling.

CRISPR-Cas9 moved the boundaries of genome editing towards complex *in vivo* applications. The promise of a sudden approach of CRISPR to gene therapy is hampered by an increasing demand of more translatable and effective animal models. Mouse molecular genetics has the potential to reproduce human diseases, but conventional knockout strategies are not selective nor versatile enough to be controlled in a spatial or temporal manner. In the recent years, the need for a more flexible *in vivo* genetic tool is growing fast, especially in the field of complex modelling in adult mice.

Cas9 expression is tightly regulated in the TOICas mouse, and expression of Cas9 was observed in the majority of the tissues only upon doxycycline treatment. The TOICas mouse has an immunocompetent background, therefore can be used for immuno-oncology applications, for example. Here, immunocompetent TOICas mice (6 per cohort) were administered 2 mg/ml of doxycycline (Dox) in drinking water *ad libitum* for 3 days. After the treatment, a TOICas mouse was euthanized together with an untreated (-Dox) TOICas control mouse (6 per cohort). Six tissues from both mice were immediately isolated and imaged with a epifluorescent stereomicroscope, in order to benchmark Cas9 expression (data not shown). *In vivo* TOICas activation/GFP expression was observed in the heart, lungs, trachea, liver, spleen, pancreas, and brain after 3 days of Dox induction, while no GFP was detected in tissues from the untreated control mouse (data not shown). Histopathological analysis showed that the mouse tumors resembled human tumors. Low differentiated adenocarcinoma could be observed in all areas of the lung. The growth pattern included epithelial tumor cells that showed a papillary, trabecular, glandular and compact growth pattern with local invasion in the alveolar spaces and adjacent tissues. Occasionally presence of almost squamous epithelial-like transformation was observed. There was high mitotic activity in tumor tissue as well as few areas of bronchial epithelium with presence of atypical mitoses. There was also presence of highly undifferentiated cells with polyploidy, anisocytosis and anisokaryosis, and a significant increase of alveolar macrophages. There was no evidence of blood vessel infiltrations or metastases. A bronchoalveolar growth pattern was observed as well as areas of trabecular and glandular growth pattern identified by multifocal nodules in alveolar and bronchiolar spaces. High mitotic activity was observed overall, particularly in all tumor nodules (based on an immunohistochemical assay for proliferative marker Ki67) and randomly in bronchial epithelium (data not shown).

Surprisingly, what was considered a negative control, the brain (the most difficult organ to reach by drug treatment) also resulted positive to GFP.

To generate an *in vivo* mouse model of cancer, a gRNA against *Trp53* and *Kras* together with a template to introduce KrasG12D mutation were used. Cells were infected with AAV harboring gRNA targeting *Kras* (Fig. 12, top panel). Cas9 endonuclease activity was assessed by time-course monitoring of GFP expression (Fig. 13). A surveyor assay, performed on genomic DNA collected from cells 4 or 9 weeks after infection, confirmed precise editing represented by the distinct pattern expected from the targeting strategy (indel and point mutation – Fig. 14). Both the knockout (KO) of p53 and *kras* mutagenesis were confirmed in fibroblast and adult mice.

Next, TOICas mice were infected with AAV harboring the guide RNA targeting *Trp53* gene (Fig. 12, bottom panel) via intratracheal injection. 24 hours before infection, Cas9 was activate by Dox in mice #37, #42 and #52, while mice #98 and #99 were left untreated (-Dox). Four weeks after infection, lungs were collected from all mice and genomic DNA was extracted to perform surveyor nuclease assay. The presence of precise mutation in *Trp53* gene was then confirmed in the treated and infected mice as shown in Fig. 15.

Example 6: Tissue-Specific Knockout

Several strategies were developed for spatial and temporal regulation of CRISPR in the TOICas system of the present disclosure. One strategy is to drive the expression of the rtTA under a tissue-specific promoter. Aalpha myosin heavy chain (aMHC) was used to drive the expression of rtTA and consequently the regulation of Cas9; this strategy is generally applicable using any other tissue promoter and induces the expression of Cas9 only in a particular tissue.

An alternative strategy is to introduce the gRNA cassette in a tissue-specific transcript driven by a polII promoter (*e.g.*, tissue specific). The gRNA is integrated in an intron and may have rybozyme or other RNA processing sequence to be subsequently cleaved by the original transcript.

Yet another strategy is based on the activation of gRNA upon induction of a tissue-specific recombinase or a tissue-specific orthogonal Cas9/CRISPR. In this case, the expression of the gRNA is prevented by the presence of a stop cassette present between the promoter and the functional part of the gRNA. The presence of orthogonal Cas9/CRISPR or a site-specific recombinase induces the removal of the stop cassette, therefore enabling the

expression of the gRNA only in the tissue where the orthogonal Cas9/CRISPR or site specific recombinase are expressed.

The examples described above enable the spatial and temporal regulation of the Cas9/CRISPR system, therefore it is possible to generate tissue-specific Knock Out, Knock
5 Down, and Base Editing in adult animals as well as in differentiated cells, starting from embryonic stem cells or induced pluripotent stem cells.

One of the application of the TOICas system is the inducible ablation of tissues or cells of interest to study the function of a cell/tissue or to mimic disease status. In this case, the temporal expression of Cas9 in TOICas system (regulated by Dox) is combined with a
10 tissue-specific expression of a gRNA targeting multiple repeats in the target genome. Two non-limiting examples of repeats as target sites are the repeats B1 in the murine genome (CTCACTATGTAGACCAGGC (SEQ ID NO: 10)) and the repeats AluI in the human genome (CCTGTAATCCCAGCACTTTCACTTTGGGAGGCCGAGGCGAGTCTCGC
TCTGTCGCCC (SEQ ID NO: 11)). The tissue-specific and temporal activation of the system
15 promotes the cleavage in multiple sites and, therefore, the degradation of the target genome only in the tissue where the TOICas system is expressed and only upon Dox treatment.

Example 7: gRNA Cloning Methods

Three main strategies were developed to clone the gRNA for a particular gene in a
20 cell or a plasmid expressing Cas9. T

The first strategy was to use recombineering to insert a cassette containing a site-specific gRNA plus a bacterial selection marker and a polIII promoter driving the expression of the gRNA. The gRNA can be inserted in a high throughput way by selecting in bacteria for cells that received the gRNA and the selection marker.

25 The second strategy was to use Cas9 from *Neisseria Meningitis* to target the *neo* gene in cells containing the TOICas construct and introduce a gRNA plus a mammalian selection marker within the neo selection marker. This strategy is not restricted to the *neo* selection marker and to the Cas9_{NM} since any other orthogonal Cas9 and insertion point can be used for gRNA integration.

30 The third strategy was to use Cas9 RiboNucleoProtein to deliver the gRNA by homologous recombination method or NHEJ based method in a precise locus. Alternative methods are insertions by lentivirus and transposons.

Example 8: Further analysis of TOICas system

As shown in Figs. 16-19, although there is the occasional cell expressing Cas9 in the absence of doxycycline, Cas9 induction is very clean. Maximal amount of Cas9 is expressed after just 24 hours of doxycycline exposure. Further, induction of Cas9 expression in A549 cells does not appear to effect: cell growth, incidence of DNA damage (as measured by gH2AX foci), or persistence of DNA damage (as measured by micronuclei formation). Generally, Cas9 expression does not appear to increase sensitivity to IR or increase DNA damage caused by IR.

SEQUENCES

Additional example of a mammalian insulator sequence for use in accordance with the present disclosure:

CTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCA
GCAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGT
GCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGC
TCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTTAG
AATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCC
TGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGA
ATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTG
GGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCT
TGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTTCAG
TGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGAGGACAGCATGGACGTGGGACATGC
AGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAG
ATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAGAGGGCCACAG
ACACTGCTGGTCCCTGTGCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGG
AAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCA
TGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCA
ACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATC
ACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGA
GAAGGTAAATCTTGCTAAATCCAGCCCGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCAT
CCAATCCAGGACGGAGTCAGTGAGGATGGGGCT (SEQ ID NO: 6)

Cas9 ERT2:

ATGGCTCTCGAGCCATCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAAGCCCGCTCATGATC
AAACGCTCTAAGAAGAAGACAGCCTGGCCTTGTCCCTGACGGCCGACCAGATGGTCAGTGCTTGTG
GATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTACCAGACCCTTCAGTGAAGCTTCGATG
ATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTCACATGATCAACTGGGCGAAGAGGGT
GCCAGGCTTTGTGGATTTGACCCCTCCATGATCAGGTCCACCTTCTGGAATGTGCCTGGCTAGAGAT
CCTGATGATTGGTCTCGTCTGGCGCTCTATGGAGCACCCAGTGAAGCTACTGTTTGCTCCTAACTTG
CTCTTGACAGGAACCAGGGAAAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGC
TACATCATCTCGGTTCCGCATGATGAATCTGCAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTAT
TTTGCTTAATTCTGGAGTGACACATTTCTGTCCAGCACCCCTGAAGTCTCTGGAAGAGAAGGACCA
TATCCACCGAGTCCTGGACAAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTGA
CCCTGCAGCAGCAGCAGCAGCGGCTGGCCAGCTCCTCCTCATCCTCTCCACATCAGGCACATGA
GTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAGAACGTCGAGTGGTGGCCCTCATGACCTG
CTGCTGGAGGCGGCGGACGCCCACCGCCTACATGCGCCCACTAGCCGTGGAGGGGCATCCGTGGA
GGAGACGGACCAAAGCCACTTGGCCACTGCGGGCTCTACTTCATCGCATTCCTTGCAAAAGTATTA
CATCACGGGGGAGGCAGAGGGTTTCCCTGCCACAGCTGACAAGAAATACTCAATCGGGCTGGACA
TCGGAATAACTCAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAATTT
AAGGTCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTTTCGA
CTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCGAA

AGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCT
TTCATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTT
GGAAACATTGTGGACGAGGTCGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAA
5 ACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAA
GTTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTGT
TCATCCAGCTGGTCCAGACATAACAATCAGCTGTTTGAGGAAAACCTATTAATGCATCTGGCGTGG
ACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCGCTCAG
CTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGAAAATCTGATTGCACTGTCAGTGGGACTGACACC
10 TAACTTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATG
ACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTCTGGCCGCTA
AGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCCC
CTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGCA
CTGGTGCGCCAGCAGCTGCCCAGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGTA
15 CGCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGGA
GAAGATGGACGGCACCGAGGAACTGCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAAGCAG
CGGACATTTGATAACGGCTCCATCCCCACAGATTTCATCTGGGAGAGCTGCACGCAATCCTGCGA
CGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATCTGACCTT
CCGGATCCCTTACTATTGTGGGGCCCTGGCAAGGGTAATTCCCGCTTTGCTGGATGACACGGAA
20 ATCTGAGGAAAACATCACTCCTTGGAACCTCGAGGAAAGTGGTCGATAAGGGAGCTTCCGCACAGT
CTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGTCTGCCTAAGCAC
AGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGG
GATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAAA
CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAATTGAATGTTTCGAT
25 TCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
AAGATCATTAAAGGATAAAGACTTCCTGGACAACGAGGAAAATGAGGATATCCTGGAAGACATTGT
GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTATGCCCATCT
GTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGAGCAGCTGAGC
AGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGGTGAAAGACTATCCTGGACTTCTGAAATC
30 AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGA
TATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTGCACGAGCATATCGAAACCTGGCCG
GGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTCGATGAGCTGGTGAAAGTC
ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCGAAAATCAGACCACACAGA
AAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
35 CCAGATCCTGAAAAGAGCACCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGTATCTGTACT
ATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTAC
GACGTGGATCATATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACC
CGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAATGA
AGAACTGATGGAGACAGCTGCTGAATGCGCAAGGTGATCACACAGCGAAAGTTTGATAACCTGACT
40 AAAGCTGAGAGAGGGGTCTGTGCAAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG
AGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGAT
GAGAATGACAACTGATCCGCGAAGTGAAGGTGCTTACTCTGAAGTCAAACTTGTGAGCGACTT
CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCAT
ACCTGAACGCAGTGGTTCGGGACCGCCCTGATTAAGAAATACCCCAAACCTGGAGAGCGAATTCTGTG
45 TACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTGGAAA
AGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCTGGCA
AATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
ACAAGGGTAGGGATTTTGCCACAGTGCAGCAAGGTCTGTCATGCCTCAAAGTGAATATTGTCAAGA
AAACAGAGGTGCAGACTGGCGGATTCGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAG
50 CTGATCGCCCGAAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
TTACTCTGTCTGCTGGTGGTCGCAAGGTGGAGAAGGGGAAAAGCAAGAAACTGAAATCCGTCAAGG
AGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGG
CTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAG
CTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGAGGAGAGCTGCAGAAGGGTAATGAACTGG
55 CCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCC
CCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATT
GAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCTCTGAG
CGCATACAAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAAATATCATTCTGTTCA
CTGTGACCAACCTGGGAGCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
60 ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACCCAGAGCATTACCGGCCTGTATGAA
ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGCCATCTGCTGGAGACATGAGAGCTGC
CAACCTTTGGCCAAGCCCGCTCATGATCAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCCTGAC
GGCCGACCAGATGGTCAGTGCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCC

TACCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGG
TTCACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAGGTCC
ACCTTCTGGAATGTGCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCTATGGAGCACC
CAGTGAAGCTACTGTTTGCTCCTAACTTGCTCTTGGACAGGAACCAGGGAAAAATGTGTAGAGGGC
5 ATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTCCGCATGATGAATCTGCAGGGA
GAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAATTCTGGAGTGTACACATTTCTGTCCAGCA
CCCTGAAGTCTCTGGAAGAGAAGGACCATATCCACCGAGTCCTGGACAAGATCACAGACACTTTG
ATCCACCTGATGGCCAAGGCAGGCTGACCTTGACAGCAGCAGCACCAGCGGCTGGCCAGCTCCT
CCTCATCCTCTCCACATCAGGCACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGT
10 GCAAGAACGTGGTGGCCCTCTATGACCTGCTGCTGGAGGCGGCGGACGCCACCGCTACATGCG
CCCACTAGCCGTGGAGGGGCATCCGTGGAGGAGACGGACCAAAGCCACTTGGCCACTGCGGGCTC
TACTTCATCGCATTCTTGCAAAAGTATTACATCACGGGGGAGGCAGAGGGTTCCCTGCCACAGC
TTGA (SEQ ID NO: 7)

15

Sequence of TOIC construct depicted in Fig. 2

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTTCTAAATACATTCAAATAT
GTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
GTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTCGGCATTTTGCCTTCTGTTTGTCTC
20 CCAGAAACGTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGA
ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTCCAATGATGAG
CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
GGATGGCATGACAGTAAGAGAATTATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCA
25 ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGACAACATGGGGGATC
ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
30 CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAAGCATTGGT
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAG
GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
35 TGCTGCTTGCAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTG
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGATAAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
40 GACCTAACCGAAGTATGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCGGAAGGGA
GAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACAGGAGAGCGCAGGAGGAGCTTCC
AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
CCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
45 GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
50 TTAACCCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTGTGACG
TACGTTAAAGATAATCATGCGTAAATTTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
TATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCACAAAA
CAATTTATTTATGTTTATTTATTTATTAATAAAAAAACAACAACTCAAAATTTCTTCTATAAAGTAAC
AAAATTTTAAACATTCTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
55 TATTATAAATAAGTAATTAGCTTAACCTTATACATAATAGAAACAAATTATACTTATTAATCGCAT
TGATTATTGACTAGTCGTATTAAGGGTTCGGATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCG
CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
GCTGTCCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
60 GGCAACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
TTGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG

CGATACCGTAAAGCACGAGGAAGCGGTACAGCCATTTCGCCGCCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCCACAGTCGATGAATCCAGA
AAAGCGGCCATTTTCCACCATGATATTTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTC
5 GCGTTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCCTGATGCTCTTC
GTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
CGCTTGGTGGTCAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCA
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCC
AATAGCAGCCAGTCCCCTTCCCGCTTCAGTGACAACGTGAGTCGAGCACAGCTGCGCAAGGAACGCCCGT
10 CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTACAGGGCACCGGACAGGTTCGT
CTTGACAAAAAGAACC GGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGC
AATCCATCTTGTTC AATGGCCGATCCCATGGTTTAGTTCCTCACCTTGTCGTATTATACTATGCCGA
TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTGCGAAAGGCCCGGAGATGAGGA
15 AGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGTGCAGAATGCCGGGCCCTCCGGAGGACC
TTCGGGCGCCCCGCCCGCCCCCTGAGCCCCGCCCTGAGCCCCGCCCGGACCCACCCCTTCCAGCC
TCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCAAAGGCCTACCCGCTTC
CATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTACGTC
CTGCACGAGCGCCCTGCGGGGCGGGGGGAACTTCTGACTAGGGGAGGAGTGAAGAGGTGGCG
20 CGAAGGGGCCACCAAAGAACGGAGCCGGTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCC
AGAGGCCACTTGTGTAGCGCCAAGTGCCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTT
GGGAAAAGCGCCTCCCCTACCCGCTAGACACCCACAGTGGGTGGCCTAGGGACAGGATTGCAAC
TCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAG
TTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGC
25 CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGT
ACATGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGT
30 TTTATTTATTTTTTAATTATTTTGTGACGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCG
GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGGCAGCCAATCAGA
GCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGA
AGCGCGCGGGCGGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCTCGCGC
CGCCCCCGCCGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCT
35 CCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTA
AAGGGCTCCGGGAGGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
CGTGGGGAGCGCCGCTGCGGCCCGCTGCCCCGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGG
GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGGCGGCGGGGCGGTGCCCGCGGTGCGGGG
40 GGCTGCGAGGGGAACAAAGGCTGCGTGCAGGGGTGTGTGCGTGGGGGGGTGACGAGGGGTGTGG
GCGCGGCGGTGCGGCTGTAAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGCT
TCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA
GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCGGGAGGGCTCGGGGGAGGGGCGCGGC
GGCCCCGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCG
TGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
45 CGCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGCGCAGGAAGGAAATGGGCGGGGA
GGGCCTTCGTGCGTGC CGCGCGCCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGG
GACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTC
TAGAGCTCTGCTAACCATTGTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
ATTGTGCTGTCTATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTC
50 TCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAGC
TGAAGAACAAGCGGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCAC
TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGT
GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTA
55 CGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
CGCTCTGTCCCGCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
AAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
TGTTTCGACCGGCAGGGAGCCGAACCTGCCCTTCTTTTCGGCCTGGAACATAATCATATGTGGCCTGG
AGAAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGCCCTTGACGATTTTGACTTAGACATG
CTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTG
60 ACCTTGACATGCTCCCCGGGTAAAGCGGGCGGCGACTCTAGATCATAATCAGCCATACCACATTTGT
AGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA

TTTCACAAATAAAGCATT TTTTCTACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT
TAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCC
GCTAGGGGACAGCAGCGAGCCGCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG
5 CCGGCAGCGTGCAGGGGACAGCCCGGGACAGGGGAAGGTGGCAGGGGATCGCTTTCCTCTGAACGC
TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAG
AGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGA
TCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCC
TGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCCTCAC
10 CACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCC
ATTCCCCCTTGTCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGC
CAATTCAGTGTCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGG
GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA
TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAG
15 GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAGGCCCTCCAAGATCCCC
AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATCACCTCCAGGGACGGTGACCCCCACCTCCGTGGGCAGCTGTGCCATGCACACCGCT
20 CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAAGGCCATTAT
CTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGACCTAGAGGGACAGCCCCC
CCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCGGGG
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCAGGGGACAGCCGGGCA
CGGGGAAGGTGGCAGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACAC
25 CTGGGGGGATACGGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAA
CGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATC
AACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGC
ATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCCTCACCACCTCTGGGGGAAAACTGCCTCCTC
ATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCTTGTCTATCAAGGGGAGT
30 TTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGTCATCACGGAGAGGCAGA
TCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
CACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
GAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGAT
35 ACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAGGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTG
CCCCTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCC
CCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCCTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCGACCTCCCCCTGGCACAACCTAAGGCCATTATCTCATCCAACCTCCAGGACGGAGTCA
40 GTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTA
CTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC
AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
ATAAGCTTTAGGCGTGTACGGTGGGCGCTATAAAAGCAGAGCTCGTTTGTGAACCGTCAGATC
45 GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCTACCCCTCGTAAA
AAGCTTGTCCACCATGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGG
ACATCGGAACCTAAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAA
TTTAAGGTCCTGGGCAACACCGATAGGCACCTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTT
GACTGTGGAGAGAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGCGCTATACCCGGCG
50 AAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTT
CTTTCATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTT
TGGAACATTGTGGACGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGA
AACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTA
AGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTG
55 TTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTATTAATGCATCTGGCGTG
GACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCGCTCA
GCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGGAATCTGATTGCACTGTCACTGGGACTGACAC
CTAAGTTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTAT
GACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGACAGACCTGTTCTGGCCGCT
60 AAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCC
CCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGC
ACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGT
ACGCCGTTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG

AGAAGATGGACGGCACCGAGGAAGTCTGGTGAAGTGAATCGCGAGGACCTGCTGCGCAAGCA
GCGGACATTTGATAACGGCTCCATCCCCACCAGATTCATCTGGGAGAGCTGCACGCAATCCTGCG
ACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCT
TCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGGGTAAATCCCGCTTTGCCTGGATGACACGGA
5 AATCTGAGGAAACAATCACTCCTTGGAACTTCGAGGAAGTGGTCGATAAGGGAGCTTCCGCACAG
TCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGTGCTAAGCA
CAGTCTGCTGTACGAGTATTTACAGTCTATAACGAAGTACTAAGGTGAAATACGTCACCGAGGG
GATGAGGAAGCCCGCTTCTGAGCGGTGAACAGAAAGCTATCGTGGACCTGCTGTTTAAAA
CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATTGAATGTTTCGAT
10 TCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
AAGATCATTAAAGGATAAAGACTTCCTGGACAACGAGGAAAATGAGGATATCCTGGAAGACATTGT
GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTTATGCCATCT
GTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAGC
AGAAAGCTGATCAACGGCATTTCGGGACAAAACAGAGTGGAAAGACTATCCTGGACTTTCTGAAATC
15 AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTCAAAGAGGA
TATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCG
GGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTCGATGAGCTGGTGAAAGTC
ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCAAAATCAGACCACACAGA
AAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
20 CCAGATCCTGAAAAGAGCACCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGTATCTGTACT
ATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTAC
GACGTGGATCATATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACC
CGGAGTGACAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGA
AGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGCGAAAGTTTGATAACCTGACT
25 AAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG
AGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGAT
GAGAATGACAACTGATCCGCGAAGTGAAGGTCACTTCTGAAGTCAAACTTGTGAGCGACTT
CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATTGCTCATGACGCAT
ACCTGAACGCAGTGGTCCGGACCGCCCTGATTAAGAAATACCCCAAACTGGAGAGCGAATTTCGTG
30 TACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTGGA
AGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCTGGCA
AATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
ACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCTGTCCATGCCTCAAGTGAATATTGTCAAGA
AAACAGAGGTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAG
35 CTGATCGCCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
TTACTCTGTCTGGTGGTTCGCAAAGGTGGAGAAGGGGAAAAGCAAGAACTGAAATCCGTCGAAGG
AGCTGTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGG
CTAAAGGCTATAAGGAAGTGAAGAAAGACCTGTGATTAAGTGAAGTACGCAAGTAACTGTTGAG
CTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAACTGG
40 CCCTGCCTTCTAAGTACGTGAACCTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCC
CCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATT
GAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAATCTGGATAAGGTCCTGAG
CGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAATATCATTATCTGTTCA
CTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
45 ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACACAGAGCATTACCGGCCTGTATGAA
ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTC
TGCTAACATGCGGTGACGTGAGGAGAAATCCTGGCCAGCACCAGGATCCATGGTGAGCAAGGGC
GAGGAGCTGTTCACCGGGTGGTGCCCATCTGTTCCAGCTGGACGGGACGTAAACCGGCCACAA
GTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCT
50 GCACCACCGCAAGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCTTACCTACGGCGTGCAGT
GCTTCGCCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAG
TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGCACAAAGCTGGAGTACAACATAACAGCCACAAGGTCTATATCACCGCCGACAAGC
55 AGAAGAACGGCATCAAGGTGAACCTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAGACCCCAACGAGAAGCGCGATCACATGGTCTGTGG
AGTTCTGTACCGCCGCGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGC
AGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTG
60 CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCTTTTCTTAATAAAATGAGGAAATTGCATCG
CATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA

CCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCATATGAATATCCTCCTTAGTTCCTATT
CCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCT
AGGGGCAGCAGCGAGCCGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCG
GCAGCGTGCGGGGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTC
5 TCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAAGAGA
GATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCC
AACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTG
GCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCAACCA
CCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCAT
10 TCCCCCTGTCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCA
ATTTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGG
ACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGAT
AAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
15 GGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGTGGAAGGCCCTCCAAGATCCCC
AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATACCTCCAGGGACGGTGACCCCCACCTCCGTTGGGACAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
20 CTCTCATCCAATCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTAC
ACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAATAT
ATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATC
TTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTATAGTCAAAATCA
GTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAA
25 ATGCACAGCGACGGATTTCGCGCTATTTAGAAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAA
TTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACAC
GTCATAGAATCTTCGTTTGACAAAAACACATGTGTGGGTACCGAGCTCTTAATTAAGGCGCGCCG
GGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCCGCCCTATAGTGAGTCGTATTACAATTCACGT
GCCGTGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGACGCA
30 CATCCCCCTTTCCGACGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAAATTTCGCGTTAAATTTT
TGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAA
TAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGA
CTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTT
35 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCCGA
TTTAGAGCTTGACGGGGAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAG
CGGGCGCTAGGGCGCTGGCAAGTGAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTT
AATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 8)

40 AAV gRNA Sequence

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCGGGCG
ACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGCCAACTCCATCAC
TAGGGGTTCTGCGGCCGACGCGTGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACG
ATACAAGGCTGTTAGAGAGATAAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAA
45 ATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGG
ACTATCATATGCTTACCGTAACCTTGAAAGTATTTGATTTCTTGGCTTTATATATCTTGTGGAAAGG
ACGAAACACCGCAGCGTTACCTCTATCGTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTGGATCCGAGGGCCTATTTT
CCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGA
50 CTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAAGTAATAATTTCTTGGGTAGTTTG
CAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACCTTGAAAGTATTTTCGATTT
CTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTGTAATAGCTCCTGCATGGGTTTTAGAG
CTAGAAATAGCAAGTTAAAATAAGGCTAGTCGGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT
GCTTTTTTCTAGAAAGGTACCAGGTCTTGAAGGAGTGGAATTGGCTCCGGTCCCGTCAGTGCGG
55 CAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTGCGCAATTGAACCGGTGCC
TAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCTTTTTTCCCGA
GGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGC
CGCCAGAACACAGGCGTACGGCCACCATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACGG
ATGATAACTGGTCCGCAGTGGTGGGCCAGATGTAAACAAATGAATGTTCTTGATTCATTTATTAAT
60 TATTATGATTCAGAAAAACATGCAGAAAATGCTGTTATTTTTTACATGGTAACGCGGCCCTCTCTT

ATTTATGGCGACATGTTGTGCCACATATTGAGCCAGTAGCGGGTGTATTATACCAGACCTTATTG
GTATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATATCTTA
CTGCATGGTTTGAACCTCTTAATTTACCAAAGAAGATCATTTTTGTGCGCCATGATTGGGGTGCTTG
TTTGGCATTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTGT
5 AGTAGATGTGATTGAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTGATCAAATC
TGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCCATCAAAAA
TCATGAGAAAAGTTAGAACCCAGAAGAATTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAA
GTTTCGTCTGACCAACATTATCATGGCCTCGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTT
GTACAAATTGTTAGGAATTATAATGCTTATCTACGTGCAAGTGATGATTTACCAAAAAATGTTTTATT
10 GAATCGGACCCAGGATTCTTTTCCAATGCTATTGTTGAAGGTGCCAAGAAGTTTCTAATACTGAA
TTTGTCAAAGTAAAAGGTCTTCATTTTTTCGCAAGAAGATGCACCTGATGAAATGGGAAAATATATC
AAATCGTTTCGTTGAGCGAGTTCTCAAAAAATGAACAATAAAGCGCTAATAAAAAGATCTTTATTTTCA
TTAGATCTGTGTGTTGGTTTTTTGTGTAAGCTTTGGCTCCAACACAGATGTTCTTAGGCTACCTAAC
TTCTAACTTTTAATATCCAGTCAACAAAGAATACCGCAAGGGTAGGTGTTGGGATAGCTGTGCGACA
15 AGCTCATGCGGGTGTGTCCACAGGGTATAGCGTACTATGCAGAATATTTGTAAGTGAAGTCAT
GATACATTCTTTGAGAGCCATTAGCTGCTACAAAACAGTAATCTGGCTGTTTAGATCAACAAGCT
AAATGATAGAAGATGAAAGTACTGGTTTCCATGTATTTTTTATTAAGTGTGATGAGAAAGTTGTAA
GTGCTTACAGGTTACTCTGTACATCTGTAGTCATGTAATTCGGAATATCTTAGAGTTTTACACACA
AAGGTGAGTGTTAAAATATTGATAAAGTTTTTGATAATCTTGTGTGAGACATGTTCTAATTTAGTTG
20 TATTTTTATTATTTTTATTGTAAGGCCTGCTGAAAATGACTGAGTATAAACTTGTGGTCTGGGCGCC
GACGGCGTGGGCAAGAGCGCTTTGACGATACAGCTAATTCAGAATCACTTTGTGGATGAGTATGA
TCCAACCATCGAGGTAACGCTGCTCTACAGTCTGCGTGCGCTTGTAAGGACGGCAGCCAGCCGCT
TTGAAAAAGATATCATTTTTATATTTATTAGAAAATTATATTGAAAGTTATTCAGTTATATGTGAT
GTCCTTTAGTTCCAAGGCTTTAACTGGGTGTTAGGGAACCATAGGTGCAAGAAAGTCCACTTCTC
25 ATGAGAGCTCACCACAGAGAAAGAAAGTCCACTTCTCAGGTAACCACGTGCGGACCGAGCGGGCG
CAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCGGGG
CGACCAAGGTGCGCCGACGCCCCGGGCTTTGCCGGGCGGCTCAGTGAGCGAGCGAGCGCGCAG
CTGCCTGCAGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACACCGCATACGTCAAAGCAACCA
TAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGC
30 TACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCACGTTTCGCCG
GCTTTCCCGCTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGATTTAGTGCTTTACGGCACCT
CGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTT
TCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACTC
AACCCTATCTCGGGCTATTCTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAA
35 ATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAATATTAACGTTTACAATTTTATGGT
GCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCCGACACCCGCCAACACCCG
CTGACGCGCCCTGACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCG
GGAGCTGCATGTTGTACAGAGTTTTTACCGTCTACCCGAAACGCGCGAGACGAAAGGCGCTCGTG
ATACGCCTATTTTTATAGGTTAATGTCATGATAAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTC
40 GGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCAT
GAGACAATAACCCGTATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATT
TCCGTGTGCGCCCTTATTCCTTTTTTTCGGGCATTTTGCCTTCTGTTTTGCTCAGCCAGAAACGCTG
GTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAA
CAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
45 TCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGCATACA
CTATTCTCAGAATGACTTGGTTGAGTACTACCAAGTCACAGAAAAGCATCTTACGGATGGCATGAC
AGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAAGTGCAGGCAACTTACTTCTGAC
AACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTTCGACACAACATGGGGGATCATGTAACCTGCC
TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCT
50 GTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAA
CAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCT
GGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTCGAGCACTG
GGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGA
TGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACC
55 AAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGAA
GATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGAC
CCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAA
CAAAAAAACCCGCTACCAAGCGGTGGTTTTGTTGCCGATCAAGAGCTACCAACTTTTTCCGA
AGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGAGCCGTAGTTAGGCC
60 ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTG
CTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGA

TOIC_ApoL1_wt_PgkPuro

10

15

20

25

30

35

40

45

50

55

60

CCGGGCTCCGGAGGACCTTCGGGCGCCCCGCCCCGCCCCCTGAGCCCGCCCCCTGAGCCCGCCCCCG
ACCCACCCCTTCCCAGCCTCTGAGCCAGAAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCC
CAAAGGCCTACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCT
ACTTCCATTTGTCACGTCTCTGCACGACGCGAGCTGCGGGGCGGGGGGAACTTCCTGACTAGGGG
5 AGGAGTAGAAGGTGGCGCGAAGGGGCCACCAAGAACGGAGCCGGTTGGCGCCTACCGGTGGAT
GTGGAATGTGTGCGAGCCAGAGGCCACTTGTGTAGCGCCAAGTGCCAGCGGGGTGCTAAAGCG
CATGCTCCAGACTGCCTTGGGAAAAGCGCCTCCCCTACCCGGTAGACACCCACAGTGGGTGGCCT
AGGGACAGGATTGCAACTCCAGTCTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAAT
CAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG
10 GCGCGCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAG
TAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCACTTGG
CAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
CATCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTC
15 CCCACCCCAATTTTGTATTTATTTATTTTAAATATTTTGTGCGAGCGATGGGGGCGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGAGAGGTGC
GGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGGCGGGCGGCG
GGCCCTATAAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCTGCTGCGTTGCCTTCGCCCCGTGCCCG
CTCCGCGCGCCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGG
20 CGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTCTGT
GGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGG
TGCGTGCGTGTGTGTGTGCGTGGGGAGCGCCGCGTGC GGCCCCGCGCTGCCCGGCGGCTGTGAGCG
CTGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGG
TGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGG
25 GGTGAGCAGGGGTGTGGGCGCGGCGGTGCGGCTGTAACCCCCCTGCACCCCTCCCCGAGT
TGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGGCTCGCCGTGC
CGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCTCGGGCCGGGGAGGG
CTCGGGGAGGGGCGCGGCGGCCCGGAGCGCGGCGGCTGTGCGAGGCGCGGCGAGCCGACGCC
ATTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCC
30 GAAATCTGGGAGGCGCCGCCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCGGGCAG
GAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCCGCGCGCCGCGTCCCTTCTCCATCTCCAGCC
TCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTG
GCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTC
CTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCC
35 AAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACT
CAATGGAGTCGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGC
CTACCTGTACTGGCAGTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTG
GACAGGCATCATACCACTCCTGCCCCGTGAAGGCGAGTCAATGGCAAGACTTCTGCGGAACAA
CGCCAAGTCATAACCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCG
40 CCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCT
CCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACTGTTGGCTGCGTATTGGAGG
AACAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTT
CTGAAACAAGCAATTGAGCTGTTTCGACCGGCAGGGAGCCGAACCTGCCTTCTTTTCGGCCTGGAA
CTAATCATATGTGGCCTGGAGAAACAGCTAAAGTGCAGAAAGCGGCGGGCCGACCGACGCCCTTGA
45 CGATTTTGACTTAGACATGCTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCT
GACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAAGCGGCGCGACTCTAGATCATAA
TCAGCCATAACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCT
GAAAGCATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGACGTTATAATGGTTACAAATA
AAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTGTCC
50 AAATCATCAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCAGGGATG
TAATTACGTCCCTCCCCGCTAGGGGCAGCAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCT
CCCCCGCATCCCCGAGCCGCGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGA
TCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAA
AAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAG
55 CACAGTGCTCATCCAGATCCAACCCCTGCTATGTGACGGGTGATCAACCAGCAGCCAGGCTGCC
CAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCA
ACCTGTTCAGTGCGTCAACACCCTCTGGGGGAAAAAATGCCTCCTCATATCCAACCCAAACCTCCC
CTGTCTCAGTGTAAGGCCATTCCCCCTTGTCTATCAAGGGGAGTTGTGCTGTGACATTGTTGGTCT
GGGGTGACACATGTTTGCCAATTCACTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGC
60 AGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGT
TCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCAT
GGAGAGGAGCACAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGG

TGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGCTAGGAC
 TGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGA
 AAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCA
 5 GTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCT
 GTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCCCTCCCCTGG
 CACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGCA
 CCTAGAGGGACAGCCCCCCCCCAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGG
 AGCAGCGAGCCCGGGGCTCCGCTCCGGTCCCGCGCTCCCCCGCATCCCCGAGCCGGCAGCG
 10 TGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTG
 CTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTA
 GAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCC
 CTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGA
 ATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTG
 15 GGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCT
 TGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTGAG
 TGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGC
 AGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCCAGCGTTCAGAACAGCCTTAAGGATAAGAAG
 ATAGATAGAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAG
 20 ACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGG
 AAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCA
 TGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCA
 ACCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATC
 ACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGA
 25 GAAGGTAAATCTTGCTAAATCCAGCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCAT
 CCAACTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAA
 CGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
 AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATC
 AGTGATAGAGAACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTC
 30 CCTATCAGTGATAGAGAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCTATAAAAGCAGAGCT
 CGTTTAGTGAACCGTCAGATCGCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGT
 ACCACTTCCCTACCCTCGTAAAAAGCTTGTCCACCATGAGATTCAAAAGCCACACTGTGGAATTGAG
 GAGGCCCTGCAGCGACATGGAGGGAGCTGCTTTGCTGAGAGTCTCTGTCTCTGCATCTGGATGAG
 TGCATTTTCCCTTGGTGTGGGAGTGAGGGCAGAGGAAGCTGGAGCGAGGGTGCAACAAAACGTTT
 CAAGTGGGACAGATACTGGAGATCCTCAAAGTAAGCCCCCTCGGTGACTGGGCTGCTGGCACCATG
 35 GACCCAGAGAGCAGTATCTTTATTGAGGATGCCATTAAGTATTTCAAGGAAAAAGTGAGCACACA
 GAATCTGCTACTCCTGCTGACTGATAATGAGGCCTGGAACGGATTCTGTGGCTGCTGCTGAACCTGCC
 CAGGAATGAGGCAGATGAGCTCCGTAAAGCTCTGGACAACCTTGCAAGACAAATGATCATGAAAG
 ACAAAAACCTGGCAGATAAAGGCCAGCAAGTACAGAACTGGTTTCTGAAAGAGTTTCTCGGTTG
 AAAAGTGAGCTTGAGGATAACATAAGAAGGCTCCGTGCCCTTGACAGATGGGGTTTCAAGAGGTCCA
 40 CAAAGGCACCACCATCGCCAATGTGGTGTCTGGCTCTCTCAGCATTTCTCTGGCATCCTGACCCTC
 GTCGGCATGGGTCTGGCACCCTTACAGAGGGAGGCAGCCTTGTACTCTTGGAACCTGGGATGGA
 GTTGGGAATCACAGCCGCTTTGACCGGGATTACCAGCAGTACCATGGACTACGGAAAGAAGTGGT
 GGACACAAGCCCAAGCCACGACCTGGTCATCAAAGCCTTGACAAATTGAAGGAGGTGAGGGA
 GTTTTGGGTGAGAACATATCCAACCTTTCTTTCCTTAGCTGGCAATACTTACCAACTCACACGAGGC
 45 ATTGGGAAGGACATCCGTGCCCTCAGACGAGCCAGAGCCAATCTTCAGTCAGTACCGCATGCCCTC
 AGCCTCACGCCCACGAGTCACTGAGCCAATCTCAGCTGAAAGCGGTGAACAGGTGGAGAGGGTTA
 ATGAACCCAGCATCCTTGGAATGAGCAGGAGTCAAGCTCACGGATGTGGCCCTGTAAGCTTC
 TTTCTTGTGCTGGATGTAGTCTACCTCGTGTACGAATCAAAGCACTTACATGAGGGGGTCAAAGTCA
 GAGACAGCTGAGGAGCTGAAGAAGGTGGCTCAGGAGCTGGAGGAGAAGCTAAACATTCTCAACA
 50 ATAATTATAAGATTCTGCAGGCGGACCAAGAACTGTGAAATTCTAAAAATACAGCATAGCAAACT
 TTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGG
 GGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGT
 ATTTTCCCAAGGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCC
 55 TTTTATAGTAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAG
 GCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAA
 AGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTCGAGATGGTCCATATGAA
 TATCCTCCTTAGTTCCCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCCAAGCCCCAGGGATGT
 AATTACGTCCCTCCCGCTAGGGGACGAGCGAGCCGCGGGGCTCCGCTCCGGTCCGGCGCTC
 60 CCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGAT
 CGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAA
 AGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGC
 ACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCC

AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAA
CCTGTTTCAGTGCGTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCC
TGTCCTCAGTGTAAGCCATTCCCCCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTG
GGGTGACACATGTTTGCCAATTTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCA
5 GGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTT
CAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATG
GAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGT
GTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGACT
GGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAA
10 AAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAG
TGCCACATCCCCACAGTTCTTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTG
TGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCCACCCTCCCCTGGC
ACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCC
GAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATA
15 TCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAA
TTATCGTATGAGTTAAATCTTAAAGTCACGTAAAGATAATCATGCGTCATTTTGACTCACGCGG
TCGTTATAGTTCAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCG
GCGACTGAGATGTCCTAAATGCACAGCGACGGATTTCGCGCTATTTAGAAAAGAGAGAGCAATATTT
CAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACT
20 CGACCTAAACTTTAAACACGTTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGAGC
TCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGA
GTCGTATTACAATTCAGTGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCA
ACTTAATCGCCTTGACGACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGA
TCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGT
25 AAAATTCGCGTTAAATTTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAAT
CCCTTATAAATCAAAAGATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTC
CACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCA
CTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAAC
CCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAG
30 GGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAAGCGGTACGCTGCGCGTAAC
CACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 12)

TOIC-Cas9_Obl_r26_AAVS_SANeo

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATAT
35 GTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
GTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGCGGCATTTTGCCTCCTGTTTTTGCTCAC
CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTCCAATGATGAG
CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
40 TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCA
ACTTAATCTGACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGATC
ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
45 GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAG
50 GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
TGCTGCTTGCAACAACAAAAACCCAGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGTACCA
ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGGATACCAAATACTGTTCTTCTAGTGAG
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTG
55 TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
GAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
60 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCTGATTCTGTGGATAACC

GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAAGCGGGCAGTGAGCGCAACGCAATTAATG
5 TGAGTTAGCTCACTCATTAGGACACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
TTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTGTGACG
TACGTTAAAGATAATCATGCGTAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
TATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCACAAA
CAATTTATTTATGTTTATTTATTTATTTATTTAAAAAAAACAAAAACTCAAAATTTCTTCTATAAAGTAAC
10 AAAACTTTTTAAACATTCTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
TATTATAAAATAAGTAATTAGCTTAACTTATACATAATAGAAACAAATTATACTTATTAATCGCAT
TGATTATTGACTAGTCACAATATGATTATCTTTCTAGGGTTAATTAAGATATCTGAAGTTCTTATAC
TTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCAAGCTAGAGACCATTAAAGGTTCCG
GATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCGCCCTCAGAAGCCATAGAGCCCCACCGCATCCC
15 CAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTTGCTGTCTGCCCCACCCCCACCCCCAGAAT
AGAATGACACCTACTCAGACAATGCGATGCAATTTCTCATTTTATTAGGAAAGGACAGTGGGAGT
GGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAGATGGCTGGCAACTAGAA
GGCACAGTCGAGGCTGATCAGCGAGCTAGAGAATTGATCCCCTCAGAAGAATACTCGTCAAGAAAG
GCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCA
20 GCCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCTGATAGCGGTCC
GCCACACCCAGCCGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCCG
CAAGCAGGCATCGCCATGGGTACGACGAGATCCTCGCCGTCGGGCATGCGCGCCTTGAGCCTGG
CGAACAGTTCGGCTGGCGCGAGCCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGG
CTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTTCGAATGGGCAGGTAGCCG
25 GATCAAGCGTATGCAGCCGCCGCAATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTG
ACAACGTCGAGCACAGCTGCGCAAGGAACGCCGCTCGTGGCCAGCCACGATAGCCGCGCTGCCTC
GTCTGCGAGTTCATTTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCCTGCG
CTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCGAAT
30 AGCCTCTCCACCCAAGCGGCCGGAAGAACCTGCGTGCAATCCATCTTGTTCAATGGCCGATCCCATG
GCGGTATCGATAAGCTAGCTTGGGCTGCAGGTGAGGGACCTAATTAAGGGTTCGGGATCCACTA
GTTCTAGAGCGGCCTCGACTCTACGATACCGTCGATCCCCACTGGAAAGACCGCGAAGAGTTTGTG
CTCAACCGCGAGCTGTGGAAGAAAAAGGGACAGGATAAGTATGACATCATCAAGGAAACCTG
ACTACTGCGCCCTACAGATCCCTGAAGTTCCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGA
35 ACTTCAAAGATGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCTACTAGTTATTAATAGTAATC
AATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTTACATAAATTACGGTAAATGG
CCCGCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGAGCTCTAGAGGTGGACTATTTACGGTAAACTGCCCATTTGGC
AGTACATCAAGTGATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
40 CTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC
ATCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCC
CCACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGACGCGATGGGGGCGGGGGGGGGG
GGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGAGAGGTGCG
GCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCTTTATGGCGAGGCGGCGGGCGGGCG
45 GCCCTATAAAAAGCGAAGCGCGCGGGCGGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGC
TCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGC
GGGACGGCCCTTCTCTCCGGGCTGTAATTAGCGCTTGTTTAAATGACGGTCGTTTCTTTCTGTG
GCTGCGTAAAGGCTTAAAGGGCTCCGGGAGTGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGT
GCGTGCGTGTGTGTGCGTGCGTGCGGAGCGCCGCTGCGGCCCGCGCTGCCCGCGGCTGTGAGCGC
50 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGCGGT
GCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGCGGGG
GTGAGCAGGGGGTGTGGGCGCGGCGGTGCGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTT
GCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCTCGGGCCGGGGAGGGC
55 TCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGCAGCCA
TTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAATCTGGCGGAGCCG
AAATCTGGGAGGCGCCGCCGACCCCCCTAGCGGGCGCGGCGAAGCGGTGCGGCGCGCGGAGG
AAGAAATGGGCGGGGAGGGCCTTCGTGCTGCGCGCGCGCCGCTCCCTTCTCCATCTCCAGCCT
CGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGG
60 CGTGTGACCGGGCGGCTCTAGAGCCTCTGCTAACCATGTTTATGCTTCTTCTTTTCTTCTACAGCTCC
TGGGCAACGTGCTGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCA
AGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTC

AATGGAGTCGGTATCGAAGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCC
TACCCTGTACTGGCACGTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGG
ACAGGCATCATACCCACTCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTCTGCGGAACAAC
5 GCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGC
CCAACAGAGAAACAGTACGAAACCCTGGAATAATCAGCTCGCGTTCCTGTGTGTCAGCAAGGCTTCTC
CCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGA
ACAGGAGCATCAAGTAGCAAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCCACTTC
TGAAACAAGCAATTGAGCTGTTCCGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAAC
10 TAATCATATGTGGCCTGGAGAAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGAC
GATTTTGACTTAGACATGCTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTG
ACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAAGCGGCCGCGACTCTAGATCATAAT
CAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTG
AAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTTATTGCAGCTTATAATGGTTACAAATAA
15 AGCAATAGCATCACAAATTTCAAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTTGTCCA
AACTCATCAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAGCCCCCAGGGATGT
AATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCGGGGCTCCGCTCCGGTCCGGCGCTC
CCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCAGGGGAAGGTGGCAGCGGAT
CGCTTTCCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAA
20 AGCTTTAGGCTGAAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGC
ACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAA
CCTGTTCACTGCGTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCC
TGCTCAGTGTAAGGCCATTCCCCCTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTG
25 GGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCA
GGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTT
CAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAAGAGCAAGTTAAAACCCAGCATG
GAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGT
GTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCCTGGAGAGATACAGCTGGGTGAGTAGGACT
30 GGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAA
AAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAG
TGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTG
TGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCCAGCCCTCCCCTGGC
ACAACGTAAGGCCATTATCTCTCATCAACTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGAC
35 CTAGAGGGACAGCCCCCCCCCAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCA
GCAGCGAGCCGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGT
GCGGGGACAGCCCCGGGCACGGGGAAGGTGGCAGGGATCGCTTTCCCTCTGAACGCTTCTCGCTGC
TCTTTGAGCCTGCAGACACTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGAAATTTAG
AATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCAGATGCTCATCCAGATCCAACCCCC
40 TGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCCAGAGCCACATCCAGCCTGGCCTTGA
ATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCACTGCGTCACCACCCTCTG
GGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCT
TGCTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAG
TGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGAGGACAGCATGGACGTGGGACATGC
45 AGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAG
ATAGGATAGAAGGACAAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAG
ACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGG
AAGAGCTTGCTGGAGAGATACAGCTGGGTACAGTAGGACTGGGACAGGCAGCTGGGAAGATTGCCA
TGATAGATGTTCATACAATCGTCAAATCGTAGGAGTGGAAGGCCCTCCAAGATCCCCAAGGCCA
50 ACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATC
ACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGA
GAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCAT
CCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAA
CGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
55 AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATC
AGTGATAGAGAACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTC
CCTATCAGTGATAGAGAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCT
CGTTTAGTGAACCGTCAGATCGCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGT
ACCTTCTCCTACCCTCGTAAAAAGCTTGTCACCACTTAGGCTCCTAAGAAAAAGCGGAAGGTGGACA
60 AGAAATACTCAATCGGGCTGGACATCGGAACCTCACTCAGTGGGGTGGGCAGTCATTACTGACGAG
TACAAAGTGCCAAGCAAGAAATTTAAGGTCTGGGCAACACCGATAGGCACTCCATCAAGAAAAA
TCTGATTGGGGCCCTGCTGTTGACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTG
CTAGAAGGCGCTATACCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAA

ATGGCCAAGGTGGACGATAGTTTCTTTCATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAA
GAAACACGAGAGACATCCTATCTTTGGAACATTGTGGACGAGGTCGCTTATCACGAAAAATACC
CCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCT
5 ATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGGCCATTTTCTGATCGAGGGAGATCTGAACCCAG
ACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAA
AACCCTATTAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAG
AAGGCTGGAGAACCTGATCGCTCAGCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGGAATCTGA
TTGCATGTCTCAGTGGGACTGACACCTAACTTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAAAC
10 TGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAAACCTGCTGGCTCAGATCGGAGATCAG
TACGCAGACCTGTTCTGCGGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGG
GTGAACACCGAGATTACAAAAGCCCCTCTGTCTGAGCTAGCATGATCAAGAGATATGACGAGCACC
TCAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCAGAGAAGTACAAGGAAATCT
TCTTTGATCAGAGTAAGAACGGGTACGCCGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCT
15 ACAAGTTTATCAAACCTATTCTGGAGAAGATGGACGGCACCGAGGAACTGCTGGTGAAACTGAAT
CGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGATAACGGCTCCATCCCCACCAGATTTCATCTG
GGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGA
GAAGATCGAAAAAATTCTGACCTTCCGATCCCTTACTATGTGGGGCCCCTGGCAAGGGGTAATTC
CCGCTTTGCTGGATGACACGGAAATCAGATGGCTAGGAAACAATCACTCCTTGGAATTCGAGGAAGTGG
20 TCGATAAGGGAGCTTCCGCACAGTCTTTCATCGAGAGAATGACAAACTTCGACAAAAACCTGCCA
AATGAGAAAGTGCTGCCTAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAAGTACT
AAGGTGAAATACGTCACCGAGGGGATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAG
CTATCGTGGACCTGCTGTTTAAACCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTAC
TTCAAGAAAATTGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCT
25 CTGGGAACCTACCACGATCTGCTGAAGATCATTAAAGGATAAAGACTTCTGGACAACGAGGAAAA
TGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGG
AACGGCTGAAAACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGG
TACACCGCTGGGGACGACTGAGCAGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGAA
AGACTATCCTGGACTTTCTGAAATCAGATGGCTTCGCTAACAGAAAATTTTATGCAGCTGATTACAG
ATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTG
30 CACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAA
GGTGGTTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGG
CCCGCGAAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGA
GGAAGGCATTAAGGAACTGGGATCCAGATCCTGAAAGAGCACCCCTGTGGAAAACACTCAGCTGC
AGAATGAGAAGCTGTATCTGTACTATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTG
35 GATATTAACCGACTGTCTGATTACGACGTGGATCATATCGTCCACAGTCATTCTGAAAAGATGAC
AGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTC
AGAGGAAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAAACTGATCACA
CAGCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTGCAGAACTGGACAAAGCAGG
CTTCATCAAGCGACAGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATA
40 GCAGGATGAACACAAAGTACGATGAGAATGACAACTGATCCGCGAAGTGAAGGTCAATTACTCTG
AAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCAGTTCTACAAAGTCAGGGAGATCAACAA
TTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTTCGGGACCGCCCTGATTAAGAAATACCC
CAAACCTGGAGAGCGAATTCGTGTACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCA
AGAGTGAGCAGGAAATTGGAAGGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCT
45 TTAAGACTGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGC
GAGACCGGAGAAATCGTGTGGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCTGTCCAT
GCCTCAAGTGAATATTGTCAAGAAAAACAGAGGTGCAGACTGGCGGATTTCAGTAAGGAATCAATTC
TGCCGAAACGAACTCTGATAAGCTGATCGCCCGGAAAGAAAGACTGGGATCCCAAGAAATATGGG
GGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTTCGCAAGGTGGAGAAGGGGAAAAAG
50 CAAGAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGA
AGAACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAAA
CTGCCAAAGTACAGCCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGAGGAGA
GCTGCAGAAGGGTAATGAACTGGCCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCA
CTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACA
55 AGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGAC
GCAAATCTGGATAAGGTCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGG
CCGAAAATATCATCTGTTCACTCTGACCACTGGGAGCCCCCTGCAGCCTTCAAGTATTTTG
ACACTACCATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTCAC
CAGAGCATTACCGGCCTGTATGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGG
60 AAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTGAGGAGAATCCTGGCCCAGCA
CCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCT
GGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG

GCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCCACCCTCGTGA
 CCACCTTCACCTACGGCGTGAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCAGGACTTCT
 TCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAAC
 5 TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG
 CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACATAACAGCCACA
 AGGTCTATATACCCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGACCCGCCACAAC
 ATCGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCC
 CGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAACGAGA
 AGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAG
 10 CTGTACAAGTAAACCTAATCTAGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCC
 ATCTGTTGTTTGGCCCTCCCCCGTGCTTCCCTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCCT
 AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG
 GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTC
 TATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCATA
 15 TGAATATCCTCCTTAGTTTCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGG
 ATGTAATTACGTCCCTCCCCCGTAGGGGACAGCGAGCCGCCCGGGGCTCCGCTCCGGTCCGGC
 GCTCCCCCGCATCCCCGAGCCGGCAGCGTGCAGGGGACAGCCGGGCACGGGGAAAGGTGGCAGCG
 GATCGCTTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGA
 AAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTGCAAAGG
 20 AGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTG
 CCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGC
 AACCTGTTCACTGCGTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCC
 CCTGTCTCAGTGTAAGGCCATTCCCCCTTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTC
 TGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTG
 25 CAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCG
 TTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCA
 TGGAGAGGAGCACAAGGAGGCGACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATG
 GTGCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGA
 CTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGG
 30 AAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCACTGGCCATGTCCCTC
 AGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGC
 TGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCCACCCTCCCCTG
 GCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGAT
 CCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGA
 35 TATCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATAT
 AATTATCGTATGAGTTAAATCTTAAAGTACAGTAAAGATAATCATGCGTCATTTTGACTCACGC
 GGTGCTTATAGTTCAAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAG
 CGGCTAGTGAATGTCCTAAATGCACAGCGACGGATTTCGCGCTATTTAGAAAGAGAGCAATAT
 TTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTA
 40 CTCGACCTAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGA
 GCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGT
 GAGTCGTATTACAATTCAGTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACC
 CAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGGCCCGCACC
 GATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTG
 45 TTAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAA
 ATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAG
 TCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCC
 CACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGA
 ACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGA
 50 AGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGATAGCGGTACAGCTGCGCGTA
 ACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 13)

TOIC_Cas9Dead_Obl_r26_AAV_PgkNeo

GTGGCACTTTTCGGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATAT
 55 GTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
 GTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGGCGCATTTTGCCTTCCTGTTTTTGCTCAC
 CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
 ACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTTCGCCCGAAGAACGTTTTCCAATGATGAG
 CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
 60 TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTACCAGTCACAGAAAAGCATCTTAC
 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCA

ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATC
ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
5 GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAAAG
10 GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCTGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAACCCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTG
15 TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAAGTGAATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
GAAAGGCGGACAGGTATCCGGTAAGCGGCGAGGTGCGAACAGGAGAGCGCAGGAGGAGCTTCC
AGGGGAAACGCCTGGTATCTTTATAGCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
20 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
CCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGCTGGCAGCAGAGGTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
25 TGAGTTAGCTCACTCATTAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
TTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCTAGAAAGATAATCATATTGTGACG
TACGTTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
TATTAATTTGAATAGATATTAAGTTTTATTATATTATTTACACTTACATACTAATAATAAATTCACAAA
30 CAATTTATTTATGTTTATTTATTTATTTATTTAAAAAACAACAAACTCAAAATTTCTTCTATAAAGTAAC
AAAACCTTTTAAACATTCTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
TATTATAAATAAGTAATTAGCTTAACTTATACATAATAGAAACAAATTATACTTATTAATCGCAT
TGATTATTGACTAGTCGTATTAAGGGTTCGGATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCG
CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
35 GCTGTCTGCCCAACCCACCCCCAGAAATAGAATGACACCTACTCAGACAATGCGATGCAATTTT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
GGCAAACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
TTGATCCCCTCAGAAGAAGCTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG
CGATACCGTAAAGCACAGGAGGAAGCGGTACGCCATTTCGCCGCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCGCCACAGTCGATGAATCCAGA
40 AAAGCGGCCATTTTCCACCATGATATTTCGGCAAGCAGGCATCGCCATGGGTACGACGAGATCCTC
GCCGTCCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCCTGATGCTCTTC
GTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
CGCTTGGTGGTCAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCA
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCACTTCCGCC
45 AATAGCAGCCAGTCCCTTCCCGCTTCACTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCCG
CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTACAGGCACCCGACAGGTCCGT
CTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGCTAGTTGTGCCAGTCATAGCCGAATAGCCCTCTCCACCCAAAGCGCCGAGAACCTGCGTGC
AATCCATCTTGTTCATGCGCATCCCATGGTTTAGTTCTCTACCTTGTCTGATTATACTATGCCGA
50 TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTGCGAAAGGCCCGGAGATGAGGA
AGAGGAGAACAGCGCGGCAGACGTGCGCTTTGAAGCGTGCGAATGCCGGGCTCCGGAGGACC
TTCGGGCGCCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCCCGGACCCACCCCTTCCAGCC
TCTGAGCCCAGAAAGCGAAGGAGCAAGCTGCTATTGGCCGCTGCCCAAAGGCCCTACCCGCTTC
CATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTACGTC
55 CTGCACGACGCGAGCTGCGGGGCGGGGGGGAACCTTCTGACTAGGGGAGGAGTAGAAGGTGGCG
CGAAGGGGCCACCAAGAACGGAGCCGTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCC
AGAGGCCACTTGTGTAGCGCCAAGTGCCAGCGGGGTGCTAAAGCGCATGCTCCAGACTGCCTT
GGGAAAAGCCGCTCCCTACCCGCTAGACACCCACAGTCAGTGGGTGGCCTAGGGACAGGATTGCAAC
TCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAG
60 TTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC

ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGT
ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCAATTTTGT
5 TTTATTTATTTTTTAATTATTTTGTGACGCGATGGGGGCGGGGGGGGGGGGGGGCGCGCGCCAGGCG
GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGGCAGCCAATCAGA
GCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGGGCGGCCCTATAAAAAGCGA
AGCGCGCGGCGGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCTCGCGC
CGCCCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCT
10 CCGGGCTGTAATTAGCGCTTGGTTTAATGACGGCTCGTTTCTTTCTGTGGCTGCGTGAAAGCCTTA
AAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
CGTGGGGAGCGCCGCTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGG
GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGGG
GGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGG
15 GCGCGGCGGTGCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCCGGCT
TCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA
GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCGGGAGGGCTCGGGGAGGGGCGCGGC
GGCCCCGAGCGCCGGCGGTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGTTAATC
TGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
20 CGCACCCCCCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGCGCAGGAAGGAAATGGGCGGGGA
GGGCCTTCGTGCGTCGCCGCGCCGCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGG
GACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTGTGACCGGCGGCTC
TAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
ATTGTGCTGTCTCATATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTC
25 TCTAGACTGGACAAGAGCAAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAG
TGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCAC
TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGT
GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTA
30 CGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
CGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
AAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
TGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGG
AGAAACAGCTAAAGTGCGAAAGCGGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATG
35 CTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTG
ACCTTGACATGCTCCCCGGGTAAAGCGGGCGCGACTCTAGATCATAATCAGCCATACCACATTTGT
AGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
AATTGTTGTTGTTAACTTGTTTATTGACGCTTATAATGGTTACAAATAAAGCAATAGCATCACAA
TTTCAGAAATAAAGCATTTTTTTCATGCACTTCTAGTTGAGTTGTGGTTTGTCCAAACTCATGATCT
40 TAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCC
GCTAGGGGACAGCAGCGAGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG
CCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGC
TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAG
AGAGATTTAGAATGACAGAATCATAGAACGGCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGA
45 TCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCC
TGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCAC
CACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCC
ATTCCCCTTGTCCTATCAAGGGGAGTTTGTCTGTGACATTGTTGGTCTGGGAGTACACATGTTTGC
CAATCCAGTGCATCAGAGAGGCAGATCTTGGGATAAGGAAGTGCAGGACAGCATGACGTGG
50 GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA
TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGTGTTGTTGTTGTTGTTGTTGTTG
GGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCC
55 AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCGAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
CTCTCATCCAACCTCCAGGACGGAGTCAGTAGGATGGGGCTGTGACCTAGAGGGACAGCCCCC
CCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCGCTAGGGGACAGCAGCGAGCCCGGGG
60 CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCA
CGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACAC
CTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAA
CGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATC

AACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGC
ATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGGGAAAAACTGCCTCCTC
ATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGCTTATCAAGGGGGAGT
5 TGTCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGA
TCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
CACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
GAGCCTGCATGTTTGTGATGGTGTCTGGATGCAAGCAGAAAGGGGTGGAAGAGCTTGCCTGGAGAGAT
10 ACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTG
CCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCC
CCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCCCGACCCTCCCCGGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCA
15 GTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTA
CTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC
AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
ATAAGCTTTAGGCGTGTACGGTGGGCGCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATC
20 GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAA
AAGCTTGTCCACCATGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGG
CCATCGGAACCTAAGTGCAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAA
TTTAAGGTCCTGGGCAACACCGATAGGCCTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTTT
GACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCG
25 AAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTT
CTTTTCATCGGCTGGAGGAATCATTCTGCTGAGGAAGATAAGAAACACGAGAGACATCCTATCTT
TGGAACATTGTGGACGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGA
AACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTA
AGTTAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTG
30 TTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTATTAATGCATCTGGCGTG
GACGCAAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCCTGATCGCTCA
GCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGAAATCTGATTGCACTGTCAGTGGGACTGACAC
CTAACTTCAAGAGCAATTTTGTCTGCGCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTAT
GACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTTCTGGCCGCT
35 AAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCC
CCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGC
ACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGT
ACGCCGTTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACCGCACCAGGAACTGCTGGTGAACAGAGTGAATCGCGAGGACCTGCTGCGCAAGCA
40 GCGGACATTTGATAACGGCTCCATCCCCACCAGATTCATCTGGGAGAGCTGCACGCAATCCTGCG
ACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCT
TCCGGATCCCTTACTATGTGGGGCCCTGGCAAGGGGTAAATCCCGCTTTGCCTGGATGACACGGA
AATCTGAGGAAACAATCACTCCTTGAACTTCGAGGAAGTGGTTCGATAAGGGAGCTTCCGCACAG
TCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCCTAAGCA
45 CAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGG
GATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAAA
CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAATTGAATGTTTCGAT
TCTGTGAGATCAGTGGCGTCAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
AAGATCATTAAAGGATAAAGACTTCTGAGACACGAGGAAAAATGAGGATATCCTGGAAGACATTGT
50 GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTTATGCCATCT
GTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAGC
AGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATC
AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGA
TATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTGCACGAGCATATCGAAACCTGGCCG
55 GGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTTCGATGAGCTGGTGAAAGTC
ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCAAAATCAGACCACACAGA
AAGGACAGAAGAAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
CCAGATCCTGAAAGAGCACCCCTGTGGAACACCTCAGCTGCAGAAATGAGAAGCTGTATCTGTACT
ATCTGCAGAAATGGCGGGGATATGACGTGGACCAAGGAGCTGGATATTAACCGACTGTCTGTATAC
60 GACGTGGATGCCATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGAC
CCGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATG
AAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGCGAAAGTTTGATAACCTGAC
TAAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG

AGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGAT
GAGAATGACAAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTT
CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCAT
ACCTGAACGCAGTGGTCGGGACCGCCCTGATTAAGAAATACCCCAAACTGGAGAGCGAATTTCGTG
5 TACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTGGA
AGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCTGGCA
AATGGGGAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
ACAAGGGTAGGGATTTTGCCACAGTGCAGCAAGGTCCCTGTCCATGCCTCAAGTGAATATTGTCAAGA
AAACAGAGGTGCAGACTGGCGGATTCAAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAG
10 CTGATCGCCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
TTACTCTGTCCTGGTGGTTCGCAAAGGTGGAGAAGGGGAAAAGCAAGAACTGAAATCCGTCAAGG
AGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGG
CTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAG
CTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAACTGG
15 CCCTGCCTTCTAAGTACGTGAACCTTCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCC
CCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATT
GAACAGATTAGCGAGTTCTCCAAAAGAGTATCCTGGCTGACGCAATCTGGATAAGGTCTGTGAG
CGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCGCAAAATATCATTCTGTTTCA
CTCTGACCAACCTGGGAGCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
20 ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACCCAGAGCATTACCGGCCTGTATGAA
ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTC
TGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGC
GAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAA
GTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCT
25 GCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTTACCTACGGCGTGCAGT
GCTTCGCCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGGCT
ACGTCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGAGGTGAAG
TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAAGGTCTATATCACCGCCGACAAGC
30 AGAAGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGG
AGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGC
AGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTG
35 CCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCTTAATAAAATGAGGAAATTGCATCG
CATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATGTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA
CCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCAATGTTCCATATGAATATCCTCCTTAGTTCCTATT
CCGCTAGCCTAGAGGGACAGCCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCT
40 AGGGGCAGCAGCGAGCCGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCG
GCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTCTCTGAACGCTTC
TCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGA
GATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCC
AACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTG
45 GCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTACCA
CCCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCAT
TCCCCCTGTCTTATCAAGGGGGAGTTTCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCA
ATTCAGTGGGGCTCAGGAGAGGCAGATTTGTTGGGATAAGGAAGTGCAGGACAGCATGGACGTGG
ACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTTCAAGAACAGCCTTAAGGAT
50 AAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGAAGAGCTTGCCTGGAGAGATAACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCC
AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
55 CTTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
CTCTCATCCAATCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTAC
ACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAAATAT
ATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATC
60 TTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTATAGTTCAAAATCA
GTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAA
ATGCACAGCGACGGATTTCGCGCTATTTAGAAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAA

TTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACAC
GTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCG
GGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTACAATTCAGTG
GCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGACGCA
5 CATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTT
TGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAA
TAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGA
CTCCAACGTCAAAGGGCGAAAAACCCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCT
10 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGA
TTTAGAGCTTGACGGGGAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAG
CGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTT
AATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 14)

15 TOIC_Cas9_KRAB_Obl_r26_AAVS_PGKNeo

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATAT
GTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
GTATTCACATTTCCGTGTCGCCCTTATTCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCAC
20 CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAG
CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
GGATGGCATGACAGTAAGAGAATTATGCAGTGTGTCGCATAACCATGAGTGATAACACTGCGGCCA
ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGACAACATGGGGGATC
25 ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
CATTCGAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
30 GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
AACTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTAAATTTAAAG
GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
35 ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGACCCGCTACATACCTCGCTCTGCTAATCCTG
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
40 GAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
AGGGGGAACCGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGACCTATGGAACAAACGCCAGCAACGCGGCTTTTACGGTT
CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCA
45 GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGCTGGCAGCAGAGGTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
TTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTGTGACG
50 TACGTTAAAGATAATCATGCGTAAATTTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
TATTAATTTGAATAGATATTAAGTTTTATTATTTTACACTTACATACTAATAATAAATTCACAAA
CAATTTATTTATGTTTATTTATTTATTAATAAAAAAAACAAAAACTCAAAATTTCTTCTATAAAGTAAC
AAAACCTTTTAAACATTCTCTTTTACAAAAATAAACTTATTTGTACTTTAAAAACAGTCATGTTG
TATTATAAAATAAGTAATTAGCTTAACCTTATACATAATAGAAACAAATTATACTTATTAATCGCAT
55 TGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCG
CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
GCTGTCCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
GGCAAACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
60 TTGATCCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG
CGATACCGTAAAGCACGAGGAAGCGGTGAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGA

AAAGCGGCCATTTTCCACCATGATATTTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTC
 GCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTC
 GTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
 5 CGCTTGGTGGTCAATGGGACAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATCAGCCA
 TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCC
 AATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAACGCCCCGT
 CGTGCCACAGCCAGATAGCCGCGCTGCCTCGCTGCAGTTCATTAGGGACCCGGACAGGTCCGT
 CTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
 10 ATTGCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGC
 AATCCATCTTGTTCATATGGCCGATCCCATGGTTTAGTTCCTCACCTTGTCGTATTATACTATGCCGA
 TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTGCGAAAGGCCCGAGATGAGGA
 AGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGTGCAGAATGCCGGGCCTCCGGAGGACC
 TTCGGGCGCCCCGCCCCGCCCCCTGAGCCCCGCCCTGAGCCCCGCCCCCGGACCCACCCCTTCCAGCC
 15 TCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCAAAGGCCCTACCCGCTTC
 CATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCACGTC
 CTGCACGACGCGAGCTGCGGGGCGGGGGGGAACCTTCTGACTAGGGGAGGAGTAGAAGGTGGCG
 CGAAGGGCCACCAAAGAACGGAGCCGCTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCC
 AGAGGCCACTTGTGTAGCGCAAGTGCCGAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTT
 20 GGGAAAAAGCGCCTCCCCTACCCGCTAGACACCCACAGTGGGTGGCCTAGGGACAGGATTGCAAC
 TCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAG
 TTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGC
 CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
 TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
 25 ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGT
 ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGG
 TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGTA
 TTTATTTATTTTAAATATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCG
 GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGCGAGCCAATCAGA
 30 GCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAAGCGA
 AGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCCTCGCGC
 CGCCCCGCCCCGCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCT
 CCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTA
 AAGGGCTCCGGGAGGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
 35 CGTGGGGAGCGCCGCTGCGGCCCGCGCTGCCCGCGGGCTGTGAGCGCTGCGGGGCGGCGCGGGG
 GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGGG
 GGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGG
 GCGCGGCGGTGCGGTGTAAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCAGGCCCGGCT
 TCGGGTGCGGGGCTCCGTGCGGGGCTGCGCGCGGGGCTGCGCGTGCCTGCGGGGCGGGGTGCGGCA
 40 GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCGCGGC
 GGCCCCGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCG
 TGCGAGAGGGGCGAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
 CGCACCCCTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGA
 GGGCCTTCGTGCGTCCCGCGCCCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGG
 45 GACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTC
 TAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
 ATTGTGCTGTCTCATATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTC
 TCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
 50 AGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAGC
 TGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCAC
 TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGT
 GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTA
 CGAAACCCTGGAATAATCAGCTCGCGTTCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
 55 CGCTCTGTCCCGCTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
 AAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
 TGTTGACCGGCAGGGAGCCGAACCTGCCTTCTTTTCGGCCTGGAACATAATCATATGTGGCCTGG
 AGAAACAGCTAAAGTGCGAAAGCGGCGGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATG
 CTCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTG
 60 ACCTTGACATGCTCCCCGGGTAAAGCGGCGCGACTCTAGATCATAATCAGCCATACCACATTTGT
 AGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
 AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA
 TTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT
 TAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAGCCCCCAGGGATGTAATTACGTCCCTCCCCC

GCTAGGGGACAGCAGCGAGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG
CCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGTTTCTCTGAACGC
TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGAAAAAGCTTTAGGCTGAAAG
5 AGAGATTTAGAATGACAGAATCATAGAACGGCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGA
TCCAACCCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCC
TGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTAC
CACCCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCAACCTCCCCTGTCTCAGTGTAAGGCC
ATTCCCCCTTGTCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGC
10 CAATTCAGTGTCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGG
GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA
TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCC
15 AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCCCTCCCCTGGCACAACGTAAGGCCATTAT
CTCTCATCCAACCTCAGGACGGAGTCAGGATGGGCTGGGACCTGACCTAGAGGGACGCCCCCC
CCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCCGGGGG
20 CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCA
CGGGGAAGGTGGCACGGGATCGCTTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACAC
CTGGGGGGATACGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAA
CGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATC
AACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGC
25 ATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCACCAACCTCTGGGGGAAAAACTGCCTCCTC
ATATCCAACCAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCTGTCTATCAAGGGGGAGT
TTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAAATTCAGTGTCATCAGGAGAGGCAGA
TCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
CACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
30 GCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
GAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAAGGGGTGGAAGAGCTTGCCTGGAGAGAT
ACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTG
CCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCC
35 CCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCA
GTGAGTAGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATGAGGAGTTTA
CTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC
40 AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
ATAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAA
AAGCTTGTCCACCATGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGG
ACATCGGAATAACTCAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAA
45 TTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTTT
GACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCG
AAAGAATCGCATCTGCTACCTGCAGGATATTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTT
CTTTCATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCTTCTT
TGGAACATTGTGGACGAGGTTCGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGA
50 AACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTA
AGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTG
TTCATCCAGCTGGTCCAGACATAACAATCAGCTGTTTGAGGAAAAACCTATTAATGCATCTGGCGTG
GACGCAAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCGCTCA
GCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGGAATCTGATTGCACTGTCACTGGGACTGACAC
55 CTAACCTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTAT
GACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTTCTGGCCGCT
AAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCC
CCTCTGTCACTAGATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCTTCTGAAGCT
ACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGT
60 ACGCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAAGTCTGCTGGTGAAGTGAATCGCGAGGACCTGCTGCGCAAGCA
GCGGACATTTGATAACGGCTCCATCCCCACAGATTCTGTTGGAGAGCTGCACGCAATCTGCG

ACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCT
 TCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGGGTAAATCCCGCTTTGCCTGGATGACACGGA
 AATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGATAAGGGAGCTTCCGCACAG
 5 TCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCCTAAGCA
 CAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGG
 GATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAAA
 CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAATTGAATGTTTCGAT
 TCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
 AAGATCATTAAGGATAAAGACTTCCTGGACAACGAGGAAAATGAGGATATCCTGGAAGACATTGT
 10 GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTTATGCCCATCT
 GTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGTGGGGACGACTGAGC
 AGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAAGACTATCCTGGACTTTCTGAAATC
 AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTACGATGACAGCCTGACCTTCAAAGAGGA
 TATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCG
 15 GGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTCGATGAGCTGGTGAAAGTC
 ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCGAAAATCAGACCACACAGA
 AAGGACAGAGAAGACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
 CCAGATCCTGAAAAGAGCACCCCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGTATCTGACT
 ATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTAC
 20 GACGTGGATCATATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACC
 CGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAATGA
 AGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGCGAAAGTTTGATAACCTGACT
 AAAGCTGAGAGAGGGGGTCTGTGCAAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG
 AGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGAT
 25 GAGAATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTT
 CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCAT
 ACCTGAACGCAGTGGTCGGGACCGCCCTGATTAAGAAATACCCCAAACCTGGAGAGCGAATTCGTG
 TACGGTGACTATAAAGGTGTACGATGTGCAAAAAAATGATCGCCAAGAGTGAGCAGGAAATTGGAAA
 AGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCCCTGGCA
 30 AATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
 ACAAGGGTAGGGATTTTGCCACAGTGCAGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGA
 AAACAGAGGTGCAGACTGGCGGATTTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTTGATAAG
 CTGATCGCCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
 TTAATCTGTCTGCTGGTGGTGCAGAAAGGTGGAGAAGGGGAAAAGCAAGAACTGAAATCCGTCAAGG
 35 AGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGG
 CTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAG
 CTGGAACCGAAGGAAGCAAGTGTGCTGACCTCCGAGGAGCTGCAGAAGGGTAATGAACCTGG
 CCTGCCTTCTAAGTACGTGAACCTCTGTATCTGGCTAGCCACTACGAGAAGTGAAGAGGCTCCC
 CCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATT
 40 GAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCTGAG
 CGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAATATCATTATCTGTTCA
 CTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
 ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACACAGAGCATTACCGGCCTGTATGAA
 ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGATGGCGGTGGCGCGCTGTCCCCGCA
 45 GCACTCCGCCGTGACCCAGGGGAGTATAATCAAAAACAAAGAGGGCATGGATGCTAAGAGCCTTA
 CCGCCTGGTCCCGAACACTGGTCACGTTAAGGATGTGTTTCGTCGATTTTACCCGGGAGGAGTGGA
 AACTGCTCGACACCGCAGCAGATCGTGTAACCGAATGTCATGCTCGAAAATTACAAAAACTTG
 GTGAGCTCGGGTACCAATTGACCAAACTCAGATGCAATGCTCGACTGGAAAAAGGAGAGGAACC
 CTGGCTCGTCGAGCGCGAAATTCATCAAGAAACACACCCGATTCTGAAACCGCCTTCGAGATTA
 50 AGAGCAGTGTGCCTAGGCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGA
 CGTCGAGGAGAATCCTGGCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCG
 GGGTGGTGGCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTACGCGTGTCCGGC
 GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCT
 GCCCCTGCCCTGGCCACCCCTCGTGACCACCTTACCTACGGCGTGCAAGTCTCGCCCGCTACCC
 55 CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCA
 CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
 CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAA
 GCTGGAGTACAACCTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAGAAGAACGGCATCA
 AAGTGAACCTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAG
 60 CAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTC
 CGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCG
 CCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGCAGCTCGCTGATCAGCC

TCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGG
 AAGGTGCCACTCCCCTGTCCTTTCCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGT
 GTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAG
 CAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGA
 5 TCCTCTAGTTGGCGCGTCATGGTCCATATGAATATCCTCCTTAGTTTCTATTCCGCTAGCCTAGAGG
 GACAGCCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCGCTAGGGGCAGCAGCGA
 GCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTCCGGGA
 CAGCCCCGGGCACGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGTGTCTTTGAG
 CCTGCAGACACCTGGGGGGATACGGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACA
 10 GAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATG
 TGCAGGGTCAACAACAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTG
 CAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTGGGGGAA
 AAAGTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCTTA
 TCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGTCATCA
 15 CGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTT
 GAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTTCAAGACAGCCTTAAGGATAAGAAGATAGGAT
 AGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGC
 TGGTCCCTGTGTCTGAGCCTGTCATGTTTGTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCT
 TGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCATGTAGAT
 20 GTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAA
 CCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAG
 GGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTA
 AATCTTGCTAAATCCAGCCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTC
 CAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGA
 25 CGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAAATATATATATAATAAGTT
 ATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCTTAAAGTCACGTA
 AAAGATAATCATGCGTCATTTTGACTCACGCGGTGTTATAGTTCAAAATCAGTGACACTTACCGC
 ATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGG
 ATTCGCGCTATTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTAT
 30 CTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACACGTCATAGAATCTTC
 GTTTGACAAAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGGAGGTTCCCTTT
 AGTGAGGGTTAATTGCGGGTCCGCCTATAGTGAGTCGTATTACAATTCAGTGCCGTCGTTTACAC
 ACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACCTAATCGCCTTGACAGCACATCCCCCTTTCGC
 CAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG
 35 GCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC
 ATTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGATAG
 GGTGAGTGTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAAGCGTGGAAGTCCAAAGTCA
 GCGAAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTGTGG
 GGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCCCCGATTTAGAGCTTGACGG
 40 GGAAAGCCGGCGAAGCTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCG
 CTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAACACACCCGCGCGCTTAATGCGCCGCTACA
 GGGCGCGTCAG (SEQ ID NO: 15)

TOIC_Cas9_Dead_VP64_Obl_r26_AAVS_SANeo

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATAT
 GTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
 GTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTGGCGCATTTTGCCTTCCTGTTTTTGTCTAC
 CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
 ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTCCAATGATGAG
 50 CACTTTTAAAGTTCTGCTATGTGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
 TCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTAC
 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCA
 ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATC
 ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
 55 ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
 GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
 GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
 CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
 GGCAACTATGGATGAACGAAATAGACAGATCGTGTAGATAGGTGCCTCACTGATTAAGCATTTGGT
 60 AACTGTAGACCAAGTTTACTCATATATACTTATAGATTGATTTAAACTTCAATTTTAAAG
 GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC

TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAACACCGCTACCAGCGGTGGTTTGTGGCCGATCAAGAGCTACCA
ACTCTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
5 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTG
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGATAAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
GAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACAGAGGGAGCTTCC
10 AGGGGGAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
CCTGGCCTTTTGTGTCCTTTTGTGTCACATGTTCTTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACGCCTCTCCCCGCGCGTTGGCCGATTCA
15 TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
TTAACCCCTACTAAAGGGAACCTCCCCTAGCTTAATTAACCCCTAGAAAGATAATCATATTGTGACG
TACGTTAAAGATAAATCATGCGTAAATTTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
20 TATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCACAAAA
CAATTTATTTATGTTTATTTATTTATTAATAAAAAAACAAAACTCAAAATTTCTTCTATAAAGTAAC
AAAATTTTAAACATTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
TATTATAAATAAGTAATTAGCTTAACCTTATACATAATAGAAACAAATTATACTTATTAATCGCAT
TGATTATTGACTAGTCGTATTAAGGGTTCGGATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCG
25 CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
GCTGTCCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
GGCAAAACAACAGATGGCTGGCACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
TTGATCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAAGGCGATGCGCTCGGAATCGGGAGCGG
30 CGATACCGTAAAGCACGAGGAAGCGGTGAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCCACAGTCGATGAATCCAGA
AAAGCGGCCATTTTCCACCATGATATTGCGCAAGCAGGCATCGCCATGGGTACGACGAGATCCTC
GCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCCTGATGCTCTT
GTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
35 CGCTTGGTGGTGAATGGGCAGGTAGCCGATCAAGCGTATGCAGCCGCCGCAATTGCATCAGCCA
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCC
AATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAACGCCCCG
CGTGCCAGCCAGCATAGCCGCGCTGCCTGCTGCACTTCAGTTCATTACAGGGCACCGGACAGGTGCGT
40 CTTGACAAAAAGAACCGGGCGCCCTGCGCTGACAGCGGGAACACGGCGGCAGCGGACGCGG
ATTGCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGC
AATCCATCTTGTTCATATGGCCGATCCCATGGCGGTATCGATAAGCTAGCTTGGGCTGCAGGTGAG
GGACCTAATTAAGGGTTCCGGATCCACTAGTTCTAGAGCGGCCCTCGACTCTACGATACCGTCGATC
CCCACTGGAAAGACCGCGAAGAGTTTGTCTCAACCGCGAGCTGTGGAAAAAAGGGACAGGAT
AAGTATGACATCATCAAGGAAACCCTGGACTACTGCGCCCTACAGATCCCTGAAGTTCTATACTT
45 TCTAGAGAATAGGAACTTCGGAATAGGAACCTCAAAGAACCGGTACCCACAGTGGGTGGCCTAG
GGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGC
CCGCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTA
ACGCCAATAGGACTTTCCATTGACGTCAATGGGTGGACTATTACGGTAAACTGCCCACTTGGCA
50 GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCG
TGGCATTATGCCAGTACATGACCTTATGGGACTTTCTTACTTGGCAGTACATCTACGTATTAGTCA
TCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCC
CACCCCCAATTTTGTATTTATTTATTTTAAATTTTGTGTCAGCGATGGGGGCGGGGGGGGGGGG
GGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGAGAGGTGCGG
55 CGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCCTTTATGGCGAGGCGGCGGCGGCGG
CCCTATAAAAAGCGAAGCGCGCGGCGGGGAGTGCCTGCGTTGCCCTTCGCCCCGTGCCCCGCT
CCGCGCCGCTCGCGCCCGCCCGCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGGCG
GGACGGCCCTTCTCTCCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTCTGTG
CTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGT
60 CGTGCGTGTGTGTGCGTGGGGAGCGCCGCGTGCAGGCCGCGCTGCCCGGCGGCTGTGAGCGCT
GCGGGCGCGGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTG
CCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGG
TGAGCAGGGGTGTGGGCGCGGCGGTGCGGCTGTAACCCCCCTGCACCCCCCTCCCGAGTTGC

TGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCCGG
GCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCGGGGAGGGCTC
GGGGGAGGGGCGCGCGGCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGCAGCCATT
5 GCCTTTTATGGTAATCGTGCAGAGGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAA
ATCTGGGAGGCGCCGCCGCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCAGGAA
GGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCGCTCCCCTTCTCCATCTCCAGCCTCG
GGGCTGCCGACGGGGACGGCTGCCTTCGGGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCG
TGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTG
10 GGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAG
AAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAA
TGGAGTCGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTA
CCCTGTACTGGCACGTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGAC
AGGCATCATACCCACTCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTCTGCGGAACAACGC
15 CAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCC
AACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGTCAGCAAGGCTTCTCCC
TGGAGAACGCACTGTACGCTCTGTCCGCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAAC
AGGAGCATCAAGTAGCAAAAAGAGGAAAGAGAGACACTACCACCGATTCTATGCCCCCACTTCTG
AAACAAGCAATTGAGCTGTTTCGACCGGACGGGAGCCGAACCTGCCTTCTTTTCTGCGCTGGAACCTA
20 ATCATATGTGGCCTGGAGAAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGA
TTTTGACTTAGACATGCTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGAC
GCTCTTGACGATTTTACCTTGACATGCTCCCCGGGTAAAGCGGCCGCGACTCTAGATCATAATCA
GCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGA
AACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAA
25 GCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCAA
ACTCATCAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCCAGGGATGTA
ATTACGTCCCTCCCCGCTAGGGGACAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCC
CCCCGATCCCCGAGCCGCGAGCGTGCGGGACAGCCCGGGCACGGGGAAGGTGGCAGCGGATC
GCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTCGAGACACCTGGGGGGATACGGGGAATAAA
30 GCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCA
CAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCCA
GAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAAC
CTGTTCAAGTGCGTCAACACCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCT
GTCTCAGTGTAAGGCCATTCCCCCTTGCTCTATCAAGGGGAGTTTGCTGTGACATTGTTGGTCTGG
35 GGTGACACATGTTTGCCAAATCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAG
GACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTC
AGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGG
AGAGGAGCAGAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGTGTTG
TCTGGATGCAAGCAGAAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTAGGACTG
40 GGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAA
AGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGT
GCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGT
GCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCA
CAACGTAAGGCCATTATCTCTCATCAACTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGCACCT
45 AGAGGGACAGCCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGACAGC
AGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGC
GGGGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCT
TTGAGCCTGCAGACACTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAAT
GACAGAATCATAGCAAGGCCTGGGTTGCAAAAGGAGACAGTGCTCATCCAGATCCAACCCCTGC
50 TATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATG
CCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCACCACCCTCTGGGG
GAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCTTGTC
CTATCAAGGGGGAGTTTGTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCA
TCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGT
55 GTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAG
GATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACAC
TGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGA
GCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTA
GATGTTTCATACAATCGTCAAATCATGAAGGTGGAAGGAGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTC
60 CAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAG
GTAAATCTTGCTAAATCCAGCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAA
CTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTAT

GAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGA
ACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGA
TAGAGAACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTAT
5 CAGTGATAGAGAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTCGTTT
AGTGAACCGTCAGATCGCCTGGAGCAATTCACAACACTTTTGTCTTATACCAACTTTCCGTACCA
CTTCCTACCCCTCGTAAAAAGCTTGTCCACCATTGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAA
ATACTCAATCGGGCTGGCCATCGGAACCTAAGTGGGGTGGGCAGTCATTACTGACGAGTACA
AAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTG
10 ATTGGGGCCCTGCTGTTCTGACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAG
AAGGCGCTATACCCGGCGGAAAGAAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGG
CCAAGGTGGACGATAGTTTCTTTCATCGGCTGGAGGAATCATTCCTGGTCGAGGAAGATAAGAAA
CACGAGAGACATCCTATCTTTGGAAACATTGTGGACGAGGTGCTTATCACGAAAAATACCCCAAC
ATCTATCATCTGCGCAAGAAACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTG
15 GCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAAT
AGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATAACAATCAGCTGTTTGAGGAAAAACCT
ATTAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCT
GGAGAACCTGATCGCTCAGCTGCCAGGCGAAAAAGAAACGGCCTGTTTGGAAATCTGATTGCAC
TGTCACTGGGACTGACACCTAAGCTCAAGGAAACAATTTTATCTGAGTGGCCGAGGACCTAAACTGCAGC
20 TGAGCAAGGACACTTATGACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGCA
GACCTGTTCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAAC
ACCGAGATTACAAAAGCCCTCTGTGACCTAGCATGATCAAGAGATATGACGAGCACCATCAGGA
TCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCGAGAAGTACAAGGAAATCTTCTTTGA
TCAGAGTAAGAACGGGTACGCCGTTATATTGACGGCGGAGCTTCACAGGAGGAATCTACAAGT
25 TTATCAAACCTATTCTGGAGAAGATGGACGGCACCGAGGAAGTCTGGTGAAACTGAATCGCGAG
GACCTGCTGCGCAAGCAGCGGACATTTGATAACGGCTCCATCCCCCACCAGATTCTCTGGGAGA
GCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGA
TCGAAAAAATCTGACCTTCCGGATCCCTTACTATGTGGGGCCCTGGCAAGGGGTAATTCCCGCT
TTGCCTGGATGACACGGAATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGAT
30 AAGGGAGCTTCCGCACAGTCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGA
GAAAGTGCTGCCTAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGT
GAAATACGTACCGAGGGGATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCG
TGGACCTGCTGTTTAAAACCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAG
AAAATTGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCAAGACAGATTTAACGCTTCTCTGGGA
35 ACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCCTGGACAACGAGGAAAAATGAGGA
TATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCT
GAAAACCTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCG
GCTGGGGACGACTGAGCAGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAAGACTAT
CCTGACTTTTCTGAAATCAGATGGCTTCGCTAACAGAAATTTTATGACAGCTGATTACAGATACAG
40 CCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTGCACGAGC
ATATCGCAAAACCTGGCCGGGTCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTC
GATGAGCTGGTGAAGTTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCGA
AAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGG
CATTAAAGGAAGTGGGATCCCAGATCCTGAAAGAGCACCCCTGTGGAAAACACTCAGCTGCAGAATG
45 AGAAGCTGTATCTGTACTATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATT
AACCGACTGTCTGATTACGACGTGGATGCCATCGTCCACAGTCATTTCCTGAAAGATGACAGCATT
GACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGG
AAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCAGACAGCGA
AAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGTCTGTGCAAACTGGACAAAGCAGGCTTCAT
50 CAAGCGACAGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGA
TGAACACAAAGTACGATGAGAATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCA
AACTTGTGAGCGACTTCAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCA
CCATGCTCATGACGCATACCTGAACGCAGTGGTCCGGACCGCCCTGATTAAGAAATACCCCAAAC
TGGAGAGCGAATTCGTGTACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGT
55 GAGCAGGAAATTGGAAGGACCCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAG
ACTGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGAC
CGGAGAAATCGTGTGGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCTGTCCATGCCTC
AAGTGAATATTGTCAAGAAAAACAGAGGTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCC
AAACGGAACCTGATGAAGCTGATCGCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGTTT
60 CGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTTCGCAAAAGGTGGAGAAGGGGAAAAAGCAAGA
AACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAAC
CCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCC
AAAGTACAGCCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGC

AGAAGGGTAATGAACTGGCCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACG
AGAAGCTGAAAGGCTCCCCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCAT
TATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAA
TCTGGATAAAGGTCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAA
5 AATATCATTTCATCTGTTCACTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTA
CCATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACCCAGAGC
ATTACCGCCTGTATGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGCCGAAAAA
GAAACGCAAAGTTGGGCGCGCCGACGCGTGGACGATTTCGATCTCGACATGCTGGGCGAGCGACG
CCCTGGATGACTTCGACCTGGATATGCTGGGCTCTGATGCCCTGGACGACTTTGACTTGGACATGT
10 TGGGATCCGACGCTCTCGATGATTTTGACCTTGACATGCTGATCAACGGCAGCGGCGAGGGCAGA
GGCAGCCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAG
CAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACG
GCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG
TTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCTCGTGACCACCTTCACCTACGGC
15 GTGCAGTGCTTCGCCCCGTACCCCGACCATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGA
GGTGAAGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG
ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAAGGTCTATATCACCGCC
GACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCG
20 TGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC
AACCCTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGT
CCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTA
ATCTAGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTC
CCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTCTTAATAAAATGAGGAAAT
25 TGCATCGCATTGTCTGAGTAGGTGTCTATTCTTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGG
GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCG
GAAAGAACCAGCTGGGGCTCGATCCTCTAGTTGGCGCTCATGGTCCATATGAATATCCTCCTTAG
TTCCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCT
CCCCCGCTAGGGGCAGCAGCGAGCCGCCGGGGTCCGCTCCGGTCCGGCGCTCCCCCGCATCCC
30 CGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTCTCTGA
ACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGA
AAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCC
AGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCC
AGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGC
35 TCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTA
AGCCATTCCCCCTTGTCTATCAAGGGGGAGTTTGTCTGTGACATTGTTGGTCTGGGGTGACACAT
TTTGCCAAATTCAGTGACATACGGAGAGGAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGA
CGTGGGACATGACAGGTGTTGAGGGCTTGGGACACTTCCAAGTACAGCGTTTACAGACGCCTT
AAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACA
40 AAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAG
CAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGC
TGGAGAATTGCCATGTAGATGTTTATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAG
ATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCC
ACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGC
45 ACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGC
CATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCA
GCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCAGTATCTATAACAAGA
AAATATATATATAATAAGTTATCAGTAAGTAAAGTAAATATAACAATATAATTATCGTATGAGT
TAAATCTTAAAGTCACGTAAAGATAATCATGCGTCATTTTGACTCACGCGGTGCTTATAGTTCA
50 AAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATG
TCCTAAATGCACAGCGACGGATTTCGCGCTATTTAGAAAGAGAGCAATATTTCAAGAATGCATG
CGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTACTCGACCTAAACTTT
AAACACGTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGC
GCGCCGGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTACAAT
55 TCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTTAATCGCCTT
GCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTA
AATTTTGTAAATCAGCTCATTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCA
AAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAA
60 CGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCAT
CACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGC
CCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA

AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCC
GCGCTTAATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 16)

TOIC_Bsd_BE_Cas9_Obl_r26_AAVS_PgkNeo

5 GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATAT
GTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
GTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTTTGCTCAC
10 CCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
ACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAG
CACTTTTAAAGTTCTGCTATGTGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTACCAGTCACAGAAAAGCATCTTAC
GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCA
ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATC
15 ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
20 GGCACTATGGATGAACGAAATAGACAGATCGTGAGATAGGTGCCTCACTGATTAAGCATTTGGT
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAG
GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAACCCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
25 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTG
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
GAAAGGCGGACAGGTATCCGGTAAGCGGCGAGGTGCGAACAGGAGCGCAGGAGGAGCTTCC
30 AGGGGGAACGCGCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
35 TTAATGCAGCTGGCAGCAGAGGTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TGAGTTAGCTCACTCATTAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
TTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCTAGAAAGATAATCATATTGTGACG
TACGTTAAAGATAATCATGCGTAAATGACGATGTGTTTTATCGATCTGTATATCGAGGTTTATT
40 TATTAATTTGAATAGATATTAAGTTTTATTATATTACACTTACATACTAATAATAAATTAACAAA
CAATTTATTTATGTTTATTTATTTATTAATAAAAAAAACAAAAACTCAAAATTTCTTCTATAAAGTAAC
AAAATTTTAAACATTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
TATTATAAAATAAGTAATTAGCTTAACCTATACATAATAGAAACAAATTATACTTATTAATCGCAT
TGATTATTGACTAGTCGTATTAAGGGTTCGGATCAGCTTGATTGAGCCCCAGCTGGTCTTTCCG
45 CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCGTGCTATTGTCTTCCCAATCCTCCCCCTT
GCTGTCCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
GGCAAACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
TTGATCCCCTCAGAAGAAGCTCGTCAAGTAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG
50 CGATACCGTAAAGCACGAGGAAGCGGTGAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCCACAGTCGATGAATCCAGA
AAAGCGGCCATTTTCCACCATGATATTGCGCAAGCAGGCATCGCCATGGGTACGACGAGATCCTC
GCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCCTGATGCTCTTC
GTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
55 CGCTTGGTGGTGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCA
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCCGCC
AATAGCAGCCAGTCCCTTCCCGCTTCACTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGT
CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTGCAAGTTCATTAGGGCACCGGACAGGTCCGT
CTTGACAAAAAGAACCGGGCGCCCTGCAGCTGACAGCCGGAACACGGCGGATCAGAGCAGCCG
60 ATTGCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGC
AATCCATCTTGTTCATGCGCGATCCCATGGTTTAGTTCCTCACCTTGTGCTATTATACTATGCCGA

TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTCGAAAGGCCCGGAGATGAGGA
AGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGTGCAGAATGCCGGGCCTCCGGAGGACC
TTCGGGCGCCCCGCCCCGCCCTGAGCCCCGCCCTGAGCCCCGCCCGGACCCACCCCTTCCCAGCC
TCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCAAAGGCCTACCCGCTTC
5 CATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCACGTC
CTGCACGACGCGAGCTGCGGGGCGGGGGGAACTTCCTGACTAGGGGAGGAGTAGAAGGTGGCG
CGAAGGGGCCACCAAAGAACGGAGCCGGTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCC
AGAGGCCACTTGTGTAGCGCCAAGTGCCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTT
GGGAAAAGCGCCTCCCCTACCCGCTAGACACCCACAGTGGGTGGCCTAGGGACAGGATTGCAAC
10 TCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATAG
TTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT
15 ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGT
TTTATTTATTTTAAATATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCG
GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGGGGCGGAGAGGTGCGGCGCGCAATCGAGA
GCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAAGCGA
20 AGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCTCGCGC
CGCCCCGCCCCGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCT
CCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTA
AAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
CGTGGGGAGCGCCGCTGCGGCCGCGCTGCCCGCGGGCTGTGAGCGCTGCGGGGCGGCGCGGGG
25 GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGGG
GGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGG
GCGCGGCGGTGCGGCTGTAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCAGGCCCCGGCT
TCGGGTGCGGGGCTCCGTGCGGGGCTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGTGGCGGCA
GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCGCGGC
30 GGCCCCGAGCGCCGCGGCTGTGCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCG
TGCGAGAGGGGCGCAGGGAATTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
CGCACCCCTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGA
GGGCCTTCGTGCGTCCCGCGCCCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGG
GACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTC
35 TAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
ATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTC
TCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCCTGACGACAAAGGAACTCGTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGACTGACGACG
TGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCAC
40 TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGT
GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTA
CGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
CGCTCTGTCCGCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
AAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
45 TGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGG
AGAAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATG
CTCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTTTGACGATTTG
ACCTTGACATGTCCCCGGGTAAAGCGGCGGACTCTAGATCATAATCAGCCTACACATTTGT
AGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
50 AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA
TTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT
TAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCC
GCTAGGGGACAGCAGCGAGCCGCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG
CCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGC
55 TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAG
AGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGA
TCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCC
TGGCCTTGAATGCCTGACAGGATGGGCGATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCTCAC
CACCCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCC
60 ATTCCCCCTTGTCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGC
CAATTCAAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGG
GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA

TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGGAAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCC
5 AAGACCAACCCCAACCCACCCACCGTGGCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCCCTCCCCTGGCACAACGTAAAGGCCATTAT
CTCTCATCCAACTCCAGGACGGAGTCAGTGAGGATGGGGCTGTCGACCTAGAGGGACAGCCCCC
CCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCGGGG
10 CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCA
CGGGGAAGGTGGCACGGGATCGCTTTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACAC
CTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAA
CGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATC
AACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGC
15 ATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCCTCTGGGGGAAAAACTGCCTCCTC
ATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTGTCTATCAAGGGGGAGT
TTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAAATTCAGTGATCAGGAGAGGCAGA
TCTTGGGGATAAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
CACTCTCCAAGTCACAGCGTTTCAAGAACGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
20 GCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
GAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGAT
ACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTG
CCCCTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCC
25 CCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCA
GTGAGGATGGGGTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTA
CTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC
30 AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
ATAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTGTGAAACCGTCAGATC
GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAA
AAGCTTGTCCACTCGAGATTCTCTAGACATCATTAATTCCCTAATTTTTGTTGACACTCTATCATTGA
TAGAGTTATTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAATGGCCAAGCCTTTGTCTCAAG
35 AAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTAC
AGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTT
ACTGGGGACCTTGTGCAGAACTCGTGGTGTCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTG
ACTTGATCGCTCGGATCGGAAATGAGAACAGGGGCACTCTTGAGCCCTGCGGACGGTGTGAGAA
GGTGCTTCTCGATCTGCATCCTGGGATCAAAGCGATAGTGAAGGACAGTGATGGACAGCCGACGG
40 CAGTTGGGATTCTGTAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTAACTCGAGATGAGCTCAG
AGACTGGCCAGTGGCTGTGGACCCACATTGAGACGGCGGATCGAGCCCCATGAGTTTGAGGTA
TTCTTCGATCCGAGAGAGCTCCGCAAGGAGACCTGCCTGCTTACGAAATTAATTGGGGGGGCGG
CACTCCATTTGGCGACATACATCAGAACACTAACAAGCACGTCGAAGTCAACTTCATCGAGAA
GTTACAGACAGAAAGATATTTCTGTCCGAACACAAGGTGCAGCATTACCTGGTTTCTCAGCTGGAG
45 CCCATGCGGCGAATGTAGTAGGGCCATCACTGAATTCTGTCAAGGTATCCCCACGTCACCTCTGTT
TATTTACATCGCAAGGCTGTACCACCACGCTGACCCCCGCAATCGACAAGGCCTGCGGGATTTGAT
CTCTTCAGGTGTGACTATCCAAATTATGACTGAGCAGGAGTCAGGATACTGCTGGAGAAACTTTGT
GAATTATAGCCCCGAGTAATGAAGCCCACTGGCCTAGGTATCCCCATCTGTGGTACGAGTACGT
TCTTGAACGTGACTGCATCATACTGGGCCTGCCTCCTTGTCTCAACATTCTGAGAAGGAAGCAGCC
50 ACAGCTGACATTCTTTACCATCGCTCTTCAGTCTTGTGATTACCAGCGACTGCCCCACACATTCTC
TGGGCCACCGGGTTGAAAAGCGGCAGCGAGACTCCCGGGACCTCAGAGTCCGCCACACCCGAAAG
TGATAAAAAGTATTCTATTGGTTTAGCCATCGGCACTAATTCCGTTGGATGGGCTGTCATAACCGA
TGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAA
AGAATCTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGA
55 ACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAA
TGAGATGGCCAAAGTTGACGATTCTTTCTTTACCGTTTGAAGAGTCCTTCTTGTGCAAGAGGA
CAAGAAACATGAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGT
ACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTA
ATCTACTTGGCTCTTGGCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATC
60 CGGACAACCTCGGATGTCGACAAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAG
AGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCC
GACGGCTAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGGGTTGTTCCGTAACCTT

ATAGCGCTCTACTAGGCCTGACACCAAATTTTAAAGTCGAACTTCGACTTAGCTGAAGATGCCAAA
 TTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCACAAATTGGAGATCA
 GTATGCGGACTTATTTTTGGCTGCCAAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAG
 5 AGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATCA
 CCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAATATAAGGAAATATT
 CTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCT
 ACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAAACTCAAT
 CGCGAAGATCTACTGCGAAAAGCAGCGGACTTTTCGACAAACGGTAGCATTCCACATCAAATCCACTT
 10 AGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGA
 AAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTC
 TCGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAGGAAGTTG
 TCGATAAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGA
 ACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTCACAGTGTACAATGAACTCACGA
 15 AAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCCTTCTAAGCGGAGAACAGAAAGAAAGCA
 ATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTT
 TAAGAAAATTGAATGCTTCGATTCTGTGAGATCTCCGGGGTAGAAGATCGATTAAATGCGTCACT
 TGGTACGTATCATGACCTCCTAAAGATAATTAAGAGACTTCCTGGATAACGAAGAGAATG
 AAGATATCTTAGAAGATATAGTGTTGACTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAA
 20 GACTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTAT
 ACGGGCTGGGGACGATTGTGCGGAAACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAA
 CTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGATG
 ACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCAC
 GAACATATTGCGAATCTTGCTGGTTCCCGAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGT
 25 AGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAAACATTGTAATCGAGATGGCAC
 GCGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAG
 AGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAATTGCAG
 AACGAGAAACTTTACCTCTATTACCTACAAAAATGGAAGGGACATGTATGTTGATCAGGAACCTGGA
 CATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTGAAGGACGATTCA
 30 ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCAAGCGA
 GGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAA
 GAAAGTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGATTT
 ATTAACGTCAGCTCGTGGAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCG
 AATGAATACGAAATACGACGAGAACGATAAGCTGATTCCGGGAAGTCAAAGTAATCACTTTAAAGT
 35 CAAAATTGGTGTCGGAATTCAGAAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACC
 ACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTAAAGAAATACCCGAAGC
 TAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGC
 GAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGAATTTCTTTAAG
 ACGGAAATCACTCTGGCAAAACGGAGAGATACGCAAAACGACCTTTAATTGAAACCAATGGGGAGAC
 40 AGGTGAAATCGTATGGGATAAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCC
 AAGTCAACATAGTAAAGAAAACAGGAGTGCAGACCGGAGGGTTTTCAAAGGAATCGATTCTTCCA
 AAAAGGAATAGTGATAAGCTCATCGTCTGTAAGGACTGGGACCCGAAAAAGTACGGTGGCTT
 CGATAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGCAAAAAGTTGAGAAGGGGAAAATCCAAGA
 AACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTGAAAAGAAC
 45 CCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAATACTACC
 AAAGTATAGTCTGTTTGTAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTC
 AAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTACG
 AGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAACTTTTGTGAGCAGCACAAACAT
 TATCTCGAGAAATCATAGAGCAAAATTCGGAATTACAGTAAGAGAGTCACTCCTAGATGCCAAT
 50 CTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAAACCCATACGTGAGCAGGCGGAAA
 ATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGATTCAAGTATTTTGACACAAC
 GATAGATCGAAACGATACACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCA
 TCACGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGACTCTGGTGGTTCTACTA
 ATCTGTCAGATATTATTGAAAAGGAGACCGGTAAGCAACTGGTTATCCAGGAATCCATCCTCATGC
 55 TCCCAGAGGAGGTGGAAGAAGTCATTGGGAACAAGCCGGAAAGCGATATACTCGTGACACCCGCC
 TACGACGAGAGCACCGACGAGAATGTCATGCTTCTGACTAGCGACGCCCCGTAATACAAGCCTTG
 GGCTCTGGTCATACAGGATAGCAACGGTGAGAACAAGATTAAGATGCTCTCTGGTGGTTCTCCAA
 GAAGAAGAGGAAAAGTCTAAAAATTCTAAAAATACAGCATAGCAAACTTTAACTCCAAATCAAGC
 CTCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCA
 60 TTAGCTGTTTGACGCCTCACCTTCTTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGGTTTGAA
 CTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTTAGTAAAATATTCA
 GAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTC
 AAGGCCCTTCATAATATCCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAA

ATTGACAGCAAGAAAGCGAGCTTCTAGATGGTCCATATGAATATCCTCCTTAGTTCCTATTCCGC
 TAGCCTAGAGGGACAGCCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGG
 GGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCA
 GCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCG
 5 CTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGAT
 TTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAAC
 CCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCCT
 TGAATGCCTGACGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCCT
 CTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCC
 10 CTTGTCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTG
 AGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACAT
 GCAGGTGTTGAGGGCTCTGGGACACTCTCAAAGTCACAGCGTTCAGAACAGCCCTTAAGGATAAGA
 AGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCAC
 AGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGT
 15 GGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGC
 CATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGAC
 CAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCA
 TCACCTCCAGGGACGGTGACCCCCCACTCCGTGGGCGAGCTGTGCCACTGACCGCTGCTTTG
 GAGAAGGTAAATCTTGCTAAATCCAGCCCCACCTCCCCTGGCACAACGTAAGGCCATTATCTCTC
 20 ATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAAT
 CGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAAATATATATA
 TAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCTTAAA
 AGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGTCTGTTATAGTTCAAAATCAGTGAC
 ACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCA
 25 CAGCGACGGATTGCGGCTATTTAGAAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTAC
 GCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAACTTTAAACACGTCAT
 AGAATCTTCGTTTGACAAAAACACATGTGTTGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGA
 GGTTCCCTTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCG
 TCGTTTTACAACGTGCTGACTGGGAAAAACCCTGGCGTTACCCAACTTAATCGCCTTGACGCACATC
 30 CCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA
 GCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTTGTTA
 AATCAGCTCATTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGAC
 CGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA
 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCA
 35 AGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCGATTTAG
 AGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGG
 CGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATG
 CGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 17)

40 TOIC_Cas9_Bsd_p300_Obl_r26_AAVS_PgkNeo

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATAT
 GTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
 GTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGGCGCATTTTGCCTTCCTGTTTTGTCTCAC
 CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
 45 ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAG
 CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
 TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCA
 ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGACAACATGGGGGATC
 50 ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
 ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTA
 GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
 GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAACTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
 CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
 55 GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
 AACTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAG
 GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
 TGAGCGTCAGACCCCGTAGAAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
 TGCTGCTTGCAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
 60 ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGACCGCCTACATACCTCGCTCTGCTAATCCTG

TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAAGGCGCAGCGGTGCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCACGCTTCCCGAAGGGA
5 GAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
AGGGGGAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
CCTGGCCTTTTGTGCTTTTGTCTACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCA
10 GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
TTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCTAGAAAGATAATCATATTGTGACG
15 TACGTTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
TATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCACAAA
CAATTTATTTATGTTTATTTATTTATTAATAAAAAAACAACAACTCAAAATTTCTTCTATAAAGTAAC
AAAACCTTTTAAACATTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
TATTATAAAAATAAGTAATTAAGCTTAACCTTATACATAATAGAAACAAATTATACCTTATTAATCGCAT
20 TGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTTCGAGCCCCAGCTGGTTCTTTCCG
CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
GCTGTCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
GGCAAACAACAGATGGCTGGCACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
25 TTGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG
CGATACCGTAAAGCACGAGGAAGCGGTGAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGA
AAAGCGGCCATTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTACAGACGAGATCCTC
GCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCTGATGCTCTTC
30 GTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
CGCTTGGTGGTCAATGGGCAGGTAGCCGATCAAGCGTATGCAGCCGCCGCAATTGCATCAGCCA
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCC
AATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAACGCCCGT
CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTGAGTTTATTTCAGGGCACCGGACAGGTTCGGT
35 CTTGACAAAAAGAACCAGGGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCGTGTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGC
AATCCATCTTGTTCATGCGCGATCCCATGGTTTAGTTCTCACCTTGTCTGATTATACTATGCCGA
TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTCGAAAGGCCCGGATGAGGA
40 AGAGGAGAACACGCGCGCAGAGCTGCGCTTTTGAAGCGTGCAGAAATGCCGGGCTCCGAGGACTT
TTCGGGCGCCCCGCCCCGCCCCCTGAGCCCCGCCCCGAGCCACCCCTTCCAGCC
TCTGAGCCCAGAAAGCGAAGGAGCAAGCTGCTATTGGCCGCTGCCCCAAAGGCCATCCCGCTTC
CATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCACGTC
CTGCACGACGCGAGCTGCGGGGCGGGGGGGAACCTTCTGACTAGGGGAGGAGTAGAAGGTGGCG
CGAAGGGGCCACCAAGAACGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCC
45 AGAGGCCACTTGTGTAGCGCCAAGTGCCCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTT
GGGAAAAGCGCCTCCCCTACCCGCTAGACACCCACAGTGGGTGGCCTAGGGACAGGATTGCAAC
TCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAG
TTCATAGCCCATATATGAGTTCCGCGTTACATAAAGTACGTTAAATGGCCCGCTGGCTGACCGC
CCAACGACCCCATGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGAGCTT
50 TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGT
ACATGACCTTATGGGACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGT
TTTATTTATTTTAAATTATTTGTGACGATGGGGGCGGGGGGGGGGGGGGGCGCGCGCCAGGCG
55 GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGA
GCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGA
AGCGCGCGGCGGGGCGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCTCGCGC
CGCCGCCCCCGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGGCGGAGCGGCCCTTCTCT
CCGGCTGTGAATTAGCGCTTGGTTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTA
AAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
60 CGTGGGGAGCGCCGCTGCGGCCGCGCTGCCCGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGG
GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGGGGTGGCCCGCGGTGCGGGGG
GGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGG

CGCGGGCGGTTCGGGCTGTAACCCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCCGGCT
TCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA
GGTGGGGGTGCCGGGCGGGGCGGGGCCCTCGGGCCGGGAGGGCTCGGGGGAGGGGCGCGGC
GGCCCCGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCG
5 TGCAGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
CGCACCCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGA
GGGCTTCGTGCGTCGCGCGCCGCGCCGTCCCTTCTCCATCTCCAGCCTCGGGGTGCCGAGGGG
GACGGTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTC
TAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
10 ATTTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTC
TCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCCTGACGACAAGGAACTCGCTCAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAGC
TGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCAC
TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCCGTGT
15 GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTA
CGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
CGCTCTGTCCGCGGTGGGCCACTTTTACTGGGTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
AAAAGAGGAAAGAGAGACACCTTACACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
TGTTGACCGGCAGGGAGCCGAACCTGCCTTCTTTTCGGCCTGGAACATAATCATATGTGGCCTGG
20 AGAAACAGCTAAAGTGCAGAAAGCGGCGGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATG
CTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTG
ACCTTGACATGCTCCCCGGGTAAAGCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGT
AGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA
25 TTTCACAAATAAAGCATTTTTTTCCTGCACTTCTAGTTGTGGTTTGTCCAAACCTCATCAATGTATCT
TAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCC
GCTAGGGGACGAGCGAGCCGCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG
CCGGCAGCGTGCGGGGACAGCCCGGGCAGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGC
TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAG
30 AGAGATTTAGAATGACAGAATCATAGAACGGCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGA
TCCAACCCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCC
TGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCAC
CACCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCC
ATTCCCCCTGTCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGC
35 CAATTCAAGTGACATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGG
GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA
TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCAGACATGCTGGTCCCTGTGTCTGAGCCTGATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
40 AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCC
AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
CTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGACCTAGAGGGACAGCCCCC
45 CCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCCGGG
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCA
CGGGGAAGGTGGCAGGGATCGCTTTCCTGTAACGCTTCTCGCTGCTCTTGAGCCTGCAGACAC
CTGGGGGATACGGGAAAAAGCTTTAGGCTGAAAGAGAGATTAGAAATGACAGAATCAGACAGAA
CGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCCTGCTATGTGCAGGGTCATC
50 AACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGC
ATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCACCAACCTCTGGGGGAAAAACTGCCTCCTC
ATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTGTCTATCAAGGGGGAGT
TTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGATCACGGAGAGGCAGA
TCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
55 CACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
GAGCCTGCATGTTTATGTTGTCTGGATGCAAGCAGAGGGGTGGAAGAGCTTGCTGGAGAGAT
ACAGCTGGGTGAGTGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTG
60 CCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCC
CCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCCCACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCA

GTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTT
CTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC
AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
5 ATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTGTGAACCGTCAGATC
GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCTACCCTCGTAAA
AAGCTTGTCCACCATTGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGG
CCATCGGAACCTAACTCAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAA
TTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTTT
10 GACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCG
AAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTT
CTTTTCATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTT
TGGAACATTGTGGACGAGGTTCGCTTATCACGAAAAATACCCACCCTATCTATCTGCGCAAGA
AACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTA
15 AGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTG
TTCATCCAGCTGGTCCAGACATAACAATCAGCTGTTTGAGGAAAACCCCTATTAATGCATCTGGCGTG
GACGCAAAAAGCCATCCTGAGTGCCAGGTGTCTAAGAGTAGAAGGCTGGAGAAGCTGATCGCTCA
GCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGGAATCTGATTGCACTGTCACTGGGACTGACAC
CTAACTTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAACTGCAGCTGAGCAAGGACACTTAT
20 GACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGACGACCTGTTTCTGGCCGCT
AAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCC
CCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGC
ACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGT
ACGCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
25 AGAAGATGGACGGCACCGAGGAAGTCTGCTGGTGAAGTGAATCGCGAGGACCTGCTGCGCAAGCA
GCGGACATTTGATAACGGCTCCATCCCCACCAGATTCTGCTGGGAGAGCTGCACGCAATCCTGCG
ACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCT
TCCGGATCCCTTACTATGTGGGGCCCTGGCAAGGGTAAATCCCGCTTTGCTGGATGACACGGA
AATCTGAGGAAACAATCACTCCTTGGAAGTTCGAGGAAGTGGTTCGATAAGGGAGCTTCCGCACAG
30 TCTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGTGCTGCTAAGCA
CAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGG
GATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAAA
CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATTGAATGTTTCGAT
TCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
35 AAGATCATTAAAGGATAAAGACTTCTGGAACAACGAGGAAAATGAGGATATCCTGGAAGACATTGT
GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTATGCCCATCT
GTTTCGATGACAAGGTGATGAACAGCTGAAGCGAAGAGGTACACCGGCTGGGACGCTGAGC
AGAAAGCTGATGACGCGCATTCGGGACAAACAGAGTGGAAAGACTATCCTGGATTTCTGAAATC
AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGA
40 TATCCAGAAGGCACAGGTGTCCGGCAGGGTGAAGTCTCTGCACGAGCATATCGCAAACCTGGCCG
GGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTTCGATGAGCTGGTGAAAGTC
ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCGAAAATCAGACCACACAGA
AAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
CCAGATCCTGAAAAGAGCACCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGTATCTGTACT
45 ATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTAC
GACGTGGATGCCATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGAC
CCGAGTGACAAAAACCGAGGAAAGAGTGATAAATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATG
AAGAACTGATGAGACGAGCTGCTGAATGCCAAACTGATCACACAGCGAAAGTTTGATAAACCTGAC
TAAAGCTGAGAGAGGGGTCTGTGCAGAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG
50 AGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGAT
GAGAATGACAACTGATCCGCGAAGTGAAGGTGCTTACTCTGAAGTCAAACTTGTGAGCGACTT
CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCAT
ACCTGAACGCACTGGTTCGGGACCGCCCTGATTAAGAAATACCCCAAACCTGGAGAGCGAATTTCGTG
TACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTGGAAA
55 AGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCTGGCA
AATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
ACAAGGGTAGGGATTTTGGCACAGTGCAGGAGGTCTGCTGCTCAAGTGAATATTGTCAAGA
AAACAGAGGTGAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTGATAAG
CTGATCGCCCGAAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
60 TTAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAG

CTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAACTGG
CCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCC
CCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAAGCATTATCTGGACGAGATCATT
5 GAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAAATCTGGATAAAGTCTCTGAG
CGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAAATATCATTTCATCTGTTCA
CTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACACAGAGCATTACCGGCCTGTATGAA
ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGATGGCCAAGCCCCTGAGCCAAGAGGA
AAGCACCTGATCGAGCGGGCCACCGCCACCATCAACAGCATCCCCATCAGCGAGGACTACAGCG
10 TGGCCTCTGCCGCCCTGAGCAGCGACGGCAGAATCTTACCGGCGTGAACGTGTACCACTTCACAG
GCGGCCCTTGCGCCGAGCTGGTGGTGCTGGGAACAGCTGCCGCCGCTGCCGCTGGCAACCTGACCT
GTATCGTGGCCATCGGCAACGAGAACCAGGGGCATCCTGAGCCCCTGCGGCAGATGCAGACAGGTG
CTGCTGGACCTGCACCCCGGCATCAAGGCCATCGTGAAGGACAGCGACGGCCAGCCCACCGCCGT
GGGCATTAGAGAGCTGCTGCCCAGCGGCTACGTGTGGGAGGGCTGACTCGAGATTTTCAAACCAG
15 AAGAACTACGACAGGCACTGATGCCCACCCTGGAAGCCCTGTACCGGCAGGACCCCGAGAGCCTG
CCCTTCAGACAGCCCCTGGATCCCCAGCTGCTGGGCATCCCCGACTACTTCGACATCGTGAAGTCC
CCCATGGACCTGAGCACCATCAAGCGGAAGCTGGACACCGGCCAGTACCAAGAGCCCTGGCAGTA
CGTGGACGACCTGGCTGATGTTCAACAACGCGCTGGCTGTACAACAGAAAGACCGCCGGTGT
ACAAGTACTGCAGCAAGCTGAGCGAGGTGTTTCGAGCAAGAGATCGACCCCGTGTATGCAGAGCCTG
20 GGCTACTGCTGCGGCAGAAAGCTGGAATTCAGCCCCCAGACCCTGTGCTGCTACGGCAAGCAGCT
GTGCACCATCCCCCGGACGCCACCTACTACAGCTACCAGAACAGATAACCACTTCTGCGAGAAGT
GCTTCAACGAGATCCAGGGCGAGAGCGTGTCCCTGGGCGACGACCCTAGCCAGCCCCAGACCACA
ATCAACAAAGAGCAGTTCAGCAAGCGGAAGAACGACACCCTGGACCCCGAGCTGTTCTGTGAATG
CACCGAGTGCGGCCGGAAGATGCACCAGATCTGCGTGCTGCACCACGAGATCATCTGGCCTGCCG
25 GCTTCGTGTGCGACGGCTGCCTGAAGAAGTCCGCCCGGACCCGGAAGAGAAAGTTCAGCGCC
AAGCGGCTGCCCTCTACCCGGCTGGGCACCTTCCTGGAAAACAGAGTGAACGACTTCCTGCGGCG
GCAGAACCAACCCGAGTCCGGCGAAGTGACAGTGCGGGTGGTGCACGCCAGCGACAAGACCGTG
GAAGTGAAGCCTGGCATGAAGGCCAGATTCTGTGACAGCGGCGAGATGGCCGAGAGCTTCCCTTA
CCGGACCAAGGCCCTGTTCCGCTTCGAAGAGATCGATGGCGTGGACCTGTGCTTCTTCGGCATGCA
30 CGTGCAAGAGTACGGCAGCGACTGCCCCCACCACAGCGGCGGGTGTACATCAGCTACCTGG
ACAGCGTGCCTTCTTCGGCCCAAGTGCTGCGGACCGCCGTGTATCACGAGATCCTGATCGGCT
ACCTGGAATACGTGAAGAAGCTGGGCTACACCACCGGCCACATCTGGGCCTGTCTCCCAGCGAG
GGCGACGACTACATCTTCCACTGCCACCCCCCGACCAGAAGATCCCCAAGCCCAAGAGACTGCA
AGAGTGGTACAAGAAGATGCTGGACAAGGCCGTGTCCGAGCGGATCGTGACGACTACAAGGAC
35 ATCTTCAAGCAGGCCACCGAGGACCGGCTGACCAGCGCCAAAGAGCTGCCCTACTTCGAGGGCGA
CTTCTGGCCCAACGTGCTGGAAGAGAGCATCAAAGAGCTGGAACAAGAGGAAGAGGAACGCAAG
CGGAAGAGAACACCAGCAACGAGAGCACCAGCTGACCAAGGGCGACAGCAAGAACGCCAAGA
AGAAGAACAACAAGAAAACCAGCAAGAACAAGAGCAGCCTGAGCCGGGGAACAAGAAAAGC
CCGGCATGCCCAACGTGTCCAACGACCTGAGCCAGAACTGTACGCCACCATGGAAGGACACAAA
40 GAGGTGTTCTTCGTTCATCCGGCTGATCGCCGACCTGCCGCCAACAGCCTGCCCCCATCGTGGAC
CCCGACCCCTGATCCCCTGCGACCTGATGGACGGCAGGGACGCCTTCTGACCCTGGCCCGGGAC
AAGCACCTGGAATTCTCCAGCCTGCGGAGAGCCAGTGGTCCACCATGTGCATGCTGGTGGAACT
GCACACCCAGAGCCAGGACGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTGAGGAGAAT
CCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGGCCAT
45 CCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCG
ATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGC
CCACCCTGTTGACCACCTTACCTACGGCTGACGTGCTTCGCCCGCTACCCCGACCACATGGAAG
AGCACCGATTCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTAAGG
ACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATC
50 GAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAATA
CAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGA
CCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGA
CCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCG
55 GCATGGACGAGCTGTACAAGTAAACCTAATCTAGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCT
AGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCTTGACCCTGGAAGGTGCCACTCCCA
CTGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCTATTCTATTCTGGG
GGGTGGGGTGGGCGAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT
GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATCCTCTAGTTGGCGCGT
60 CATGGTCCATATGAATATCCTCCTTAGTTTCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCCAA
AGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCGAGCGAGCCGCCCCGGGGCTCCG
CTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCCGGGCACGGGG

AAGGTGGCACGGGATCGCTTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGG
 GGGATACGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCC
 TGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCA
 GCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCA
 5 CAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCACCACCCTCTGGGGGAAAAACTGCCTCCTCATATC
 CAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCTTGTCTATCAAGGGGGAGTTTGCT
 GTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAAGTGATCACGGAGAGGCAGATCTTG
 GGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTC
 TCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAA
 10 GTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGC
 CTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAG
 CTGGGTCAAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTATACAATCGTCAA
 ATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCA
 CTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCC
 15 ACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAATCTTGCTAAATCCAG
 CCCGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACTCCAGGACGGAGTCAGTGA
 GGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTA
 TATATTAGTCACGATATCTATAACAAGAAAAATATATATATAATAAGTTATCACGTAAGTAGAACAT
 GAAATAACAATATAATTATCGTATGAGTTAAATCTTAAAAAGTCACGTAAAAAGATAATCATGCGTCA
 20 TTTTGACTCACGCGGTGCTTATAGTTCAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCA
 CGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGGATTGCGCTATTTAGAAA
 GAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAA
 GATTGCGCTTTACTCGACCTAACTTTAAACACGTATAGAATCTTCGTTTGACAAAAACCACAT
 TGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGG
 25 GTCGCCCTATAGTGAGTCGTATTACAATTCAGTGCCGCTCGTTTTACAACGTCGTGACTGGGAAAA
 CCCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGA
 AGAGCCCCGACCCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATGTAA
 GCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTTTAACCAATAGGC
 CGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAG
 30 TTTGGAACAAGAGTCCACTATTAAGAAGCTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTAT
 CAGGGCGATGGCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAA
 GCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTAGAGCTTGACGGGGAAAGCCGGCGAACGT
 GGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGATGCGGT
 35 ACGCTGCGCGTAACCACCACACCCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAG (SEQ ID
 NO: 18)

TOIC_Cas9_Obl_r26_AAVS_PgkNeo_Bsd_tracR

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTTCTAAATACATTCAAATAT
 GTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
 40 GTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTGGCGCATTTTGCCTTCCTGTTTTGCTCAC
 CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
 ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAG
 CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
 TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
 45 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCA
 ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGACACAACATGGGGGATC
 ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAACCAACGACGAGCGTGAC
 ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTA
 GCTTCCCGGCAACAATTAAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
 50 GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGTGAGCGTGGGTCTCGCGGTAT
 CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
 GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
 AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAG
 GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
 55 TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
 TGCTGCTTGCAAAACAAAAACCACCGCTACCAGCGGTGGTTTGTGTCGGGATCAAGAGCTACCA
 ACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGTGTAG
 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTG
 TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
 60 CCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
 GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA

GAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCCTTTTACGGTT
CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACC
5 GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGAAACCGCCTCTCCCCGCGGTTGGCCGATTCA
TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
10 TTAACCCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTGTGACG
TACGTTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATT
TATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAA
CAATTTATTTATGTTTATTTATTTATTAATAAAAAAACAAAACTCAAAATTTCTTCTATAAAGTAAC
AAAACCTTTTAAACATTCTCTCTTTTACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTG
15 TATTATAAAATAAGTAATTAGCTTAACTTATACATAATAGAAAACAAATTATACTTATTAATCGCAT
TGATTATTGACTAGTCGTATTAAGGGTTCGGATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCG
CCTCAGAAGCCATAGAGCCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
GCTGCTGCCCCACCCCAAGGTAAGATGACACCTACTCAGACAATGCGATGCAATTTTC
CTCATTTTATTAGGAAAGGACAGTGAGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
20 GGCAACAACAGATGGCTGGCACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
TTGATCCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG
CGATACCGTAAAGCACGAGGAAGCGGTGAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGG
GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCCACAGTCGATGAATCCAGA
AAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTC
25 GCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCCCTGATGCTCTTC
GTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
CGTTGGTGGTGAATGGGACAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCTTGCATCAGCCA
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACACCTACTCAGAGAGATCCTGCCCGGCACTTCGCCC
AATAGCAGCCAGTCCCTTCCCGCTTCACTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCCGT
30 CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTACAGGGCACCGGACAGGTGCGT
CTTGACAAAAAGAACC GGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGC
AATCCATCTTGTTCATGCGCGATCCCATGGTTTAGTTCCTCACCTTGTCGTATTATACTATGCCGA
TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTGCGAAAGGCCCGGAGATGAGGA
35 AGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGTGCAGAATGCCGGGCCCTCCGGAGGACC
TTCGGGCGCCCCGCCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCCCGGACCCACCCCTTCCAGCC
TCTGAGCCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCAAGGCCATACCCGCTT
CATTGCTCAGCGGTGCTGTCCATCTGCACGAGTACATGAGTACGAGACGTGCTACTTCCATTGTTCACGTC
CTGCACGACGCGAGCTGCGGGGCGGGGGGGAACCTTCTGACTAGGGGAGGAGTAGAAGGTGGCG
40 CGAAGGGGCCACCAAGAACGGAGCCGGTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCC
AGAGGCCACTTGTGTAGCGCCAAGTGCCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTT
GGGAAAAGCGCCTCCCCTACCCGCTAGACACCCACAGTGGGTGGCCTAGGGACAGGATTGCAAC
TCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAG
TTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGC
45 CCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGT
ACATGACCTTATGGACTTTCTTACTTGCGAGTACATGCTACGTATTAGTCATCGCTATTACCATGGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTA
50 TTTATTTATTTTTTAATTATTTTGTGACGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCG
GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGA
GCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAGCGA
AGCGCGCGGGCGGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCTCGCGC
CGCCCCCCCCGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCT
55 CCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTCTGTGGCTGCGTGAAAGCCTTA
AAGGGCTCCGGGAGGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
CGTGGGGAGCGCCCGTGC GGCCCCGCGTGC CGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGG
GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCGGGGCGGTGCCCGCGGTGCGGGG
GGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGG
60 GCGCGGCGGTGCGGCTGTAAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGT
TCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA
GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCGGGAGGGCTCGGGGGAGGGGCGCGGC

GGCCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCG
TGCGAGAGGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
CGCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGA
GGGCCTTCGTGCGTCGCCGCGCCGCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGCAGGGG
5 GACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTGTGACCGGCGGGCTC
TAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
ATTGTGCTGTCTCATCTTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTCT
TCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAGC
10 TGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCAC
TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGT
GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTA
CGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
CGCTCTGTCCGCGCTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
15 AAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
TGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGG
AGAAACAGCTAAAGTGCAGAAAGCGGGCGGCGGACCGACGCGCTTGACGATTTTGACTTAGACATG
CTCCACCCGATGCCCTTGACGACTTTGACCTTGATGCTGCCTGCTGACGCTCTTGACGATTTTG
ACCTTGACATGCTCCCCGGGTAAAGCGGCGCGACTCTAGATCATAATCAGCCATACCACATTTGT
20 AGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
AATTGTTGTTGTTAACTTGTTATTGTCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA
TTTCACAAATAAAGCATTTTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT
TAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCC
GCTAGGGGACAGCAGCGAGCCGCCCCGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG
25 CCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGC
TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAG
AGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCAGTGCTCATCCAGA
TCCAACCCCTGCTATGTGCAGGGTCATCAACCCAGCAGCCAGGCTGCCAGAGCCACATCCAGCC
TGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCCTAC
30 CACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCC
ATTCCCCCTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGC
CAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGG
GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA
TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAG
35 GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTATACATCAATCGTCAAATCATGAAGGCTGGAAGAGCCCTCAAGATCCCC
AAGACCAACCCCAACCCACCCAGCTGCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
CTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
40 CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
CTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGACCTAGAGGGACAGCCCCC
CCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGACAGCAGCGCCGCGGGG
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCA
CGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACAC
45 CTGGGGGGATACGGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAA
CGGCCTGGGTTGCAAAGGAGCAGAGTGTCTATCCAGATCCAACCCCTGCTATGTGCAGGGTCATC
AACCAGAGCCAGGCTGCCAGAGGCCATCCAGCTGGCCTTGAATGCCCTGCAAGGATGGGGC
ATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGTCCATCACCACCTCTGGGGGAAAAACTGCCTCCTC
ATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGCTCTATCAAGGGGGAGT
50 TTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAAGTGCATCACGGAGAGGCAGA
TCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
CACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
GAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAAGGGGTGGAAGAGCTTGCTGGAGAGAT
55 ACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTATACAAATCG
TCAAATCATGAAGGCTGGAAGAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCAGCTG
CCCACCTGGCATGTCCCTCAGTGCCACATCCCCACAGTTCTTATCACCTCCAGGGACGGTGACCC
CCCCACCTCCGTGGGCGAGCTGTGCCACTGCAGCACCCTCTTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCCCGACCTCCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCA
60 GTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTA
CTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC

AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
 ATAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATC
 GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCTACCCTCGTAAA
 AAGCTTGTCCACCATGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGG
 5 ACATCGGAATAACTCAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAA
 TTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTT
 GACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCG
 AAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTT
 CTTTCATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTT
 10 TGGAACATTGTGGACGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGA
 AACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTA
 AGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTG
 TTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCCATTAAATGCATCTGGCGTG
 GACGCAAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCCTGATCGCTCA
 15 GCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGGAATCTGATTGCACTGTCAGTGGGACTGACAC
 CTAACCTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAACTGCAGCTGAGCAAGGACACTTAT
 GACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGAGACCTGTTCTGAGGCGCT
 AAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGGATTACAAAAGCC
 CCTCTGTCTGAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGC
 20 ACTGGTGCGCCAGCAGCTGCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGT
 ACGCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
 AGAAGATGGACGGCACCGAGGAAGTCTGCTGTTGAACTGAATCGCGAGGACCTGCTGCGCAAGCA
 GCGGACATTTGATAACGGCTCCATCCCCACCAGATTCATCTGGGAGAGCTGCACGCAATCCTGCG
 ACGACAGGAAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCT
 25 TCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGGGTAAATTCGCTTTGCTGGATGACACGGA
 AATCTGAGGAAACAATCACTCCTTGAACTTCGAGGAAGTGGTCGATAAGGGAGCTTCCGCACAG
 TCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGTGCTGCCTAAGCA
 CAGTCTGTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGGAGG
 GATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAAA
 30 CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAATTGAATGTTTCGAT
 TCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
 AAGATCATTAAAGGATAAAGACTTCCTGGACAACGAGGAAAATGAGGATATCCTGGAAGACATTGT
 GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTATGCCATCT
 GTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAGC
 35 AGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAAGACTATCCTGGACTTTCTGAAATC
 AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTACGATGACAGCCTGACCTCAAAGAGGA
 TATCCAGAACGACAGGTGTCCGGGACGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCG
 GGTCCTCCGCCATCAAGAAAGGTATTCTGCAGACCTGAAAGGTGGTCGATGAGTGGTGAAAGTC
 ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCAAAATCAGACCACACAGA
 40 AAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
 CCAGATCCTGAAAGAGCACCCCTGTGGAACAACTCAGCTGCAGAATGAGAAGCTGTATCTGTACT
 ATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTAC
 GACGTGGATCATATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACC
 CGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGA
 45 AGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGCGAAAGTTTGATAACCTGACT
 AAAGCTGAGAGAGGGGGTCTGTGCAAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG
 AGACCAGACAGATCACAAAGCACGTGCTGATCTGATAGCAGGATGAACACAAAGTACGAT
 GAGAATGACAAACTGATCCGCGAAGTGAAGTCACTTACTCTGAAGTCAAAACTTGTGGAGCGACTT
 CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCAT
 50 ACCTGAACGCAGTGGTCGGGACCGCCCTGATTAAGAAATACCCCAAACTGGAGAGCGAATTCTGT
 TACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTGGAAA
 AGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCTGGCA
 AATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
 ACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCTGTCCATGCCTCAAGTGAATATTGTCAAGA
 55 AAACAGAGGTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAG
 CTGATCGCCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
 TTACTCTGTCTGGTGGTTCGCAAGGTGGAGAAGGGGAAAAGCAAGAACTGAAATCCGTCGAGG
 AGCTGTGGGTACACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGG
 CTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAG
 60 CTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGAGGAGAGCTGCAGAAGGGTAATGAACTGG
 CCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCC
 CCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATT

GAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAATCTGGATAAGGTCCTGAG
CGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAATATCATTTCATCTGTTCA
CTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
5 ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACCCAGAGCATTACCGGCCTGTATGAA
ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTC
TGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGC
GAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACCGGCCACAA
GTTACGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCT
10 GCACCACCGGCAAGCTGCCCCTGGCCCTGGCCACCCCTCGTGACCACCTTCACCTACGGCGTGACGT
GCTTCGCCCCGTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAG
TTCGAGGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGCACAAGCTGGAGTACAATAACAAGCCACAAGGTCTATATCACCGCCGACAAGC
15 AGAAGAACGGCATCAAGGTGAAGTTCAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGG
AGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGC
AGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGTTCCCTCCCCGTG
20 CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCTTTTCTTAATAAAAATGAGGAAATTGCATCG
CATTGTCTGAGTAGGTGTCTTCTATTTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA
CCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCTGTACAAAAAGCAGGCTTTAAAGGAACCAATT
CAGTCGACTGGATCCGGTACCAAGGTGGGCGAGGAAGAGGGCCTATTTCCCATGATTCCTTCATAT
25 TTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGAT
ATAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGAGTTTAAAATTATG
TTTTAAAATGGACTATCATATGCTTACCCTAAGTAAAGTATTTTCGATTTCTTGGCTTTATATATC
TTGTGGAAGGACGAAACACCGAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACCAGTCGGTCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTAAATCT
CTAGACATCATTAATTCCTAATTTTTTGTGACACTCTATCATTGATAGAGTTATTTTACCCTCCCT
30 ATCAGTGATAGAGAAAAGTGAAATGGCCAAGCCTTTGTCTCAAGAAAGTCCACCCTCATTGAAA
GAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCT
CTAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTTACTGGGGGACCTTGTGCAGAAC
TCGTGGTGTGTTGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAA
ATGAGAACAGGGGCATCTTGAGCCCCCTGCGGACGGTGTGACAGGTGCTTCTCGATCTGCATCCTG
35 GGATCAAAGCGATAGTGAAGGACAGTGATGGACAGCCGACGGCAGTTGGGATTTCGTGAATTGCTG
CCCTCTGGTTATGTGTGGGAGGGCTAAATGGTCCATATGAATATCCTCCTTAGTTCTTATCCGCTA
GCCTAGAGGGACAGCCCCCCCCCAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGG
CAGCAGCGAGCCGCGGGGCTCCGCTCCGGCTCCGCGCTCCCCCGCATCCCCAGCGCGGACG
GTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCTCTGAACGCTTCTCGCT
40 GCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTT
AGAATGACAGAATCATAGAACGGCCTGGGTTGCAAGGAGCACAGTGCTCATCCAGATCCAACCC
CCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTG
AATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCAACCCCTCT
GGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCGTGTCTCAGTGTAAGGCCATTCCCCC
45 TTGTCCTATCAAGGGGGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAAATCA
GTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATG
CAGGTGTTGAGGGCTCTGGGACACTTCCAAGTACAGCGTTCAGAACAGCCTTAAGGATAAGCAA
GATAGGATAGAAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCACA
GACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTG
50 GAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCC
ATGTAGATGTTTATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACC
AACCCCAACCCACCCACCGTGGCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCAT
CACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCGAGCTGTGCCACTGCAGCACCGCTCTTTGG
AGAAGGTAAATCTTGCTAAATCCAGCCCCGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCA
55 TCCAACCTCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATC
GCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAATATATATAT
AATAAGTTATACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCTTAAAA
GTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGTCTGTTATAGTTCAAAATCAGTGACA
CTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCAC
60 AGCGACGGATTTCGCGCTATTTAGAAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTAC
GCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACACGTCAT
AGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGGA

GGTTCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTACAATTCCTGGCCG
 TCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATC
 CCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA
 5 GCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTTGTAA
 AATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGAC
 CGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA
 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCA
 AGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCCGATTAG
 10 AGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGCGGG
 CGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCGCGCTTAATG
 CGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 19)

TOIC_Cas9_Nickase_Obl_r26_AAVS_PgkNeo

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATAT
 15 GTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA
 GTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGGCGCATTTTGCCTTCCTGTTTTTGCTCAC
 CCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGA
 ACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAGAAGCTTTTCCAATGATGAG
 20 CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG
 TCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
 GGATGGCATGACAGTAAGAGAATTATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCA
 ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGATC
 ATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
 25 ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTA
 GCTTCCCGGCAACAATTAAGACTGGATGGAGGCGGATAAAGTTGCAGGACCCTTCTGCGCTC
 GGCCCTTCCGGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
 CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA
 GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
 30 AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTAAATTTAAAG
 GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCAC
 TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATC
 TGCTGTTGCAACAACAAAACACCGCTACCAGCGGTGGTTGTTTGCCGGATCAAGAGCTACCA
 ACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGATG
 35 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTG
 TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
 CCGGATAAAGGCGCAGCGGTGCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
 GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
 GAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
 40 AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT
 CCTGGCCTTTTGTGCTGCTTTTGTCTACATGTTCTTTCCCTGCGTTATCCCTGATTCTGTGGATAACC
 GTATTACCGCCTTTGAGTGAGCTGATACCGTCTCGCCGAGCCGAACGACCGGAGCGCAGTGATCA
 GTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
 45 TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
 TGAGTTAGCTCACTCATTAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG
 AATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAA
 TTAACCCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTGTGACG
 TACGTTAAAGATAATCATGCGTAAATTTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTATT
 50 TATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCACAAA
 CAATTTATTTATGTTTATTTATTTATTAATAAAAAAACAACAACTCAAAATTTCTTCTATAAAGTAAC
 AAAACTTTTAAACATTCTCTTTTACAAAAATAAACTTATTTGTACTTTAAAAACAGTCATGTTG
 TATTATAAATAAGTAATTAGCTTAACCTTATACATAATAGAAACAAATTATACTTATTAATCGCAT
 TGATTATTGACTAGTCGTATTAAGGGTTCGGATCAGCTTGATTTCGAGCCCCAGCTGGTTCTTTCCG
 55 CCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTT
 GCTGTCCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTT
 CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGG
 GGCAACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAA
 TTGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGG
 60 CGATACCGTAAAGCACGAGGAAGCGGTACGCCATTCGCCGCCAAGCTCTTCAGCAATATCACGG
 GTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGA
 AAAGCGGCCATTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTC

GCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTC
GTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTT
CGCTTGGTGGTCAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCA
5 TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCC
AATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAACGCCCCGT
CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTACAGGGCACCGGACAGGTGCGT
CTTGACAAAAAGAACCGGGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGCTGTGTGTCGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGC
10 AATCCATCTTGTTCATGCGCGATCCCATGGTTTATGTTCCCTCACCTTGTCGTATTATACTATGCCGA
TATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTGCGAAAGGCCCGGAGATGAGGA
AGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGTGCGAATGCCGGGCTCCGGAGGACC
TTCGGGCGCCCCGCCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCCCGGACCCACCCCTTCCAGCC
TCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCAAAGGCCTACCCGCTTC
15 CATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCACGTC
CTGCACGACGCGAGCTGCGGGGCGGGGGGGAACCTTCCTGACTAGGGGAGGAGTAGAAGGTGGCG
CGAAGGGGCCACCAAGAACGGAGCCGTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCC
AGAGGCCACTTGTGTAGCGCAAGTGCCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTT
GGGAAAAGCGCTCCCTACCCGCTAGACACCCACAGTGGGTGGCCTAGGGAGACAGATTGCAAC
20 TCCAGTCTTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAG
TTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
TCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGT
25 ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTA
TTTATTTATTTTTAATTATTTTGTGACGCGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCG
GGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGA
GCGGCGCGCTCCGAAAGTTCTTTTATGGCGAGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGC
30 AGCGCGCGGGCGGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCGCCCTCGCGC
CGCCCCGCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCT
CCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTCTGTGGCTGCGTGAAAGCCTTA
AAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTG
CGTGGGAGCGCCGCGTGCGGCCGCGCTGCCCGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGG
35 GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGGG
GGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGGTGTGG
GCGCGGCGGTGCGGCTGTAAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCCGCT
TCGGGTGCGGGCTCCGTGCGGGGCGTGGCGGGGCTCGCGTGCCGGGCGGGGGGTGGCGGCA
GGTGGGGGTGCGGGCGGGGCGGGGCGGGGCGGGGAGGGGCTCGGGGAGGGGTCGGGGCGCGGCG
40 GGCCCCGAGCGCGCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCG
TGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGC
CGCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGCGCAGGAAGGAAATGGGCGGGGA
GGGCCTTCGTGCGTGC CGCGCGCCGCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGG
GACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTCCGCTTCTGGCGTGTGACCGGCGGCTC
TAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTT
45 ATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAAGTC
TCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCTGACGACAAGGAAACTCGCTCAAAGCTGGGAGTTGAGCAGCCTAACCTGTACTGCCACG
TGAAGAACAAGCGGGCCCTGCTCGATGCCCTCGGAATCGAGATGCTGGACAGGGATCATACCCAC
TCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGT
50 GCTCTCCTCTCACATCGCGACGGGGCTAAAGTGATCTCGGCACCCGCCCAACAGAGAAACAGTA
CGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCACTGTA
CGCTCTGTCCGCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGC
AAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGC
TGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAATAATCATATGTGGCCTGG
55 AGAAACAGCTAAAGTGCGAAAGCGGCGGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATG
CTCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTG
ACCTTGACATGCTGCCCCGGTAAAGCGGCGCGACTCTAGATCATAATCAGCCATACCACATTTGT
AGAGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
AATTGTTGTTGTTAACTTGTATTGTCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA
60 TTTCACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT
TAAGGGATCCCTAGAGGGACAGCCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCC
GCTAGGGGACGAGCGAGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAG

CCGGCAGCGTGCAGGGGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCTCTGAACGC
TTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAG
AGAGATTTAGAATGACAGAATCATAGAACGGCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGA
TCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCCAGAGCCACATCCAGCC
5 TGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCAC
CACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCC
ATTCCCCCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGC
CAATTCAGTGTCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGG
GACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGA
10 TAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAG
GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCC
AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
15 CTTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
CTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
CTCTCATCCAATCCAGGACGGAGTCAGTGAGGATGGGGCTGTGACCTAGAGGGACAGCCCCC
CCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCGCGGG
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGCA
20 CGGGGAAGGTGGCACGGGATCGCTTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACAC
CTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAA
CGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATC
AACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGC
ATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCAACCTCTGGGGGAAAAACTGCCTCCTC
25 ATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGGCCATTCCCCCTGTCTATCAAGGGGGAGT
TTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTTCAGTGCATCACGGAGAGGCAGA
TCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGA
CACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCT
30 GAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAAGGGGTGGAAGAGCTTGCTGGAGAGAT
ACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTG
CCCCTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCC
CCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAAT
35 CCAGCCCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCA
GTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCT
ATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTA
CTCCCTATCAGTGATAGAGAACGTATGACAGTTTACTCCCTATCAGTGATAGAGAACGTATCTAC
AGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT
40 ATAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTGTGAACCGTCAGATC
GCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAA
AAGCTTGTCACCATGGCTCCTAAGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGG
CCATCGGAATAACTCAGTGGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAA
TTTAAGGTCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGATTGGGGCCCTGCTGTTT
45 GACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCG
AAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTT
CTTTCATCGGCTGGAGGAATCATTCTGTCGAGGAAGATAAGAAACACGAGAGATCCTATCCTT
TGGAACATTGTGGACGAGGTGCTTATCACGAAAAATACCCACCATCTATGCTGCGCAAGT
AACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTA
50 AGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTG
TTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTATTAATGCATCTGGCGTG
GACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCGCTCA
GCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGGAATCTGATTGCACTGTCACTGGGACTGACAC
CTAACTTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTAT
55 GACGATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTTCTGGCCGCT
AAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCC
CCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGC
ACTGTGCGCCGAGAGCTGCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACCGGT
ACGCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
60 AGAAGATGGACGGCACCGAGGAAGTCTGGTGAAGTGAATCGCGAGGACCTGCTGCGCAAGCA
GCGGACATTTGATAACGGCTCCATCCCCACAGATTCTGTTGGAGAGCTGCACGCAATCCTGCG
ACGACAGGAAGACTTCTACCCATTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCT

TCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGGGTAATTCCCGCTTTGCCTGGATGACACGGA
AATCTGAGGAAACAATCACTCCTTGAACTTCGAGGAAGTGGTCGATAAGGGAGCTTCCGCACAG
TCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCCTAAGCA
5 CAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGG
GATGAGGAAGCCCGCCTTCCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAAA
CCAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAATTGAATGTTTCGAT
TCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTG
AAGATCATTAAAGGATAAAGACTTCCTGGACAACGAGGAAAAATGAGGATATCCTGGAAGACATTGT
10 GCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTTATGCCCATCT
GTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAGC
AGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAAGACTATCCTGGACTTTCTGAAATC
AGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTACGATGACAGCCTGACCTCAAAGAGGA
TATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCG
15 GGTCCTCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTCGATGAGCTGGTGAAAGTC
ATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCAAAATCAGACCACACAGA
AAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATC
CCAGATCCTGAAAGAGCACCCCTGTGGAACAACTCAGCTGCAGAATGAGAAGCTGTATCTGTATT
ATCTGCAGAATGGGCGGGATATGTACGTGGACACGAGCTGGATATTAACCGCTGTCTGATTAC
20 GACGTGGATCATATCGTCCCACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACC
CGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGA
AGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGCGAAAGTTTGATAACCTGACT
AAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGG
AGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGAT
25 GAGAATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTT
CAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCAT
ACCTGAACGCAGTGGTCGGGACCGCCCTGATTAAGAAATACCCCAAACTGGAGAGCGAATTCTGTG
TACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTTGAAA
AGCCACCGCTAAGTATTTCTTTACTCAAAACATCATGAATTTCTTTAAGACTGAGATCACCTGGCA
30 AATGGGGAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGG
ACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGA
AAACAGAGGTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAG
CTGATCGCCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGC
TTACTCTGTCTGGTGGTTCGCAAAGGTGGAGAAGGGGAAAAGCAAGAACTGAAATCCGTCAAGG
35 AGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGG
CTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAG
CTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAAGTGG
CCCTGCCTTCTAAGTACGTGAACCTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCC
CCGAGGATAACGAAACAGAAACAGCTGTTTGTGGACGACGACAAGCATTATCTGGACGAGATCATT
40 GAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCTCTGAG
CGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAAATATCATTATCTGTTCA
CTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAACGAT
ACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACACAGAGCATTACCGGCCTGTATGAA
ACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTC
45 TGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGC
GAGGAGCTGTTACCGGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACCGGCCACAA
GTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTTCATCT
GCACCACCGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTTACCTACGGCGGAGCT
GCTTCGCCCCGCTACCCCGACCACATGAAGCAGCAGCTTCTTCAAGTCCGCCATCGCCGAGGCT
50 ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAG
TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAAGGTCTATATCACCGCCGACAAGC
AGAAGAACGGCATCAAGGTGAACCTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCATCGGGCAGGGCCCCGTGCTGCTGCCCGACAACCACTA
55 CCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTCTGCTGG
AGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGC
AGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTG
CCTTCCTTGACCCTGGAAGGTGCCACTCCCAGTGTCCTTTCCCTAATAAAATGAGGAAATGTCATCG
CATTGCTGAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
60 TTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA
CCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCATATGAATATCCTCCTTAGTTCTTATT
CCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCT
AGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCG

GCAGCGTGCGGGGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTC
 TCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAAGAGA
 GATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCC
 AACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTG
 5 GCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCGTCACCA
 CCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCAT
 TCCCCCTTGCTCATCAAGGGGGAGTTTGTCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCA
 ATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGG
 ACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGAT
 10 AAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAG
 GCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
 GGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
 AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCC
 AAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTT
 15 CTTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCT
 CTTTGAGAAAGGTAAATCTTGCTAAATCCAGCCGACCTCCCCTGGCACAACGTAAGGCCATTAT
 CTCTCATCCAATCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTAC
 ACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAAATAT
 ATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATC
 20 TTAAGTACGTAAGATAATCATGCGTCATTTTGACTCACGCGTTCGTTATAGTTCAAAATCA
 GTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAA
 ATGCACAGCGACGGATTTCGCGCTATTTAGAAAAGAGAGCAATATTTCAAGAATGCATGCGTCAA
 TTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACAC
 GTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCG
 25 GGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTTCGTATTACAATTCAGTG
 GCCGTGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGACGCA
 CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
 CGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTT
 TGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAA
 30 TAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGTGGA
 CTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT
 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCCGA
 TTTAGAGCTTGACGGGGAAAGCCGCGCAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAG
 CGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTT
 35 AATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO: 20)

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

40 The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method
 45 are recited.

In the claims, as well as in the specification above, all transitional phrases such as “comprising,” “including,” “carrying,” “having,” “containing,” “involving,” “holding,” “composed of,” and the like are to be understood to be open-ended, *i.e.*, to mean including but not limited to. Only the transitional phrases “consisting of” and “consisting essentially
 50 of” shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

What is claimed is:

CLAIMS

1. A engineered nucleic acid construct comprising:

- 5 (a) a promoter operably linked to a nucleic acid encoding a regulatory protein;
(b) an inducible promoter operably linked to a nucleic acid encoding an enzyme that cleaves nucleic acid, a nucleic acid encoding an enzyme that nicks nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid, wherein activity of the inducible promoter is modulated by the regulatory protein;
10 (c) at least two insulators located downstream from (a) and upstream from (b);
(d) at least one insulator located downstream from (b) and upstream from (a);
and
(e) at least one deoxyribonucleic acid (DNA)-binding domain recognition sequence located downstream from (b) and upstream from (a).

15 2. The engineered nucleic acid construct of claim 1, comprising at least two DNA-binding domain recognition sequences.

3. The engineered nucleic acid construct of claim 2, wherein each of the at least two
20 DNA-binding domain recognition sequences are obtained from different species.

4. The engineered nucleic acid construct of claim 3, wherein at least one of the DNA-binding domain recognition sequences corresponds to a nucleotide sequence in a AAVS1 locus of a human genome.

25 5. The engineered nucleic acid construct of claim 3, wherein at least one of the DNA-binding domain recognition sequences corresponds to a nucleotide sequence in a Rosa26 locus of a mouse genome.

30 6. The engineered nucleic acid construct of claim 2, wherein at least one DNA-binding domain recognition sequence is located upstream of and adjacent to (a), and wherein at least one DNA-binding domain recognition sequence is located downstream of and adjacent to (b).

35 7. The engineered nucleic acid construct of any one of claims 1-6, further comprising a promoter operably linked to a nucleic acid encoding a selectable marker protein.

8. The engineered nucleic acid construct of any one of claims 1-7, wherein the enzyme is a nuclease, a nickase or a recombinase.

5 9. The engineered nucleic acid construct of claim 8, wherein the enzyme is a nuclease.

10. The engineered nucleic acid construct of claim 9, wherein the nuclease is a Cas nuclease.

10 11. The engineered nucleic acid construct of claim 10, wherein the Cas nuclease is a Cas9 nuclease.

12. The engineered nucleic acid construct of claim 11, wherein the Cas9 nuclease is a wild-type Cas9 nuclease.

15

13. The engineered nucleic acid construct of claim 11 wherein the Cas9 nuclease is a catalytically inactive Cas9 nuclease.

14. The engineered nucleic acid construct of claim 13, wherein the catalytically inactive
20 Cas9 nuclease is fused to a transcriptional activator peptide, transcriptional repressor peptide, or an epigenomic regulator peptide.

15. The engineered nucleic acid construct of claim 9, wherein the nuclease is a CpfI nuclease.

25

16. The engineered nucleic acid construct of any one of claims 10-15, wherein the construct further comprises a nucleic acid encoding a guide RNA.

17. The engineered nucleic acid construct of claim 16, wherein the guide RNA is located
30 downstream of (b) and upstream of (d).

18. The engineered nucleic acid construct of claim 8, wherein the enzyme is a recombinase.

19. The engineered nucleic acid construct of claim 18, wherein the recombinase is Cre recombinase.

20. The engineered nucleic acid construct of any one of claims 1-19, wherein the
5 promoter of (b) is a tissue-specific inducible promoter.

21. The engineered nucleic acid construct of any one of claims 1-19, wherein the promoter of (b) is a developmental-specific inducible promoter.

10 22. The engineered nucleic acid construct of any one of claims 1-21, wherein the regulatory protein is a tetracycline-controlled transactivator (tTA) protein, a reverse tetracycline-controlled transactivator (rtTA) protein, or a Lac repressor protein.

15 23. The engineered nucleic acid construct of claim 22, wherein the regulatory protein is a rtTA protein.

24. The engineered nucleic acid construct of any one of claims 1-23, wherein the promoter of (a) is an inducible promoter.

20 25. The engineered nucleic acid construct of claim 24, wherein the promoter of (a) is a tissue-specific inducible promoter.

26. The engineered nucleic acid construct of claim 24, wherein the promoter of (a) is a developmental-specific inducible promoter.

25

27. The engineered nucleic acid construct of any one of claims 1-26, wherein the at least one DNA-binding domain recognition sequence is cleaved by a nuclease having a FokI nuclease domain.

30 28. The engineered nucleic acid construct of claim 27, wherein the nuclease is a zinc finger nuclease, a Tale nuclease, a Cpf1 nuclease or a hybrid dCas9-FokI nuclease.

29. A vector comprising the engineered nucleic acid construct of any one of claims 1-28.

35 30. The vector of claim 29, wherein the vector has a size of at least 20 kilobases.

31. The vector of claim 29 or 30, wherein the vector is a baculovirus vector, a bacterial artificial chromosome, or a yeast artificial chromosome.

5 32. A cell comprising the engineered nucleic acid construct of any one of claims 1-31.

33. A cell comprising the vector of any one of claims 29-32.

34. The cell of claim 32 or 33, wherein the cell is a stem cell.

10

35. The cell of claim 34, wherein the stem cell is a pluripotent stem cell.

36. The cell of claim 35, wherein the pluripotent stem cell is an induced pluripotent stem cell (iPSC).

15

37. A method of modifying a cell genome, comprising:

introducing into a cell an engineered nucleic acid construct comprising

(a) a promoter operably linked to a nucleic acid encoding a regulatory protein;

20

(b) an inducible promoter operably linked to a nucleic acid encoding an enzyme that cleaves nucleic acid, a nucleic acid encoding an enzyme that nicks nucleic acid, or a nucleic acid encoding an enzyme that catalyzes exchange of nucleic acid, wherein activity of the inducible promoter is modulated by the regulatory protein;

25

(c) at least two insulators located downstream from (a) and upstream from (b);

(d) at least one insulator located downstream from (b) and upstream from (a); and

30

(e) at least one deoxyribonucleic acid (DNA)-binding domain recognition sequence located downstream from (b) and upstream from (a);

introducing into the cell an engineered nucleic acid comprising a promoter operably linked to a nucleic acid encoding a guide RNA (gRNA) that targets a genomic region of the cell;

introducing into the cell an engineered nucleic acid comprising a promoter operably

linked to a nucleic acid encoding a hybrid nuclease that binds to the at least one DNA-binding domain recognition sequence; and

incubating the cell in the presence of an effector substance that modulates activity of the regulatory protein under conditions that result in modification of the cell genome.

5

38. A method of modifying a cell genome, comprising:

introducing into a cell an engineered nucleic acid construct comprising

(a) a promoter operably linked to a nucleic acid encoding a regulatory protein;

10

(b) an inducible promoter operably linked to a nucleic acid encoding an enzyme that cleaves nucleic acid, nicks nucleic acid, or catalyzes exchange of nucleic acid, wherein activity of the inducible promoter is modulated by the regulatory protein;

(c) at least two insulators located downstream from (a) and upstream from (b);

15

(d) at least one insulator located downstream from (b) and upstream from (a);

(e) a promoter operably linked to a nucleic acid encoding a guide RNA (gRNA) that targets a genomic region of the cell; and

20

(f) at least one deoxyribonucleic acid (DNA)-binding domain recognition sequence located downstream from (b) and upstream from (a);

introducing into the cell an engineered nucleic acid comprising a promoter operably linked to a nucleic acid encoding a hybrid nuclease that binds to the at least one DNA-binding domain recognition sequence; and

25

incubating the cell in the presence of an effector substance that modulates activity of the regulatory protein under conditions that result in modification of the cell genome.

39. An engineered nucleic acid comprising:

(a) a promoter operably linked to a nucleic acid encoding a Cas9 nuclease that does not comprise a nuclear localization signal, wherein the nucleic acid encoding the enzyme is flanked by estrogen receptor (ERT2) sequences; and

30

(b) a deoxyribonucleic acid (DNA)-binding nuclease recognition sequence.

40. A method of modifying a cell genome, comprising:

35

introducing into a cell an engineered nucleic acid comprising

(a) a promoter operably linked to a nucleic acid encoding a Cas9 nuclease that does not comprise a nuclear localization signal, wherein the nucleic acid encoding the enzyme is flanked by estrogen receptor (ERT2) sequences; and

(b) a deoxyribonucleic acid (DNA)-binding nuclease recognition
5 sequence;

introducing into the cell an engineered nucleic acid comprising a promoter operably linked to a nucleic acid encoding a guide RNA (gRNA) that targets a genomic region of the cell; and

incubating the cell in the presence of tamoxifen under conditions that result in
10 modification of the cell genome.

41. A transgenic mouse comprising in the genome of the mouse an engineered nucleic acid construct of any one of claims 1-31.

15 42. The transgenic mouse of claim 41, wherein the engineered nucleic acid construct is integrated in the Rosa26 locus of the mouse genome.

43. The transgenic mouse of claim 41 or 42, wherein the mouse is immunocompetent.

20 44. The transgenic mouse of any one of claims 41-43, wherein expression of the enzyme is not detectable in the absence of induction of the inducible promoter.

45. An engineered nucleic acid comprising the sequence of any one of SEQ ID NO: 8 or 12-20.

25 46. A transgenic mouse comprising in the genome of the mouse the engineered nucleic acid construct of claim 45.

47. The transgenic mouse of claim 46, wherein the engineered nucleic acid construct is
30 integrated in the Rosa26 locus of the mouse genome.

48. The transgenic mouse of claim 45 or 46 wherein the mouse is immunocompetent.

49. The transgenic mouse of any one of claims 46-48, wherein expression of the enzyme
35 is not detectable in the absence of induction of the inducible promoter.

Fig. 1

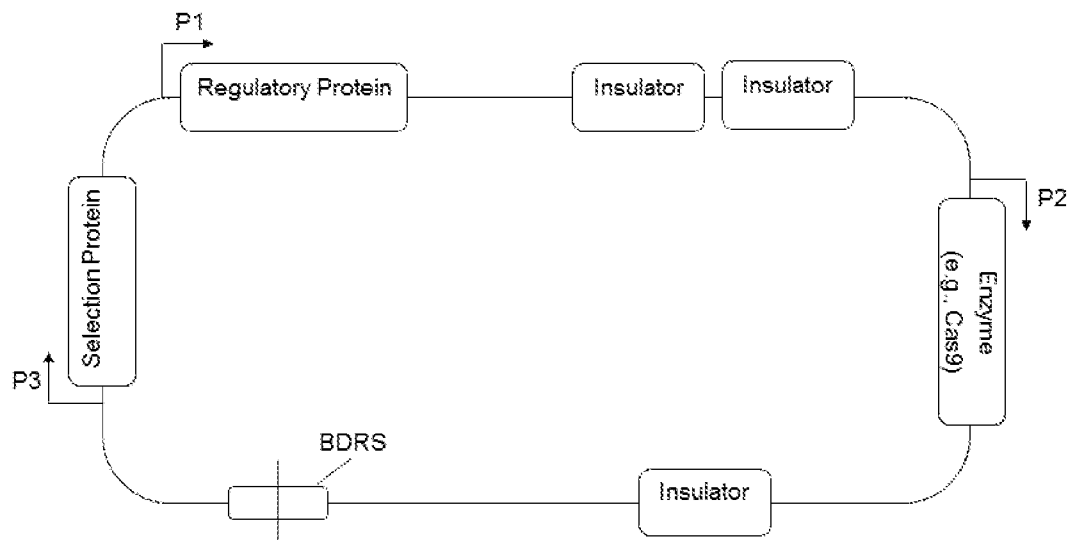


Fig. 2

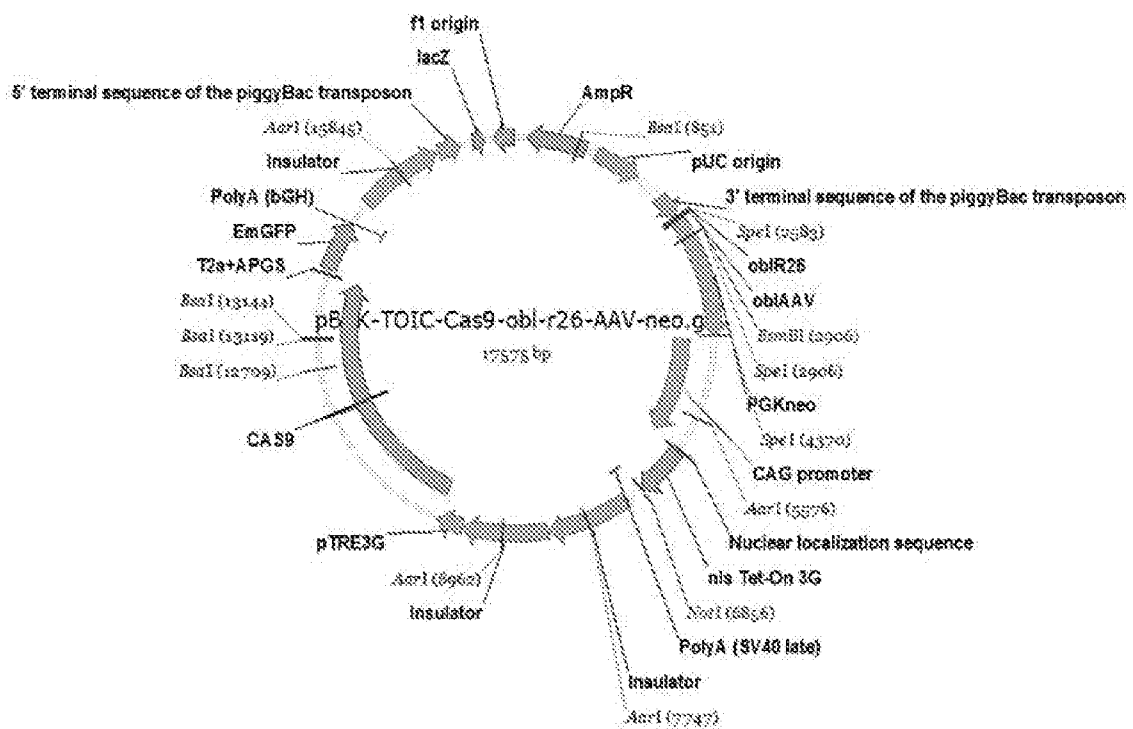


Fig. 3A

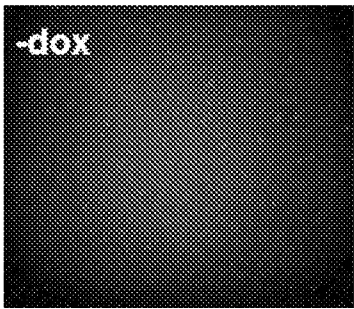


Fig. 3B

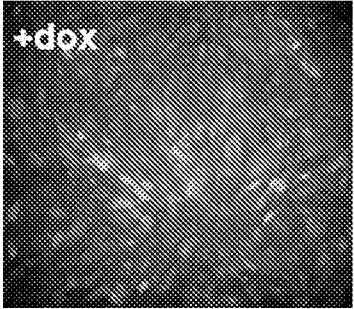


Fig. 4

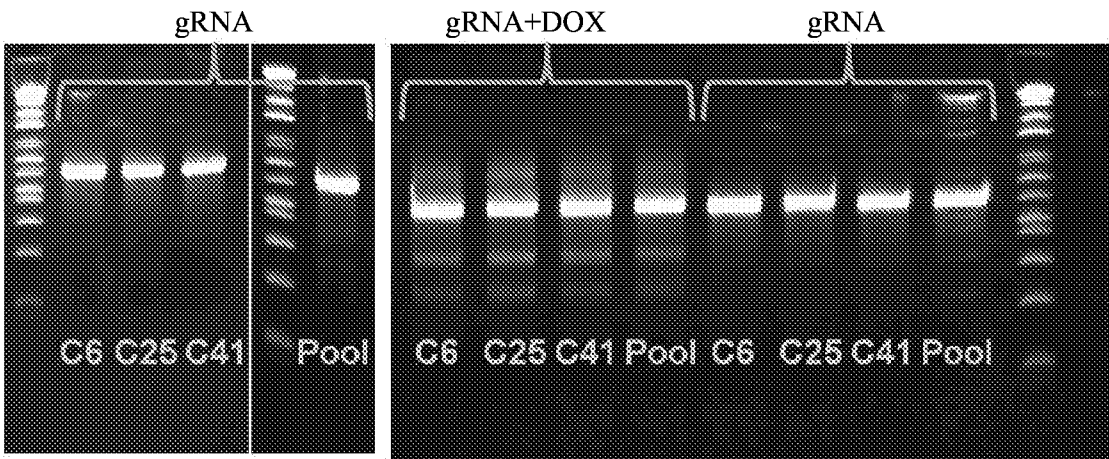


Fig. 5

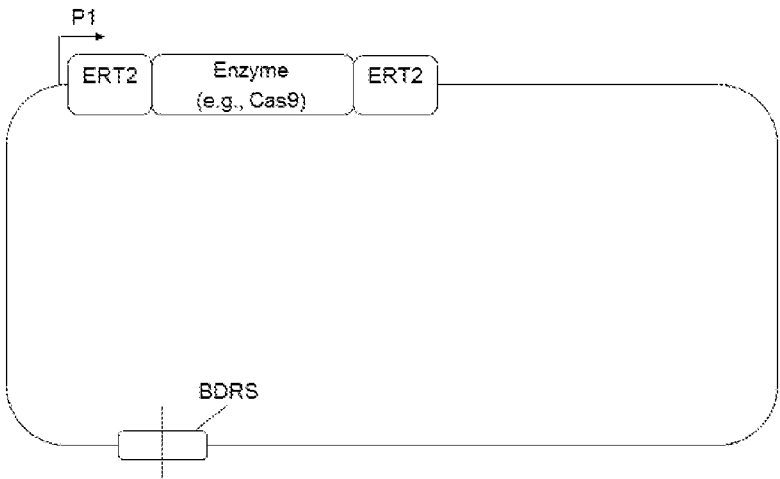


Fig. 6

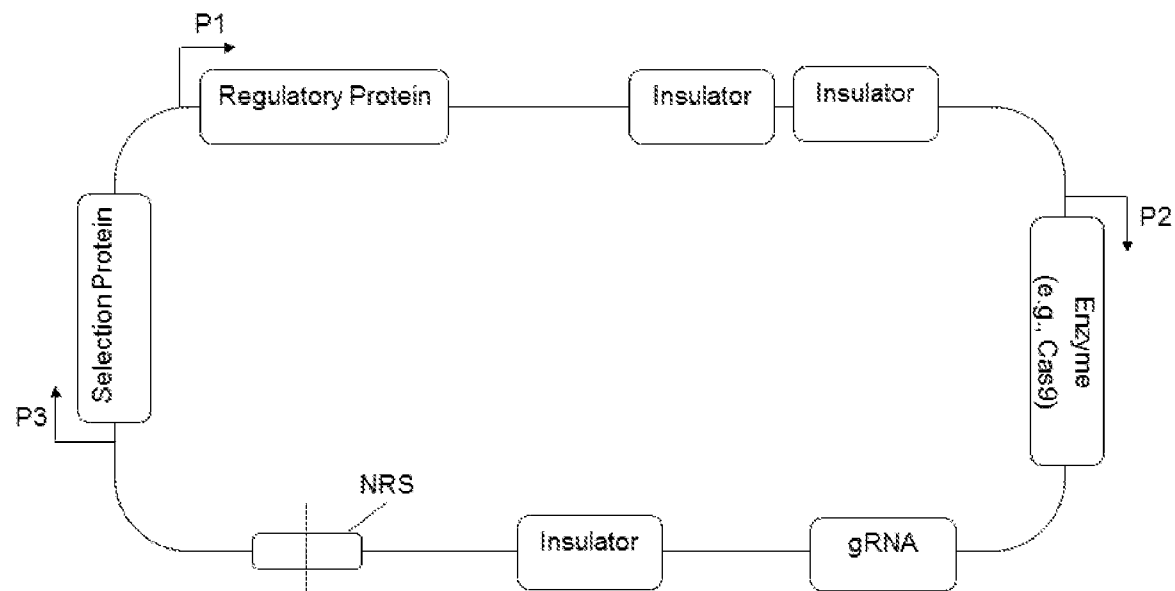
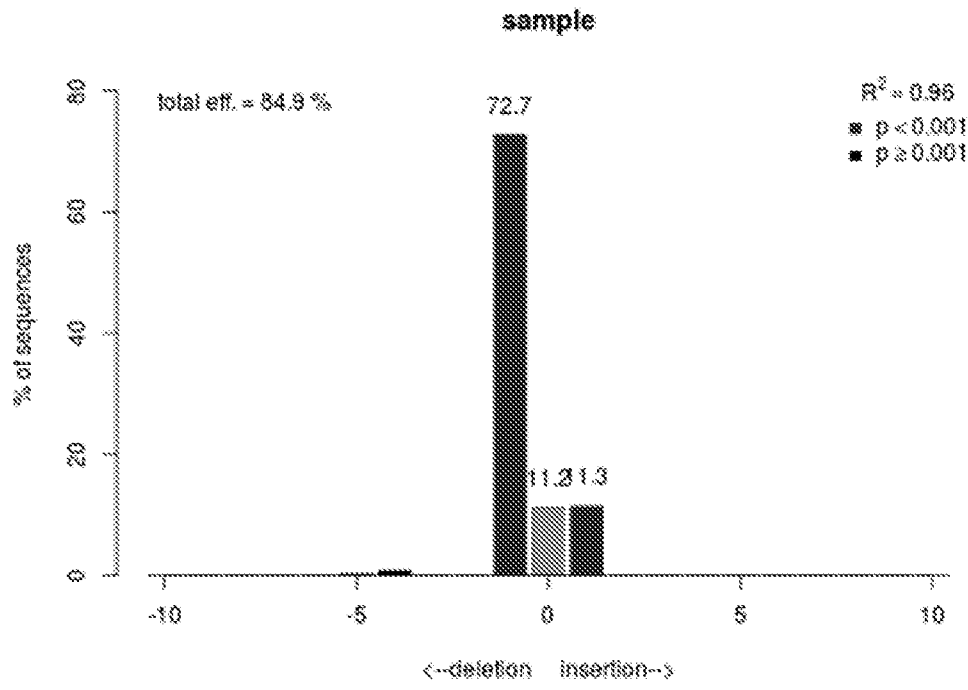


Fig. 7



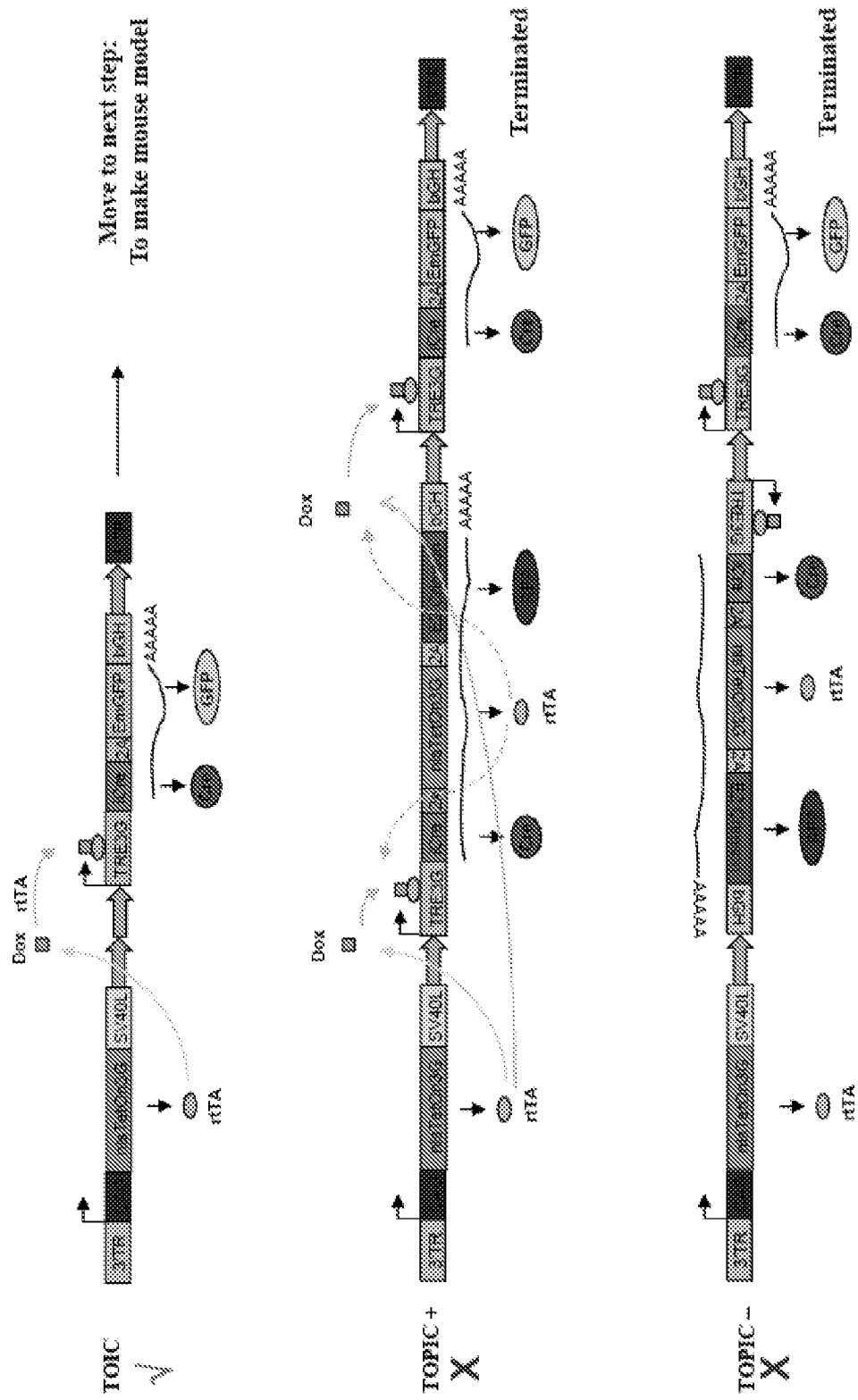


Fig. 8

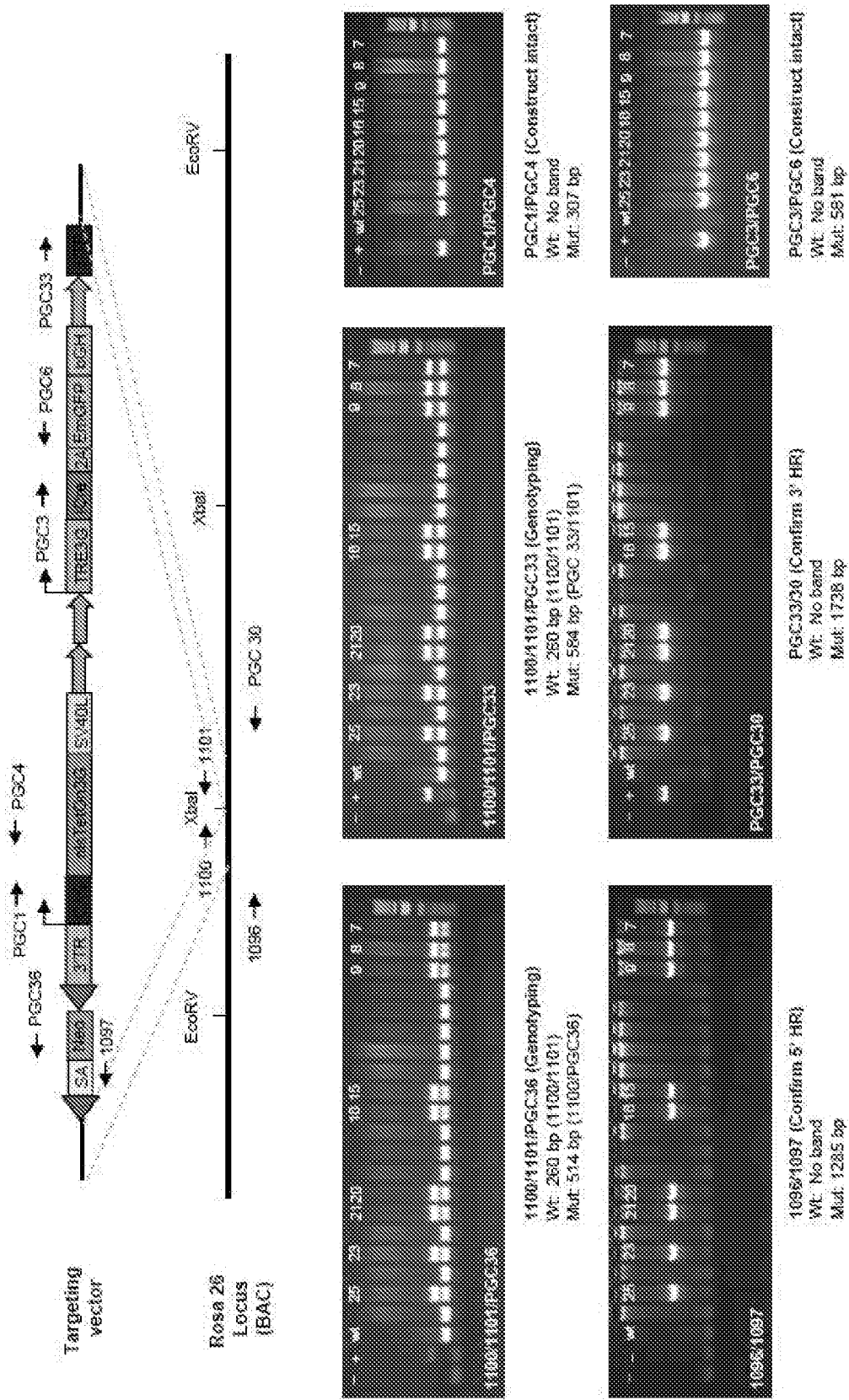


Fig. 9

Fig. 10



Fig. 11

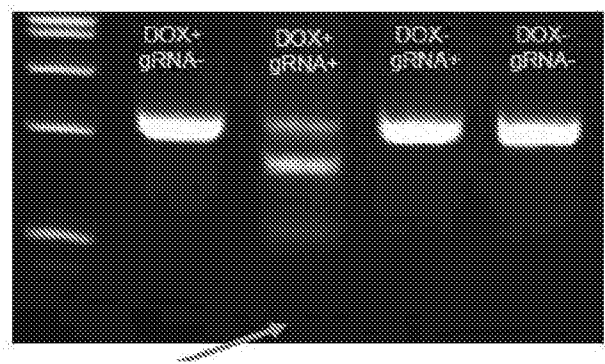


Fig. 12

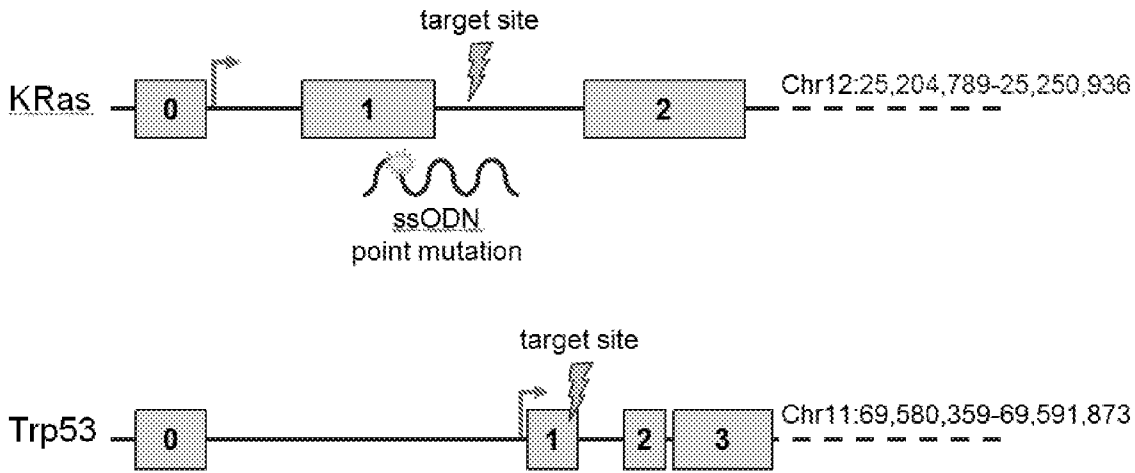


Fig. 13

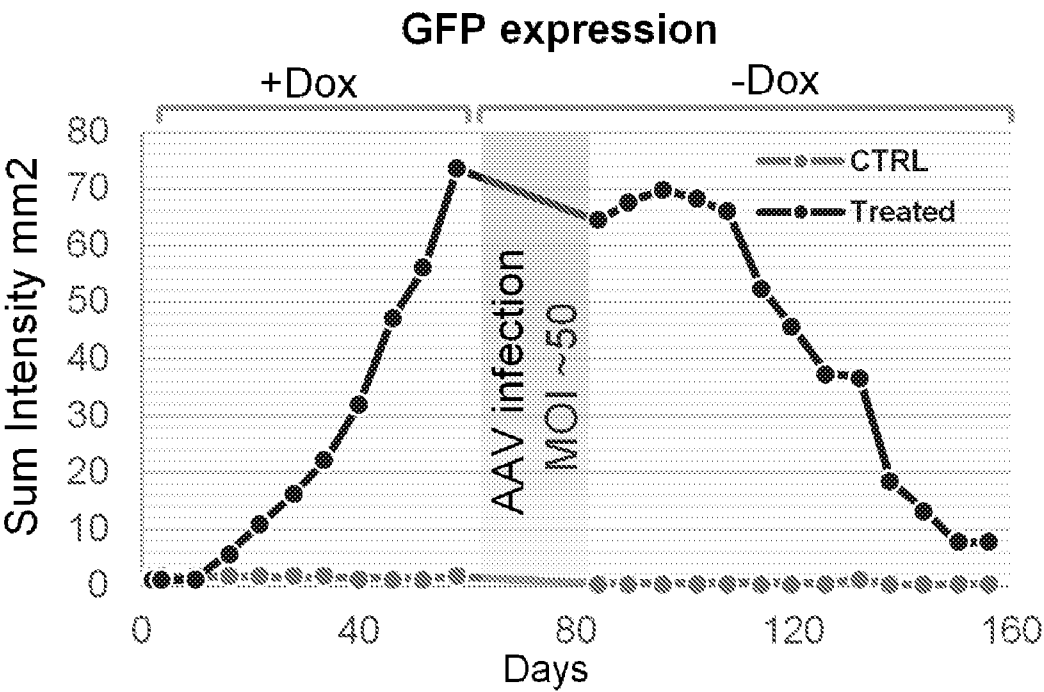


Fig. 14

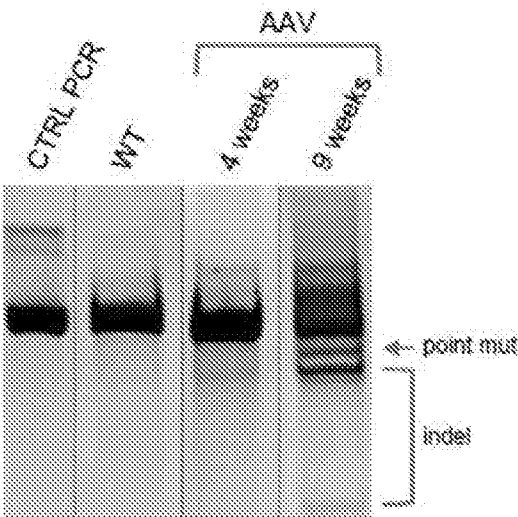
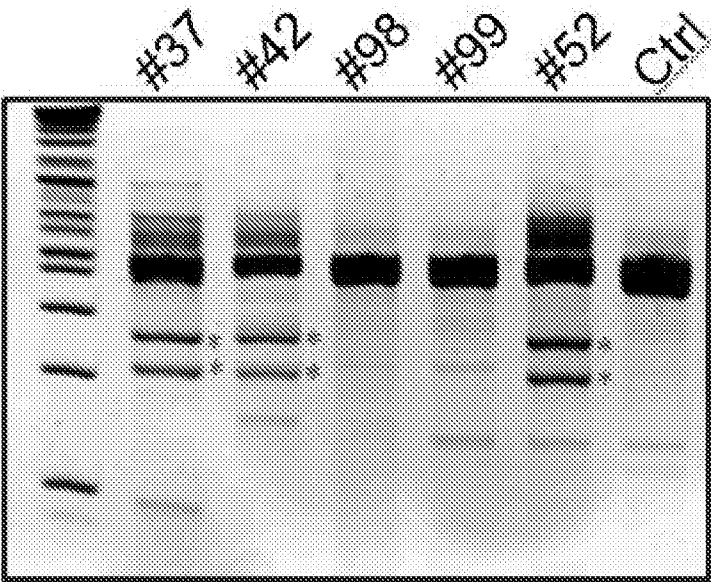


Fig. 15



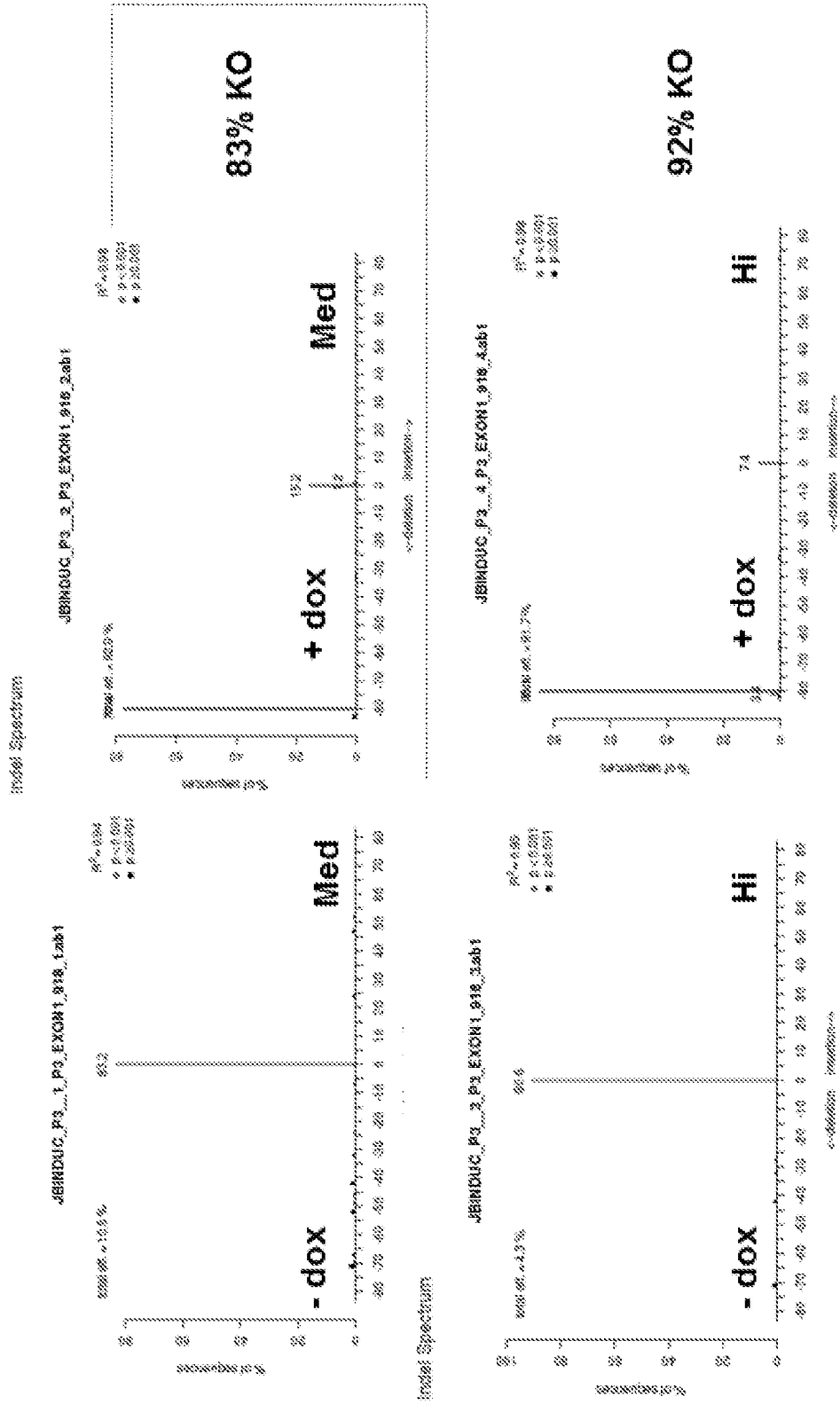


Fig. 16

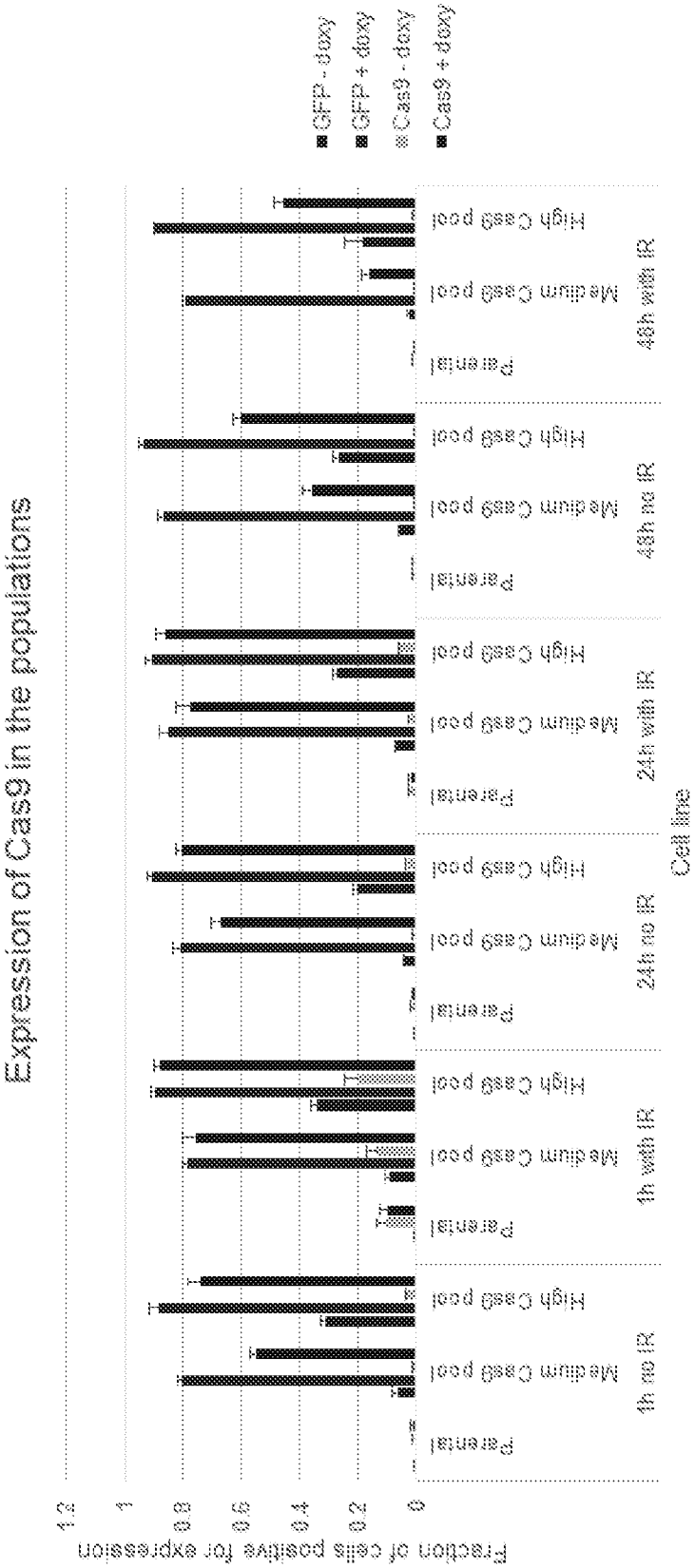


Fig. 17

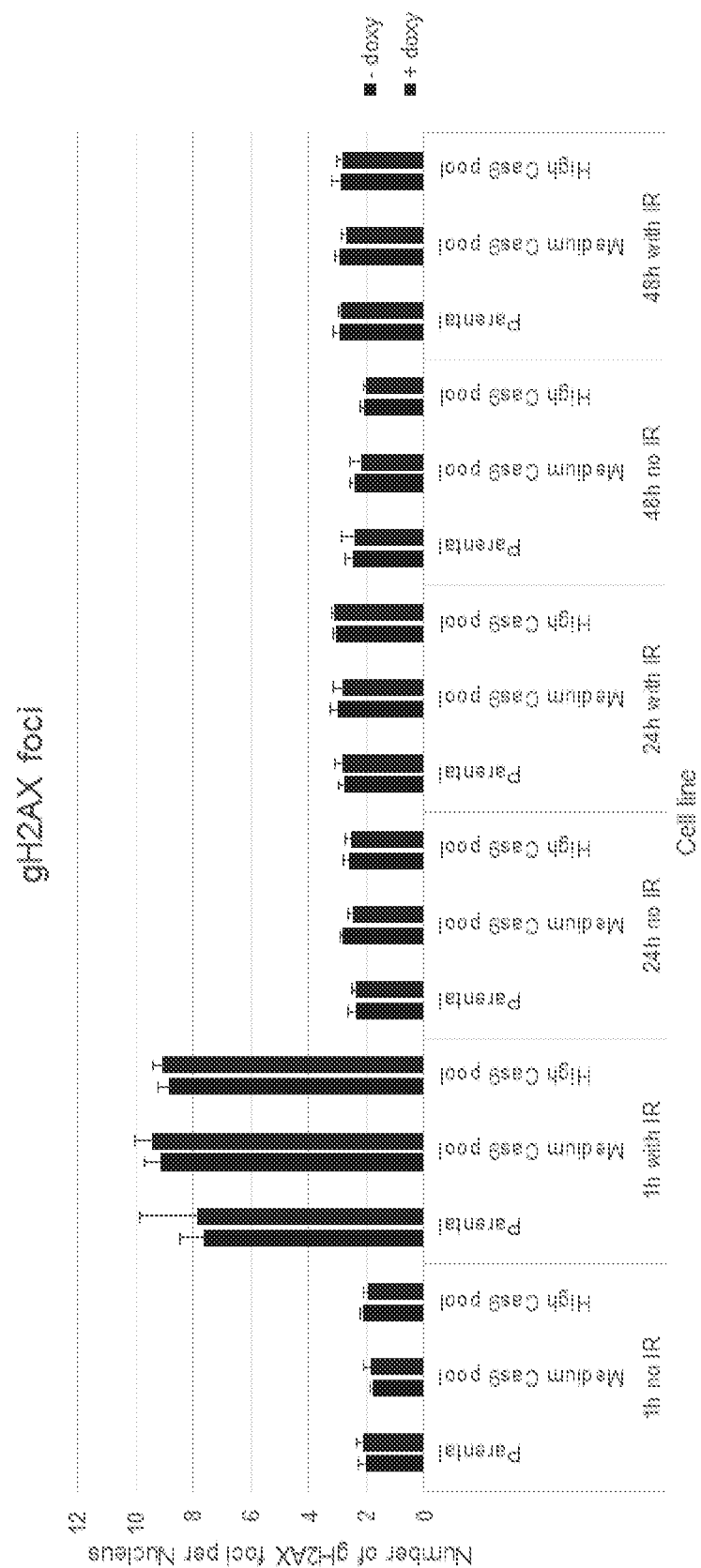


Fig. 18

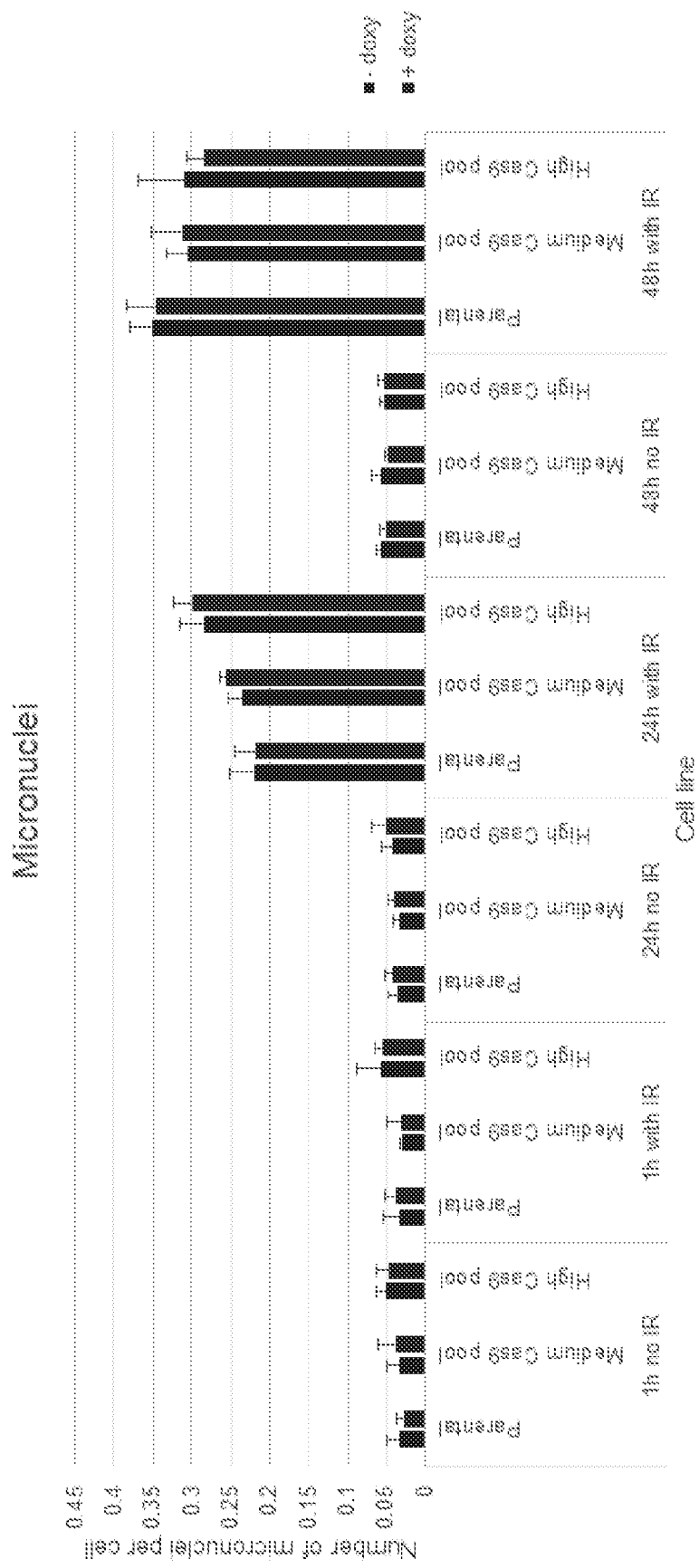


Fig. 19

Fig. 20

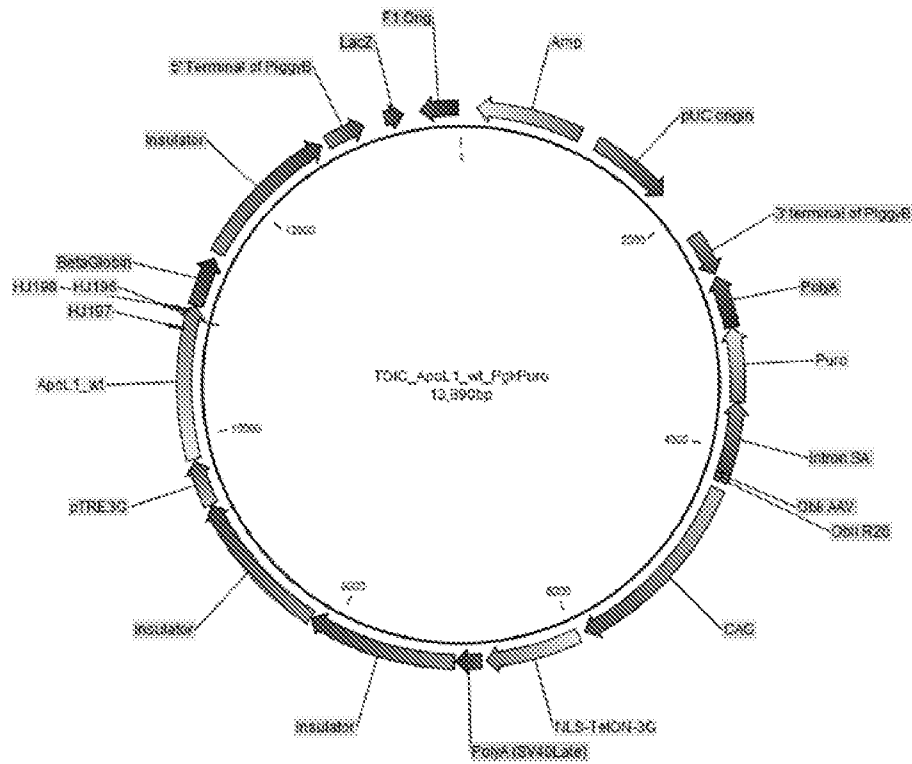


Fig. 21

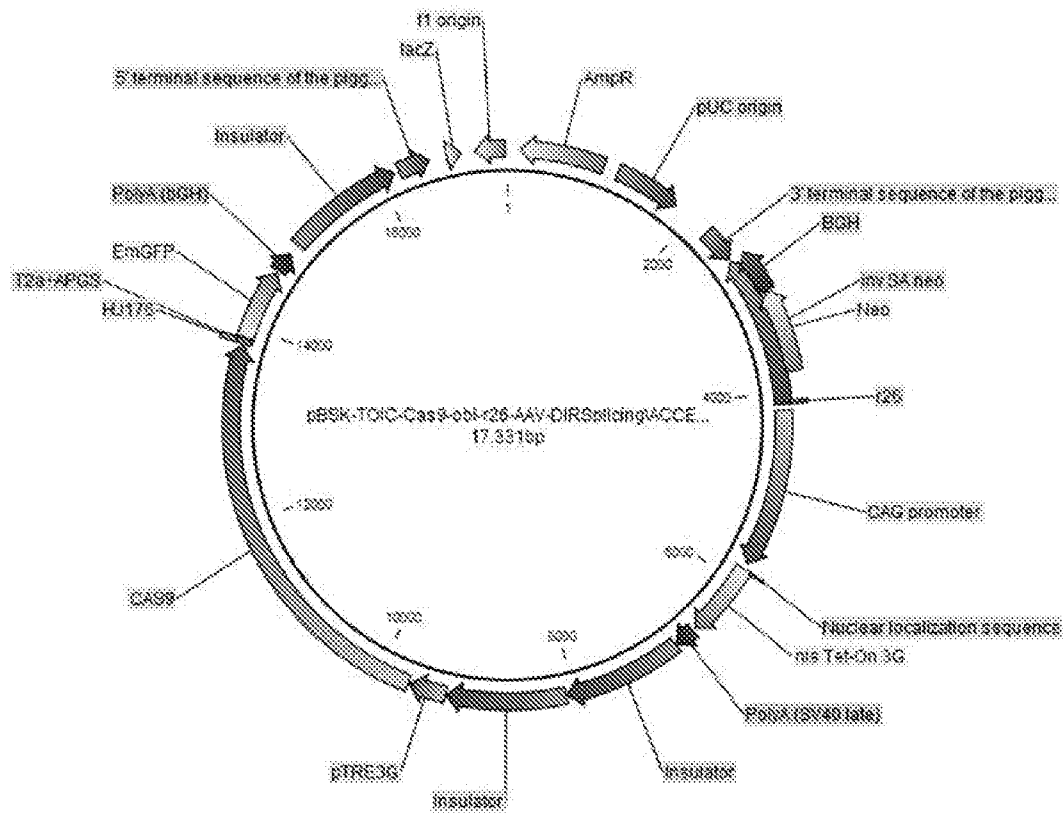


Fig. 22

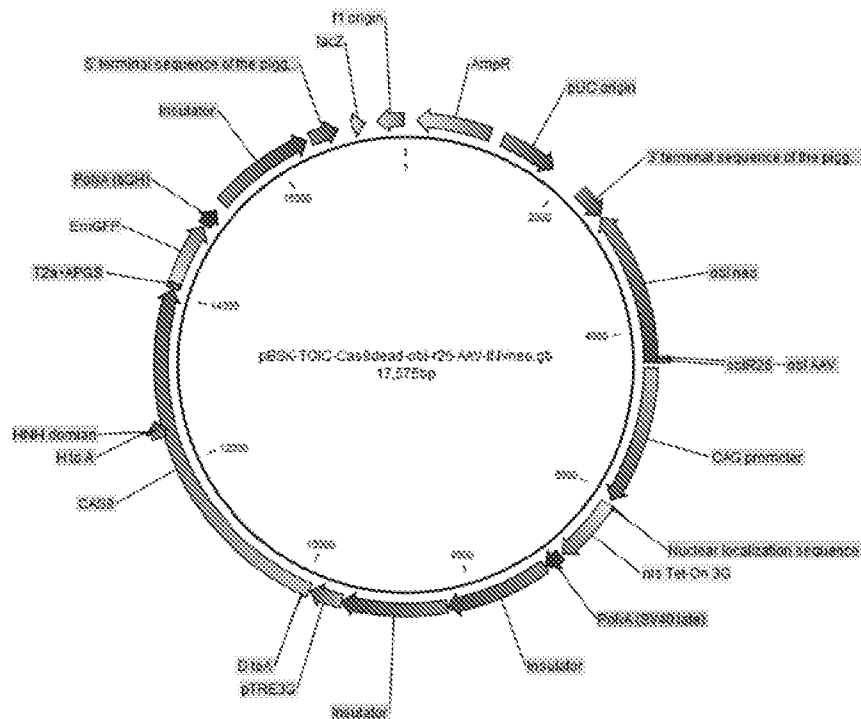


Fig. 23

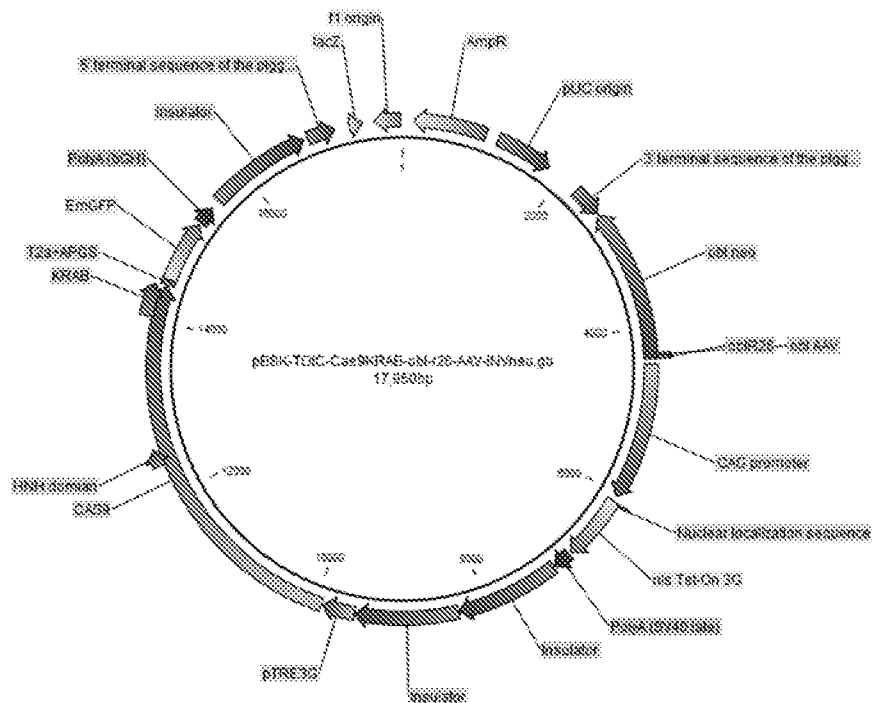


Fig. 24

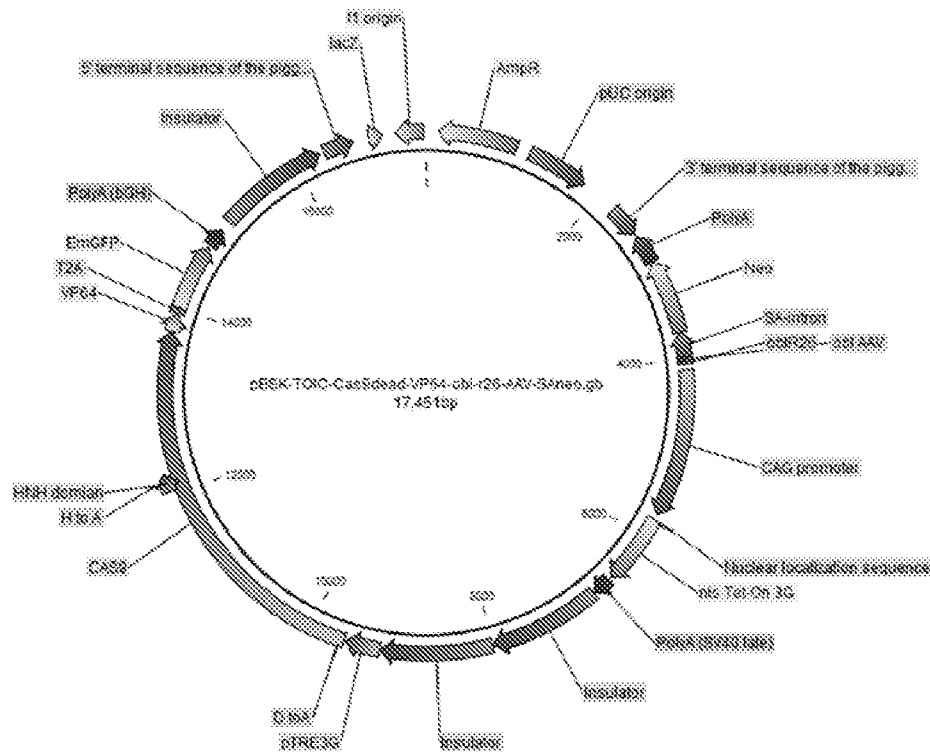


Fig. 25

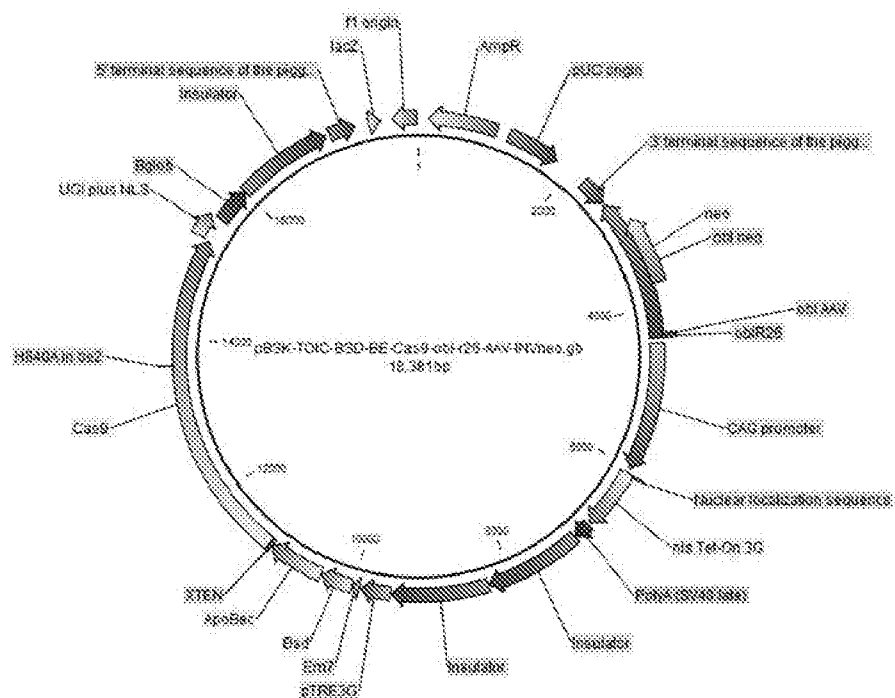


Fig. 26

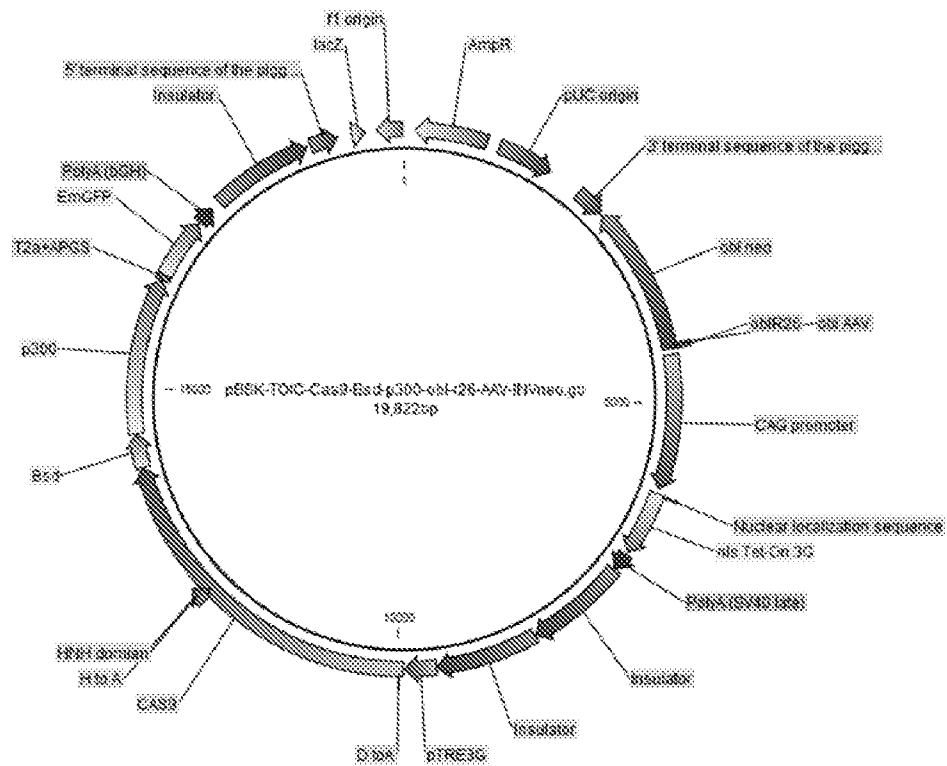


Fig. 27

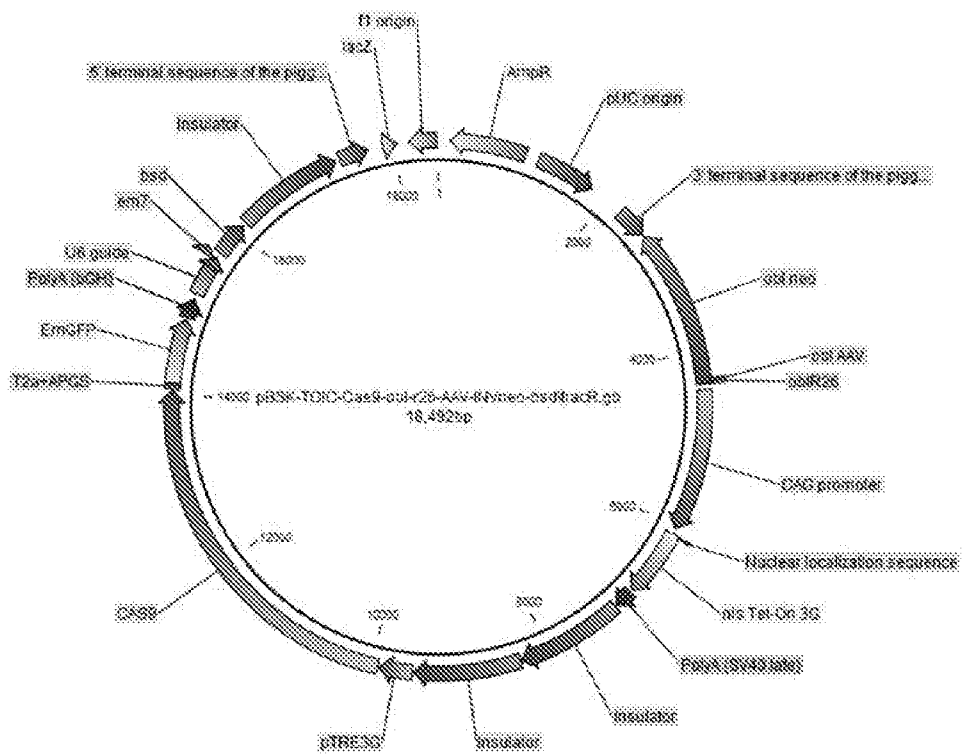


Fig. 28

