(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 May 2024 (23.05.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/107985 A2

(51) International Patent Classification:
C12N 15/861 (2006.01) *C12N 7/00* (2006.01)

(21) International Application Number:
PCT/US2023/080063

(22) International Filing Date:
16 November 2023 (16.11.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/426,028 16 November 2022 (16.11.2022) US
63/596,201 03 November 2023 (03.11.2023) US

(71) Applicant: **FORGE BIOLOGICS, INC.** [US/US]; 3900 Gantz Road, Grove City, Ohio 43123 (US).

(72) Inventors: **ADSERO, Angela M.**; 4869 Arbor Grove Court, Groveport, Ohio 43125 (US). **CHESTNUT, Brendan**; 3945 Blakewood Place, Unit 203, Groveport, Ohio 43125 (US). **PADEGIMAS, Linas**; 1214 Bonnie Lane, Mayfield Heights, Ohio 44124 (US). **DISMUKE, David**; 414 Selwood Place, Cary, North Carolina 27519 (US).

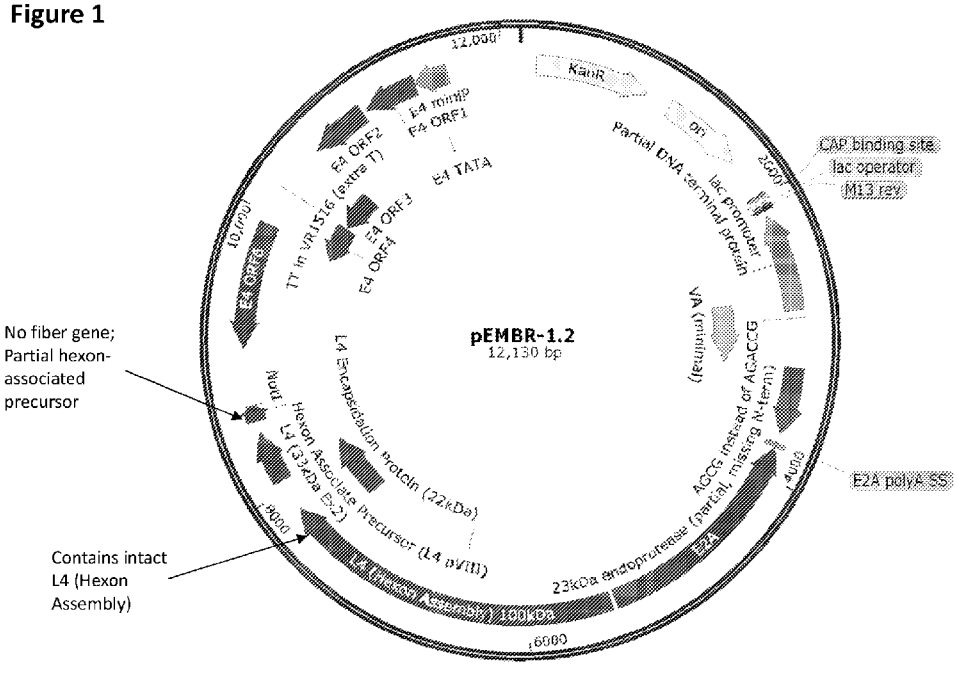
(74) Agent: **VETTER, Michael L.** et al.; CHOATE, HALL & STEWART LLP, Two International Place, Boston, Massachusetts 02110 (US).

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

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

(54) Title: ADENOVIRAL HELPER PLASMID

Figure 1



(57) Abstract: The present disclosure provides improved adenoviral helper plasmids for the production of recombinant adeno-associated viruses.

[Continued on next page]

WO 2024/107985 A2

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

ADENOVIRAL HELPER PLASMID

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to and the benefit of U.S. Provisional Patent Application Nos. 63/426,028, filed on November 16, 2022 and 63/596,201, filed on November 3, 2023, the entire contents of which are hereby incorporated by reference in their entirety.

BACKGROUND

[0002] Adeno-associated virus (AAV) technology has quickly become a dominant form of gene therapy for genetic diseases. AAVs can be produced in large scale in a variety of host cell systems, including mammalian cells, such as HEK293 cells. Traditionally, AAV production in mammalian cells involves the introduction of multiple plasmids to the host cells, the plasmids encoding, for example, a human gene or genes of interest, and various viral genes critical for viral replication and packaging. Due to the number of genes required for proper replication, these are traditionally delivered on two or three separate plasmids.

[0003] One such plasmid, termed an "adenoviral helper" plasmid, contains genes critical for AAV production from a host cell. Adenoviral helper plasmids containing E2a, VA RNA, and E4 genes have been shown to be critical to promoting AAV production in mammalian host cell systems.

[0004] Despite much advancement over the last two decades, concerns regarding the cost and safety of AAV production continue to limit the therapeutic potential of AAV technology. These concerns are due in part to the large size of many helper plasmids, which is due to the provision of a large number of genes on a single helper plasmid to support AAV production. The safety concerns are due in part to the production, albeit at low levels, of potentially cytotoxic and/or inflammatory viral proteins that are not necessary for AAV replication.

SUMMARY

[0005] In some embodiments, the present disclosure, provides, among other things, an adenoviral helper plasmid. In some embodiments, the present disclosure provides an adenoviral helper plasmid with reduced size relative to those known in the art. In some embodiments, the present disclosure provides an adenoviral helper plasmid comprising nucleotide sequences encoding E2a, VA RNA, E4; and an L4 region. In some embodiments, an adenoviral helper plasmid as described herein comprises nucleotide sequences encoding proteins from other viruses. In some embodiments, an adenoviral helper plasmid as described herein comprises nucleotide sequences encoding proteins from other viruses, including HSV-1 UL30, HSV-1 UL42, and/or HSV-1 UL29.

[0006] In some embodiments, the present disclosure provides an adenoviral helper plasmid that does not comprise one or more nucleotide sequences encoding one or more of fiber protein; L1-52/55K (Packaging Protein 3), peripentonal Hexon-Associated protein, and an L4 region. In some embodiments, the present disclosure provides an adenoviral helper plasmid comprising a fragment, portion, or partial form of E2a protein, VA RNA, E4, and an L4 region. In some embodiments, the present disclosure provides an adenoviral helper plasmid that does not comprise one or more nucleotide sequences encoding one or more of Hexon Associated Precursor (L4 pVIII) protein, DNA Terminal Protein, and 23kDa endoprotease. In some embodiments, the present disclosure provides an adenoviral helper plasmid that does not comprise one or more nucleotide sequences encoding one or more of E4orf1, E4orf2, E4orf3, and E4orf7. In some embodiments, an adenoviral helper plasmid provides herein comprises a kanamycin resistance gene.

[0007] In some embodiments, the present disclosure provides an adenoviral helper plasmid in which expression of E2a protein is under the control of one or more of an E2a promoter, chicken β-actin promoter, and SV40 promoter. In some embodiments, the present disclosure provides an adenoviral helper plasmid in which expression of E4 open reading frames (ORFs) is under the control of one or more of a native E4 promoter and SV40 promoter.

[0008] In some embodiments, the present disclosure provides an adenoviral helper plasmid comprising a nucleotide sequence that is at least 80% identical to SEQ ID NO: 1-3, 5, 7, 9, 11-12, 14-20, 22, 24, 26-29, 31, 33, 35-37, 39-53, 56, 59, 62, 65, 68 or 80-105. In

some embodiments, the present disclosure provides an adenoviral helper plasmid comprising a nucleotide sequence that encodes for an amino acid sequence that is at least 80% identical to SEQ ID NO: 4, 6, 8, 10, 13, 21, 23, 25, 30, 32, 34, 38, 55, 58, 61, 64, 67, or 69. In some embodiments, the present disclosure provides an adenoviral helper plasmid comprising a nucleotide sequence that is at least 80% identical to any one of SEQ ID NO: 41-49, 70, or 80-105.

[0009] In some embodiments, the present disclosure provides for provision of the L4 region in *trans* relative to an adenoviral helper plasmid. In some embodiments the present disclosure provides an L4 trans plasmid comprising a nucleotide sequence encoding an engineered L4 region. In some embodiments, an L4 *trans* plasmid encodes for an L4 region (e.g., L4 33K and L4 22K . In some embodiments, an L4 *trans* plasmid encodes for L4 33K. In some embodiments, an L4 *trans* plasmid encodes for L4 22K. In some embodiments the present disclosure provides a composition comprising an adenoviral helper plasmid described herein and an an L4 *trans* plasmid described herein.

[0010] In some embodiments, the present disclosure provides methods of producing recombinant adenoviral associated viral vectors (rAAV). In some embodiments, a method of producing a rAAV comprises transfecting a producer cell with an AAV vector plasmid, and an adenoviral helper plasmid as described herein. In some embodiments a method of producing a rAAV comprises transfecting a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, and an adenoviral helper plasmid described herein. In some embodiments a method of producing a rAAV further comprises transfecting a producer cell with an L4 *trans* plasmid. In some embodiments, a producer cell stably expresses Rep-Cap. In some embodiments, a producer cell comprises a nucleotide sequence encoding an L4 region or portion thereof (e.g., L4 33K and/or L4 22K).

BRIEF DESCRIPTION OF THE DRAWING

[0011] Figure 1. shows a plasmid map illustrating adenoviral helper plasmid pEMBR-1.2.

[0012] Figure 2. shows vector yields obtained using pEMBR-1.2 and commercially available pX80 as the adenoviral helper plasmid.

[0013] Figure 3. shows vector transgene purity and vector capsid purity obtained using pEMBR-1.2 or commercially available pX80 as the adenoviral helper plasmid.

[0014] Figure 4. shows a comparison between the GFP expression levels obtained following the transformation of HEK293 cells with recombinant AAV RH.10, ssCMV-GFP transgene, and either pX80 or pEMBR helper plasmid.

[0015] Figure 5. shows a plasmid map illustrating adenoviral helper plasmid pEMBR-1.3 and pEMBR-1.3B.

[0016] Figure 6. shows a plasmid map illustrating adenoviral helper plasmid pEMBR-1.4 and pEMBR-1.4B.

[0017] Figure 7. shows a plasmid map illustrating adenoviral helper plasmid pEMBR-1.5.

[0018] Figure 8. shows a plasmid map illustrating adenoviral helper plasmid pEMBR-1.5A.

[0019] Figure 9. shows an exemplary sequence map of native Adenovirus Type 5 genome including, among other things, VA RNA, E2a, and E4 regions (Panel A). Panels B, C, and D show the approximate sizes from pEMBR-1.2B2 of our VA RNA region (Panel B), E2a region (Panel C), and E4 region (Panel D) in comparison to the region from Matsushita 1998 annotated on the native Adenovirus sequence.

[0020] Figure 10. shows an exemplary plasmid map containing VA RNA, E2a, and E4 regions, for example, as defined by restriction enzyme sites, e.g., as in plasmids described herein.

[0021] Figure 11. shows a sequence map illustrating E2a region of pEMBR-1.2B2.

[0022] Figure 12. shows sequences maps for various promoters upstream of E2a regions within various pEMBR plasmids as described herein.

[0023] Figure 13. shows sequence maps for various truncations to the NotI end of the pEMBR1.2B2 E2a region, with one truncation construct having the addition of a promoter, as described herein.

[0024] Figure 14. shows a table summarizing plasmid size (bp), VA RNA region modifications, E2a region modifications, and E4 region modifications in various pEMBR plasmids as described herein.

[0025] Figure 15. shows average AAV9:eGFP titer (VG/mL) as measured by qPCR (Panel A) and ddPCR (Panel B) and protein expression of Rep proteins and E2a as measured by Western Blot (Panel C) using various pEMBR plasmids as described herein.

[0026] Figure 16. shows average AAV9:eGFP titer (VG/mL) as measured by qPCR using various pEMBR plasmids as described herein.

[0027] Figure 17. shows a table summarizing the inclusion of JEP, GEP, and native L4 (33K & 22K) and expression of E2a protein and production of AAV in various pEMBR plasmids as described herein.

[0028] Figure 18. shows a table summarizing plasmid size (bp), VA RNA region modifications, E2a region modifications, and E4 region modifications in various pEMBR plasmids as described herein.

[0029] Figure 19. shows a table summarizing plasmid size (bp), VA RNA region modifications, E2a region modifications, E4 region modifications, and AAV production relative to pEMBR-1.2.B2 in various pEMBR plasmids as described herein.

[0030] Figure 20. shows sequence maps for native E4 region (Panel A) and SV40 promoter driven E4 ORF4 and E4 ORF6 with SV40 poly(A) signal (Panel B) (e.g., as present in pEMBR-1.7B2).

[0031] Figure 21. shows a sequence map illustrating E2a region of pEMBR-1.2B2. The native L4 region (circled) illustrates sequences for 33k and 22k protein, contains 50bp upstream of native promoter, and the presumably native poly(a) signal sequence. The L4 expression construct shown below reserves all the same features as the native L4 region

except the native polyA signal has been replaced with a synthetic 49bp polyA signal sequence.

[0032] Figure 22. shows average AAV9:eGFP titer (VG/mL) using various pEMBR plasmids as described herein.

[0033] Figure 23. shows average AAV titer compared to pEMBR-1.2B2 titer and plasmid size (bp) using various pEMBR plasmids as described herein.

[0034] Figure 24. shows a plasmid map illustrating adenoviral helper plasmid pEMBR-1.37B2+L4-minKanR-v2.

[0035] Figure 25. shows a sequence map for an exemplary E2a region which has been engineered to reduce plasmid size while maintaining AAV production.

[0036] Figure 26. shows the location of the sequences reducing the E2a region in comparison to native Adenovirus Type 5 genome.

[0037] Figure 27 shows a wild-type AAV E2a region with the additional depiction of the L4 region ORFs for 22K and 33K protein. The partial JEP sequence is depicted below JEP by the smallest open arrow, to draw attention to how this natural E2a promoter sequence can be added plasmids described herein when the L4 33K protein is added.

[0038] Figure 28A and Figure 28B show rescue of AAV titer by L4 region addition. Figure 28A shows verage rAAV titer (VG/mL) from triplicate 75 mL quadruple plasmid transfected cultures determined by qPCR for the pEMBR-1.3 plasmid containing the partial JEP (PJ) E2a construct plus the pUC57 vector containing the specified L4 construct in comparison to wild type plus empty vector (pEMBR-1.2B2 +pUC57) production. The Partial JEP construct is shown as a reference. Figure 28B shows Rep and E2a western blots of RIPA lysates from transfected cultures in 28A.). Beta-actin is shown as a loading control. Adenovirus type 5 infected HEK293/HEK293T cells are shown as a positive control and uninfected/nontransfected HEK293 cells are shown as a negative control.

[0039] Figures 29A, 29B, and 29C show increase of AAV titer by provision of additional copies of L4 and L33K. Figure 29A shows the E2a regions of the helper plasmids

used in quadruple plasmid transient transfection. Figure 29B shows the average rAAV titer (VG/mL) from triplicate 75 mL quadruple plasmid transfected cultures determined by qPCR for the E2a constructs indicated in the figure plus the pUC57 vector containing the specified L4 construct to compare L4 gene dosage of 1X L4 genes, 2X L4 genes, 1X 33K with 2X 22K, and 2X 33K with 1X 22K against the control condition (wild type plus empty vector; 1.2B2 +pUC57). Figure 29C shows the Rep and E2a western blots of RIPA lysates from transfected cultures in Figure 29A. Beta-actin is shown as a loading control. Adenovirus type 5 infected HEK293/HEK293T cells are shown as a positive control and uninfected/nontransfected HEK293 cells are shown as a negative control.

DEFINITIONS

[0040] ***Agent*** : In general, the term "agent", as used herein, is used to refer to an entity (e.g., for example, a lipid, metal, nucleic acid, polypeptide, polysaccharide, small molecule, etc, or complex, combination, mixture or system [e.g., cell, tissue, organism] thereof), or phenomenon (e.g., heat, electric current or field, magnetic force or field, etc). In appropriate circumstances, as will be clear from context to those skilled in the art, the term may be utilized to refer to an entity that is or comprises a cell or organism, or a fraction, extract, or component thereof. Alternatively or additionally, as context will make clear, the term may be used to refer to a natural product in that it is found in and/or is obtained from nature. In some instances, again as will be clear from context, the term may be used to refer to one or more entities that is man-made in that it is designed, engineered, and/or produced through action of the hand of man and/or is not found in nature. In some embodiments, an agent may be utilized in isolated or pure form; in some embodiments, an agent may be utilized in crude form. In some embodiments, potential agents may be provided as collections or libraries, for example that may be screened to identify or characterize active agents within them. In some cases, the term "agent" may refer to a compound or entity that is or comprises a polymer; in some cases, the term may refer to a compound or entity that comprises one or more polymeric moieties. In some embodiments, the term "agent" may refer to a compound or entity that is not a polymer and/or is substantially free of any polymer and/or of one or more particular polymeric moieties. In some embodiments, the

term may refer to a compound or entity that lacks or is substantially free of any polymeric moiety.

[0041] ***Approximately/about***: As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" or "about" refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

[0042] ***Comparable***: As used herein, the term "comparable" refers to two or more agents, entities, situations, sets of conditions, etc., that may not be identical to one another but that are sufficiently similar to permit comparison therebetween so that one skilled in the art will appreciate that conclusions may reasonably be drawn based on differences or similarities observed. In some embodiments, comparable sets of conditions, circumstances, individuals, or populations are characterized by a plurality of substantially identical features and one or a small number of varied features. Those of ordinary skill in the art will understand, in context, what degree of identity is required in any given circumstance for two or more such agents, entities, situations, sets of conditions, etc to be considered comparable. For example, those of ordinary skill in the art will appreciate that sets of circumstances, individuals, or populations are comparable to one another when characterized by a sufficient number and type of substantially identical features to warrant a reasonable conclusion that differences in results obtained or phenomena observed under or with different sets of circumstances, individuals, or populations are caused by or indicative of the variation in those features that are varied.

[0043] ***Corresponding to***: As used herein, the term "corresponding to" may be used to designate the position/identity of a structural element in a compound or composition through comparison with an appropriate reference compound or composition. For example, in some embodiments, a monomeric residue in a polymer (e.g., an amino acid residue in a polypeptide or a nucleic acid residue in a polynucleotide) may be identified as "corresponding to" a residue in an appropriate reference polymer. For example, those of

ordinary skill will appreciate that, for purposes of simplicity, residues in a polypeptide are often designated using a canonical numbering system based on a reference related polypeptide, so that an amino acid "*corresponding to*" a residue at position 190, for example, need not actually be the 190th amino acid in a particular amino acid chain but rather corresponds to the residue found at 190 in the reference polypeptide; those of ordinary skill in the art readily appreciate how to identify "*corresponding*" amino acids. For example, those skilled in the art will be aware of various sequence alignment strategies, including software programs such as, for example, BLAST, CS-BLAST, CUSASW++, DIAMOND, FASTA, GGSEARCH/GLSEARCH, Genoogle, HMMER, HHpred/HHsearch, IDF, Infernal, KLAST, USEARCH, parasail, PSI-BLAST, PSI-Search, ScalaBLAST, Sequilab, SAM, SSEARCH, SWAPHI, SWAPHI-LS, SWIMM, or SWIPE that can be utilized, for example, to identify "corresponding" residues in polypeptides and/or nucleic acids in accordance with the present disclosure.

[0044] ***Downstream***: As used herein, the term "downstream" refers to the location or position of a nucleic acid sequence relative to a reference nucleic acid sequence, particularly a position that, during RNA transcription, is closer to the 3' end of the transcribed RNA molecule encoded by the reference sequence. For example, for two sequences, A and B, such that sequence A is downstream of sequence B, transcription of sequence B proceeds toward sequence A.

[0045] ***Nucleic acid***: As used herein, in its broadest sense, the term "nucleic acid" refers to any compound and/or substance that is or can be incorporated into an oligonucleotide chain. In some embodiments, a nucleic acid is a compound and/or substance that is or can be incorporated into an oligonucleotide chain via a phosphodiester linkage. As will be clear from context, in some embodiments, "*nucleic acid*" refers to an individual nucleic acid residue (e.g., a nucleotide and/or nucleoside); in some embodiments, "*nucleic acid*" refers to an oligonucleotide chain comprising individual nucleic acid residues. In some embodiments, a "*nucleic acid*" is or comprises RNA; in some embodiments, a "*nucleic acid*" is or comprises DNA. In some embodiments, a nucleic acid is, comprises, or consists of one or more natural nucleic acid residues. In some embodiments, a nucleic acid is, comprises, or consists of one or more nucleic acid analogs. In some embodiments, a

nucleic acid analog differs from a nucleic acid in that it does not utilize a phosphodiester backbone. For example, in some embodiments, a nucleic acid is, comprises, or consists of one or more "*peptide nucleic acids*", which are known in the art and have peptide bonds instead of phosphodiester bonds in the backbone, are considered within the scope of the present invention. Alternatively or additionally, in some embodiments, a nucleic acid has one or more phosphorothioate and/or 5'-N-phosphoramidite linkages rather than phosphodiester bonds. In some embodiments, a nucleic acid is, comprises, or consists of one or more natural nucleosides (e.g., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxy guanosine, and deoxycytidine). In some embodiments, a nucleic acid is, comprises, or consists of one or more nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3 -methyl adenosine, 5-methylcytidine, C-5 propynyl-cytidine, C-5 propynyl-uridine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5 -propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, 0(6)-methylguanine, 2-thiocytidine, methylated bases, intercalated bases, and combinations thereof). In some embodiments, a nucleic acid comprises one or more modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose) as compared with those in natural nucleic acids. In some embodiments, a nucleic acid has a nucleotide sequence that encodes a functional gene product such as an RNA or protein. In some embodiments, a nucleic acid includes one or more introns. In some embodiments, nucleic acids are prepared by one or more of isolation from a natural source, enzymatic synthesis by polymerization based on a complementary template (*in vivo* or *in vitro*), reproduction in a recombinant cell or system, and chemical synthesis. In some embodiments, a nucleic acid is at least 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 1 10, 120, 130, 140, 150, 160, 170, 180, 190, 20, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 or more residues long. In some embodiments, a nucleic acid is partly or wholly single stranded; in some embodiments, a nucleic acid is partly or wholly double stranded. In some embodiments a nucleic acid has a nucleotide sequence comprising at least one element that encodes, or is the complement of a sequence that encodes, a polypeptide. In some embodiments, a nucleic acid has enzymatic activity.

[0046] ***Operably linked***: As used herein, the term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A control element "*operably linked*" to a functional element is associated in such a way that expression and/or activity of the functional element is achieved under conditions compatible with the control element. In some embodiments, "*operably linked*" control elements are contiguous (e.g., covalently linked) with the coding elements of interest; in some embodiments, control elements act in *trans* to or otherwise at a distance from the functional element of interest.

[0047] ***Producer cell:*** As used herein, the term "producer cell" refers to any cell used to produce recombinant AAV (rAAV). In some embodiments, a producer cell is a mammalian cell. In some embodiments, a producer cell is a transformed mammalian cell. In some embodiments, a producer cell is a Vero, HeLa, HEK293, HEK293T cell or derivative thereof.

[0048] ***Transformation:*** As used herein, the term "transformation" refers to any process by which exogenous DNA is introduced into a host cell. Transformation may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. In some embodiments, a particular transformation methodology is selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, mating, lipofection. In some embodiments, a "*transformed*" cell is stably transformed in that the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. In some embodiments, a transformed cell transiently expresses introduced nucleic acid for limited periods of time.

[0049] ***Upstream:*** As used herein, the term "upstream" refers to the location or position of a nucleic acid sequence relative to a reference nucleic acid sequence, particularly a position that, during RNA transcription, is closer to the 5` end of the transcribed RNA molecule encoded by the reference sequence. For example, for two sequences, A and B, such that sequence A is upstream of sequence B, transcription of sequence B proceeds away from sequence A.

[0050] ***Vector***: As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "*plasmid*", which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "*expression vectors*." Standard techniques may be used for recombinant DNA, oligonucleotide synthesis, and tissue culture and transformation (e.g., electroporation, lipofection). Enzymatic reactions and purification techniques may be performed according to manufacturer's specifications or as commonly accomplished in the art or as described herein. The foregoing techniques and procedures may be generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification. See e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)), which is incorporated herein by reference for any purpose.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

[0051] The helper functions adenoviruses provide for AAV replication have been previously described. Without wishing to be bound by any particular hypothesis, adenoviral E1A protein has been described to activate AAV gene expression by binding and activating the AAV P5 rep promoter. Similarly, E2A, another adenoviral protein, has been described to activate AAV P5 promoter transcription. E2A has also been described to cooperate with virus associated RNA I (VA RNAI) to enhance the translation of AAV RNAs. Adenoviral E4orf4 has been shown to induce cell-cycle arrest at the G2/M border, as well as to aid in AAV production. Adenoviral E4orf6 has been described to enhance the conversion of

single-stranded recombinant AAV genomes into double-stranded genomes, a rate-limiting step of viral DNA-replication both in vitro and in vivo. VA RNAI has also been described to support AAV replication. It has been described that VA RNAI physically interacts with the double-stranded RNA-activated protein kinase (PKR), which would otherwise elicit an antiviral immune response blocking viral protein production.

[0052] Prior studies have suggested that in HEK293 cells, which provide the El gene, the minimal set of genes in trans for efficient recombinant AAV production is E2a, E4orf6 and the VA RNAI gene. A helper plasmid named pXX6-80, containing this set of genes as well as additional genes that may not be necessary for production, is used for the production of adenovirus-free recombinant AAV.

[0053] One major ongoing challenge in the development and optimization of AAV vectors for clinical applications is to increase the amount of viruses being produced. Due to their non-proliferative nature, their production depends solely on the transfection efficiency of the parvoviral genomic components into the packaging cell lines (for example human embryonic kidney cells, HEK293 or HEK293T, or insect cells e.g., Sf9). Thus, it remains of high importance to develop means to increase recombinant AAV (rAAV) production.

[0054] Other major challenges related to the production of rAAVs for clinical applications are those related to the cost to produce such rAAVs in large quantities, and also to the safety of the final products themselves. For example, commercially available helper plasmids, such as pXX6-80, appear to transcribe low levels of the Ad fiber protein. Importantly, the fiber protein is not required for AAV production, and could be immunogenic in humans. In addition, the size of pXX6-80 is rather large, at over 18kb. This large plasmid size increases the difficulty and cost of its manufacturing, which can be highly impactful when sourcing GMP plasmids for the manufacturing of clinical-grade AAV.

[0055] Different versions of adenoviral helper plasmids have been derived by others, including, for example, pFAdDeltaF6 (derived at the University of Pennsylvania) and pHelper (Agilent). The pFAdDeltaF6 plasmid is about 3kb smaller than pXX6-80, but retains the fiber gene sequence. The pHelper plasmid, which is available from Agilent, is

smaller than pXX6-80, at about 11.6kb. However, it contains an ampicillin resistance gene, which is generally discouraged for plasmids used in AAV production.

[0056] The present disclosure addresses the above-described technical challenges by providing compositions and methods described herein.

[0057] In some embodiments, the present disclosure relates to adenovirus derived helper plasmids (an adenoviral helper plasmid) comprising adenoviral DNA sequences encoding viral helper proteins. In some embodiments, adenoviral helper plasmids of the present invention are used in methods of production of recombinant adeno-associated viruses (rAAVs). In some embodiments, adenoviral helper plasmids of the present disclosure increase production of rAAVs.

[0058] In some embodiments, the present disclosure provides an adenoviral helper plasmid comprising nucleotide sequences encoding proteins derived from sources that are not adenovirus. In some embodiments, the present disclosure provides an adenoviral helper plasmid comprising nucleotide sequences encoding proteins derived from viruses other than adenovirus. In some embodiments, an adenoviral helper plasmid comprises all or a portion of an adenoviral nucleotide sequence encoding adenoviral proteins E2a and E4, L4 region, as well as non-coding RNA VA RNA. In some embodiments, present disclosure describes improved adenoviral helper plasmids that are smaller than the leading commercially available adenoviral helper plasmids, and that allow for safer and less costly production of rAAVs in producer cell expression systems.

[0059] In some embodiments, the present disclosure provides an adenoviral helper plasmid that has reduced overall size relative to presently avaliable adenoviral helper plasmids (e.g., pXX6-80 at 18.932 kbp; pALD-X80 at 18.876 kbp; pHelper at 11.635 kbp; pFAdDeltaF6 at 15.420 kbp).

[0060] In some embodiments, the present disclosure provides adenoviral helper plasmids having a smaller size. In some embodiments, an adenoviral helper plasmid of the present disclosure is approximately between 6.5 kb and 15.5 kb. In some embodiments, an adenoviral helper plasmid of the present disclosure has a size that is approximately 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, or 16 kb. In some embodiments, an

adenoviral helper plasmid of the present disclosure has a size that is approximately 6-7 kb; 6.5-7.5 kb; 7-8 kb; 7.5-8.5 kb; 8-9 kb; 8.5-9.5 kb; 9-10 kb; 9.5-10.5 kb; 10-11 kb; 10.5-11.5 kb; 11-12 kb; 11.5-12.5 kb; 12-13 kb; 12.5-13.5 kb; 13-14 kb; 13.5-14.5 kb; 14-15 kb; 14.5-15.5 kb; 15-16 kb. The smaller size of the adenoviral helper plasmids of the present disclosure enables the simpler and less costly production of AAV at the quantities necessary for large-scale manufacturing of AAV. In some embodiments, removing genes and/or portions of genes makes an adenoviral helper plasmid of the present disclosure safer, since the producing cells would not produce the adenovirus structural proteins (e.g., fiber, 100K and hexon assembly), that could co-purify with AAV during downstream processing and would therefore present a lower risk of inadvertently introducing adenovirus structural proteins to patients.

[0061] In some embodiments, removal of adenoviral helper genes resulting in a smaller adenoviral helper plasmid enables addition of supplementary genes to further improve AAV quality and yield. Although these supplementary genes increase the size of the plasmid relative to the smallest versions, they enable comparable or higher AAV productivity and are therefore worth the additional cost to produce. Importantly, these plasmids are still smaller than commercially available helper plasmids such as, for example, pALD-X80.

Adenoviral helper plasmids

Helper genes and resistance genes

[0062] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises one or more nucleotide sequence(s) encoding proteins selected from the group consisting of E2b, E2a, E4orf4, E1B55K, E1b19K, E1a, E4orf6, VA RNA, and combinations thereof.

[0063] In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding an E4 region and a VA RNA region. In some embodiments, an E4 region comprises one or more of E4orf1, E4orf2, E4orf3, E4orf4, E4orf5, E4orf6, and E4orf7. In some embodiments, E4orf1 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 53. In some embodiments, E4orf1 has an amino acid

sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 55. In some embodiments, E4orf2 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 56. In some embodiments, E4orf2 has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 58. In some embodiments, E4orf3 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 59. In some embodiments, E4orf3 has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 61. In some embodiments, E4orf4 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 62. In some embodiments, E4orf4 has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 64. In some embodiments, E4orf6 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 65. In some embodiments, E4orf6 has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 67. In some embodiments, E4orf7 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 68. In some embodiments, E4orf7 has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 69. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence comprising E4orf1. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence comprising E4orf2. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence comprising E4orf1 and does not comprise a nucleotide sequence comprising E4orf2. In some embodiments, expression of the E4 region is under the control of an E4 mini promoter. In some embodiments, an E4 region is operably linked to an E4 mini promoter. In some embodiments, an E4 mini promoter has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 1. In some embodiments, an E4 region is operably linked to an SV40 promoter. In some embodiments, expression of the E4 region is under the control of an SV40 promoter. In some embodiments, an SV40 promoter has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 2.

[0064] In some embodiments, an adenoviral helper plasmid of the present invention comprises a resistance gene. In some embodiments, an adenoviral helper plasmid of the

present invention comprises an ampicillin resistance gene (e.g., a nucleotide sequence encoding a protein conferring resistance to ampicillin). In some embodiments, an adenoviral helper plasmid of the present invention does not comprise an ampicillin resistance gene. In some embodiments, an adenoviral helper plasmid of the present invention comprises a kanamycin resistance gene (e.g., a nucleotide sequence encoding a protein conferring resistance to kanamycin). In some embodiments, an adenoviral helper plasmid of the present invention does not comprise a kanamycin resistance gene.

E2a region

[0065] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an E2a region. In some embodiments an E2a region, as used herein, comprise(s) one or more nucleotide sequences encoding an E2a protein and an engineered L4 region. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an E2a region that has been modified. In some embodiments, the E2a region of an adenoviral helper plasmid of the present disclosure comprises deletions to reduce plasmid size while maintaining production of E2a protein and AAV.

[0066] In some embodiments, the E2a region has been modified to comprise deletions while keeping JEP promoter sequence that corresponds to that identified by Jing et al., 2001. Inhibition of adenovirus cytotoxicity, replication, and E2a gene expression by adeno-associated virus. Virology, 291(1), pp.140-151, which is incorporated herein by reference for any purpose. In some embodiments, an E2a region has been modified to comprise deletions while keeping GEP promoter sequence that corresponds to that identified by Guilfoyle et al., 1985. Two functions encoded by adenovirus early region 1A are responsible for the activation and repression of the DNA-binding protein gene. The EMBO Journal, 4(3), pp.707-713, which is incorporated herein by reference for any purpose.

[0067] In some embodiments, the E2a region has been modified to comprise deletions while keeping a certain amount of nucleotides upstream of E2a (e.g., 1 kbp or 50 bp upstream of E2a).

[0068] In some embodiments, the E2a region has been modified to comprise truncations starting from the NotI end of adenoviral helper plasmids of the present disclosure.

[0069] In some embodiments, the E2a region has been modified as illustrated in Figs. 12-14 and/or 18-19 and/or 25, 29A.

L4 region

[0070] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an L4 region. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises one or more nucleotide sequence(s) in an L4 region which encodes one or more protein(s) (e.g., 33k, 22k, etc.).

[0071] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an L4 region that has been modified. In some embodiments, the L4 region of an adenoviral helper plasmid of the present disclosure comprises deletions to reduce plasmid size while maintaining production of L4 33k and L422k protein and AAV.

[0072] In some embodiments, an adenoviral helper plasmid of the present disclosure encodes for but does not express L4 33k and L422k protein. In some embodiments, the L4 region of an adenoviral helper plasmid of the present disclosure comprises modifications to abolish expression of L4 33k and L422k protein while maintaining nucleotide sequences that encode for L4 33k and L422k.

[0073] In some embodiments, the L4 region of an adenoviral helper plasmid of the present disclosure comprises deletions which abolish expression of L4 100K. In some embodiments, the L4 region of an adenoviral helper plasmid of the present disclosure comprises deletions which removes nucleotides encoding L4 100K abolish expression of L4 100K.

[0074] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 22K. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a

nucleotide sequence encoding L4 33K and L4 22K. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K and L4 22K modified such that L4 33K is not expressed. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K and L4 22K modified such that L4 22 K is not expressed.

[0075] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a native L4 region (~1379 bp in length). In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an L4 expression construct (~1184 bp), referred to as +L4 in adenoviral helper plasmid designs (e.g., those as described herein), which has replaced the native polyA signal with the 49bp synthetic polyA signal (SEQ ID NO: 108). In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 22K. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K and L4 22K. In some embodiments, an adenoviral helper plasmid of the present disclosure expresses L4 genes 33K and L4 22K from a native L4 promoter. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding a native L4 promoter, 50bp nucleotides upstream of the native promoter, L4 33K and L4 22K. In some embodiments the 50bp nucleotides upstream of the native promoter are determined relative to a reference adenovirus sequence. In some embodiments a reference sequence is GenBank: M73260.1(Ad5 sequence).

[0076] In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise an L4 region. In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise one or more nucleotide sequence(s) in an L4 region which encode one or more protein(s) (e.g., 33k, 22k, etc.). In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise a nucleotide sequence encoding L4 33K. In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise a nucleotide sequence encoding L4 22K.

[0077] In some embodiments an L4 region is provided in *trans* relative to an adenoviral helper plasmid. In some embodiments, an L4 region is provided in a producer

cell. In some embodiments, a producer cell comprises a nucleotide sequence encoding an L4 region. In some embodiments, a producer cell comprises a nucleotide sequence encoding L4 proteins L4 33K and L4 22K. In some embodiments, a producer cell comprises a nucleotide sequence encoding L4 33K. In some embodiments, a producer cell comprises a nucleotide sequence encoding L4 22K.

[0078] In some embodiments, an L4 region or portions thereof is/are provided by a plasmid other than an adenoviral helper plasmid. In some embodiments the present disclosure provides an L4 *trans* plasmid. In some embodiments an L4 *trans* plasmid comprises an adenoviral L4 region. In some embodiments, an L4 *trans* plasmid of the present disclosure comprises one or more nucleotide sequence(s) in an L4 region which encode one or more protein(s) (e.g., 33k, 22k, etc.). In some embodiments, an L4 *trans* plasmid of the present disclosure provides one or more nucleotide sequence(s) in an L4 region which encode one or more protein(s) (e.g., 33k, 22k, etc.) on a plasmid other than an adenoviral helper plasmid. As in some embodiments the one or more nucleotide sequence(s) in an L4 region which encode one or more protein(s) (e.g., 33k, 22k, etc.) are not provided on the same plasmid (e.g., *cis*) as other adenoviral helper components the plasmid is referred to as an L4 *trans* plasmid.

[0079] In some embodiments, an L4 *trans* plasmid of the present disclosure comprises one or more nucleotide sequence(s) in an L4 region which encode and express one or more protein(s) (e.g., 33k, 22k, etc.). In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence that encodes for and expresses L4 33k and L422k protein. In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence that encodes for and expresses L4 33k protein. In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence that encodes for and expresses L4 22k protein.

[0080] In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K. In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence encoding L4 22K. In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K and L4 22K. In some embodiments, an L4 *trans* plasmid of the

present disclosure comprises a nucleotide sequence encoding L4 33K and L4 22K modified such that L4 33K is not expressed. In some embodiments, an L4 *trans* plasmid of the present disclosure comprises a nucleotide sequence encoding L4 33K and L4 22K modified such that L4 22 K is not expressed.

[0081] In some embodiments, expression of proteins encoded by an L4 *trans* plasmid of the present disclosure is under the control of a promoter. In some embodiments, expression of proteins encoded by an L4 *trans* plasmid of the present disclosure is under the control of an EF1alpha promoter.

[0082] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an E2a region modified as illustrated in Figs. 18 and/or 19 and/or 25.

Fiber gene

[0083] In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise a nucleotide sequence encoding adenoviral fiber protein. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding a full-length adenoviral fiber protein. In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding a portion or fragment of adenoviral fiber protein. In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to that of pXX6-80, excluding the nucleotide sequence encoding an adenoviral fiber protein.

L1-52/55K (Packaging Protein 3) gene

[0084] In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise a nucleotide sequence encoding a L1-52/55K (Packaging Protein 3) protein. In some embodiments, an adenoviral helper plasmid of the present invention does not comprise a nucleotide sequence encoding Peripentonal Hexon-Associated genes.

L4 region

[0085] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a complete L4 (hexon assembly) gene. In some embodiments, an adenoviral

helper plasmid of the present disclosure comprises a nucleotide sequence encoding a complete L4 (hexon assembly). In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 3. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 4. In some embodiments, an adenoviral helper plasmid of the present invention comprises a complete L4 (33 kDa Ex2) gene. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding a complete L4 (33 kDa Ex2). In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 5. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 6.

[0086] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a complete L4 Encapsidation Protein gene. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding a complete L4 Encapsidation Protein. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 7. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 8.

[0087] In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise an L4 (hexon assembly) gene. In some embodiments, an adenoviral helper plasmid does not comprise an L4 Encapsidation Protein gene. In some embodiments, an adenoviral helper plasmid does not comprise an L4 (hexon assembly) gene and does not comprise an L4 Encapsidation Protein gene. In some embodiments, an adenoviral helper plasmid of the present disclosure does not comprise a nucleotide sequence encoding L4 (hexon assembly). In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding L4 Encapsidation Protein. In some embodiments, an

adenoviral helper plasmid does not comprise a nucleotide sequence encoding L4 (hexon assembly) and does not comprise a nucleotide sequence encoding L4 Encapsidation Protein gene. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding a fragment of L4 33 kDa Ex2. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 9. In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding a fragment of L4 33 kDa Ex2. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 10. In some embodiments, a nucleotide sequence encoding a fragment of L4 33 kDa Ex2 comprises an E2a promoter region (see, for example, Casper et al., "Identification of an adeno-associated virus Rep protein binding site in the adenovirus E2a promoter." Journal of virology 79.1 (2005)). In some embodiments, an E2a promoter region has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 11. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding a fragment of L4 33 kDa Ex2. In some embodiments, an adenoviral helper plasmid does not comprise an E2a promoter region.

[0088] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding a fragment of hexon-associated precursor (L4 pVIII). In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 12. In some embodiments, an adenoviral helper plasmid comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 13. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding hexon-associated precursor (L4 pVIII). In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding a fragment of partial hexon-associated precursor (L4 pVIII).

VA RNA region

[0089] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a VA RNA region having a nucleotide sequence that is at least 80%, 85%, 90%,

95%, 99%, or 100% identical to SEQ ID NO: 14. In some embodiments, an adenoviral helper plasmid comprises a VA RNA region having a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 15. In some embodiments, a VA RNA region comprises a VA RNAI gene having a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 16. In some embodiments, a VA RNA region comprises a VA RNAI gene having a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 17. In some embodiments, a VA RNA region comprises a VA RNAII gene having a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 18. In some embodiments, a VA RNA region comprises a VA RNAII gene having a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 19.

[0090] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding a fragment of DNA Terminal Protein. In some embodiments, a nucleotide sequence encoding a fragment of DNA Terminal Protein is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 20. In some embodiments, a fragment of DNA Terminal Protein has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 21. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding DNA Terminal Protein. In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding a fragment of 23kDa endoprotease. In some embodiments, an adenoviral helper plasmid comprise a nucleotide sequence at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 22. In some embodiments, a fragment of 23kDa endoprotease region has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 23. In some embodiments, an adenoviral helper plasmid does not comprise a nucleotide sequence encoding 23kDa endoprotease region.

Introduction of genes encoding supplementary features

[0091] In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an E2a gene. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence encoding E2a. In some embodiments, an

adenoviral helper plasmid of the present disclosure comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 24. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 25. In some embodiments, expression of E2a is under the control of a promoter. In some embodiments, a nucleotide sequence encoding E2a is operably linked to a promoter. In some embodiments, a promoter is, for example, a CMV promoter, a PGK promoter, an SV40 promoter, an EF-1α promoter, a Ubc promoter, a CAG promoter, or a β-actin promoter. In some embodiments, a nucleotide sequence encoding E2a is operably linked to a transcriptional enhancer. In some embodiments, a transcriptional enhancer is, for example, a CMV enhancer. In some embodiments, a nucleotide sequence encoding E2a is operably linked to a regulatory intron. In some embodiments, expression of E2a is under the control of a chicken β-actin promoter. In some embodiments, a nucleotide sequence encoding E2a is operably linked to a chicken β-actin promoter. In some embodiments, a chicken β-actin promoter has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 26. In some embodiments, a chicken β-actin promoter is positioned upstream of a nucleotide sequence encoding E2a. In some embodiments, expression of E2a is under the control of an E2a promoter and a chicken β-actin promoter. In some embodiments, a nucleotide sequence encoding E2a is operably linked to an E2a promoter and a chicken β-actin promoter. In some embodiments, a chicken β-actin promoter is positioned upstream of an E2a promoter. In some embodiments, expression of E2a is under the control of chicken β-actin promoter and a CMV enhancer. In some embodiments, a nucleotide sequence encoding E2a is operably linked to a chicken β-actin promoter and a CMV enhancer. In some embodiments, a chicken β-actin promoter and a CMV enhancer are positioned upstream of an E2a promoter. In some embodiments, an adenoviral helper plasmid comprises an E2a polyadenylation signal. In some embodiments, an E2a polyadenylation signal is positioned downstream of a nucleotide sequence encoding E2a. In some embodiments, an E2a polyadenylation signal has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 27. In some embodiments, an adenoviral helper plasmid comprises an SV40 polyadenylation signal. In some embodiments, an SV40 polyadenylation signal is positioned downstream of a nucleotide sequence encoding E2a. In some embodiments, an SV40 polyadenylation signal is

positioned downstream of an E2a polyadenylation signal. In some embodiments, an SV40 polyadenylation signal has a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 28.

[0092] In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding UL30 derived from HSV-1. In some embodiments, a nucleotide sequence encoding UL30 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 29. In some embodiments, an amino acid sequence UL30 is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 30. In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding UL42 derived from HSV-1. In some embodiments, a nucleotide sequence encoding UL42 has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 31. In some embodiments, an amino acid sequence of UL42 is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 32. In some embodiments, an adenoviral helper plasmid comprises a nucleotide sequence encoding UL30 derived from HSV-1, and a nucleotide sequence encoding UL42 derived from HSV-1. In some embodiments, a nucleotide sequence encoding UL30 and a nucleotide sequence encoding UL42 are separated by a P2a cleavage site. In some embodiments, a P2a cleavage site has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 33. In some embodiments, a P2a cleavage site has an amino acid sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 34. In some embodiments, expression of UL30 and/or UL42 gene is/are under the control of an EF-1α promoter. In some embodiments, a nucleotide sequence encoding UL30 is operably linked to a promoter. In some embodiments, a nucleotide sequence encoding UL30 is operably linked to a CMV promoter, a PGK promoter, an SV40 promoter, an EF-1α promoter, a Ubc promoter, a CAG promoter, or a β-actin promoter. In some embodiments, a nucleotide sequence encoding UL30 is operably linked to a transcriptional enhancer. In some embodiments, a transcriptional enhancer is, for example, a CMV enhancer. In some embodiments, a nucleotide sequence encoding UL30 is operably linked to a regulatory intron. In some embodiments a nucleotide sequence encoding UL42 and/or a nucleotide sequence encoding UL30 are operably linked to an EF-1α promoter. In some embodiments, an EF-1α promoter has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or

100% identical to SEQ ID NO: 35. In some embodiments, expression of UL30 and/or UL42 is/are under the control of an SV40 promoter. In some embodiments, a nucleotide sequence encoding UL42 and/or a nucleotide sequence encoding UL30 are operably linked to an SV40 promoter. In some embodiments, an SV40 promoter has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 51.

[0093] In some embodiments, an adenoviral helper plasmid comprises a polyadenylation signal. In some embodiments, a polyadenylation signal is synthetic. In some embodiments, a synthetic polyadenylation signal is a 49bp synthetic polyadenylation signal (SEQ ID NO. 108).

[0094] In some embodiments, a polyadenylation signal is a β-globin polyadenylation signal, SV40 polyadenylation signal, or a Bovine Growth Hormone (bGH) polyadenylation signal. In some embodiments, an adenoviral helper plasmid comprises a polyadenylation signal downstream of a nucleotide sequence encoding UL42. In some embodiments, an adenoviral helper plasmid comprises a β-globin polyadenylation signal downstream of a nucleotide sequence encoding UL42. In some embodiments, a β-globin polyadenylation signal has a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 36. In some embodiments, an adenoviral helper plasmid comprises a Bovine Growth Hormone (bGH) polyadenylation signal downstream of a nucleotide sequence encoding UL42. In some embodiments, a bGH polyadenylation signal has a sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 52.

Exemplary adenoviral helper plasmids

[0095] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 41. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (Hexon Assembly) (SEQ ID NO: 3; SEQ ID NO: 4), L4 (33kDa Ex2) (SEQ ID NO: 5; SEQ ID NO: 6), L4 Encapsidation Protein (22 kDa) (SEQ ID NO: 7; SEQ ID NO: 8), L4 pVIII Hexon-Associated Precursor (SEQ ID NO: 12; SEQ ID NO: 13), VA

RNA region A (SEQ ID NO: 14), VA RNAI-A (SEQ ID NO: 16), VA RNAII-A (SEQ ID NO: 18), partial DNA Terminal Protein (SEQ ID NO: 20; SEQ ID NO: 21), 23kDa endoprotease fragment region (SEQ ID NO: 22; SEQ ID NO: 23), and E2a (SEQ ID NO: 24; SEQ ID NO: 25), and does not comprise the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, and Peripentonal Hexon-Associated genes.

[0096] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 42. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region A (SEQ ID NO: 14), VA RNAI-A (SEQ ID NO: 16), VA RNAII-A (SEQ ID NO: 18), partial DNA Terminal Protein (SEQ ID NO: 20; SEQ ID NO: 21), 23kDa endoprotease fragment region (SEQ ID NO: 22; SEQ ID NO: 23), and E2a (SEQ ID NO: 24; SEQ ID NO: 25), and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, and L4 pVIII Hexon-Associated Precursor.

[0097] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 43. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region B (SEQ ID NO: 15), VA RNAI-B (SEQ ID NO: 17), VA RNAII-B (SEQ ID NO: 19), and E2a (SEQ ID NO: 24; SEQ ID NO: 25), and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, L4 pVIII Hexon-Associated Precursor, DNA Terminal Protein, and 23kDa endoprotease fragment region.

[0098] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 44. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region A (SEQ ID NO: 14), VA RNAI-A (SEQ ID NO: 16), VA RNAII-A (SEQ ID NO: 18), partial DNA Terminal Protein (SEQ ID NO: 20; SEQ ID NO: 21), 23kDa endoprotease fragment region (SEQ ID NO: 22; SEQ ID NO: 23), E2a (SEQ ID NO: 24; SEQ ID NO: 25), and a chicken β-actin promoter upstream of E2a, and does not comprise the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, and L4 pVIII Hexon-Associated Precursor.

[0099] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 45. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region B (SEQ ID NO: 15), VA RNAI-B (SEQ ID NO: 17), VA RNAII-B (SEQ ID NO: 19), E2a (SEQ ID NO: 24; SEQ ID NO: 25), and a chicken β-actin promoter upstream of E2a, and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, L4 pVIII Hexon-Associated Precursor, DNA Terminal Protein, and 23kDa endoprotease fragment region.

[0100] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 46. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter

(SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region B (SEQ ID NO: 15), VA RNAI-B (SEQ ID NO: 17), VA RNAII-B (SEQ ID NO: 19), E2a (SEQ ID NO: 24; SEQ ID NO: 25), SV40 polyadenylation signal downstream of E2a (SEQ ID NO: 28), and a chicken β-actin promoter upstream of E2a, and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, L4 pVIII Hexon-Associated Precursor, DNA Terminal Protein, and 23kDa endoprotease fragment region.

[0101] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 47. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region A (SEQ ID NO: 14), VA RNAI-A (SEQ ID NO: 16), VA RNAII-A (SEQ ID NO: 18), partial DNA Terminal Protein (SEQ ID NO: 20; SEQ ID NO: 21), 23kDa endoprotease fragment region (SEQ ID NO: 22; SEQ ID NO: 23), E2a (SEQ ID NO: 24; SEQ ID NO: 25), a chicken β-actin promoter upstream of E2a, an HSV-1-derived UL30 gene (SEQ ID NO: 29; SEQ ID NO: 30), an HSV-1-derived UL42 gene (SEQ ID NO: 31; SEQ ID NO: 32), EF-1α promoter (SEQ ID NO: 35) upstream of UL30, and a β-globin polyadenylation signal (SEQ ID NO: 36) downstream of UL42, and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, and L4 pVIII Hexon-Associated Precursor.

[0102] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 48. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: E4 mini promoter (SEQ ID NO: 1), L4 (33kDa Ex2) (SEQ ID NO: 9; SEQ ID NO: 10), VA RNA region A

(SEQ ID NO: 14), VA RNAI-A (SEQ ID NO: 16), VA RNAII-A (SEQ ID NO: 18), partial DNA Terminal Protein (SEQ ID NO: 20; SEQ ID NO: 21), 23kDa endoprotease fragment region (SEQ ID NO: 22; SEQ ID NO: 23), E2a (SEQ ID NO: 24; SEQ ID NO: 25), a chicken β-actin promoter upstream of E2a, an HSV-1-derived UL30 gene (SEQ ID NO: 29; SEQ ID NO: 30), an HSV-1-derived UL42 gene (SEQ ID NO: 31; SEQ ID NO: 32), SV40 promoter (SEQ ID NO: 51) upstream of UL30, and a Bovine Growth Hormone (bGH) polyadenylation signal (SEQ ID NO: 52) downstream of UL42, and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, and L4 pVIII Hexon-Associated Precursor.

[0103] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 49. In some embodiments, an adenoviral helper plasmid of the present disclosure comprises the following components having nucleotide sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to the indicated sequences: SV40 promoter upstream of E4 region (SEQ ID NO: 2),VA RNA region B (SEQ ID NO: 15), VA RNAI-B (SEQ ID NO: 17), VA RNAII-B (SEQ ID NO: 19), E2a (SEQ ID NO: 24; SEQ ID NO: 25), SV40 polyadenylation signal downstream of E2a (SEQ ID NO: 28), SV40 polyadenylation signal downstream of E4orf6 (SEQ ID NO: 50), and a chicken β-actin promoter upstream of E2a, and does not comprise or encode the following components: a fiber gene, an L1-52/55K (Packaging Protein 3) gene, Peripentonal Hexon-Associated genes, full-length L4 (Hexon Assembly) gene, L4 Encapsidation Protein, L4 pVIII Hexon-Associated Precursor, L4 (33kDa Ex2), DNA Terminal Protein, and 23kDa endoprotease fragment region, E4 mini promoter upstream of E4 region, a gene encoding E4orf1, a gene encoding E4orf2, and a gene encoding E4orf3.

[0104] In some embodiments, an adenoviral helper plasmid of the present disclosure has a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 94. In some embodiments, an adenoviral helper plasmid is designed to include "B2" designs comprising a SV40 polyA site to potentially increase expression of E2A and a synthesized sequence of a smaller VA region (contains Ad2 VA RNA I and VA RNA II)

that does not contain the flanking Ad Terminal Protein nor Endoprotease gene sequences. In some embodiments, this region was synthesized with flanking StuI and BsrGI sites and the insert was cloned into pEMBR-1.2 to make pEMBR-1.2B2 (e.g., SEQ ID NO: 94).]

[0105] In some embodiments, adenoviral helper plasmids as described herein are 11000 bp or fewer. In some embodiments, adenoviral helper plasmids as described herein are 10000 bp or fewer. In some embodiments, adenoviral helper plasmids as described herein are 9500 bp or fewer. In some embodiments, adenoviral helper plasmids as described herein are 9000 bp or fewer, In some embodiments, adenoviral helper plasmids as described herein are 8000 bp or fewer. In some embodiments, adenoviral helper plasmids as described herein are 7500 bp or fewer. In some embodiments, adenoviral helper plasmids as described herein are 7000 bp or fewer.

[0106] In some embodiments, an L4 *trans* plasmid comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 115. In some embodiments, an L4 *trans* plasmid comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 119. In some embodiments, an L4 *trans* plasmid comprises a nucleotide sequence that is at least 80%, 85%, 90%, 95%, 99%, or 100% identical to SEQ ID NO: 8.

Methods of Production

[0107] In some embodiments, an adenoviral helper plasmid of the present disclosure is useful in methods of producing rAAV. In some embodiments, an adenoviral helper plasmid and an L4 *trans* plasmid of the present disclosure are useful in methods of producing rAAV. In some embodiments, rAAV is produced by transfection of a producer cell. In some embodiments, a producer cell is a mammalian cell. In some embodiments, a producer cell is a transformed mammalian cell. In some embodiments, a producer cell is a Vero, HeLa, HEK293, HEK293T cell or derivative thereof.

[0108] In some embodiments, an L4 region is provided in a producer cell. In some embodiments, a producer cell comprises a nucleotide sequence encoding an L4 region. In some embodiments, a producer cell comprises a nucleotide sequence encoding L4 proteins L4 33K and L4 22K. In some embodiments, a producer cell comprises a nucleotide

sequence encoding L4 33K. In some embodiments, a producer cell comprises a nucleotide sequence encoding L4 22K.In some embodiments, a method of producing a rAAV comprises transfection of a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, and an adenoviral helper plasmid. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, an adenoviral helper plasmid, and an L4 *trans* plasmid. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, an adenoviral helper plasmid comprising an L4 region, and an L4 *trans* plasmid. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, an adenoviral helper plasmid comprising an L4 region, and an L4 *trans* plasmid comprising a nucleotide sequence encoding L4 22K. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, an adenoviral helper plasmid comprising an L4 region, and an L4 *trans* plasmid comprising a nucleotide sequence encoding L4 33K. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, and an adenoviral helper plasmid comprising an L4 region, wherein the producer cell comprises a nucleotide sequence encoding an L4 region or portions thereof (e.g., L33K and/or L22K).

[0109] In some embodiments, an AAV vector plasmid comprises AAV inverted terminal repeats (ITRs) and a transgene of interest. In some embodiments, an adenoviral helper plasmid is any adenoviral helper plasmid described herein. In some embodiments, an L4 *trans* plasmid is any L4 *trans* plasmid as described herein.

[0110] In some embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap with an AAV vector plasmid and an adenoviral helper plasmid. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap with an AAV vector plasmid, an adenoviral helper plasmid and an L4 *trans* plasmid. In some

embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap with an AAV vector plasmid, an adenoviral helper plasmid comprising an L4 region, and an L4 *trans* plasmid. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap with an AAV vector plasmid, an adenoviral helper plasmid comprising an L4 region, and an L4 *trans* plasmid comprising a nucleotide sequence encoding L4 22K. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap with an AAV vector plasmid, an adenoviral helper plasmid comprising an L4 region, and an L4 *trans* plasmid comprising a nucleotide sequence encoding L4 33K. In some embodiments, a method of producing a rAAV comprises transfection of a producer cell stably expressing Rep-Cap and comprising a nucleotide sequence encoding an L4 region or portions thereof (e.g., L33K and/or L22K) with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, and an adenoviral helper plasmid comprising an L4 region.

[0111] In some embodiments, an AAV vector plasmid comprises AAV inverted terminal repeats (ITRs) and a transgene of interest. In some embodiments, an adenoviral helper plasmid is any adenoviral helper plasmid described herein. In some embodiments, an L4 *trans* plasmid is any L4 *trans* plasmid as described herein.

EXEMPLIFICATION

[0112] The main purpose of the work described in this disclosure is to develop novel adenoviral helper plasmids for rAAV production that are smaller, contain fewer non-necessary adenoviral genes, and that function as well or better than the most commonly used adenoviral helper plasmids.

[0113] The plasmids provided in this disclosure were synthesized *de novo*, were sequence-verified, and were scaled up for use in large-scale rAAV manufacturing. Production of rAAV studies were performed to compare vector yields when using the provided plasmids versus other commercially available adenoviral helper plasmids. Vector quality and activity were also assessed from rAAV produced with the different adenoviral helper plasmids to confirm that rAAV produced with the provided plasmids is at least

equivalent, if not superior in quality. Taken together, these following examples demonstrate that provided adenoviral helper plasmids generate rAAV of high yield and quality, in a potentially safer and more cost-effective design.

Example 1: Exemplary methods for the production of rAAVs using adenoviral helper plasmids described herein

[0114] HEK293 cells were transfected with a control adenoviral helper plasmid (e.g., a commercially available plasmid, such as pALD-X80, or an adenoviral helper plasmid described in herein). The adenoviral helper plasmid was co-transfected along with pAAVrep2cap9 and pAAV-CMV-GFP or pAAV-CAG-GFP plasmids using PEI transfection in order to generate AAV9/GFP. Four days post-transfection, the HEK293 cells were harvested via 0.5% Triton X-100 lysis and nuclease addition (to degrade RNA, cellular genomic DNA, and remaining plasmid DNA). After 3-4 hours of lysis/nuclease treatment, the cell lysate was sampled and submitted for qPCR titer analysis. Samples were treated with another nuclease, then EDTA and heat-treated, followed by qPCR of diluted samples to determine vector genome copy number per sample. As a metric of transfection efficiency, cells positive for GFP were quantified using fluorescence microscopy.

Example 2: Adenoviral helper plasmid lacking Fiber, L1-52/55K, and Peripentonal Hexon-Associated genes, and having a partial L4 Hexon-Associated Precursor

[0115] To reduce the size of the adenoviral helper plasmid, an adenoviral helper plasmid (pEMBR-1.2: SEQ ID NO: 41) was designed which lacks the fiber gene, the L1-52/55K (Packaging Protein 3) gene, and most of the Hexon Associated Precursor, as well as the Peripentonal Hexon-Associated protein. These deletions were made relative to commercially available helper plasmids, such as pXX6-80. The Adenoviral helper genes were synthesized and assembled into a kanamycin-resistant plasmid backbone. The resulting plasmid is approximately 6.7 kb smaller than pXX6-80.

[0116] The adenoviral helper plasmid described above enabled the production of AAV in HEK293 cells. No major difference in AAV vector yield was observed between cells transfected with pALD-X80, and cells transfected with pEMBR-1.2 as measured by

qPCR (See Fig. 2). rAAV vector produced with pEMBR-1.2 produces a normal vector with the correct ratios of the VP proteins ,as observed when assessing vector capsid purity by SDS-PAGE (See Fig. 3), and the correct size of packaged transgene ,as observed when assessing vector transgene purity by alkaline gel electrophoresis (See Fig. 3). Further, pEMBR-1.2 enabled the production of a fully functional vector capable of transfecting cells. No difference was observed in the transfection of HEK293 cells to generate AAVRH.10/ssCMV-GFP produced with pALD-X80 or pEMBR-1.2 (See Fig. 4).

Example 3: Modifications to the E2a region of adenoviral helper plasmids to reduce plasmid size while maintaining production of AAV

[0117] In order to further reduce the size of adenoviral helper plasmids (e.g., those as described herein), adenoviral helper plasmids were designed with modifications to the E2a region. These plasmids were designed to contain the VA RNA region from pEMBR-1.2B2 (SEQ ID NO: 70) and the E4 region from pEMBR-1.2 (SEQ ID NO: 41) with various modifications to the E2a region based on the promoter regions identified by Guilfoyle (GEP) and Jing (JEP) (See Fig. 14). The sequence modifications to the E2a region based on the GEP and JEP promoters are depicted in Fig. 12 and the sequence modifications to the E2a region based on truncations to the NotI end of the pEMBR1.2B2 E2a region are depicted in Fig. 13.

[0118] These plasmids were tested for the ability to produce AAV in a large volume of 500 mL culture (See Fig. 15A-B). In large volume, pEMR-1.2B2-GEP+1kbp (SEQ ID NO: 81) outperformed pEMBR-1.2B2-GEP (SEQ ID NO: 80). Rep52 and E2a protein expression was also assessed in these plasmids (See Fig. 15C). These data demonstrate that Rep52 is present at a similar level in all constructs tested, and E2a protein expression is likely correlated with AAV production. Although the JEP+1kbp plasmid had E2a protein produced at normal levels, it is missing sequence encoding the L4 (33K and 22K) region.

[0119] Further testing was performed to analyze AAV production with various constructs in 75 mL culture (See Fig. 16). These data show that there is likely little difference in AAV production between pEMBR-1.2B2 (SEQ ID NO:70), pEMBR-1.2B2-

GEP+50bp (SEQ ID NO: 82), and pEMBR-1.2B2-GEP+50bp+Kozak (SEQ ID NO: 83). The JEP+NotIΔ1.5kbp which does not contain the GEP promoter nor L4 (33K and 22K) region produced AAV at a low level.

[0120] A table depicting which elements are included in various plasmids (e.g., those as described herein) and ability to produce E2a protein and AAV is shown in Fig. 17. Without wishing to be bound by any particular theory, these data indicate that the L4 region (containing 33K and 22K ORFs), JEP/GEP regional requirements, and adequate E2a production is important for the production of AAV.

Example 4: Exemplary engineered adenoviral helper plasmids to reduce plasmid size while maintaining production of AAV

[0121] In order to further reduce the size of adenoviral helper plasmids (e.g., those as described herein), adenoviral helper plasmids were designed with modifications to the L4 region, in particular. These plasmids were designed with various modifications to the L4 region (See Figs. 18-19). The sequence modifications to the L4 region are depicted in Fig. 21. The ability of these plasmids to produce AAV from 75 mL culture were assessed (See Fig. 22). Without wishing to be bound by any particular theory, these data demonstrate that only a construct comprising nucleotide sequence for E4 ORFs 4&6 may be required to maintain AAV production. Without wishing to be bound by any particular theory, these data also demonstrate that adding the sequence for the L4 region (containing 33K and 22K ORFs) with the synthetic polyA tail rescues AAV production in constructs which previously did not produce AAV.

[0122] Further testing was performed to analyze AAV production in various constructs with and without the L4 region (containing 33K and 22K ORFs) with the synthetic polyA tail, in comparison to AAV production with pEMBR-1.2B2 where plasmid size is also shown (See Fig. 23). The addition of the L4 region (containing 33K and 22K ORFs) with the synthetic polyA tail to plasmids which previously did not produce AAV resulted in rescued AAV production. An exemplary adenoviral plasmid map is depicted in

Fig. 24 which has been engineered (e.g., with modifications as described herein) to reduce plasmid size while maintaining production of AAV.

[0123] Additionally, this Example describes adenoviral helper plasmids that have been engineered to reduce plasmid size while maintaining production of AAV. The E2a region has been engineered in plasmids described in this example to include native E2a ORF and native E2a polyA tail, the L4 region consisting of 50bp upstream of native L4 promoter, native L4 promoter, and 33 kDa and 22 kDa ORFs (See Fig. 25). A fragment of the L4 33 kDa Ex 2 comprising an E2a promoter region has been repeated for E2a expression. The E2a region has been engineered in plasmids described in this example to not contain L3 23 kDa viral endoprotease, Hexon Assembly, L4 100K region, nor L4 pVIII Hexon Associated Precursor proteins.

Example 5: Introduction of supplementary accessory genes to modified adenoviral helper plasmids

[0124] This Example describes the removal of adenoviral helper genes which results in a smaller adenoviral helper plasmid and allows for the addition of supplementary genes to further improve AAV quality and yield. Specifically, various pEMBR plasmids of various sizes and comprising various supplementary genes (e.g., UL30, UL42, etc.) will be designed from pEMBR backbone plasmids (e.g., those as described herein) and tested AAV production. The HSV-1 DNA polymerase genes (UL30 and UL42) will be added back to plasmids to help replicate the AAV transgene, even when cells are not in S phase. The supplementary genes will be designed to be made as a single transcript. Any number of promoters could be used, including CBA, CMV, PGK, etc. and any number of polyA sites could be used. The direction in which the construct is cloned into the plasmid should theoretically not affect expression because this region will contain both promoter and polyA signal which induces expression of supplementary genes independent from the rest of the plasmid.

[0125] Vector yields for AAV (e.g., AAV9) in clarified lysate will be measured by qPCR with various pEMBR plasmids designed to comprise various supplementary genes (e.g., UL30, UL42, etc.; See pEMBR-1.5A+L4 in Fig. 23).

Example 6: Provision of L4 region sequence *in trans* rescues AAV production

[0126] This Example demonstrates that provision of L4 region sequences, e.g., L4 33K and/or L4 22K, *in trans* (i.e., not included in the adenoviral helper plasmid containing VA RNA, an E4 region, and an E2a region) can rescue and/or increase production of AAV.

[0127] In the present example, a quadruple transfection was performed to add the L4 *trans* plasmid encoding the L4 region to cells *in trans*. Adding the L4 region *in trans* restored rAAV production to wild-type (pEMBR-1.2B2) levels regardless of whether the native promoter driven L4 (L4_pUC57) or an exogenously driven (by human elongation factor 1 alpha, Ef1αL4_pUC57) L4 was added (Figure 28A).

[0128] To determine which L4 protein was more essential, we also utilized plasmids with natively driven codon-optimized 22K or 33K (22K_pUC57 and 33K_pUC57). Our results indicated that 22K was more essential to rAAV production; however, it appeared that the L4 region proteins' effect on production was additive (Figure 28A). Analysis of E2a protein by western blot revealed a slight decrease in E2a protein when 22K was present (Figure 28B).

[0129] To further test the rescue and/or increase AAV production, L4 gene dosage by adding additional copies of L4 region proteins was performed. The current example used for rAAV production was achieved by quadruple transfection of producer cells. The E2a region of the helper plasmids used in the quadruple transfection are shown in Figure 29A. Figure 29B shows average rAAV titer (VG/mL) from triplicate 75 mL quadruple plasmid transfected cultures determined by qPCR for the E2a constructs indicated plus the pUC57 vector containing the specified L4 construct to compare L4 gene dosage of 1X L4 genes, 2X L4 genes, 1X 33K with 2X 22K, and 2X 33K with 1X 22K against the control condition (wild type plus empty vector; 1.2B2 +pUC57). Rep and E2a expression levels resulting from

the transfections described in Figure 29B are shown in Figure 29C. Beta-actin is shown as a loading control. Adenovirus type 5 infected HEK293/HEK293T cells are shown as a positive control and uninfected/nontransfected HEK293 cells are shown as a negative control. Notably, in some conditions the provision of an additional copy of the L4 region (e.g., 1.2B2+L4_pUC57) resulted in increased rAAV titer. Further, provision of additional copies of 33K resulted in increased rAAV titer in several varying conditions. See Figure 29B.

Example 7: Table of Sequences

[0130] The below Table of Sequences lists and describes the various sequences discussed herein. Unless stated otherwise, all sequences are recited with 5' to 3' directionality of the positive strand of a plasmid. This directionality is preserved irrespective of the orientation of a gene or element described to be associated with a sequence. Asterisks as used herein indicate a stop codon.

SEQ ID. No.	Description	Sequence
1	E4 miniP (3' to 5')	ACAGTCAGCCTTACCAGTAAAAAAGAAAACCTATTAAA AAAACACCACTCGACACGGCACCAGCTCAATCAGTCA CAGTGTAAAAAAGGGCCAAGTGCAGAGCGAGTATATA TAGGACTAAAAAATGACGTAACGGTTAAAGTCCACAAA AAACACCCAGAAAACCGCACGCGAACCTACGCCCAGA AACGAAAGCCAAAAAACCCCACAACTTCCTCAAATCGTC ACTTCCGTTTTCCC
73	E4 miniP	GGGAAAACGGAAGTGACGATTTGAGGAAGTTGTGGGT TTTTTGGCTTTCGTTTCTGGGCGTAGGTTCGCGTGCG GTTTTCTGGGTGTTTTTTTGTGGACTTTAACCGTTACGT CATTTTTTAGTCCTATATATACTCGCTCTGCACTTGGC CCTTTTTTTACACTGTGACTGATTGAGCTGGTGCCGTGT CGAGTGGTGTTTTTTTTAATAGGTTTTCTTTTTTTACTGGT AAGGCTGACTGT
2	SV40 Promoter (3' to 5')	TTTGCAAAAGCCTAGGCCTCCAAAAAAGCCTCCTCAC TACTTCTGGAATAGCTCAGAGGCCGAGGCGGCCTCG GCCTCTGCATAAATAAAAAAAATTAGTCAGCCATGGG GCGGAGAATGGGCGGAACTGGGCGGAGTTAGGGGC GGGATGGGCGGAGTTAGGGGCGGGACTATGGTTGCT GACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGC TGGGGAGCCTGGGGACTTTCCACACCTGGTTGCTGA CTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTG GGGAGCCTGGGGACTTTCCACACCCTAACTGACACAC

74	SV40 Promoter	GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCC CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTA GTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCC AGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAG TCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCC CGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCC CCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCG AGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTG AGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAA
3	L4 (Hexon Assembly)	ATGCCCTTCTCCCACGCAGACACGATCGGCACACTCA GCGGGTTCATCACCGTAATTTCACTTTCCGCTTCGCT GGGCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGC GCCACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGC GCTTACCTCCTTTGCCATGCTTGATTAGCACCGGTGG GTTGCTGAAACCCACCATTTGTAGCGCCACATCTTCTC TTTCTTCCTCGCTGTCCACGATTACCTCTGGTGATGGC GGGCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTC TTCTTGGGCGCAATGGCCAAATCCGCCGCCGAGGTC GATGGCCGCGGGCTGGGTGTGCGCGGCACCAGCGC GTCTTGTGATGAGTCTTCCTCGTCCTCGGACTCGATA CGCCGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGA GGCGGCGGCGACGGGGACGGGGACGACACGTCCTC CATGGTTGGGGGACGTCGCGCCGCACCGCGTCCGCG CTCGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTG GCCATTTCCTTCTCCTATAGGCAGAAAAAGATCATGGA GTCAGTCGAGAAGAAGGACAGCCTAACCGCCCCCTCT GAGTTCGCCACCACCGCCTCCACCGATGCCGCCAAC GCGCCTACCACCTTCCCCGTCGAGGCACCCCCGCTT GAGGAGGAGGAAGTGATTATCGAGCAGGACCCAGGT TTTGTAAGCGAAGACGACGAGGACCGCTCAGTACCAA CAGAGGATAAAAAGCAAGACCAGGACAACGCAGAGG CAAACGAGGAACAAGTCGGGCGGGGGGACGAAAGGC ATGGCGACTACCTAGATGTGGGAGACGACGTGCTGTT GAAGCATCTGCAGCGCCAGTGCGCCATTATCTGCGAC GCGTTGCAAGAGCGCAGCGATGTGCCCCTCGCCATA GCGGATGTCAGCCTTGCCTACGAACGCCACCTATTCT CACCGCGCGTACCCCCCAAACGCCAAGAAAACGGCA CATGCGAGCCCAACCCGCGCCTCAACTTCTACCCCGT ATTTGCCGTGCCAGAGGTGCTTGCCACCTATCACATC TTTTTCCAAAACTGCAAGATACCCCTATCCTGCCGTGC CAACCGCAGCCGAGCGGACAAGCAGCTGGCCTTGCG GCAGGGCGCTGTCATACCTGATATCGCCTCGCTCAAC GAAGTGCCAAAAATCTTTGAGGGTCTTGGACGCGACG AGAAGCGCGCGGCAAACGCTCTGCAACAGGAAAACA GCGAAAATGAAAGTCACTCTGGAGTGTTGGTGGAACT CGAGGGTGACAACGCGCGCCTAGCCGTACTAAAACG CAGCATCGAGGTCACCCACTTTGCCTACCCGGCACTT

		AACCTACCCCCCAAGGTCATGAGCACAGTCATGAGTG AGCTGATCGTGCGCCGTGCGCAGCCCCTGGAGAGGG ATGCAAATTTGCAAGAACAAACAGAGGAGGGCCTACC CGCAGTTGGCGACGAGCAGCTAGCGCGCTGGCTTCA AACGCGCGAGCCTGCCGACTTGGAGGAGCGACGCAA ACTAATGATGGCCGCAGTGCTCGTTACCGTGGAGCTT GAGTGCATGCAGCGGTTCTTTGCTGACCCGGAGATGC AGCGCAAGCTAGAGGAAACATTGCACTACACCTTTCG ACAGGGCTACGTACGCCAGGCCTGCAAGATCTCCAAC GTGGAGCTCTGCAACCTGGTCTCCTACCTTGGAATTT TGCACGAAAACCGCCTTGGGCAAAACGTGCTTCATTC CACGCTCAAGGGCGAGGCGCGCCGCGACTACGTCCG CGACTGCGTTTACTTATTTCTATGCTACACCTGGCAGA CGGCCATGGGCGTTTGGCAGCAGTGCTTGGAGGAGT GCAACCTCAAGGAGCTGCAGAAACTGCTAAAGCAAAA CTTGAAGGACCTATGGACGGCCTTCAACGAGCGCTCC GTGGCCGCGCACCTGGCGGACATCATTTTCCCCGAA CGCCTGCTTAAAACCCTGCAACAGGGTCTGCCAGACT TCACCAGTCAAAGCATGTTGCAGAACTTTAGGAACTTT ATCCTAGAGCGCTCAGGAATCTTGCCCGCCACCTGCT GTGCACTTCCTAGCGACTTTGTGCCCATTAAGTACCG CGAATGCCCTCCGCCGCTTTGGGGCCACTGCTACCTT CTGCAGCTAGCCAACTACCTTGCCTACCACTCTGACA TAATGGAAGACGTGAGCGGTGACGGTCTACTGGAGT GTCACTGTCGCTGCAACCTATGCACCCCGCACCGCTC CCTGGTTTGCAATTCGCAGCTGCTTAACGAAAGTCAA ATTATCGGTACCTTTGAGCTGCAGGGTCCCTCGCCTG ACGAAAAGTCCGCGGCTCCGGGGTTGAAACTCACTCC GGGGCTGTGGACGTCGGCTTACCTTCGCAAATTTGTA CCTGAGGACTACCACGCCCACGAGATTAGGTTCTACG AAGACCAATCCCGCCCGCCTAATGCGGAGCTTACCGC CTGCGTCATTACCCAGGGCCACATTCTTGGCCAATTG CAAGCCATCAACAAAGCCCGCCAAGAGTTTCTGCTAC GAAAGGGACGGGGGGTTTACTTGGACCCCCAGTCCG GCGAGGAGCTCAACCCAATCCCCCCGCCGCCGCAGC CCTATCAGCAGCAGCCGCGGGCCCTTGCTTCCCAGG ATGGCACCCAAAAAGAAGCTGCAGCTGCCGCCGCCA CCCACGGACGAGGAGGAATACTGGGACAGTCAGGCA GAGGAGGTTTTGGACGAGGAGGAGGAGGACATGATG GAAGACTGGGAGAGCCTAGACGAGGAAGCTTCCGAG GTCGAAGAGGTGTCAGACGAAACACCGTCACCCTCG GTCGCATTCCCCTCGCCGGCGCCCCAGAAATCGGCA ACCGGTTCCAGCATGGCTACAACCTCCGCTCCTCAGG CGCCGCCGGCACTGCCCGTTCGCCGACCCAACCGTA G
4	L4 (Hexon Assembly)	MPFSHADTIGTLSGFITVISLSASLGSSSSSCVRIPRATGS SSFSRRTVRLPPLPCLISTGGLLKPTICSATSSLSSSLSTI

		TSGDGGRSGLGEGRFFFFLGAMAKSAAEVDGRGLGVR GTSASCDESSSSSDSIRRLIRFFGGARGGGGDGDGDDT SSMVGGRRAAPRPRSGVVSRCSSSRLAISFSYRQKKIM ESVEKKDSLTAPSEFATTASTDAANAPTTFPVEAPPLEE EEVIIEQDPGFVSEDDEDRSVPTEDKKQDQDNAEANEE QVGRGDERHGDYLDVGDDVLLKHLQRQCAIICDALQER SDVPLAIADVSLAYERHLFSPRVPPKRQENGTCEPNPRL NFYPVFAVPEVLATYHIFFQNCKIPLSCRANRSRADKQLA LRQGAVIPDIASLNEVPKIFEGLGRDEKRAANALQQENSE NESHSGVLVELEGDNARLAVLKRSIEVTHFAYPALNLPP KVMSTVMSELIVRRAQPLERDANLQEQTEEGLPAVGDE QLARWLQTREPADLEERRKLMMAAVLVTVELECMQRFF ADPEMQRKLEETLHYTFRQGYVRQACKISNVELCNLVSY LGILHENRLGQNVLHSTLKGEARRDYVRDCVYLFLCYTW QTAMGVWQQCLEECNLKELQKLLKQNLKDLWTAFNER SVAAHLADIIFPERLLKTLQQGLPDFTSQSMLQNFRNFIL ERSGILPATCCALPSDFVPIKYRECPPPLWGHCYLLQLA NYLAYHSDIMEDVSGDGLLECHCRCNLCTPHRSLVCNS QLLNESQIIGTFELQGPSPDEKSAAPGLKLTPGLWTSAYL RKFVPEDYHAHEIRFYEDQSRPPNAELTACVITQGHILG QLQAINKARQEFLLRKGRGVYLDPQSGEELNPIPPPPQP YQQQPRALASQDGTQKEAAAAAATHGRGGILGQSGRG GFGRGGGGHDGRLGEPRRGSFRGRRGVRRNTVTLGRI PLAGAPEIGNRFQHGYNLRSSGAAGTARSPTQP**
5	L4 (33kDa Ex2)	GCCCATACTGCACCGGCGGCAGCGGCAGCAACAGCA GCGGCCACACAGAAGCAAAGGCGACCGGATAGCAAG ACTCTGACAAAGCCCAAGAAATCCACAGCGGCGGCA GCAGCAGGAGGAGGAGCGCTGCGTCTGGCGCCCAAC GAACCCGTATCGACCCGCGAGCTTAGAAACAGGATTT TTCCCACTCTGTATGCTATATTTCAACAGAGCAGGGG CCAAGAACAAGAGCTGAAAATAAAAAACAGGTCTCTG CGATCCCTCACCCGCAGCTGCCTGTATCACAAAAGCG AAGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGG CTCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGAC TAG
6	L4 (33kDa Ex2)	AHTAPAAAAATAAATQKQRRPDSKTLTKPKKSTAAAAAG GGALRLAPNEPVSTRELRNRIFPTLYAIFQQSRGQEQEL KIKNRSLRSLTRSCLYHKSEDQLRRTLEDAEALFSKYCAL TLKD*
7	L4 Encapsidation Protein (22 kDa)	ATGGCACCCAAAAAGAAGCTGCAGCTGCCGCCGCCA CCCACGGACGAGGAGGAATACTGGGACAGTCAGGCA GAGGAGGTTTTGGACGAGGAGGAGGAGGACATGATG GAAGACTGGGAGAGCCTAGACGAGGAAGCTTCCGAG GTCGAAGAGGTGTCAGACGAAACACCGTCACCCTCG GTCGCATTCCCCTCGCCGGCGCCCCAGAAATCGGCA ACCGGTTCCAGCATGGCTACAACCTCCGCTCCTCAGG CGCCGCCGGCACTGCCCGTTCGCCGACCCAACCGTA

		GATGGGACACCACTGGAACCAGGGCCGGTAAGTCCA AGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGC GCCAAGGCTACCGCTCATGGCGCGGGCACAAGAACG CCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACAT CTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGC GTGGCCTTCCCCCGTAACATCCTGCATTACTACCGTC ATCTCTACAGCCCATACTGCACCGGCGGCAGCGGCA GCAACAGCAGCGGCCACACAGAAGCAAAGGCGACCG GATAG
8	L4 Encapsidation Protein (22 kDa)	MAPKKKLQLPPPPTDEEEYWDSQAEEVLDEEEEDMME DWESLDEEASEVEEVSDETPSPSVAFPSPAPQKSATGS SMATTSAPQAPPALPVRRPNRRWDTTGTRAGKSKQPP PLAQEQQQRQGYRSWRGHKNAIVACLQDCGGNISFAR RFLLYHHGVAFPRNILHYYRHLYSPYCTGGSGSNSSGHT EAKATG*
9	L4 (33kDa Ex2) (Partial)	CACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAA GACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCTGA CTCTTAAGGACTAG
10	L4 (33kDa Ex2) (Partial)	HKSEDQLRRTLEDAEALFSKYCALTLKD*
11	E2a promoter region (3' to 5')	AGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA
75	E2a promoter region	TAAGAGTCAGCGCGCAGTATTTACTGAAGAGAGCCT
12	L4 pVIII Hexon-Associated Precursor	ATGAGCAAGGAAATTCCCACGCCCTACATGTGGAGTT ACCAGCCACAAATGGGACTTGCGGCTGGAGCTGCCC AAGACTACTCAACCCGAATAAACTACATGAGCGCGGG
13	L4 pVIII Hexon-Associated Precursor	MSKEIPTPYMWSYQPQMGLAAGAAQDYSTRINYMSA
14	VA RNA Region A	TATCCGTAGATGTACCTGGACATCCAGGTGATGCCGG CGGCGGTGGTGGAGGCGCGCGGAAAGTCGCGGACG CGGTTCCAGATGTTGCGCAGCGGCAAAAAGTGCTCCA TGGTCGGGACGCTCTGGCCGGTCAGGCGCGCGCAAT CGTTGACGCTCTAGCGTGCAAAAGGAGAGCCTGTAAG CGGGCACTCTTCCGTGGTCTGGTGGATAAATTCGCAA GGGTATCATGGCGGACGACCGGGGTTCGAGCCCCGT ATCCGGCCGTCCGCCGTGATCCATGCGGTTACCGCC CGCGTGTCGAACCCAGGTGTGCGACGTCAGACAACG GGGGAGTGCTCCTTTTGGCTTCCTTCCAGGCGCGGC GGCTGCTGCGCTAGCTTTTTTGGCCACTGGCCGCGC GCAGCGTAAGCGGTTAGGCTGGAAAGCGAAAGCATTA

		AGTGGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCA AGGGTTGAGTCGCGGGACCCCCGGTTCGAGTCTCGG ACCG
15	Ad2 VA RNA region-B	ATCCGTAGATGTACCTGGACATCCAGGTGATGCCGGC GGCGGTGGTGGAGGCGCGCGGAAAGTCGCGGACGC GGTTCCAGATGTTGCGCAGCGGCAAAAAGTGCTCCAT GGTCGGGACGCTCTGGCCGGTGAGGCGTGCGCAGTC GTTGACGCTCTAGACCGTGCAAAAGGAGAGCCTGTAA GCGGGCACTCTTCCGTGGTCTGGTGGATAAATTCGCA AGGGTATCATGGCGGACGACCGGGGTTCGAACCCCG GATCCGGCCGTCCGCCGTGATCCATGCGGTTACCGC CCGCGTGTCGAACCCAGGTGTGCGACGTCAGACAAC GGGGGAGCGCTCCTTTTGGCTTCCTTCCAGGCGCGG CGGCTGCTGCGCTAGCTTTTTTGGCCACTGGCCGCG CGCGGCGTAAGCGGTTAGGCTGGAAAGCGAAAGCAT TAAGTGGCTCGCTCCCTGTAGCCGGAGGGTTATTTTC CAAGGGTTGAGTCGCAGGACCCCCGGTTCGAGTCTC GGGCCGGCCGGACTGCGGCGAACGGGGGTTTGCCT CCCCGTCATGCAAGACCCCGCTTGCAAATTCCTCCGG AAACAGGGACGAGCCCCTTTTTTGCTTTTCCCAGATG CATCCGGTGCTGCGGCAGATGCGCCCCCCTCCTCAG CAGCGGCAAGAGCAAGAGCAGCGGCAGACATGCAGG GCACCCTCCCCTTCTCCTACCGCGTCAGGAGGGGCA ACATC
16	VA RNAI-A	AGCGGGCACTCTTCCGTGGTCTGGTGGATAAATTCGC AAGGGTATCATGGCGGACGACCGGGGTTCGAGCCCC GTATCCGGCCGTCCGCCGTGATCCATGCGGTTACCG CCCGCGTGTCGAACCCAGGTGTGCGACGTCAGACAA CGGGGGAGTGCTCCTTTT
17	VA RNAI-B	AGCGGGCACTCTTCCGTGGTCTGGTGGATAAATTCGC AAGGGTATCATGGCGGACGACCGGGGTTCGAACCCC GGATCCGGCCGTCCGCCGTGATCCATGCGGTTACCG CCCGCGTGTCGAACCCAGGTGTGCGACGTCAGACAA CGGGGGAGCGCTCCTTTT
18	Partial VA RNAII-A	GGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAAGG GTTGAGTCGCGGGACCCCCGGTTCGAGTCTCGGACC G
19	VA RNA II-B	GGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAAGG GTTGAGTCGCAGGACCCCCGGTTCGAGTCTCGGGCC GGCCGGACTGCGGCGAACGGGGGTTTGCCTCCCCGT CATGCAAGACCCCGCTTGCAAATTCCTCCGGAAACAG GGACGAGCCCCTTTT
20	Partial DNA terminal protein (3’ to 5’)	CAGAAGCACCATGTCCTTGGGTCCGGCCTGCTGAATG CGCAGGCGGTCGGCCATGCCCCAGGCTTCGTTTTGA CATCGGCGCAGGTCTTTGTAGTAGTCTTGCATGAGCC TTTCTACCGGCACTTCTTCTTCTCCTTCCTCTTGTCCT GCATCTCTTGCATCTATCGCTGCGGCGGCGGCGGAG

		TTTGGCCGTAGGTGGCGCCCTCTTCCTCCCATGCGTG TGACCCCGAAGCCCCTCATCGGCTGAAGCAGGGCTA GGTCGGCGACAACGCGCTCGGCTAATATGGCCTGCT GCACCTGCGTGAGGGTAGACTGGAAGTCATCCATGTC CACAAAGCGGTGGTATGCGCCCGTGTTGATGGTGTAA GTGCAGTTGGCCATAACGGACCAGTTAACGGTCTGGT GACCCGGCTGCGAGAGCTCGGTGTACCTGAGACGCG AGTAAGCCCTCGAGTCAAATACGTAGTCGTTGCAAGT CCGCACCAGGTACTGGTATCCCACCAAAAAGTGCGGC GGCGGCTGGCGGTAGAGGGGCCAGCGTAGGGTGGC CGGGGCTCCGGGGGCGAGATCTTCCAACATAAGGCG ATGATATCCGTAGATGTACCTGGACATCCAGGTGATG CCGGCGGCGGTGGTGGAGGCGCGCGGAAAGTCGCG GACGCGGTTCCAGATGTTGCGCAGCGGCAAAAAGTG CTCCAT
76	Partial DNA terminal protein	ATGGAGCACTTTTTGCCGCTGCGCAACATCTGGAACC GCGTCCGCGACTTTCCGCGCGCCTCCACCACCGCCG CCGGCATCACCTGGATGTCCAGGTACATCTACGGATA TCATCGCCTTATGTTGGAAGATCTCGCCCCCGGAGCC CCGGCCACCCTACGCTGGCCCCTCTACCGCCAGCCG CCGCCGCACTTTTTGGTGGGATACCAGTACCTGGTGC GGACTTGCAACGACTACGTATTTGACTCGAGGGCTTA CTCGCGTCTCAGGTACACCGAGCTCTCGCAGCCGGG TCACCAGACCGTTAACTGGTCCGTTATGGCCAACTGC ACTTACACCATCAACACGGGCGCATACCACCGCTTTG TGGACATGGATGACTTCCAGTCTACCCTCACGCAGGT GCAGCAGGCCATATTAGCCGAGCGCGTTGTCGCCGA CCTAGCCCTGCTTCAGCCGATGAGGGGCTTCGGGGT CACACGCATGGGAGGAAGAGGGCGCCACCTACGGCC AAACTCCGCCGCCGCCGCAGCGATAGATGCAAGAGA TGCAGGACAAGAGGAAGGAGAAGAAGAAGTGCCGGT AGAAAGGCTCATGCAAGACTACTACAAAGACCTGCGC CGATGTCAAAACGAAGCCTGGGGCATGGCCGACCGC CTGCGCATTCAGCAGGCCGGACCCAAGGACATGGTG CTTCTG
21	Partial DNA terminal protein	MEHFLPLRNIWNRVRDFPRASTTAAGITWMSRYIYGYHR LMLEDLAPGAPATLRWPLYRQPPPHFLVGYQYLVRTCN DYVFDSRAYSRLRYTELSQPGHQTVNWSVMANCTYTIN TGAYHRFVDMDDFQSTLTQVQQAILAERVVADLALLQP MRGFGVTRMGGRGRHLRPNSAAAAAIDARDAGQEEGE EEVPVERLMQDYYKDLRRCQNEAWGMADRLRIQQAGP KDMVLL
22	23kDa endoproteas e (partial sequence)	CCGAGACTGGGGGCGTACACTGGATGGCCTTTGCCT GGAACCCGCACTCAAAAACATGCTACCTCTTTGAGCC CTTTGGCTTTTCTGACCAGCGACTCAAGCAGGTTTAC CAGTTTGAGTACGAGTCACTCCTGCGCCGTAGCGCCA

		TTGCTTCTTCCCCCGACCGCTGTATAACGCTGGAAAA GTCCACCCAAAGCGTACAGGGGCCCAACTCGGCCGC CTGTGGACTATTCTGCTGCATGTTTCTCCACGCCTTTG CCAACTGGCCCCAAACTCCCATGGATCACAACCCCAC CATGAACCTTATTACCGGGGTACCCAACTCCATGCTC AACAGTCCCCAGGTACAGCCCACCCTGCGTCGCAAC CAGGAACAGCTCTACAGCTTCCTGGAGCGCCACTCGC CCTACTTCCGCAGCCACAGTGCGCAGATTAGGAGCG CCACTTCTTTTTGTCACTTGAAAAACATGTAA
23	23kDa endoproteas e (partial sequence)	MGSSEQELKAIVKDLGCGPYFLGTYDKRFPGFVSPHKLA CAIVNTAGRETGGVHWMAFAWNPRSKTCYLFEPFGFSD QRLKQVYQFEYESLLRRSAIASSPDRCITLEKSTQSVQG PNSAACGLFCCMFLHAFANWPQTPMDHNPTMNLITGVP NSMLNSPQVQPTLRRNQEQLYSFLERHSPYFRSHSAQI RSATSFCHLKNM
24	E2a (3’ to 5’)	TTAAAAATCAAAGGGGTTCTGCCGCGCATCGCTATGC GCCACTGGCAGGGACACGTTGCGATACTGGTGTTTAG TGCTCCACTTAAACTCAGGCACAACCATCCGCGGCAG CTCGGTGAAGTTTTCACTCCACAGGCTGCGCACCATC ACCAACGCGTTTAGCAGGTCGGGCGCCGATATCTTGA AGTCGCAGTTGGGGCCTCCGCCCTGCGCGCGCGAGT TGCGATACACAGGGTTGCAGCACTGGAACACTATCAG CGCCGGGTGGTGCACGCTGGCCAGCACGCTCTTGTC GGAGATCAGATCCGCGTCCAGGTCCTCCGCGTTGCT CAGGGCGAACGGAGTCAACTTTGGTAGCTGCCTTCCC AAAAAGGGCGCGTGCCCAGGCTTTGAGTTGCACTCG CACCGTAGTGGCATCAAAAGGTGACCGTGCCCGGTCT GGGCGTTAGGATACAGCGCCTGCATAAAAGCCTTGAT CTGCTTAAAAGCCACCTGAGCCTTTGCGCCTTCAGAG AAGAACATGCCGCAAGACTTGCCGGAAAACTGATTGG CCGGACAGGCCGCGTCGTGCACGCAGCACCTTGCGT CGGTGTTGGAGATCTGCACCACATTTCGGCCCCACCG GTTCTTCACGATCTTGGCCTTGCTAGACTGCTCCTTCA GCGCGCGCTGCCCGTTTTCGCTCGTCACATCCATTTC AATCACGTGCTCCTTATTTATCATAATGCTTCCGTGTA GACACTTAAGCTCGCCTTCGATCTCAGCGCAGCGGTG CAGCCACAACGCGCAGCCCGTGGGCTCGTGATGCTT GTAGGTCACCTCTGCAAACGACTGCAGGTACGCCTGC AGGAATCGCCCCATCATCGTCACAAAGGTCTTGTTGC TGGTGAAGGTCAGCTGCAACCCGCGGTGCTCCTCGTT CAGCCAGGTCTTGCATACGGCCGCCAGAGCTTCCACT TGGTCAGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTT ATCCACGTGGTACTTGTCCATCAGCGCGCGCGCAGC CTCCATGCCCTTCTCCCACGCAGACACGATCGGCACA CTCAGCGGGTTCATCACCGTAATTTCACTTTCCGCTTC GCTGGGCTCTTCCTCTTCCTCTTGCGTCCGCATACCA CGCGCCACTGGGTCGTCTTCATTCAGCCGCCGCACT

		GTGCGCTTACCTCCTTTGCCATGCTTGATTAGCACCG GTGGGTTGCTGAAACCCACCATTTGTAGCGCCACATC TTCTCTTTCTTCCTCGCTGTCCACGATTACCTCTGGTG ATGGCGGGCGCTCGGGCTTGGGAGAAGGGCGCTTCT TTTTCTTCTTGGGCGCAATGGCCAAATCCGCCGCCGA GGTCGATGGCCGCGGGCTGGGTGTGCGCGGCACCA GCGCGTCTTGTGATGAGTCTTCCTCGTCCTCGGACTC GATACGCCGCCTCATCCGCTTTTTTGGGGGCGCCCG GGGAGGCGGCGGCGACGGGGACGGGGACGACACGT CCTCCATGGTTGGGGGACGTCGCGCCGCACCGCGTC CGCGCTCGGGGGTGGTTTCGCGCTGCTCCTCTTCCC GACTGGCCAT
77	E2a	ATGGCCAGTCGGGAAGAGGAGCAGCGCGAAACCACC CCCGAGCGCGGACGCGGTGCGGCGCGACGTCCCCC AACCATGGAGGACGTGTCGTCCCCGTCCCCGTCGCC GCCGCCTCCCCGGGCGCCCCCAAAAAAGCGGATGAG GCGGCGTATCGAGTCCGAGGACGAGGAAGACTCATC ACAAGACGCGCTGGTGCCGCGCACACCCAGCCCGCG GCCATCGACCTCGGCGGCGGATTTGGCCATTGCGCC CAAGAAGAAAAAGAAGCGCCCTTCTCCCAAGCCCGAG CGCCCGCCATCACCAGAGGTAATCGTGGACAGCGAG GAAGAAAGAGAAGATGTGGCGCTACAAATGGTGGGTT TCAGCAACCCACCGGTGCTAATCAAGCATGGCAAAGG AGGTAAGCGCACAGTGCGGCGGCTGAATGAAGACGA CCCAGTGGCGCGTGGTATGCGGACGCAAGAGGAAGA GGAAGAGCCCAGCGAAGCGGAAAGTGAAATTACGGT GATGAACCCGCTGAGTGTGCCGATCGTGTCTGCGTG GGAGAAGGGCATGGAGGCTGCGCGCGCGCTGATGG ACAAGTACCACGTGGATAACGATCTAAAGGCGAACTT CAAACTACTGCCTGACCAAGTGGAAGCTCTGGCGGCC GTATGCAAGACCTGGCTGAACGAGGAGCACCGCGGG TTGCAGCTGACCTTCACCAGCAACAAGACCTTTGTGA CGATGATGGGGCGATTCCTGCAGGCGTACCTGCAGT CGTTTGCAGAGGTGACCTACAAGCATCACGAGCCCAC GGGCTGCGCGTTGTGGCTGCACCGCTGCGCTGAGAT CGAAGGCGAGCTTAAGTGTCTACACGGAAGCATTATG ATAAATAAGGAGCACGTGATTGAAATGGATGTGACGA GCGAAAACGGGCAGCGCGCGCTGAAGGAGCAGTCTA GCAAGGCCAAGATCGTGAAGAACCGGTGGGGCCGAA ATGTGGTGCAGATCTCCAACACCGACGCAAGGTGCTG CGTGCACGACGCGGCCTGTCCGGCCAATCAGTTTTCC GGCAAGTCTTGCGGCATGTTCTTCTCTGAAGGCGCAA AGGCTCAGGTGGCTTTTAAGCAGATCAAGGCTTTTAT GCAGGCGCTGTATCCTAACGCCCAGACCGGGCACGG TCACCTTTTGATGCCACTACGGTGCGAGTGCAACTCA AAGCCTGGGCACGCGCCCTTTTTGGGAAGGCAGCTA CCAAAGTTGACTCCGTTCGCCCTGAGCAACGCGGAG

		GACCTGGACGCGGATCTGATCTCCGACAAGAGCGTG CTGGCCAGCGTGCACCACCCGGCGCTGATAGTGTTC CAGTGCTGCAACCCTGTGTATCGCAACTCGCGCGCG CAGGGCGGAGGCCCCAACTGCGACTTCAAGATATCG GCGCCCGACCTGCTAAACGCGTTGGTGATGGTGCGC AGCCTGTGGAGTGAAAACTTCACCGAGCTGCCGCGG ATGGTTGTGCCTGAGTTTAAGTGGAGCACTAAACACC AGTATCGCAACGTGTCCCTGCCAGTGGCGCATAGCGA TGCGCGGCAGAACCCCTTTGATTTTTAA
25	E2a	MASREEEQRETTPERGRGAARRPPTMEDVSSPSPSPP PPRAPPKKRMRRRIESEDEEDSSQDALVPRTPSPRPST SAADLAIAPKKKKKRPSPKPERPPSPEVIVDSEEEREDVA LQMVGFSNPPVLIKHGKGGKRTVRRLNEDDPVARGMRT QEEEEEPSEAESEITVMNPLSVPIVSAWEKGMEAARALM DKYHVDNDLKANFKLLPDQVEALAAVCKTWLNEEHRGL QLTFTSNKTFVTMMGRFLQAYLQSFAEVTYKHHEPTGC ALWLHRCAEIEGELKCLHGSIMINKEHVIEMDVTSENGQ RALKEQSSKAKIVKNRWGRNVVQISNTDARCCVHDAAC PANQFSGKSCGMFFSEGAKAQVAFKQIKAFMQALYPNA QTGHGHLLMPLRCECNSKPGHAPFLGRQLPKLTPFALS NAEDLDADLISDKSVLASVHHPALIVFQCCNPVYRNSRA QGGGPNCDFKISAPDLLNALVMVRSLWSENFTELPRMV VPEFKWSTKHQYRNVSLPVAHSDARQNPFDF*
26	Chicken B Actin promoter (3’ to 5’)	CGCCCGCCGCGCGCTTCGCTTTTTATAGGGCCGCCG CCGCCGCCGCCTCGCCATAAAAGGAAACTTTCGGAG CGCGCCGCTCTGATTGGCTGCCGCCGCACCTCTCCG CCTCGCCCCGCCCCGCCCCTCGCCCCGCCCCGCCCC GCCTGGCGCGCGCCCCCCCCCCCCCCCGCCCCCAT CGCTGCACAAAATAATTAAAAAATAAATAAATACAAAAT TGGGGGTGGGGAGGGGGGGGAGATGGGGAGAGTGA AGCAGAACGTGGGGCTCACCTCGA
78	Chicken B Actin promoter	TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCAT CTCCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTA TTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGG GGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGG GGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTG CGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGT TTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCT ATAAAAAGCGAAGCGCGCGGCGGGCG
27	E2a polyA signal (3’ to 5’)	TGATTATTTACCCCCACCCTTGCCGTCTGCGCCGT
79	E2a polyA signal	ACGGCGCAGACGGCAAGGGTGGGGGTAAATAATCA
28	SV40pA #4	GATCCAGACATGATAAGATACATTGATGAGTTTGGACA AACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTT

		GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA TAAGCTGCAATAAACAAGTT
29	UL30 DNA Polymerase Catalytic Subunit	ATGTTTTCCGGTGGCGGCGGCCCGCTGTCCCCCGGA GGAAAGTCGGCGGCCAGGGCGGCGTCCGGGTTTTTT GCGCCCGCCGGCCCTCGCGGAGCCAGCCGGGGACC CCCGCCTTGTTTGAGGCAAAACTTTTACAACCCCTACC TCGCCCCAGTCGGGACGCAACAGAAGCCGACCGGGC CAACCCAGCGCCATACGTACTATAGCGAATGCGATGA ATTTCGATTCATCGCCCCGCGGGTGCTGGACGAGGAT GCCCCCCCGGAGAAGCGCGCCGGGGTGCACGACGG TCACCTCAAGCGCGCCCCCAAGGTGTACTGCGGGGG GGACGAGCGCGACGTCCTCCGCGTCGGGTCGGGCG GCTTCTGGCCGCGGCGCTCGCGCCTGTGGGGCGGC GTGGACCACGCCCCGGCGGGGTTCAACCCCACCGTC ACCGTCTTTCACGTGTACGACATCCTGGAGAACGTGG AGCACGCGTACGGCATGCGCGCGGCCCAGTTCCACG CGCGGTTTATGGACGCCATCACACCGACGGGGACCG TCATCACGCTCCTGGGCCTGACTCCGGAAGGCCACC GGGTGGCCGTTCACGTTTACGGCACGCGGCAGTACT TTTACATGAACAAGGAGGAGGTCGACAGGCACCTACA ATGCCGCGCCCCACGAGATCTCTGCGAGCGCATGGC CGCGGCCCTGCGCGAGTCCCCGGGCGCGTCGTTCC GCGGCATCTCCGCGGACCACTTCGAGGCGGAGGTGG TGGAGCGCACCGACGTGTACTACTACGAGACGCGCC CCGCTCTGTTTTACCGCGTCTACGTCCGAAGCGGGCG CGTGCTGTCGTACCTGTGCGACAACTTCTGCCCGGCC ATCAAGAAGTACGAGGGTGGGGTCGACGCCACCACC CGGTTCATCCTGGACAACCCCGGGTTCGTCACCTTCG GCTGGTACCGTCTCAAACCGGGCCGGAACAACACGC TAGCCCAGCCGCGGGCCCCGATGGCCTTCGGGACAT CCAGCGACGTCGAGTTTAACTGTACGGCGGACAACCT GGCCATCGAGGGGGGCATGAGCGACCTACCGGCATA CAAGCTCATGTGCTTCGATATCGAATGCAAGGCGGGG GGGGAGGACGAGCTGGCCTTTCCGGTGGCCGGGCA CCCGGAGGACCTGGTCATCCAGATATCCTGTCTGCTC TACGACCTGTCCACCACCGCCCTGGAGCACGTCCTCC TGTTTTCGCTCGGTTCCTGCGACCTCCCCGAATCCCA CCTGAACGAGCTGGCGGCCAGGGGCCTGCCCACGCC CGTGGTTCTGGAATTCGACAGCGAATTCGAGATGCTG TTGGCCTTCATGACCCTTGTGAAACAGTACGGCCCCG AGTTCGTGACCGGGTACAACATCATCAACTTCGACTG GCCCTTCTTGCTGGCCAAGCTGACGGACATTTACAAG GTCCCCCTGGACGGGTACGGCCGCATGAACGGCCGG GGCGTGTTTCGCGTGTGGGACATAGGCCAGAGCCAC TTCCAGAAGCGCAGCAAGATAAAGGTGAACGGCATGG TGAACATCGACATGTACGGGATTATAACCGACAAGAT CAAGCTCTCGAGCTACAAGCTCAACGCCGTGGCCGAA

		GCCGTCCTGAAGGACAAGAAGAAGGACCTGAGCTATC GCGACATCCCCGCCTACTACGCCGCCGGGCCCGCGC AACGCGGGGTGATCGGCGAGTACTGCATACAGGATT CCCTGCTGGTGGGCCAGCTGTTTTTTAAGTTTTTGCC CCATCTGGAGCTCTCGGCCGTCGCGCGCTTGGCGGG TATTAACATCACCCGCACCATCTACGACGGCCAGCAG ATCCGCGTCTTTACGTGCCTGCTGCGCCTGGCCGACC AGAAGGGCTTTATTCTGCCGGACACCCAGGGGCGATT TAGGGGCGCCGGGGGGGAGGCGCCCAAGCGTCCGG CCGCAGCCCGGGAGGACGAGGAGCGGCCAGAGGAG GAGGGGGAGGACGAGGACGAACGCGAGGAGGGCGG GGGCGAGCGGGAGCCGGAGGGCGCGCGGGAGACC GCCGGCAGGCACGTGGGGTACCAGGGGGCCAGGGT CCTTGACCCCACTTCCGGGTTTCACGTGAACCCCGTG GTGGTGTTCGACTTTGCCAGCCTGTACCCCAGCATCA TCCAGGCCCACAACCTGTGCTTCAGCACGCTCTCCCT GAGGGCCGACGCAGTGGCGCACCTGGAGGCGGGCA AGGACTACCTGGAGATCGAGGTGGGGGGGCGACGG CTGTTCTTCGTCAAGGCTCACGTGCGAGAGAGCCTCC TCAGCATCCTCCTGCGGGACTGGCTCGCCATGCGAAA GCAGATCCGCTCGCGGATTCCCCAGAGCAGCCCCGA GGAGGCCGTGCTCCTGGACAAGCAGCAGGCCGCCAT CAAGGTCGTGTGTAACTCGGTGTACGGGTTCACGGGA GTGCAGCACGGACTCCTGCCGTGCCTGCACGTTGCC GCGACGGTGACGACCATCGGCCGCGAGATGCTGCTC GCGACCCGCGAGTACGTCCACGCGCGCTGGGCGGC CTTCGAACAGCTCCTGGCCGATTTCCCGGAGGCGGC CGACATGCGCGCCCCCGGGCCCTATTCCATGCGCAT CATCTACGGGGACACGGACTCCATCTTTGTGCTGTGC CGCGGCCTCACGGCCGCCGGGCTGACGGCCGTGGG CGACAAGATGGCGAGCCACATCTCGCGCGCGCTGTT TCTGCCCCCCATCAAACTCGAGTGCGAAAAGACGTTC ACCAAGCTGCTGCTGATCGCCAAGAAAAAGTACATCG GCGTCATCTACGGGGGTAAGATGCTCATCAAGGGCGT GGATCTGGTGCGCAAAAACAACTGCGCGTTTATCAAC CGCACCTCCAGGGCCCTGGTCGACCTGCTGTTTTACG ACGATACCGTCTCCGGAGCCGCCGCGGCGTTAGCCG AGCGCCCCGCGGAGGAGTGGCTGGCGCGACCCCTG CCCGAGGGACTGCAGGCGTTCGGGGCCGTCCTCGTA GACGCCCATCGGCGCATCACCGACCCGGAGAGGGAC ATCCAGGACTTTGTCCTCACCGCCGAACTGAGCAGAC ACCCGCGCGCGTACACCAACAAGCGCCTGGCCCACC TGACGGTGTATTACAAGCTCATGGCCCGCCGCGCGC AGGTCCCGTCCATCAAGGACCGGATCCCGTACGTGAT CGTGGCCCAGACCCGCGAGGTAGAGGAGACGGTCGC GCGGCTGGCCGCCCTCCGCGAGCTAGACGCCGCCG CCCCAGGGGACGAGCCCGCCCCCCCCGCGGCCCTG

		CCCTCCCCGGCCAAGCGCCCCCGGGAGACGCCGTC GCCTGCCGACCCCCCGGGAGGCGCGTCCAAGCCCC GCAAGCTGCTGGTGTCCGAGCTGGCCGAGGATCCCG CATACGCCATTGCCCACGGCGTCGCCCTGAACACGG ACTATTACTTCTCCCACCTGTTGGGGGCGGCGTGCGT GACATTCAAGGCCCTGTTTGGGAATAACGCCAAGATC ACCGAGAGTCTGTTAAAAAGGTTTATTCCCGAAGTGT GGCACCCCCCGGACGACGTGGCCGCGCGGCTCCGG ACCGCAGGGTTCGGGGCGGTGGGTGCCGGCGCTAC GGCGGAGGAAACTCGTCGAATGTTGCATAGAGCCTTT GATACTCTAGCA
30	UL30 DNA Polymerase Catalytic Subunit	MFSGGGGPLSPGGKSAARAASGFFAPAGPRGASRGPP PCLRQNFYNPYLAPVGTQQKPTGPTQRHTYYSECDEFR FIAPRVLDEDAPPEKRAGVHDGHLKRAPKVYCGGDERD VLRVGSGGFWPRRSRLWGGVDHAPAGFNPTVTVFHVY DILENVEHAYGMRAAQFHARFMDAITPTGTVITLLGLTPE GHRVAVHVYGTRQYFYMNKEEVDRHLQCRAPRDLCER MAAALRESPGASFRGISADHFEAEVVERTDVYYYETRPA LFYRVYVRSGRVLSYLCDNFCPAIKKYEGGVDATTRFILD NPGFVTFGWYRLKPGRNNTLAQPRAPMAFGTSSDVEF NCTADNLAIEGGMSDLPAYKLMCFDIECKAGGEDELAFP VAGHPEDLVIQISCLLYDLSTTALEHVLLFSLGSCDLPES HLNELAARGLPTPVVLEFDSEFEMLLAFMTLVKQYGPEF VTGYNIINFDWPFLLAKLTDIYKVPLDGYGRMNGRGVFR VWDIGQSHFQKRSKIKVNGMVNIDMYGIITDKIKLSSYKL NAVAEAVLKDKKKDLSYRDIPAYYAAGPAQRGVIGEYCI QDSLLVGQLFFKFLPHLELSAVARLAGINITRTIYDGQQIR VFTCLLRLADQKGFILPDTQGRFRGAGGEAPKRPAAARE DEERPEEEGEDEDEREEGGGEREPEGARETAGRHVGY QGARVLDPTSGFHVNPVVVFDFASLYPSIIQAHNLCFSTL SLRADAVAHLEAGKDYLEIEVGGRRLFFVKAHVRESLLSI LLRDWLAMRKQIRSRIPQSSPEEAVLLDKQQAAIKVVCN SVYGFTGVQHGLLPCLHVAATVTTIGREMLLATREYVHA RWAAFEQLLADFPEAADMRAPGPYSMRIIYGDTDSIFVL CRGLTAAGLTAVGDKMASHISRALFLPPIKLECEKTFTKL LLIAKKKYIGVIYGGKMLIKGVDLVRKNNCAFINRTSRALV DLLFYDDTVSGAAAALAERPAEEWLARPLPEGLQAFGA VLVDAHRRITDPERDIQDFVLTAELSRHPRAYTNKRLAHL TVYYKLMARRAQVPSIKDRIPYVIVAQTREVEETVARLAA LRELDAAAPGDEPAPPAALPSPAKRPRETPSPADPPGG ASKPRKLLVSELAEDPAYAIAHGVALNTDYYFSHLLGAAC VTFKALFGNNAKITESLLKRFIPEVWHPPDDVAARLRTAG FGAVGAGATAEETRRMLHRAFDTLA
31	UL42 DNA Polymerase Processivity Subunit	ACGGATTCCCCTGGCGGTGTGGCCCCCGCCTCCCCC GTGGAGGACGCGTCGGACGCGTCCCTCGGGCAGCC GGAGGAGGGGGCGCCCTGCCAGGTGGTCCTGCAGG GCGCCGAACTTAATGGAATCCTACAGGCGTTTGCCCC

		GCTGCGCACGAGCCTTCTGGACTCGCTTCTGGTTATG GGCGACCGGGGCATCCTTATCCATAACACGATCTTTG GGGAGCAGGTGTTCCTGCCCCTGGAACACTCGCAATT CAGTCGGTATCGCTGGCGCGGACCCACGGCGGCGTT CCTGTCTCTCGTGGACCAGAAGCGCTCCCTCCTGAGC GTGTTTCGCGCCAACCAGTACCCGGACCTACGTCGG GTGGAGTTGGCGATCACGGGCCAGGCCCCGTTTCGC ACGCTGGTTCAGCGCATATGGACGACGACGTCCGAC GGCGAGGCCGTTGAGCTAGCCAGCGAGACGCTGATG AAGCGCGAACTGACGAGCTTTGTGGTGCTGGTTCCCC AGGGAACCCCCGACGTTCAGTTGCGCCTGACGAGGC CGCAGCTCACCAAGGTCCTTAACGCGACCGGGGCCG ATAGTGCCACGCCCACCACGTTCGAGCTCGGGGTTAA CGGCAAATTTTCCGTGTTCACCACGAGTACCTGCGTC ACCTTTGCTGCCCGCGAGGAGGGCGTGTCGTCCAGC ACCAGCACCCAGGTCCAGATCCTGTCCAACGCGCTCA CCAAGGCGGGCCAGGCCGCCGCGAACGCCAAGACG GTGTACGGGGAAAATACCCATCGCACCTTCTCTGTGG TCGTCGACGATTGCAGCATGCGGGCGGTGCTCCGGC GACTGCAGGTCGGCGGGGGCACCCTCAAGTTCTTCC TCACGACCCCCGTCCCCAGTCTGTGCGTCACCGCCA CCGGTCCCAACGCGGTATCGGCGGTATTTCTCCTGAA ACCCCAGAAGATTTGCCTGGACTGGCTGGGTCATAGC CAGGGGTCTCCTTCAGCCGGGAGCTCGGCCTCCCGG GCCTCTGGGAGCGAGCCAACAGACAGCCAGGACTCC GCGTCGGACGCGGTCAGCCACGGCGATCCGGAAGAC CTCGATGGCGCTGCCCGGGCGGGAGAGGCGGGGGC CTTGCATGCCTGTCCGATGCCGTCGTCGACCACGCG GGTCACTCCCACGACCAAGCGGGGGCGCTCGGGGG GCGAGGATGCGCGCGCGGACACGGCCCTAAAGAAAC CTAAGACGGGGTCGCCCACCGCACCCCCGCCCGCAG ATCCAGTCCCCCTGGACACGGAGGACGACTCCGATG CGGCGGACGGGACGGCGGCCCGTCCCGCCGCTCCA GACGCCCGGAGCGGAAGCCGTTACGCGTGTTACTTT CGCGACCTCCCGACCGGAGAAGCAAGCCCCGGCGCC TTCTCCGCCTTCCGGGGGGGCCCCCAAACCCCGTAT GGTTTTGGATTCCCC
32	UL42 DNA Polymerase Processivity Subunit	TDSPGGVAPASPVEDASDASLGQPEEGAPCQVVLQGA ELNGILQAFAPLRTSLLDSLLVMGDRGILIHNTIFGEQVFL PLEHSQFSRYRWRGPTAAFLSLVDQKRSLLSVFRANQY PDLRRVELAITGQAPFRTLVQRIWTTTSDGEAVELASETL MKRELTSFVVLVPQGTPDVQLRLTRPQLTKVLNATGADS ATPTTFELGVNGKFSVFTTSTCVTFAAREEGVSSSTSTQ VQILSNALTKAGQAAANAKTVYGENTHRTFSVVVDDCS MRAVLRRLQVGGGTLKFFLTTPVPSLCVTATGPNAVSAV FLLKPQKICLDWLGHSQGSPSAGSSASRASGSEPTDSQ DSASDAVSHGDPEDLDGAARAGEAGALHACPMPSSTT

		RVTPTTKRGRSGGEDARADTALKKPKTGSPTAPPPADP VPLDTEDDSDAADGTAARPAAPDARSGSRYACYFRDLP TGEASPGAFSAFRGGPQTPYGFGFP*
33	P2a cleavage site	GCAACAAACTTCTCTCTGCTGAAACAAGCCGGAGATG TCGAAGAGAATCCTGGACCG
34	P2a cleavage site	ATNFSLLKQAGDVEENPGP
35	EF-1alpha core promoter	GGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGT TGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGA GAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCG TGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAG AACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCT TTTTCGCAACGGGTTTGCCGCCAGAACACAG
36	B-globin Poly A signal	AATAAAGGAAATTTATTTTCATTGCAATAGTGTGTTGG AATTTTTTGTGTCTCTCA
37	UL29 (ICP8 – ssDNA BP)	ATGGAGACAAAGCCCAAGACGGCAACCACCATCAAG GTCCCCCCCGGGCCCCTGGGATACGTGTACGCTCGC GCGTGTCCGTCCGAAGGCATCGAGCTTCTGGCGTTAC TGTCGGCACGCAGCGGCGATTCCGACGTCGCCGTGG CGCCCCTGGTCGTGGGCCTGACCGTGGAGAGCGGCT TTGAGGCCAACGTGGCCGTGGTCGTGGGTTCTCGCA CGACGGGGCTCGGGGGTACCGCGGTGTCCCTGAAAC TGACGCCCTCGCACTACAGCTCGTCCGTGTACGTCTT TCACGGCGGCCGGCACCTGGACCCCAGCACCCAGGC CCCGAACCTGACGCGACTTTGCGAGCGGGCACGCCG CCATTTTGGCTTTTCGGACTACACCCCCCGGCCCGGC GACCTCAAACACGAGACGACGGGGGAGGCGCTGTGT GAGCGCCTCGGCCTGGACCCGGACCGCGCCCTCCTG TATCTGGTCGTTACCGAGGGCTTCAAGGAGGCCGTGT GCATCAACAACACCTTTCTGCACCTGGGAGGCTCGGA CAAGGTAACCATAGGCGGGGCGGAGGTGCACCGCAT ACCCGTGTACCCGTTGCAGCTGTTCATGCCGGATTTT AGCCGTGTCATCGCAGAGCCGTTCAACGCCAACCACC GATCGATCGGGGAGAATTTTACCTACCCGCTTCCGTT TTTTAACCGCCCCCTCAACCGCCTCCTGTTCGAGGCG GTCGTGGGACCCGCCGCCGTGGCACTGCGATGCCGA AACGTGGACGCCGTGGCCCGCGCCGCCGCCCACCTG GCGTTTGACGAAAACCACGAGGGCGCCGCCCTCCCC GCCGACATTACGTTCACGGCCTTCGAAGCCAGCCAG GGTAAGACCCCGCGGGGCGGGCGCGACGGCGGCGG CAAGGGCCCGGCGGGCGGGTTCGAACAGCGCCTGG CCTCCGTCATGGCCGGAGACGCCGCCCTGGCCCTCG AGTCTATCGTGTCGATGGCCGTCTTTGACGAGCCGCC

		CACCGACATCTCCGCGTGGCCGCTGTTCGAGGGCCA GGACACGGCCGCGGCCCGCGCCAACGCCGTCGGGG CGTACCTGGCGCGCGCCGCGGGACTCGTGGGGGCC ATGGTATTTAGCACCAACTCGGCCCTCCATCTCACCG AGGTGGACGACGCCGGCCCGGCGGACCCAAAGGAC CACAGCAAACCCTCCTTTTACCGCTTCTTCCTCGTGCC CGGGACCCACGTGGCGGCCAACCCACAGGTGGACCG CGAGGGACACGTGGTGCCCGGGTTCGAGGGTCGGC CCACCGCGCCCCTCGTCGGCGGAACCCAGGAATTTG CCGGCGAGCACCTGGCCATGCTGTGTGGGTTTTCCC CGGCGCTGCTGGCCAAGATGCTGTTTTACCTGGAGC GCTGCGACGGCGGCGTGATCGTCGGGCGCCAGGAG ATGGACGTGTTTCGATACGTCGCGGACTCCAACCAGA CCGACGTGCCCTGTAACCTATGCACCTTCGACACGCG CCACGCCTGCGTACACACGACGCTCATGCGCCTCCG GGCGCGCCATCCAAAGTTCGCCAGCGCCGCCCGCGG AGCCATCGGCGTCTTCGGGACCATGAACAGCATGTAT AGCGACTGCGACGTGCTGGGAAACTACGCCGCCTTC TCGGCCCTGAAGCGCGCGGACGGATCCGAGACCGCC CGGACCATCATGCAGGAGACGTACCGCGCGGCGACC GAGCGCGTCATGGCCGAACTCGAGACCCTGCAGTAC GTGGACCAGGCGGTCCCCACGGCCATGGGGCGGCT GGAGACCATCATCACCAACCGCGAGGCCCTGCATAC GGTGGTGAACAACGTCAGGCAGGTCGTGGACCGCGA GGTGGAGCAGCTGATGCGCAACCTGGTGGAGGGGAG GAACTTCAAGTTTCGCGACGGTCTGGGCGAGGCCAA CCACGCCATGTCCCTGACGCTGGACCCGTACGCGTG CGGGCCGTGCCCCCTGCTTCAGCTTCTCGGGCGGCG ATCCAACCTCGCCGTGTACCAGGACCTGGCCCTGAGT CAGTGCCACGGGGTGTTCGCCGGGCAGTCGGTCGAG GGGCGCAACTTTCGCAATCAATTCCAACCGGTGCTGC GGCGGCGCGTGATGGACATGTTTAACAACGGGTTTCT GTCGGCCAAAACGCTGACGGTCGCGCTCTCGGAGGG GGCGGCTATCTGCGCCCCCAGCCTAACGGCGGGCCA GACGGCCCCCGCCGAGAGCAGCTTCGAGGGCGACGT TGCCCGCGTGACCCTGGGGTTTCCCAAGGAGCTGCG CGTCAAGAGCCGCGTGTTGTTCGCGGGCGCGAGCGC CAACGCGTCCGAGGCCGCCAAGGCGCGGGTCGCCA GCCTCCAGAGCGCCTACCAGAAGCCCGACAAGCGCG TGGACATCCTCCTCGGACCGCTGGGCTTTCTGCTCAA GCAGTTCCACGCGGCCATCTTCCCCAACGGCAAGCC CCCGGGGTCCAACCAGCCGAACCCGCAGTGGTTCTG GACGGCCCTCCAACGCAACCAGCTTCCCGCCCGGCT CCTGTCGCGCGAGGACATCGAGACCATCGCGTTCATT AAAAAGTTTTCCCTGGACTACGGCGCGATAAACTTTAT TAACCTGGCCCCCAACAACGTGAGCGAGCTGGCGAT GTACTACATGGCAAACCAGATTCTGCGGTACTGCGAT

		CACTCGACATACTTCATCAACACCCTTACGGCCATCAT CGCGGGGTCCCGCCGTCCCCCCAGCGTGCAGGCTG CCGCCGCGTGGTCCGCGCAGGGCGGGGCGGGCCTG GAGGCCGGGGCCCGCGCGCTGATGGACGCCGTGGA CGCGCATCCGGGCGCGTGGACGTCCATGTTCGCCAG CTGCAACCTGCTGCGGCCCGTCATGGCGGCGCGCCC CATGGTCGTGTTGGGGTTGAGCATCAGCAAGTACTAC GGCATGGCCGGCAACGACCGTGTGTTTCAGGCCGGG AACTGGGCCAGCCTGATGGGCGGCAAAAACGCGTGC CCGCTCCTTATTTTTGACCGCACCCGCAAGTTCGTCC TGGCCTGTCCCCGGGCCGGGTTTGTGTGCGCGGCCT CAAGCCTCGGCGGCGGAGCGCACGAAAGCTCGCTGT GCGAGCAGCTCCGGGGCATTATCTCCGAGGGCGGGG CGGCCGTCGCCAGTAGCGTGTTCGTGGCGACCGTGA AAAGCCTGGGGCCCCGCACCCAGCAGCTGCAGATCG AGGACTGGCTGGCGCTCCTGGAGGACGAGTACCTAA GCGAGGAGATGATGGAGCTGACCGCGCGTGCCCTGG AGCGCGGCAACGGCGAGTGGTCGACGGACGCGGCC CTGGAGGTGGCGCACGAGGCCGAGGCCCTAGTCAGC CAACTCGGCAACGCCGGGGAGGTGTTTAACTTTGGG GATTTTGGCTGCGAGGACGACAACGCGACGCCGTTC GGCGGCCCGGGGGCCCCGGGACCGGCATTTGCCGG CCGCAAACGGGCGTTCCACGGGGATGACCCGTTTGG GGAGGGGCCCCCCGACAAAAAGGGAGACCTGACGTT GGATATGCTG
38	UL29 (ICP8 – ssDNA BP)	METKPKTATTIKVPPGPLGYVYARACPSEGIELLALLSAR SGDSDVAVAPLVVGLTVESGFEANVAVVVGSRTTGLGG TAVSLKLTPSHYSSSVYVFHGGRHLDPSTQAPNLTRLCE RARRHFGFSDYTPRPGDLKHETTGEALCERLGLDPDRA LLYLVVTEGFKEAVCINNTFLHLGGSDKVTIGGAEVHRIP VYPLQLFMPDFSRVIAEPFNANHRSIGENFTYPLPFFNRP LNRLLFEAVVGPAAVALRCRNVDAVARAAAHLAFDENHE GAALPADITFTAFEASQGKTPRGGRDGGGKGPAGGFEQ RLASVMAGDAALALESIVSMAVFDEPPTDISAWPLFEGQ DTAAARANAVGAYLARAAGLVGAMVFSTNSALHLTEVD DAGPADPKDHSKPSFYRFFLVPGTHVAANPQVDREGHV VPGFEGRPTAPLVGGTQEFAGEHLAMLCGFSPALLAKM LFYLERCDGGVIVGRQEMDVFRYVADSNQTDVPCNLCT FDTRHACVHTTLMRLRARHPKFASAARGAIGVFGTMNS MYSDCDVLGNYAAFSALKRADGSETARTIMQETYRAAT ERVMAELETLQYVDQAVPTAMGRLETIITNREALHTVVN NVRQVVDREVEQLMRNLVEGRNFKFRDGLGEANHAMS LTLDPYACGPCPLLQLLGRRSNLAVYQDLALSQCHGVFA GQSVEGRNFRNQFQPVLRRRVMDMFNNGFLSAKTLTV ALSEGAAICAPSLTAGQTAPAESSFEGDVARVTLGFPKE LRVKSRVLFAGASANASEAAKARVASLQSAYQKPDKRV

		DILLGPLGFLLKQFHAAIFPNGKPPGSNQPNPQWFWTAL QRNQLPARLLSREDIETIAFIKKFSLDYGAINFINLAPNNV SELAMYYMANQILRYCDHSTYFINTLTAIIAGSRRPPSVQ AAAAWSAQGGAGLEAGARALMDAVDAHPGAWTSMFAS CNLLRPVMAARPMVVLGLSISKYYGMAGNDRVFQAGN WASLMGGKNACPLLIFDRTRKFVLACPRAGFVCAASSL GGGAHESSLCEQLRGIISEGGAAVASSVFVATVKSLGPR TQQLQIEDWLALLEDEYLSEEMMELTARALERGNGEWS TDAALEVAHEAEALVSQLGNAGEVFNFGDFGCEDDNAT PFGGPGAPGPAFAGRKRAFHGDDPFGEGPPDKKGDLT LDML
39	HSV TK promoter	AAATGAGTCTTCGGACCTCGCGGGGGCCGCTTAAGC GGTGGTTAGGGTTTGTCTGACGCGGGGGGAGGGGGA AGGAACGAAACACTCTCATTCGGAGGCGGCTCGGGG TTTGGTCTTGGTGGCCACGGGCACGCAGAAGAGCGC CGCGATCCTCTTAAGCACCCCCCCGCCCTCCGTGGA GGCGGGGGTTTGGTCGGCGGGTGGTAACTGGCGGG CCGCTGACTCGGGCGGGTCGCGCGCCCCAGAGTGTG ACCTTTTCGGTCTGCTCGCAGACCCCCGGGCGGCGC CGCCGCGGCGGCGACGGGCTCGCTGGGTCCTAGGC TCCATGGGGACCGTATACGTGGACAGGCTCTGGAGC ATCCGCACGACTGCGGTGATATTACCGGAGACCTTCT GCGGGACGAGCCGGGTCACGCGGCTGACGCGGAGC GTCCGTTGGGCGACAAACACCAGGACGGGGCACAGG TACACTATCTTGTCACCCGGAGGCGCGAGGGACTGCA GGAGCTTCAGGGAGTGGCGCAGCTGCTTCATCCCCG TGGCCCGTTGCTCGCGTTTGCTGGCGGTGTCCCCGG AAGAAATATATTTGCATGTCTTTAGTTCTATGATGACAC AAACCCCGCCCAGCGTCTTGTCATTGGCGAATTCGAA CACGCAGATGCAGTCGGGGCGGCGCGGTCCCAGGTC CACTTCGCATATTAAGGTGACGCGTGTGGCCTCGAAC ACCGAGCGACCCTGCAGCGACCCGCTTAA
40	HSV TK polyA signal	CGGCAATAAAAAGACAGAATAAAACGCACGGTGTTGG GTCGTTTGTTC
41	pEMBR-1.2 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG

		TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA

		GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA

		GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCATGGAG TCAGTCGAGAAGAAGGACAGCCTAACCGCCCCCTCTG AGTTCGCCACCACCGCCTCCACCGATGCCGCCAACG CGCCTACCACCTTCCCCGTCGAGGCACCCCCGCTTG AGGAGGAGGAAGTGATTATCGAGCAGGACCCAGGTTT

		TGTAAGCGAAGACGACGAGGACCGCTCAGTACCAACA GAGGATAAAAAGCAAGACCAGGACAACGCAGAGGCA AACGAGGAACAAGTCGGGCGGGGGGACGAAAGGCAT GGCGACTACCTAGATGTGGGAGACGACGTGCTGTTG AAGCATCTGCAGCGCCAGTGCGCCATTATCTGCGACG CGTTGCAAGAGCGCAGCGATGTGCCCCTCGCCATAG CGGATGTCAGCCTTGCCTACGAACGCCACCTATTCTC ACCGCGCGTACCCCCCAAACGCCAAGAAAACGGCAC ATGCGAGCCCAACCCGCGCCTCAACTTCTACCCCGTA TTTGCCGTGCCAGAGGTGCTTGCCACCTATCACATCT TTTTCCAAAACTGCAAGATACCCCTATCCTGCCGTGCC AACCGCAGCCGAGCGGACAAGCAGCTGGCCTTGCGG CAGGGCGCTGTCATACCTGATATCGCCTCGCTCAACG AAGTGCCAAAAATCTTTGAGGGTCTTGGACGCGACGA GAAGCGCGCGGCAAACGCTCTGCAACAGGAAAACAG CGAAAATGAAAGTCACTCTGGAGTGTTGGTGGAACTC GAGGGTGACAACGCGCGCCTAGCCGTACTAAAACGC AGCATCGAGGTCACCCACTTTGCCTACCCGGCACTTA ACCTACCCCCCCAAGGTCATGAGCACAGTCATGAGTGA GCTGATCGTGCGCCGTGCGCAGCCCCTGGAGAGGGA TGCAAATTTGCAAGAACAAACAGAGGAGGGCCTACCC GCAGTTGGCGACGAGCAGCTAGCGCGCTGGCTTCAA ACGCGCGAGCCTGCCGACTTGGAGGAGCGACGCAAA CTAATGATGGCCGCAGTGCTCGTTACCGTGGAGCTTG AGTGCATGCAGCGGTTCTTTGCTGACCCGGAGATGCA GCGCAAGCTAGAGGAAACATTGCACTACACCTTTCGA CAGGGCTACGTACGCCAGGCCTGCAAGATCTCCAAC GTGGAGCTCTGCAACCTGGTCTCCTACCTTGGAATTT TGCACGAAAACCGCCTTGGGCAAAACGTGCTTCATTC CACGCTCAAGGGCGAGGCGCGCCGCGACTACGTCCG CGACTGCGTTTACTTATTTCTATGCTACACCTGGCAGA CGGCCATGGGCGTTTGGCAGCAGTGCTTGGAGGAGT GCAACCTCAAGGAGCTGCAGAAACTGCTAAAGCAAAA CTTGAAGGACCTATGGACGGCCTTCAACGAGCGCTCC GTGGCCGCGCACCTGGCGGACATCATTTTCCCCGAA CGCCTGCTTAAAACCCTGCAACAGGGTCTGCCAGACT TCACCAGTCAAAGCATGTTGCAGAACTTTAGGAACTTT ATCCTAGAGCGCTCAGGAATCTTGCCCGCCACCTGCT GTGCACTTCCTAGCGACTTTGTGCCCATTAAGTACCG CGAATGCCCTCCGCCGCTTTGGGGCCACTGCTACCTT CTGCAGCTAGCCAACTACCTTGCCTACCACTCTGACA TAATGGAAGACGTGAGCGGTGACGGTCTACTGGAGT GTCACTGTCGCTGCAACCTATGCACCCCGCACCGCTC CCTGGTTTGCAATTCGCAGCTGCTTAACGAAAGTCAA ATTATCGGTACCTTTGAGCTGCAGGGTCCCTCGCCTG ACGAAAAGTCCGCGGCTCCGGGGTTGAAACTCACTCC GGGGCTGTGGACGTCGGCTTACCTTCGCAAATTTGTA

		CCTGAGGACTACCACGCCCACGAGATTAGGTTCTACG AAGACCAATCCCGCCCGCCTAATGCGGAGCTTACCGC CTGCGTCATTACCCAGGGCCACATTCTTGGCCAATTG CAAGCCATCAACAAAGCCCGCCAAGAGTTTCTGCTAC GAAAGGGACGGGGGGTTTACTTGGACCCCCAGTCCG GCGAGGAGCTCAACCCAATCCCCCCGCCGCCGCAGC CCTATCAGCAGCAGCCGCGGGCCCTTGCTTCCCAGG ATGGCACCCAAAAAGAAGCTGCAGCTGCCGCCGCCA CCCACGGACGAGGAGGAATACTGGGACAGTCAGGCA GAGGAGGTTTTGGACGAGGAGGAGGAGGACATGATG GAAGACTGGGAGAGCCTAGACGAGGAAGCTTCCGAG GTCGAAGAGGTGTCAGACGAAACACCGTCACCCTCG GTCGCATTCCCCTCGCCGGCGCCCCAGAAATCGGCA ACCGGTTCCAGCATGGCTACAACCTCCGCTCCTCAGG CGCCGCCGGCACTGCCCGTTCGCCGACCCAACCGTA GATGGGACACCACTGGAACCAGGGCCGGTAAGTCCA AGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGC GCCAAGGCTACCGCTCATGGCGCGGGCACAAGAACG CCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACAT CTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGC GTGGCCTTCCCCCGTAACATCCTGCATTACTACCGTC ATCTCTACAGCCCATACTGCACCGGCGGCAGCGGCA GCAACAGCAGCGGCCACACAGAAGCAAAGGCGACCG GATAGCAAGACTCTGACAAAGCCCAAGAAATCCACAG CGGCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTG GCGCCCAACGAACCCGTATCGACCCGCGAGCTTAGA AACAGGATTTTTCCCACTCTGTATGCTATATTTCAACA GAGCAGGGGCCAAGAACAAGAGCTGAAAATAAAAAAC AGGTCTCTGCGATCCCTCACCCGCAGCTGCCTGTATC ACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAAG ACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCTGAC TCTTAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAAG CGCGAAAACTACGTCATCTCCAGCGGCCACACCCGG CGCCAGCACCTGTTGTCAGCGCCATTATGAGCAAGGA AATTCCCACGCCCTACATGTGGAGTTACCAGCCACAA ATGGGACTTGCGGCTGGAGCTGCCCAAGACTACTCAA CCCGAATAAACTACATGAGCGCGGGGCGGCCGCCGT TTGTGTTATGTTTCAACGTGTTTATTTTTCAATTGCAGA AAATTTCAAGTCATTTTTCATTCAGTAGTATAGCCCCA CCACCACATAGCTTATACAGATCACCGTACCTTAATCA AACTCACAGAACCCTAGTATTCAACCTGCCACCTCCCT CCCAACACACAGAGTACACAGTCCTTTCTCCCCGGCT GGCCTTAAAAAGCATCATATCATGGGTAACAGACATAT TCTTAGGTGTTATATTCCACACGGTTTCCTGTCGAGCC AAACGCTCATCAGTGATATTAATAAACTCCCCGGGCA GCTCACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGC CACAGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGC

		GGCGAAGGAGAAGTCCACGCCTACATGGGGGTAGAG TCATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCA GCAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCG TCCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGC GATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTC CTCCGGGCACAGCAGCGCACCCTGATCTCACTTAAAT CAGCACAGTAACTGCAGCACAGCACCACAATATTGTT CAAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTC ATGGCGGGGACCACAGAACCCACGTGGCCATCATAC CACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAA ACACGCTGGACATAAACATTACCTCTTTTGGCATGTTG TAATTCACCACCTCCCGGTACCATATAAACCTCTGATT AAACATGGCGCCATCCACCACCATCCTAAACCAGCTG GCCAAAACCTGCCCGCCGGCTATACACTGCAGGGAA CCGGGACTGGAACAATGACAGTGGAGAGCCCAGGAC TCGTAACCATGGATCATCATGCTCGTCATGATATCAAT GTTGGCACAACACAGGCACACGTGCATACACTTCCTC AGGATTACAAGCTCCTCCCGCGTTAGAACCATATCCC AGGGAACAACCCATTCCTGAATCAGCGTAAATCCCAC ACTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGC ATTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGAT CCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGG AGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGA CAACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAAT GGAACGCCGGACGTAGTCATATTTCCTGAAGCAAAAC CAGGTGCGGGCGTGACAAACAGATCTGCGTCTCCGG TCTCGCCGCTTAGATCGCTCTGTGTAGTAGTTGTAGTA TATCCACTCTCTCAAAGCATCCAGGCGCCCCCTGGCT TCGGGTTCTATGTAAACTCCTTCATGCGCCGCTGCCC TGATAACATCCACCACCGCAGAATAAGCCACACCCAG CCAACCTACACATTCGTTCTGCGAGTCACACACGGGA GGAGCGGGAAGAGCTGGAAGAACCATGTTTTTTTTTT TATTCCAAAAGATTATCCAAAACCTCAAAATGAAGATC TATTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCA AACTCTACAGCCAAAGAACAGATAATGGCATTTGTAAG ATGTTGCACAATGGCTTCCAAAAGGCAAACGGCCCTC ACGTCCAAGTGGACGTAAAGGCTAAACCCTTCAGGGT GAATCTCCTCTATAAACATTCCAGCACCTTCAACCATG CCCAAATAATTCTCATCTCGCCACCTTCTCAATATATC TCTAAGCAAATCCCGAATATTAAGTCCGGCCATTGTAA AAATCTGCTCCAGAGCGCCCTCCACCTTCAGCCTCAA GCAGCGAATCATGATTGCAAAAATTCAGGTTCCTCACA GACCTGTATAAGATTCAAAAGCGGAACATTAACAAAAA TACCGCGATCCCGTAGGTCCCTTCGCAGGGCCAGCT GAACATAATCGTGCAGGTCTGCACGGACCAGCGCGG CCACTTCCCCGCCAGGAACCATGACAAAAGAACCCAC ACTGATTATGACACGCATACTCGGAGCTATGCTAACC

		AGCGTAGCCCCGATGTAAGCTTGTTGCATGGGCGGC GATATAAAATGCAAGGTGCTGCTCAAAAAATCAGGCA AAGCCTCGCGCAAAAAAGAAAGCACATCGTAGTCATG CTCATGCAGATAAAGGCAGGTAAGCTCCGGAACCACC ACAGAAAAAGACACCATTTTTCTCTCAAACATGTCTGC GGGTTTCTGCATAAACACAAAATAAAATAACAAAAAAA CATTTAAACATTAGAAGCCTGTCTTACAACAGGAAAAA CAACCCTTATAAGCATAAGACGGACTACGGCCATGCC GGCGTGACCGTAAAAAAACTGGTCACCGTGATTAAAA AGCACCACCGACAGCTCCTCGGTCATGTCCGGAGTCA TAATGTAAGACTCGGTAAACACATCAGGTTGATTCACA TCGGTCAGTGCTAAAAAGCGACCGAAATAGCCCGGG GGAATACATACCCGCAGGCGTAGAGACAACATTACAG CCCCCATAGGAGGTATAACAAAATTAATAGGAGAGAA AAACACATAAACACCTGAAAAACCCTCCTGCCTAGGC AAAATAGCACCCTCCCGCTCCAGAACAACATACAGCG CTTCCACAGCGGCAGCCATAACAGTCAGCCTTACCAG TAAAAAAGAAAACCTATTAAAAAAACACCACTCGACAC GGCACCAGCTCAATCAGTCACAGTGTAAAAAAGGGCC AAGTGCAGAGCGAGTATATATAGGACTAAAAAATGAC GTAACGGTTAAAGTCCACAAAAAACACCCAGAAAACC GCACGCGAACCTACGCCCAGAAACGAAAGCCAAAAAA CCCACAACTTCCTCAAATCGTCACTTCCGTTTTCCCAC GTTACGTCACTTCCCATTTTAAGAAAACTACAATTCCC AACACATACAAGTTACTCCGCCCTTAATTAAATCGGAT CCGATATCTAGATGTATTCGCGAGGTACCGAGCTCGA ATTCTCTGGCCGTCGTTTTACAACGTCGTGACTGGGA AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGG CCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCT GAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTT ACGCATCTGTGCGGTATTTCACACCGCATATGGTGCA CTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGAC AAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGA GGTTTTCACCGTCATCACCGAAACGCGCGA
42	pEMBR-1.3 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG

		TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA

		GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA

		GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCCACAAA AGCGAAGATCAGCTTCGGCGCACGCTGGAAGACGCG GAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA AGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGCG AAAACTACGTCATCTCCAGCGGCCACACCCGGCGCCA

		GCACCTGTTGTCAGCGCCATTGGCGCGCCGGCCGGC CGAATATCTTCATTTAAATGTTTAAACATCGATGCGGC CGCCGTTTGTGTTATGTTTCAACGTGTTTATTTTTCAAT TGCAGAAAATTTCAAGTCATTTTTCATTCAGTAGTATAG CCCCACCACCACATAGCTTATACAGATCACCGTACCTT AATCAAACTCACAGAACCCTAGTATTCAACCTGCCACC TCCCTCCCAACACACAGAGTACACAGTCCTTTCTCCC CGGCTGGCCTTAAAAAGCATCATATCATGGGTAACAG ACATATTCTTAGGTGTTATATTCCACACGGTTTCCTGT CGAGCCAAACGCTCATCAGTGATATTAATAAACTCCCC GGGCAGCTCACTTAAGTTCATGTCGCTGTCCAGCTGC TGAGCCACAGGCTGCTGTCCAACTTGCGGTTGCTTAA CGGGCGGCGAAGGAGAAGTCCACGCCTACATGGGGG TAGAGTCATAATCGTGCATCAGGATAGGGCGGTGGTG CTGCAGCAGCGCGCGAATAAACTGCTGCCGCCGCCG CTCCGTCCTGCAGGAATACAACATGGCAGTGGTCTCC TCAGCGATGATTCGCACCGCCCGCAGCATAAGGCGC CTTGTCCTCCGGGCACAGCAGCGCACCCTGATCTCAC TTAAATCAGCACAGTAACTGCAGCACAGCACCACAAT ATTGTTCAAAATCCCACAGTGCAAGGCGCTGTATCCA AAGCTCATGGCGGGGACCACAGAACCCACGTGGCCA TCATACCACAAGCGCAGGTAGATTAAGTGGCGACCCC TCATAAACACGCTGGACATAAACATTACCTCTTTTGGC ATGTTGTAATTCACCACCTCCCGGTACCATATAAACCT CTGATTAAACATGGCGCCATCCACCACCATCCTAAAC CAGCTGGCCAAAACCTGCCCGCCGGCTATACACTGCA GGGAACCGGGACTGGAACAATGACAGTGGAGAGCCC AGGACTCGTAACCATGGATCATCATGCTCGTCATGAT ATCAATGTTGGCACAACACAGGCACACGTGCATACAC TTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAACCA TATCCCAGGGAACAACCCATTCCTGAATCAGCGTAAA TCCCACACTGCAGGGAAGACCTCGCACGTAACTCACG TTGTGCATTGTCAAAGTGTTACATTCGGGCAGCAGCG GATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTCTC AAAAGGAGGTAGACGATCCCTACTGTACGGAGTGCGC CGAGACAACCGAGATCGTGTTGGTCGTAGTGTCATGC CAAATGGAACGCCGGACGTAGTCATATTTCCTGAAGC AAAACCAGGTGCGGGCGTGACAAACAGATCTGCGTCT CCGGTCTCGCCGCTTAGATCGCTCTGTGTAGTAGTTG TAGTATATCCACTCTCTCAAAGCATCCAGGCGCCCCC TGGCTTCGGGTTCTATGTAAACTCCTTCATGCGCCGC TGCCCTGATAACATCCACCACCGCAGAATAAGCCACA CCCAGCCAACCTACACATTCGTTCTGCGAGTCACACA CGGGAGGAGCGGGAAGAGCTGGAAGAACCATGTTTT TTTTTTTATTCCAAAAGATTATCCAAAACCTCAAAATGA AGATCTATTAAGTGAACGCGCTCCCCTCCGGTGGCGT GGTCAAACTCTACAGCCAAAGAACAGATAATGGCATTT

		GTAAGATGTTGCACAATGGCTTCCAAAAGGCAAACGG CCCTCACGTCCAAGTGGACGTAAAGGCTAAACCCTTC AGGGTGAATCTCCTCTATAAACATTCCAGCACCTTCAA CCATGCCCAAATAATTCTCATCTCGCCACCTTCTCAAT ATATCTCTAAGCAAATCCCGAATATTAAGTCCGGCCAT TGTAAAAATCTGCTCCAGAGCGCCCTCCACCTTCAGC CTCAAGCAGCGAATCATGATTGCAAAAATTCAGGTTCC TCACAGACCTGTATAAGATTCAAAAGCGGAACATTAAC AAAAATACCGCGATCCCGTAGGTCCCTTCGCAGGGCC AGCTGAACATAATCGTGCAGGTCTGCACGGACCAGCG CGGCCACTTCCCCGCCAGGAACCATGACAAAAGAACC CACACTGATTATGACACGCATACTCGGAGCTATGCTA ACCAGCGTAGCCCCGATGTAAGCTTGTTGCATGGGCG GCGATATAAAATGCAAGGTGCTGCTCAAAAAATCAGG CAAAGCCTCGCGCAAAAAAGAAAGCACATCGTAGTCA TGCTCATGCAGATAAAGGCAGGTAAGCTCCGGAACCA CCACAGAAAAAGACACCATTTTTCTCTCAAACATGTCT GCGGGTTTCTGCATAAACACAAAATAAAATAACAAAAA AACATTTAAACATTAGAAGCCTGTCTTACAACAGGAAA AACAACCCTTATAAGCATAAGACGGACTACGGCCATG CCGGCGTGACCGTAAAAAAACTGGTCACCGTGATTAA AAAGCACCACCGACAGCTCCTCGGTCATGTCCGGAGT CATAATGTAAGACTCGGTAAACACATCAGGTTGATTCA CATCGGTCAGTGCTAAAAAGCGACCGAAATAGCCCGG GGGAATACATACCCGCAGGCGTAGAGACAACATTACA GCCCCCATAGGAGGTATAACAAAATTAATAGGAGAGA AAAACACATAAACACCTGAAAAACCCTCCTGCCTAGG CAAAATAGCACCCTCCCGCTCCAGAACAACATACAGC GCTTCCACAGCGGCAGCCATAACAGTCAGCCTTACCA GTAAAAAAGAAAACCTATTAAAAAAACACCACTCGACA CGGCACCAGCTCAATCAGTCACAGTGTAAAAAAGGGC CAAGTGCAGAGCGAGTATATATAGGACTAAAAAATGA CGTAACGGTTAAAGTCCACAAAAAACACCCAGAAAAC CGCACGCGAACCTACGCCCAGAAACGAAAGCCAAAAA ACCCACAACTTCCTCAAATCGTCACTTCCGTTTTCCCA CGTTACGTCACTTCCCATTTTAAGAAAACTACAATTCC CAACACATACAAGTTACTCCGCCCTTAATTAAATCGGA TCCGATATCTAGATGTATTCGCGAGGTACCGAGCTCG AATTCTCTGGCCGTCGTTTTACAACGTCGTGACTGGG AAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGC ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAG GCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCC TTACGCATCTGTGCGGTATTTCACACCGCATATGGTG CACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA GCCAGCCCCGACACCCGCCAACACCCGCTGACGCGC CCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAG

		ACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAG AGGTTTTCACCGTCATCACCGAAACGCGCGA
43	pEMBR-1.3B Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA

		TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTATCCG TAGATGTACCTGGACATCCAGGTGATGCCGGCGGCG GTGGTGGAGGCGCGCGGAAAGTCGCGGACGCGGTT CCAGATGTTGCGCAGCGGCAAAAAGTGCTCCATGGTC GGGACGCTCTGGCCGGTGAGGCGTGCGCAGTCGTTG ACGCTCTAGACCGTGCAAAAGGAGAGCCTGTAAGCG GGCACTCTTCCGTGGTCTGGTGGATAAATTCGCAAGG GTATCATGGCGGACGACCGGGGTTCGAACCCCGGAT CCGGCCGTCCGCCGTGATCCATGCGGTTACCGCCCG CGTGTCGAACCCAGGTGTGCGACGTCAGACAACGGG GGAGCGCTCCTTTTGGCTTCCTTCCAGGCGCGGCGG CTGCTGCGCTAGCTTTTTTGGCCACTGGCCGCGCGC GGCGTAAGCGGTTAGGCTGGAAAGCGAAAGCATTAA GTGGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAA GGGTTGAGTCGCAGGACCCCCGGTTCGAGTCTCGGG CCGGCCGGACTGCGGCGAACGGGGGTTTGCCTCCCC GTCATGCAAGACCCCGCTTGCAAATTCCTCCGGAAAC AGGGACGAGCCCCTTTTTTGCTTTTCCCAGATGCATC CGGTGCTGCGGCAGATGCGCCCCCCTCCTCAGCAGC GGCAAGAGCAAGAGCAGCGGCAGACATGCAGGGCAC CCTCCCCTTCTCCTACCGCGTCAGGAGGGGCAACATC TGTACACTCTCGGGTGATTATTTACCCCCACCCTTGCC GTCTGCGCCGTTTAAAAATCAAAGGGGTTCTGCCGCG CATCGCTATGCGCCACTGGCAGGGACACGTTGCGATA CTGGTGTTTAGTGCTCCACTTAAACTCAGGCACAACC ATCCGCGGCAGCTCGGTGAAGTTTTCACTCCACAGGC TGCGCACCATCACCAACGCGTTTAGCAGGTCGGGCG CCGATATCTTGAAGTCGCAGTTGGGGCCTCCGCCCTG CGCGCGCGAGTTGCGATACACAGGGTTGCAGCACTG GAACACTATCAGCGCCGGGTGGTGCACGCTGGCCAG CACGCTCTTGTCGGAGATCAGATCCGCGTCCAGGTCC TCCGCGTTGCTCAGGGCGAACGGAGTCAACTTTGGTA GCTGCCTTCCCAAAAAGGGCGCGTGCCCAGGCTTTG AGTTGCACTCGCACCGTAGTGGCATCAAAAGGTGACC GTGCCCGGTCTGGGCGTTAGGATACAGCGCCTGCAT AAAAGCCTTGATCTGCTTAAAAGCCACCTGAGCCTTTG

		CGCCTTCAGAGAAGAACATGCCGCAAGACTTGCCGGA AAACTGATTGGCCGGACAGGCCGCGTCGTGCACGCA GCACCTTGCGTCGGTGTTGGAGATCTGCACCACATTT CGGCCCCACCGGTTCTTCACGATCTTGGCCTTGCTAG ACTGCTCCTTCAGCGCGCGCTGCCCGTTTTCGCTCGT CACATCCATTTCAATCACGTGCTCCTTATTTATCATAAT GCTTCCGTGTAGACACTTAAGCTCGCCTTCGATCTCA GCGCAGCGGTGCAGCCACAACGCGCAGCCCGTGGG CTCGTGATGCTTGTAGGTCACCTCTGCAAACGACTGC AGGTACGCCTGCAGGAATCGCCCCATCATCGTCACAA AGGTCTTGTTGCTGGTGAAGGTCAGCTGCAACCCGCG GTGCTCCTCGTTCAGCCAGGTCTTGCATACGGCCGCC AGAGCTTCCACTTGGTCAGGCAGTAGTTTGAAGTTCG CCTTTAGATCGTTATCCACGTGGTACTTGTCCATCAGC GCGCGCGCAGCCTCCATGCCCTTCTCCCACGCAGAC ACGATCGGCACACTCAGCGGGTTCATCACCGTAATTT CACTTTCCGCTTCGCTGGGCTCTTCCTCTTCCTCTTGC GTCCGCATACCACGCGCCACTGGGTCGTCTTCATTCA GCCGCCGCACTGTGCGCTTACCTCCTTTGCCATGCTT GATTAGCACCGGTGGGTTGCTGAAACCCACCATTTGT AGCGCCACATCTTCTCTTTCTTCCTCGCTGTCCACGAT TACCTCTGGTGATGGCGGGCGCTCGGGCTTGGGAGA AGGGCGCTTCTTTTTCTTCTTGGGCGCAATGGCCAAA TCCGCCGCCGAGGTCGATGGCCGCGGGCTGGGTGT GCGCGGCACCAGCGCGTCTTGTGATGAGTCTTCCTC GTCCTCGGACTCGATACGCCGCCTCATCCGCTTTTTT GGGGGCGCCCGGGGAGGCGGCGGCGACGGGGACG GGGACGACACGTCCTCCATGGTTGGGGGACGTCGCG CCGCACCGCGTCCGCGCTCGGGGGTGGTTTCGCGCT GCTCCTCTTCCCGACTGGCCATTTCCTTCTCCTATAGG CAGAAAAAGATCCACAAAAGCGAAGATCAGCTTCGGC GCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAATA CTGCGCGCTGACTCTTAAGGACTAGTTTCGCGCCCTT TCTCAAATTTAAGCGCGAAAACTACGTCATCTCCAGCG GCCACACCCGGCGCCAGCACCTGTTGTCAGCGCCAT TGGCGCGCCGGCCGGCCGAATATCTTCATTTAAATGT TTAAACATCGATGCGGCCGCCGTTTGTGTTATGTTTCA ACGTGTTTATTTTTCAATTGCAGAAAATTTCAAGTCATT TTTCATTCAGTAGTATAGCCCCACCACCACATAGCTTA TACAGATCACCGTACCTTAATCAAACTCACAGAACCCT AGTATTCAACCTGCCACCTCCCTCCCAACACACAGAG TACACAGTCCTTTCTCCCCGGCTGGCCTTAAAAAGCA TCATATCATGGGTAACAGACATATTCTTAGGTGTTATA TTCCACACGGTTTCCTGTCGAGCCAAACGCTCATCAG TGATATTAATAAACTCCCCGGGCAGCTCACTTAAGTTC ATGTCGCTGTCCAGCTGCTGAGCCACAGGCTGCTGTC CAACTTGCGGTTGCTTAACGGGCGGCGAAGGAGAAG

		TCCACGCCTACATGGGGGTAGAGTCATAATCGTGCAT CAGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGAAT AAACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAATA CAACATGGCAGTGGTCTCCTCAGCGATGATTCGCACC GCCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACAG CAGCGCACCCTGATCTCACTTAAATCAGCACAGTAAC TGCAGCACAGCACCACAATATTGTTCAAAATCCCACA GTGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGAC CACAGAACCCACGTGGCCATCATACCACAAGCGCAG GTAGATTAAGTGGCGACCCCTCATAAACACGCTGGAC ATAAACATTACCTCTTTTGGCATGTTGTAATTCACCAC CTCCCGGTACCATATAAACCTCTGATTAAACATGGCG CCATCCACCACCATCCTAAACCAGCTGGCCAAAACCT GCCCGCCGGCTATACACTGCAGGGAACCGGGACTGG AACAATGACAGTGGAGAGCCCAGGACTCGTAACCATG GATCATCATGCTCGTCATGATATCAATGTTGGCACAAC ACAGGCACACGTGCATACACTTCCTCAGGATTACAAG CTCCTCCCGCGTTAGAACCATATCCCAGGGAACAACC CATTCCTGAATCAGCGTAAATCCCACACTGCAGGGAA GACCTCGCACGTAACTCACGTTGTGCATTGTCAAAGT GTTACATTCGGGCAGCAGCGGATGATCCTCCAGTATG GTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGAT CCCTACTGTACGGAGTGCGCCGAGACAACCGAGATC GTGTTGGTCGTAGTGTCATGCCAAATGGAACGCCGGA CGTAGTCATATTTCCTGAAGCAAAACCAGGTGCGGGC GTGACAAACAGATCTGCGTCTCCGGTCTCGCCGCTTA GATCGCTCTGTGTAGTAGTTGTAGTATATCCACTCTCT CAAAGCATCCAGGCGCCCCCTGGCTTCGGGTTCTATG TAAACTCCTTCATGCGCCGCTGCCCTGATAACATCCA CCACCGCAGAATAAGCCACACCCAGCCAACCTACACA TTCGTTCTGCGAGTCACACACGGGAGGAGCGGGAAG AGCTGGAAGAACCATGTTTTTTTTTTTATTCCAAAAGAT TATCCAAAACCTCAAAATGAAGATCTATTAAGTGAACG CGCTCCCCTCCGGTGGCGTGGTCAAACTCTACAGCCA AAGAACAGATAATGGCATTTGTAAGATGTTGCACAATG GCTTCCAAAAGGCAAACGGCCCTCACGTCCAAGTGGA CGTAAAGGCTAAACCCTTCAGGGTGAATCTCCTCTATA AACATTCCAGCACCTTCAACCATGCCCAAATAATTCTC ATCTCGCCACCTTCTCAATATATCTCTAAGCAAATCCC GAATATTAAGTCCGGCCATTGTAAAAATCTGCTCCAGA GCGCCCTCCACCTTCAGCCTCAAGCAGCGAATCATGA TTGCAAAAATTCAGGTTCCTCACAGACCTGTATAAGAT TCAAAAGCGGAACATTAACAAAAATACCGCGATCCCG TAGGTCCCTTCGCAGGGCCAGCTGAACATAATCGTGC AGGTCTGCACGGACCAGCGCGGCCACTTCCCCGCCA GGAACCATGACAAAAGAACCCACACTGATTATGACAC GCATACTCGGAGCTATGCTAACCAGCGTAGCCCCGAT

		GTAAGCTTGTTGCATGGGCGGCGATATAAAATGCAAG GTGCTGCTCAAAAAATCAGGCAAAGCCTCGCGCAAAA AAGAAAGCACATCGTAGTCATGCTCATGCAGATAAAG GCAGGTAAGCTCCGGAACCACCACAGAAAAAGACACC ATTTTTCTCTCAAACATGTCTGCGGGTTTCTGCATAAA CACAAAATAAAATAACAAAAAAACATTTAAACATTAGAA GCCTGTCTTACAACAGGAAAAACAACCCTTATAAGCAT AAGACGGACTACGGCCATGCCGGCGTGACCGTAAAA AAACTGGTCACCGTGATTAAAAAGCACCACCGACAGC TCCTCGGTCATGTCCGGAGTCATAATGTAAGACTCGG TAAACACATCAGGTTGATTCACATCGGTCAGTGCTAAA AAGCGACCGAAATAGCCCGGGGGAATACATACCCGC AGGCGTAGAGACAACATTACAGCCCCCATAGGAGGTA TAACAAAATTAATAGGAGAGAAAAACACATAAACACCT GAAAAACCCTCCTGCCTAGGCAAAATAGCACCCTCCC GCTCCAGAACAACATACAGCGCTTCCACAGCGGCAGC CATAACAGTCAGCCTTACCAGTAAAAAAGAAAACCTAT TAAAAAAACACCACTCGACACGGCACCAGCTCAATCA GTCACAGTGTAAAAAAGGGCCAAGTGCAGAGCGAGTA TATATAGGACTAAAAAATGACGTAACGGTTAAAGTCCA CAAAAAACACCCAGAAAACCGCACGCGAACCTACGCC CAGAAACGAAAGCCAAAAAACCCACAACTTCCTCAAAT CGTCACTTCCGTTTTCCCACGTTACGTCACTTCCCATT TTAAGAAAACTACAATTCCCAACACATACAAGTTACTC CGCCCTTAATTAAATCGGATCCGATATCTAGATGTATT CGCGAGGTACCGAGCTCGAATTCTCTGGCCGTCGTTT TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCT TCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGC CTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTAT TTCACACCGCATATGGTGCACTCTCAGTACAATCTGCT CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGC CAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGC TCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCA CCGAAACGCGCGA
44	pEMBR-1.4 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG

		TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA

		GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA

		GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCCACAAA AGCGAAGATCAGCTTCGGCGCACGCTGGAAGACGCG GAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA AGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGCG AAAACTACGTCATCTCCAGCGGCCACACCCGGCGCCA

		GCACCTGTTGTCAGCGCCATTGGCGCGCCCGCCCGC CGCGCGCTTCGCTTTTTATAGGGCCGCCGCCGCCGC CGCCTCGCCATAAAAGGAAACTTTCGGAGCGCGCCG CTCTGATTGGCTGCCGCCGCACCTCTCCGCCTCGCC CCGCCCCGCCCCTCGCCCCGCCCCGCCCCGCCTGG CGCGCGCCCCCCCCCCCCCCCCCGCCCCCATCGCTGC ACAAAATAATTAAAAAATAAATAAATACAAAATTGGGG GTGGGGAGGGGGGGGAGATGGGGAGAGTGAAGCAG AACGTGGGGCTCACCTCGAGGCCGGCCGAATATCTT CATTTAAATGTTTAAACATCGATGCGGCCGCCGTTTGT GTTATGTTTCAACGTGTTTATTTTTCAATTGCAGAAAAT TTCAAGTCATTTTTCATTCAGTAGTATAGCCCCACCAC CACATAGCTTATACAGATCACCGTACCTTAATCAAACT CACAGAACCCTAGTATTCAACCTGCCACCTCCCTCCC AACACACAGAGTACACAGTCCTTTCTCCCCGGCTGGC CTTAAAAAGCATCATATCATGGGTAACAGACATATTCT TAGGTGTTATATTCCACACGGTTTCCTGTCGAGCCAAA CGCTCATCAGTGATATTAATAAACTCCCCGGGCAGCT CACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGCCAC AGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGCGG CGAAGGAGAAGTCCACGCCTACATGGGGGTAGAGTC ATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCAG CAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCGT CCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCG ATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTCC TCCGGGCACAGCAGCGCACCCTGATCTCACTTAAATC AGCACAGTAACTGCAGCACAGCACCACAATATTGTTC AAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTCA TGGCGGGGACCACAGAACCCACGTGGCCATCATACC ACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAAA CACGCTGGACATAAACATTACCTCTTTTGGCATGTTGT AATTCACCACCTCCCGGTACCATATAAACCTCTGATTA AACATGGCGCCATCCACCACCATCCTAAACCAGCTGG CCAAAACCTGCCCGCCGGCTATACACTGCAGGGAAC CGGGACTGGAACAATGACAGTGGAGAGCCCAGGACT CGTAACCATGGATCATCATGCTCGTCATGATATCAATG TTGGCACAACACAGGCACACGTGCATACACTTCCTCA GGATTACAAGCTCCTCCCGCGTTAGAACCATATCCCA GGGAACAACCCATTCCTGAATCAGCGTAAATCCCACA CTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCA TTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGATC CTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGA GGTAGACGATCCCTACTGTACGGAGTGCGCCGAGAC AACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAATG GAACGCCGGACGTAGTCATATTTCCTGAAGCAAAACC AGGTGCGGGCGTGACAAACAGATCTGCGTCTCCGGT CTCGCCGCTTAGATCGCTCTGTGTAGTAGTTGTAGTAT

		ATCCACTCTCTCAAAGCATCCAGGCGCCCCCTGGCTT CGGGTTCTATGTAAACTCCTTCATGCGCCGCTGCCCT GATAACATCCACCACCGCAGAATAAGCCACACCCAGC CAACCTACACATTCGTTCTGCGAGTCACACACGGGAG GAGCGGGAAGAGCTGGAAGAACCATGTTTTTTTTTTTA TTCCAAAAGATTATCCAAAACCTCAAAATGAAGATCTA TTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCAAA CTCTACAGCCAAAGAACAGATAATGGCATTTGTAAGAT GTTGCACAATGGCTTCCAAAAGGCAAACGGCCCTCAC GTCCAAGTGGACGTAAAGGCTAAACCCTTCAGGGTGA ATCTCCTCTATAAACATTCCAGCACCTTCAACCATGCC CAAATAATTCTCATCTCGCCACCTTCTCAATATATCTCT AAGCAAATCCCGAATATTAAGTCCGGCCATTGTAAAAA TCTGCTCCAGAGCGCCCTCCACCTTCAGCCTCAAGCA GCGAATCATGATTGCAAAAATTCAGGTTCCTCACAGAC CTGTATAAGATTCAAAAGCGGAACATTAACAAAAATAC CGCGATCCCGTAGGTCCCTTCGCAGGGCCAGCTGAA CATAATCGTGCAGGTCTGCACGGACCAGCGCGGCCA CTTCCCCGCCAGGAACCATGACAAAAGAACCCACACT GATTATGACACGCATACTCGGAGCTATGCTAACCAGC GTAGCCCCGATGTAAGCTTGTTGCATGGGCGGCGATA TAAAATGCAAGGTGCTGCTCAAAAAATCAGGCAAAGC CTCGCGCAAAAAAGAAAGCACATCGTAGTCATGCTCA TGCAGATAAAGGCAGGTAAGCTCCGGAACCACCACAG AAAAAGACACCATTTTTCTCTCAAACATGTCTGCGGGT TTCTGCATAAACACAAAATAAAATAACAAAAAAACATTT AAACATTAGAAGCCTGTCTTACAACAGGAAAAACAACC CTTATAAGCATAAGACGGACTACGGCCATGCCGGCGT GACCGTAAAAAAACTGGTCACCGTGATTAAAAAGCAC CACCGACAGCTCCTCGGTCATGTCCGGAGTCATAATG TAAGACTCGGTAAACACATCAGGTTGATTCACATCGGT CAGTGCTAAAAAGCGACCGAAATAGCCCGGGGGAATA CATACCCGCAGGCGTAGAGACAACATTACAGCCCCCA TAGGAGGTATAACAAAATTAATAGGAGAGAAAAACACA TAAACACCTGAAAAACCCTCCTGCCTAGGCAAAATAG CACCCTCCCGCTCCAGAACAACATACAGCGCTTCCAC AGCGGCAGCCATAACAGTCAGCCTTACCAGTAAAAAA GAAAACCTATTAAAAAAACACCACTCGACACGGCACC AGCTCAATCAGTCACAGTGTAAAAAAGGGCCAAGTGC AGAGCGAGTATATATAGGACTAAAAAATGACGTAACG GTTAAAGTCCACAAAAAACACCCAGAAAACCGCACGC GAACCTACGCCCAGAAACGAAAGCCAAAAAACCCACA ACTTCCTCAAATCGTCACTTCCGTTTTCCCACGTTACG TCACTTCCCATTTTAAGAAAACTACAATTCCCAACACA TACAAGTTACTCCGCCCTTAATTAAATCGGATCCGATA TCTAGATGTATTCGCGAGGTACCGAGCTCGAATTCTC TGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC

		TGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCC CCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGC ACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG GCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCA TCTGTGCGGTATTTCACACCGCATATGGTGCACTCTC AGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGC CCCGACACCCGCCAACACCCGCTGACGCGCCCTGAC GGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGC TGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTT TCACCGTCATCACCGAAACGCGCGA
45	pEMBR-1.4B Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT

		TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTATCCG TAGATGTACCTGGACATCCAGGTGATGCCGGCGGCG GTGGTGGAGGCGCGCGGAAAGTCGCGGACGCGGTT CCAGATGTTGCGCAGCGGCAAAAAGTGCTCCATGGTC GGGACGCTCTGGCCGGTGAGGCGTGCGCAGTCGTTG ACGCTCTAGACCGTGCAAAAGGAGAGCCTGTAAGCG GGCACTCTTCCGTGGTCTGGTGGATAAATTCGCAAGG GTATCATGGCGGACGACCGGGGTTCGAACCCCGGAT CCGGCCGTCCGCCGTGATCCATGCGGTTACCGCCCG CGTGTCGAACCCAGGTGTGCGACGTCAGACAACGGG GGAGCGCTCCTTTTGGCTTCCTTCCAGGCGCGGCGG CTGCTGCGCTAGCTTTTTTGGCCACTGGCCGCGCGC GGCGTAAGCGGTTAGGCTGGAAAGCGAAAGCATTAA GTGGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAA GGGTTGAGTCGCAGGACCCCCGGTTCGAGTCTCGGG CCGGCCGGACTGCGGCGAACGGGGGTTTGCCTCCCC GTCATGCAAGACCCCGCTTGCAAATTCCTCCGGAAAC AGGGACGAGCCCCTTTTTTGCTTTTCCCAGATGCATC CGGTGCTGCGGCAGATGCGCCCCCCTCCTCAGCAGC GGCAAGAGCAAGAGCAGCGGCAGACATGCAGGGCAC CCTCCCCTTCTCCTACCGCGTCAGGAGGGGCAACATC TGTACACTCTCGGGTGATTATTTACCCCCACCCTTGCC GTCTGCGCCGTTTAAAAATCAAAGGGGTTCTGCCGCG CATCGCTATGCGCCACTGGCAGGGACACGTTGCGATA CTGGTGTTTAGTGCTCCACTTAAACTCAGGCACAACC ATCCGCGGCAGCTCGGTGAAGTTTTCACTCCACAGGC TGCGCACCATCACCAACGCGTTTAGCAGGTCGGGCG CCGATATCTTGAAGTCGCAGTTGGGGCCTCCGCCCTG

		CGCGCGCGAGTTGCGATACACAGGGTTGCAGCACTG GAACACTATCAGCGCCGGGTGGTGCACGCTGGCCAG CACGCTCTTGTCGGAGATCAGATCCGCGTCCAGGTCC TCCGCGTTGCTCAGGGCGAACGGAGTCAACTTTGGTA GCTGCCTTCCCAAAAAGGGCGCGTGCCCAGGCTTTG AGTTGCACTCGCACCGTAGTGGCATCAAAAGGTGACC GTGCCCGGTCTGGGCGTTAGGATACAGCGCCTGCAT AAAAGCCTTGATCTGCTTAAAAGCCACCTGAGCCTTTG CGCCTTCAGAGAAGAACATGCCGCAAGACTTGCCGGA AAACTGATTGGCCGGACAGGCCGCGTCGTGCACGCA GCACCTTGCGTCGGTGTTGGAGATCTGCACCACATTT CGGCCCCACCGGTTCTTCACGATCTTGGCCTTGCTAG ACTGCTCCTTCAGCGCGCGCTGCCCGTTTTCGCTCGT CACATCCATTTCAATCACGTGCTCCTTATTTATCATAAT GCTTCCGTGTAGACACTTAAGCTCGCCTTCGATCTCA GCGCAGCGGTGCAGCCACAACGCGCAGCCCGTGGG CTCGTGATGCTTGTAGGTCACCTCTGCAAACGACTGC AGGTACGCCTGCAGGAATCGCCCCATCATCGTCACAA AGGTCTTGTTGCTGGTGAAGGTCAGCTGCAACCCGCG GTGCTCCTCGTTCAGCCAGGTCTTGCATACGGCCGCC AGAGCTTCCACTTGGTCAGGCAGTAGTTTGAAGTTCG CCTTTAGATCGTTATCCACGTGGTACTTGTCCATCAGC GCGCGCGCAGCCTCCATGCCCTTCTCCCACGCAGAC ACGATCGGCACACTCAGCGGGTTCATCACCGTAATTT CACTTTCCGCTTCGCTGGGCTCTTCCTCTTCCTCTTGC GTCCGCATACCACGCGCCACTGGGTCGTCTTCATTCA GCCGCCGCACTGTGCGCTTACCTCCTTTGCCATGCTT GATTAGCACCGGTGGGTTGCTGAAACCCACCATTTGT AGCGCCACATCTTCTCTTTCTTCCTCGCTGTCCACGAT TACCTCTGGTGATGGCGGGCGCTCGGGCTTGGGAGA AGGGCGCTTCTTTTTCTTCTTGGGCGCAATGGCCAAA TCCGCCGCCGAGGTCGATGGCCGCGGGCTGGGTGT GCGCGGCACCAGCGCGTCTTGTGATGAGTCTTCCTC GTCCTCGGACTCGATACGCCGCCTCATCCGCTTTTTT GGGGGCGCCCGGGGAGGCGGCGGCGACGGGGACG GGGACGACACGTCCTCCATGGTTGGGGGACGTCGCG CCGCACCGCGTCCGCGCTCGGGGGTGGTTTCGCGCT GCTCCTCTTCCCGACTGGCCATTTCCTTCTCCTATAGG CAGAAAAAGATCCACAAAAGCGAAGATCAGCTTCGGC GCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAATA CTGCGCGCTGACTCTTAAGGACTAGTTTCGCGCCCTT TCTCAAATTTAAGCGCGAAAACTACGTCATCTCCAGCG GCCACACCCGGCGCCAGCACCTGTTGTCAGCGCCAT TGGCGCGCCCGCCCGCCGCGCGCTTCGCTTTTTATA GGGCCGCCGCCGCCGCCGCCTCGCCATAAAAGGAAA CTTTCGGAGCGCGCCGCTCTGATTGGCTGCCGCCGC ACCTCTCCGCCTCGCCCCGCCCCGCCCCTCGCCCCG

		CCCCGCCCCGCCTGGCGCGCGCCCCCCCCCCCCC CCGCCCCCATCGCTGCACAAAATAATTAAAAAATAAAT AAATACAAAATTGGGGGTGGGGAGGGGGGGGAGATG GGGAGAGTGAAGCAGAACGTGGGGCTCACCTCGAGG CCGGCCGAATATCTTCATTTAAATGTTTAAACATCGAT GCGGCCGCCGTTTGTGTTATGTTTCAACGTGTTTATTT TTCAATTGCAGAAAATTTCAAGTCATTTTTCATTCAGTA GTATAGCCCCACCACCACATAGCTTATACAGATCACC GTACCTTAATCAAACTCACAGAACCCTAGTATTCAACC TGCCACCTCCCTCCCAACACACAGAGTACACAGTCCT TTCTCCCCGGCTGGCCTTAAAAAGCATCATATCATGG GTAACAGACATATTCTTAGGTGTTATATTCCACACGGT TTCCTGTCGAGCCAAACGCTCATCAGTGATATTAATAA ACTCCCCGGGCAGCTCACTTAAGTTCATGTCGCTGTC CAGCTGCTGAGCCACAGGCTGCTGTCCAACTTGCGGT TGCTTAACGGGCGGCGAAGGAGAAGTCCACGCCTAC ATGGGGGTAGAGTCATAATCGTGCATCAGGATAGGGC GGTGGTGCTGCAGCAGCGCGCGAATAAACTGCTGCC GCCGCCGCTCCGTCCTGCAGGAATACAACATGGCAG TGGTCTCCTCAGCGATGATTCGCACCGCCCGCAGCAT AAGGCGCCTTGTCCTCCGGGCACAGCAGCGCACCCT GATCTCACTTAAATCAGCACAGTAACTGCAGCACAGC ACCACAATATTGTTCAAAATCCCACAGTGCAAGGCGC TGTATCCAAAGCTCATGGCGGGGACCACAGAACCCAC GTGGCCATCATACCACAAGCGCAGGTAGATTAAGTGG CGACCCCTCATAAACACGCTGGACATAAACATTACCT CTTTTGGCATGTTGTAATTCACCACCTCCCGGTACCAT ATAAACCTCTGATTAAACATGGCGCCATCCACCACCAT CCTAAACCAGCTGGCCAAAACCTGCCCGCCGGCTATA CACTGCAGGGAACCGGGACTGGAACAATGACAGTGG AGAGCCCAGGACTCGTAACCATGGATCATCATGCTCG TCATGATATCAATGTTGGCACAACACAGGCACACGTG CATACACTTCCTCAGGATTACAAGCTCCTCCCGCGTTA GAACCATATCCCAGGGAACAACCCATTCCTGAATCAG CGTAAATCCCACACTGCAGGGAAGACCTCGCACGTAA CTCACGTTGTGCATTGTCAAAGTGTTACATTCGGGCA GCAGCGGATGATCCTCCAGTATGGTAGCGCGGGTTTC TGTCTCAAAAGGAGGTAGACGATCCCTACTGTACGGA GTGCGCCGAGACAACCGAGATCGTGTTGGTCGTAGT GTCATGCCAAATGGAACGCCGGACGTAGTCATATTTC CTGAAGCAAAACCAGGTGCGGGCGTGACAAACAGAT CTGCGTCTCCGGTCTCGCCGCTTAGATCGCTCTGTGT AGTAGTTGTAGTATATCCACTCTCTCAAAGCATCCAGG CGCCCCCTGGCTTCGGGTTCTATGTAAACTCCTTCAT GCGCCGCTGCCCTGATAACATCCACCACCGCAGAATA AGCCACACCCAGCCAACCTACACATTCGTTCTGCGAG TCACACACGGGAGGAGCGGGAAGAGCTGGAAGAACC

		ATGTTTTTTTTTTTATTCCAAAAGATTATCCAAAACCTC AAAATGAAGATCTATTAAGTGAACGCGCTCCCCTCCG GTGGCGTGGTCAAACTCTACAGCCAAAGAACAGATAA TGGCATTTGTAAGATGTTGCACAATGGCTTCCAAAAG GCAAACGGCCCTCACGTCCAAGTGGACGTAAAGGCTA AACCCTTCAGGGTGAATCTCCTCTATAAACATTCCAGC ACCTTCAACCATGCCCAAATAATTCTCATCTCGCCACC TTCTCAATATATCTCTAAGCAAATCCCGAATATTAAGTC CGGCCATTGTAAAAATCTGCTCCAGAGCGCCCTCCAC CTTCAGCCTCAAGCAGCGAATCATGATTGCAAAAATTC AGGTTCCTCACAGACCTGTATAAGATTCAAAAGCGGA ACATTAACAAAAATACCGCGATCCCGTAGGTCCCTTC GCAGGGCCAGCTGAACATAATCGTGCAGGTCTGCAC GGACCAGCGCGGCCACTTCCCCGCCAGGAACCATGA CAAAAGAACCCACACTGATTATGACACGCATACTCGG AGCTATGCTAACCAGCGTAGCCCCGATGTAAGCTTGT TGCATGGGCGGCGATATAAAATGCAAGGTGCTGCTCA AAAAATCAGGCAAAGCCTCGCGCAAAAAAGAAAGCAC ATCGTAGTCATGCTCATGCAGATAAAGGCAGGTAAGC TCCGGAACCACCACAGAAAAAGACACCATTTTTCTCTC AAACATGTCTGCGGGTTTCTGCATAAACACAAAATAAA ATAACAAAAAAACATTTAAACATTAGAAGCCTGTCTTA CAACAGGAAAAACAACCCTTATAAGCATAAGACGGAC TACGGCCATGCCGGCGTGACCGTAAAAAAACTGGTCA CCGTGATTAAAAAGCACCACCGACAGCTCCTCGGTCA TGTCCGGAGTCATAATGTAAGACTCGGTAAACACATC AGGTTGATTCACATCGGTCAGTGCTAAAAAGCGACCG AAATAGCCCGGGGGAATACATACCCGCAGGCGTAGA GACAACATTACAGCCCCCATAGGAGGTATAACAAAATT AATAGGAGAGAAAAACACATAAACACCTGAAAAACCCT CCTGCCTAGGCAAAATAGCACCCTCCCGCTCCAGAAC AACATACAGCGCTTCCACAGCGGCAGCCATAACAGTC AGCCTTACCAGTAAAAAAGAAAACCTATTAAAAAAACA CCACTCGACACGGCACCAGCTCAATCAGTCACAGTGT AAAAAAGGGCCAAGTGCAGAGCGAGTATATATAGGAC TAAAAAATGACGTAACGGTTAAAGTCCACAAAAAACAC CCAGAAAACCGCACGCGAACCTACGCCCAGAAACGA AAGCCAAAAAACCCACAACTTCCTCAAATCGTCACTTC CGTTTTCCCACGTTACGTCACTTCCCATTTTAAGAAAA CTACAATTCCCAACACATACAAGTTACTCCGCCCTTAA TTAAATCGGATCCGATATCTAGATGTATTCGCGAGGTA CCGAGCTCGAATTCTCTGGCCGTCGTTTTACAACGTC GTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCG CCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAT AGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAG TTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGG TATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCG

		CATATGGTGCACTCTCAGTACAATCTGCTCTGATGCC GCATAGTTAAGCCAGCCCCGACACCCGCCAACACCC GCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCA TCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT GCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACG CGCGA
46	pEMBR-1.4B2 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA

		AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTATCCG TAGATGTACCTGGACATCCAGGTGATGCCGGCGGCG GTGGTGGAGGCGCGCGGAAAGTCGCGGACGCGGTT CCAGATGTTGCGCAGCGGCAAAAAGTGCTCCATGGTC GGGACGCTCTGGCCGGTGAGGCGTGCGCAGTCGTTG ACGCTCTAGACCGTGCAAAAGGAGAGCCTGTAAGCG GGCACTCTTCCGTGGTCTGGTGGATAAATTCGCAAGG GTATCATGGCGGACGACCGGGGTTCGAACCCCGGAT CCGGCCGTCCGCCGTGATCCATGCGGTTACCGCCCG CGTGTCGAACCCAGGTGTGCGACGTCAGACAACGGG GGAGCGCTCCTTTTGGCTTCCTTCCAGGCGCGGCGG CTGCTGCGCTAGCTTTTTTGGCCACTGGCCGCGCGC GGCGTAAGCGGTTAGGCTGGAAAGCGAAAGCATTAA GTGGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAA GGGTTGAGTCGCAGGACCCCCGGTTCGAGTCTCGGG CCGGCCGGACTGCGGCGAACGGGGGTTTGCCTCCCC GTCATGCAAGACCCCGCTTGCAAATTCCTCCGGAAAC AGGGACGAGCCCCTTTTTTGCTTTTCCCAGATGCATC CGGTGCTGCGGCAGATGCGCCCCCCTCCTCAGCAGC GGCAAGAGCAAGAGCAGCGGCAGACATGCAGGGCAC CCTCCCCTTCTCCTACCGCGTCAGGAGGGGCAACATC GATCCAGACATGATAAGATACATTGATGAGTTTGGACA AACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTT GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA TAAGCTGCAATAAACAAGTTTGTACACTCTCGGGTGAT TATTTACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAA TCAAAGGGGTTCTGCCGCGCATCGCTATGCGCCACTG GCAGGGACACGTTGCGATACTGGTGTTTAGTGCTCCA CTTAAACTCAGGCACAACCATCCGCGGCAGCTCGGTG AAGTTTTCACTCCACAGGCTGCGCACCATCACCAACG CGTTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCA

		GTTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATA CACAGGGTTGCAGCACTGGAACACTATCAGCGCCGG GTGGTGCACGCTGGCCAGCACGCTCTTGTCGGAGAT CAGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGC GAACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAG GGCGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGT AGTGGCATCAAAAGGTGACCGTGCCCGGTCTGGGCG TTAGGATACAGCGCCTGCATAAAAGCCTTGATCTGCTT AAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAAC ATGCCGCAAGACTTGCCGGAAAACTGATTGGCCGGAC AGGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGT TGGAGATCTGCACCACATTTCGGCCCCACCGGTTCTT CACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCG CGCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCAC GTGCTCCTTATTTATCATAATGCTTCCGTGTAGACACT TAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCCACAAA AGCGAAGATCAGCTTCGGCGCACGCTGGAAGACGCG GAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA AGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGCG AAAACTACGTCATCTCCAGCGGCCACACCCGGCGCCA GCACCTGTTGTCAGCGCCATTGGCGCGCCCGCCCGC CGCGCGCTTCGCTTTTTATAGGGCCGCCGCCGCCGC CGCCTCGCCATAAAAGGAAACTTTCGGAGCGCGCCG CTCTGATTGGCTGCCGCCGCACCTCTCCGCCTCGCC

		CCGCCCCGCCCCTCGCCCCGCCCCGCCCCGCCTGG CGCGCGCCCCCCCCCCCCCCCCCGCCCCCATCGCTGC ACAAAATAATTAAAAAATAAATAAATACAAAATTGGGG GTGGGGAGGGGGGGGAGATGGGGAGAGTGAAGCAG AACGTGGGGCTCACCTCGAGGCCGGCCGAATATCTT CATTTAAATGTTTAAACATCGATGCGGCCGCCGTTTGT GTTATGTTTCAACGTGTTTATTTTTCAATTGCAGAAAAT TTCAAGTCATTTTTCATTCAGTAGTATAGCCCCACCAC CACATAGCTTATACAGATCACCGTACCTTAATCAAACT CACAGAACCCTAGTATTCAACCTGCCACCTCCCTCCC AACACACAGAGTACACAGTCCTTTCTCCCCGGCTGGC CTTAAAAAGCATCATATCATGGGTAACAGACATATTCT TAGGTGTTATATTCCACACGGTTTCCTGTCGAGCCAAA CGCTCATCAGTGATATTAATAAACTCCCCGGGCAGCT CACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGCCAC AGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGCGG CGAAGGAGAAGTCCACGCCTACATGGGGGTAGAGTC ATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCAG CAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCGT CCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCG ATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTCC TCCGGGCACAGCAGCGCACCCTGATCTCACTTAAATC AGCACAGTAACTGCAGCACAGCACCACAATATTGTTC AAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTCA TGGCGGGGACCACAGAACCCACGTGGCCATCATACC ACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAAA CACGCTGGACATAAACATTACCTCTTTTGGCATGTTGT AATTCACCACCTCCCGGTACCATATAAACCTCTGATTA AACATGGCGCCATCCACCACCATCCTAAACCAGCTGG CCAAAACCTGCCCGCCGGCTATACACTGCAGGGAAC CGGGACTGGAACAATGACAGTGGAGAGCCCAGGACT CGTAACCATGGATCATCATGCTCGTCATGATATCAATG TTGGCACAACACAGGCACACGTGCATACACTTCCTCA GGATTACAAGCTCCTCCCGCGTTAGAACCATATCCCA GGGAACAACCCATTCCTGAATCAGCGTAAATCCCACA CTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCA TTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGATC CTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGA GGTAGACGATCCCTACTGTACGGAGTGCGCCGAGAC AACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAATG GAACGCCGGACGTAGTCATATTTCCTGAAGCAAAACC AGGTGCGGGCGTGACAAACAGATCTGCGTCTCCGGT CTCGCCGCTTAGATCGCTCTGTGTAGTAGTTGTAGTAT ATCCACTCTCTCAAAGCATCCAGGCGCCCCCTGGCTT CGGGTTCTATGTAAACTCCTTCATGCGCCGCTGCCCT GATAACATCCACCACCGCAGAATAAGCCACACCCAGC CAACCTACACATTCGTTCTGCGAGTCACACACGGGAG

		GAGCGGGAAGAGCTGGAAGAACCATGTTTTTTTTTTTA TTCCAAAAGATTATCCAAAACCTCAAAATGAAGATCTA TTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCAAA CTCTACAGCCAAAGAACAGATAATGGCATTTGTAAGAT GTTGCACAATGGCTTCCAAAAGGCAAACGGCCCTCAC GTCCAAGTGGACGTAAAGGCTAAACCCTTCAGGGTGA ATCTCCTCTATAAACATTCCAGCACCTTCAACCATGCC CAAATAATTCTCATCTCGCCACCTTCTCAATATATCTCT AAGCAAATCCCGAATATTAAGTCCGGCCATTGTAAAAA TCTGCTCCAGAGCGCCCTCCACCTTCAGCCTCAAGCA GCGAATCATGATTGCAAAAATTCAGGTTCCTCACAGAC CTGTATAAGATTCAAAAGCGGAACATTAACAAAAATAC CGCGATCCCGTAGGTCCCTTCGCAGGGCCAGCTGAA CATAATCGTGCAGGTCTGCACGGACCAGCGCGGCCA CTTCCCCGCCAGGAACCATGACAAAAGAACCCACACT GATTATGACACGCATACTCGGAGCTATGCTAACCAGC GTAGCCCCGATGTAAGCTTGTTGCATGGGCGGCGATA TAAAATGCAAGGTGCTGCTCAAAAAATCAGGCAAAGC CTCGCGCAAAAAAGAAAGCACATCGTAGTCATGCTCA TGCAGATAAAGGCAGGTAAGCTCCGGAACCACCACAG AAAAAGACACCATTTTTCTCTCAAACATGTCTGCGGGT TTCTGCATAAACACAAAATAAAATAACAAAAAAACATTT AAACATTAGAAGCCTGTCTTACAACAGGAAAAACAACC CTTATAAGCATAAGACGGACTACGGCCATGCCGGCGT GACCGTAAAAAAACTGGTCACCGTGATTAAAAAGCAC CACCGACAGCTCCTCGGTCATGTCCGGAGTCATAATG TAAGACTCGGTAAACACATCAGGTTGATTCACATCGGT CAGTGCTAAAAAGCGACCGAAATAGCCCGGGGGAATA CATACCCGCAGGCGTAGAGACAACATTACAGCCCCCA TAGGAGGTATAACAAAATTAATAGGAGAGAAAAACACA TAAACACCTGAAAAACCCTCCTGCCTAGGCAAAATAG CACCCTCCCGCTCCAGAACAACATACAGCGCTTCCAC AGCGGCAGCCATAACAGTCAGCCTTACCAGTAAAAAA GAAAACCTATTAAAAAAACACCACTCGACACGGCACC AGCTCAATCAGTCACAGTGTAAAAAAGGGCCAAGTGC AGAGCGAGTATATATAGGACTAAAAAATGACGTAACG GTTAAAGTCCACAAAAAACACCCAGAAAACCGCACGC GAACCTACGCCCAGAAACGAAAGCCAAAAAACCCACA ACTTCCTCAAATCGTCACTTCCGTTTTCCCACGTTACG TCACTTCCCATTTTAAGAAAACTACAATTCCCAACACA TACAAGTTACTCCGCCCTTAATTAAATCGGATCCGATA TCTAGATGTATTCGCGAGGTACCGAGCTCGAATTCTC TGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC TGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCC CCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGC ACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG GCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCA

		TCTGTGCGGTATTTCACACCGCATATGGTGCACTCTC AGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGC CCCGACACCCGCCAACACCCGCTGACGCGCCCTGAC GGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGC TGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTT TCACCGTCATCACCGAAACGCGCGA
47	pEMBR-1.5 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG

		TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG

		GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC

		TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCCACAAA AGCGAAGATCAGCTTCGGCGCACGCTGGAAGACGCG GAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA AGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGCG AAAACTACGTCATCTCCAGCGGCCACACCCGGCGCCA GCACCTGTTGTCAGCGCCATTGGCGCGCCCGCCCGC CGCGCGCTTCGCTTTTTATAGGGCCGCCGCCGCCGC CGCCTCGCCATAAAAGGAAACTTTCGGAGCGCGCCG CTCTGATTGGCTGCCGCCGCACCTCTCCGCCTCGCC CCGCCCCGCCCCTCGCCCCGCCCCGCCCCGCCTGG CGCGCGCCCCCCCCCCCCCCCCGCCCCCATCGCTGC ACAAAATAATTAAAAAATAAATAAATACAAAATTGGGG GTGGGGAGGGGGGGGAGATGGGGAGAGTGAAGCAG AACGTGGGGCTCACCTCGAGGCCGGCCGAATATCTT CATTTAAATGGGCAGAGCGCACATCGCCCACAGTCCC CGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGG TGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAG TGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGT GGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTG AACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAC AGCACCGCGGGCCCGATCCACCGGTACTGTTGGTAA AGCCACCATGTTTTCCGGTGGCGGCGGCCCGCTGTC CCCCGGAGGAAAGTCGGCGGCCAGGGCGGCGTCCG GGTTTTTTGCGCCCGCCGGCCCTCGCGGAGCCAGCC GGGGACCCCCCGCCTTGTTTGAGGCAAAACTTTTACAA CCCCTACCTCGCCCCAGTCGGGACGCAACAGAAGCC GACCGGGCCAACCCAGCGCCATACGTACTATAGCGA ATGCGATGAATTTCGATTCATCGCCCCGCGGGTGCTG GACGAGGATGCCCCCCCGGAGAAGCGCGCCGGGGT GCACGACGGTCACCTCAAGCGCGCCCCCAAGGTGTA CTGCGGGGGGGGACGAGCGCGACGTCCTCCGCGTCG GGTCGGGCGGCTTCTGGCCGCGGCGCTCGCGCCTGT GGGGCGGCGTGGACCACGCCCCGGCGGGGTTCAAC CCCACCGTCACCGTCTTTCACGTGTACGACATCCTGG AGAACGTGGAGCACGCGTACGGCATGCGCGCGGCCC AGTTCCACGCGCGGTTTATGGACGCCATCACACCGAC

		GGGGACCGTCATCACGCTCCTGGGCCTGACTCCGGA AGGCCACCGGGTGGCCGTTCACGTTTACGGCACGCG GCAGTACTTTTACATGAACAAGGAGGAGGTCGACAGG CACCTACAATGCCGCGCCCCACGAGATCTCTGCGAG CGCATGGCCGCGGCCCTGCGCGAGTCCCCGGGCGC GTCGTTCCGCGGCATCTCCGCGGACCACTTCGAGGC GGAGGTGGTGGAGCGCACCGACGTGTACTACTACGA GACGCGCCCCGCTCTGTTTTACCGCGTCTACGTCCGA AGCGGGCGCGTGCTGTCGTACCTGTGCGACAACTTCT GCCCGGCCATCAAGAAGTACGAGGGTGGGGTCGACG CCACCACCCGGTTCATCCTGGACAACCCCGGGTTCGT CACCTTCGGCTGGTACCGTCTCAAACCGGGCCGGAA CAACACGCTAGCCCAGCCGCGGGCCCCGATGGCCTT CGGGACATCCAGCGACGTCGAGTTTAACTGTACGGC GGACAACCTGGCCATCGAGGGGGGCATGAGCGACCT ACCGGCATACAAGCTCATGTGCTTCGATATCGAATGC AAGGCGGGGGGGGAGGACGAGCTGGCCTTTCCGGT GGCCGGGCACCCGGAGGACCTGGTCATCCAGATATC CTGTCTGCTCTACGACCTGTCCACCACCGCCCTGGAG CACGTCCTCCTGTTTTCGCTCGGTTCCTGCGACCTCC CCGAATCCCACCTGAACGAGCTGGCGGCCAGGGGCC TGCCCACGCCCGTGGTTCTGGAATTCGACAGCGAATT CGAGATGCTGTTGGCCTTCATGACCCTTGTGAAACAG TACGGCCCCGAGTTCGTGACCGGGTACAACATCATCA ACTTCGACTGGCCCTTCTTGCTGGCCAAGCTGACGGA CATTTACAAGGTCCCCCTGGACGGGTACGGCCGCAT GAACGGCCGGGGCGTGTTTCGCGTGTGGGACATAGG CCAGAGCCACTTCCAGAAGCGCAGCAAGATAAAGGTG AACGGCATGGTGAACATCGACATGTACGGGATTATAA CCGACAAGATCAAGCTCTCGAGCTACAAGCTCAACGC CGTGGCCGAAGCCGTCCTGAAGGACAAGAAGAAGGA CCTGAGCTATCGCGACATCCCCGCCTACTACGCCGCC GGGCCCGCGCAACGCGGGGTGATCGGCGAGTACTG CATACAGGATTCCCTGCTGGTGGGCCAGCTGTTTTTT AAGTTTTTGCCCCATCTGGAGCTCTCGGCCGTCGCGC GCTTGGCGGGTATTAACATCACCCGCACCATCTACGA CGGCCAGCAGATCCGCGTCTTTACGTGCCTGCTGCG CCTGGCCGACCAGAAGGGCTTTATTCTGCCGGACACC CAGGGGCGATTTAGGGGCGCCGGGGGGGAGGCGCC CAAGCGTCCGGCCGCAGCCCGGGAGGACGAGGAGC GGCCAGAGGAGGAGGGGGAGGACGAGGACGAACGC GAGGAGGGCGGGGGCGAGCGGGAGCCGGAGGGCG CGCGGGAGACCGCCGGCAGGCACGTGGGGTACCAG GGGGCCAGGGTCCTTGACCCCACTTCCGGGTTTCAC GTGAACCCCGTGGTGGTGTTCGACTTTGCCAGCCTGT ACCCCAGCATCATCCAGGCCCACAACCTGTGCTTCAG CACGCTCTCCCTGAGGGCCGACGCAGTGGCGCACCT

		GGAGGCGGGCAAGGACTACCTGGAGATCGAGGTGGG GGGGCGACGGCTGTTCTTCGTCAAGGCTCACGTGCG AGAGAGCCTCCTCAGCATCCTCCTGCGGGACTGGCT CGCCATGCGAAAGCAGATCCGCTCGCGGATTCCCCA GAGCAGCCCCGAGGAGGCCGTGCTCCTGGACAAGCA GCAGGCCGCCATCAAGGTCGTGTGTAACTCGGTGTAC GGGTTCACGGGAGTGCAGCACGGACTCCTGCCGTGC CTGCACGTTGCCGCGACGGTGACGACCATCGGCCGC GAGATGCTGCTCGCGACCCGCGAGTACGTCCACGCG CGCTGGGCGGCCTTCGAACAGCTCCTGGCCGATTTC CCGGAGGCGGCCGACATGCGCGCCCCCGGGCCCTA TTCCATGCGCATCATCTACGGGGACACGGACTCCATC TTTGTGCTGTGCCGCGGCCTCACGGCCGCCGGGCTG ACGGCCGTGGGCGACAAGATGGCGAGCCACATCTCG CGCGCGCTGTTTCTGCCCCCCATCAAACTCGAGTGCG AAAAGACGTTCACCAAGCTGCTGCTGATCGCCAAGAA AAAGTACATCGGCGTCATCTACGGGGGTAAGATGCTC ATCAAGGGCGTGGATCTGGTGCGCAAAAACAACTGCG CGTTTATCAACCGCACCTCCAGGGCCCTGGTCGACCT GCTGTTTTACGACGATACCGTCTCCGGAGCGGCCGC CGCGTTAGCCGAGCGCCCCGCGGAGGAGTGGCTGG CGCGACCCCTGCCCGAGGGACTGCAGGCGTTCGGG GCCGTCCTCGTAGACGCCCATCGGCGCATCACCGAC CCGGAGAGGGACATCCAGGACTTTGTCCTCACCGCC GAACTGAGCAGACACCCGCGCGCGTACACCAACAAG CGCCTGGCCCACCTGACGGTGTATTACAAGCTCATGG CCCGCCGCGCGCAGGTCCCGTCCATCAAGGACCGGA TCCCGTACGTGATCGTGGCCCAGACCCGCGAGGTAG AGGAGACGGTCGCGCGGCTGGCCGCCCTCCGCGAG CTAGACGCCGCCGCCCCAGGGGACGAGCCCGCCCC CCCCGCGGCCCTGCCCTCCCCGGCCAAGCGCCCCC GGGAGACGCCGTCGCCTGCCGACCCCCCGGGAGGC GCGTCCAAGCCCCGCAAGCTGCTGGTGTCCGAGCTG GCCGAGGATCCCGCATACGCCATTGCCCACGGCGTC GCCCTGAACACGGACTATTACTTCTCCCACCTGTTGG GGGCGGCGTGCGTGACATTCAAGGCCCTGTTTGGGA ATAACGCCAAGATCACCGAGAGTCTGTTAAAAAGGTTT ATTCCCGAAGTGTGGCACCCCCCGGACGACGTGGCC GCGCGGCTCCGGACCGCAGGGTTCGGGGCGGTGGG TGCCGGCGCTACGGCGGAGGAAACTCGTCGAATGTT GCATAGAGCCTTTGATACTCTAGCAGAATTCGGCAGT GGAGCAACAAACTTCTCTCTGCTGAAACAAGCCGGAG ATGTCGAAGAGAATCCTGGACCGACGGATTCCCCTGG CGGTGTGGCCCCCGCCTCCCCCGTGGAGGACGCGTC GGACGCGTCCCTCGGGCAGCCGGAGGAGGGGGCGC CCTGCCAGGTGGTCCTGCAGGGCGCCGAACTTAATG GAATCCTACAGGCGTTTGCCCCGCTGCGCACGAGCC

		TTCTGGACTCGCTTCTGGTTATGGGCGACCGGGGCAT CCTTATCCATAACACGATCTTTGGGGAGCAGGTGTTC CTGCCCCTGGAACACTCGCAATTCAGTCGGTATCGCT GGCGCGGACCCACGGCGGCGTTCCTGTCTCTCGTGG ACCAGAAGCGCTCCCTCCTGAGCGTGTTTCGCGCCAA CCAGTACCCGGACCTACGTCGGGTGGAGTTGGCGAT CACGGGCCAGGCCCCGTTTCGCACGCTGGTTCAGCG CATATGGACGACGACGTCCGACGGCGAGGCCGTTGA GCTAGCCAGCGAGACGCTGATGAAGCGCGAACTGAC GAGCTTTGTGGTGCTGGTTCCCCAGGGAACCCCCGA CGTTCAGTTGCGCCTGACGAGGCCGCAGCTCACCAA GGTCCTTAACGCGACCGGGGCCGATAGTGCCACGCC CACCACGTTCGAGCTCGGGGTTAACGGCAAATTTTCC GTGTTCACCACGAGTACCTGCGTCACCTTTGCTGCCC GCGAGGAGGGCGTGTCGTCCAGCACCAGCACCCAGG TCCAGATCCTGTCCAACGCGCTCACCAAGGCGGGCC AGGCGGCCGCCAACGCCAAGACGGTGTACGGGGAAA ATACCCATCGCACCTTCTCTGTGGTCGTCGACGATTG CAGCATGCGGGCGGTGCTCCGGCGACTGCAGGTCGG CGGGGGCACCCTCAAGTTCTTCCTCACGACCCCCGTC CCCAGTCTGTGCGTCACCGCCACCGGTCCCAACGCG GTATCGGCGGTATTTCTCCTGAAACCCCAGAAGATTT GCCTGGACTGGCTGGGTCATAGCCAGGGGTCTCCTT CAGCCGGGAGCTCGGCCTCCCGGGCCTCTGGGAGC GAGCCAACAGACAGCCAGGACTCCGCGTCGGACGCG GTCAGCCACGGCGATCCGGAAGACCTCGATGGCGCT GCCCGGGCGGGAGAGGCGGGGGCCTTGCATGCCTG TCCGATGCCGTCGTCGACCACGCGGGTCACTCCCAC GACCAAGCGGGGGCGCTCGGGGGGCGAGGATGCGC GCGCGGACACGGCCCTAAAGAAACCTAAGACGGGGT CGCCCACCGCACCCCCGCCCGCAGATCCAGTCCCCC TGGACACGGAGGACGACTCCGATGCGGCGGACGGGA CGGCGGCCCGTCCCGCCGCTCCAGACGCCCGGAGC GGAAGCCGTTACGCGTGTTACTTTCGCGACCTCCCGA CCGGAGAAGCAAGCCCCGGCGCCTTCTCCGCCTTCC GGGGGGGCCCCCAAACCCCGTATGGTTTTGGATTCC CCTGATAAGATCCGACTGCAGGTAGAATAAAGGAAAT TTATTTTCATTGCAATAGTGTGTTGGAATTTTTTGTGTC TCTCAGTTTAAACGCGGCCGCCGTTTGTGTTATGTTTC AACGTGTTTATTTTTCAATTGCAGAAAATTTCAAGTCAT TTTTCATTCAGTAGTATAGCCCCACCACCACATAGCTT ATACAGATCACCGTACCTTAATCAAACTCACAGAACCC TAGTATTCAACCTGCCACCTCCCTCCCAACACACAGA GTACACAGTCCTTTCTCCCCGGCTGGCCTTAAAAAGC ATCATATCATGGGTAACAGACATATTCTTAGGTGTTAT ATTCCACACGGTTTCCTGTCGAGCCAAACGCTCATCA GTGATATTAATAAACTCCCCGGGCAGCTCACTTAAGTT

		CATGTCGCTGTCCAGCTGCTGAGCCACAGGCTGCTGT CCAACTTGCGGTTGCTTAACGGGCGGCGAAGGAGAA GTCCACGCCTACATGGGGGTAGAGTCATAATCGTGCA TCAGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGAA TAAACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAATA CAACATGGCAGTGGTCTCCTCAGCGATGATTCGCACC GCCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACAG CAGCGCACCCTGATCTCACTTAAATCAGCACAGTAAC TGCAGCACAGCACCACAATATTGTTCAAAATCCCACA GTGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGAC CACAGAACCCACGTGGCCATCATACCACAAGCGCAG GTAGATTAAGTGGCGACCCCTCATAAACACGCTGGAC ATAAACATTACCTCTTTTGGCATGTTGTAATTCACCAC CTCCCGGTACCATATAAACCTCTGATTAAACATGGCG CCATCCACCACCATCCTAAACCAGCTGGCCAAAACCT GCCCGCCGGCTATACACTGCAGGGAACCGGGACTGG AACAATGACAGTGGAGAGCCCAGGACTCGTAACCATG GATCATCATGCTCGTCATGATATCAATGTTGGCACAAC ACAGGCACACGTGCATACACTTCCTCAGGATTACAAG CTCCTCCCGCGTTAGAACCATATCCCAGGGAACAACC CATTCCTGAATCAGCGTAAATCCCACACTGCAGGGAA GACCTCGCACGTAACTCACGTTGTGCATTGTCAAAGT GTTACATTCGGGCAGCAGCGGATGATCCTCCAGTATG GTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGAT CCCTACTGTACGGAGTGCGCCGAGACAACCGAGATC GTGTTGGTCGTAGTGTCATGCCAAATGGAACGCCGGA CGTAGTCATATTTCCTGAAGCAAAACCAGGTGCGGGC GTGACAAACAGATCTGCGTCTCCGGTCTCGCCGCTTA GATCGCTCTGTGTAGTAGTTGTAGTATATCCACTCTCT CAAAGCATCCAGGCGCCCCCTGGCTTCGGGTTCTATG TAAACTCCTTCATGCGCCGCTGCCCTGATAACATCCA CCACCGCAGAATAAGCCACACCCAGCCAACCTACACA TTCGTTCTGCGAGTCACACACGGGAGGAGCGGGAAG AGCTGGAAGAACCATGTTTTTTTTTTTATTCCAAAAGAT TATCCAAAACCTCAAAATGAAGATCTATTAAGTGAACG CGCTCCCCTCCGGTGGCGTGGTCAAACTCTACAGCCA AAGAACAGATAATGGCATTTGTAAGATGTTGCACAATG GCTTCCAAAAGGCAAACGGCCCTCACGTCCAAGTGGA CGTAAAGGCTAAACCCTTCAGGGTGAATCTCCTCTATA AACATTCCAGCACCTTCAACCATGCCCAAATAATTCTC ATCTCGCCACCTTCTCAATATATCTCTAAGCAAATCCC GAATATTAAGTCCGGCCATTGTAAAAATCTGCTCCAGA GCGCCCTCCACCTTCAGCCTCAAGCAGCGAATCATGA TTGCAAAAATTCAGGTTCCTCACAGACCTGTATAAGAT TCAAAAGCGGAACATTAACAAAAATACCGCGATCCCG TAGGTCCCTTCGCAGGGCCAGCTGAACATAATCGTGC AGGTCTGCACGGACCAGCGCGGCCACTTCCCCGCCA

		GGAACCATGACAAAAGAACCCACACTGATTATGACAC GCATACTCGGAGCTATGCTAACCAGCGTAGCCCCGAT GTAAGCTTGTTGCATGGGCGGCGATATAAAATGCAAG GTGCTGCTCAAAAAATCAGGCAAAGCCTCGCGCAAAA AAGAAAGCACATCGTAGTCATGCTCATGCAGATAAAG GCAGGTAAGCTCCGGAACCACCACAGAAAAAGACACC ATTTTTCTCTCAAACATGTCTGCGGGTTTCTGCATAAA CACAAAATAAAATAACAAAAAAACATTTAAACATTAGAA GCCTGTCTTACAACAGGAAAAACAACCCTTATAAGCAT AAGACGGACTACGGCCATGCCGGCGTGACCGTAAAA AAACTGGTCACCGTGATTAAAAAGCACCACCGACAGC TCCTCGGTCATGTCCGGAGTCATAATGTAAGACTCGG TAAACACATCAGGTTGATTCACATCGGTCAGTGCTAAA AAGCGACCGAAATAGCCCGGGGGAATACATACCCGC AGGCGTAGAGACAACATTACAGCCCCCATAGGAGGTA TAACAAAATTAATAGGAGAGAAAAACACATAAACACCT GAAAAACCCTCCTGCCTAGGCAAAATAGCACCCTCCC GCTCCAGAACAACATACAGCGCTTCCACAGCGGCAGC CATAACAGTCAGCCTTACCAGTAAAAAAGAAAACCTAT TAAAAAAACACCACTCGACACGGCACCAGCTCAATCA GTCACAGTGTAAAAAAGGGCCAAGTGCAGAGCGAGTA TATATAGGACTAAAAAATGACGTAACGGTTAAAGTCCA CAAAAAACACCCAGAAAACCGCACGCGAACCTACGCC CAGAAACGAAAGCCAAAAAACCCACAACTTCCTCAAAT CGTCACTTCCGTTTTCCCACGTTACGTCACTTCCCATT TTAAGAAAACTACAATTCCCAACACATACAAGTTACTC CGCCCTTAATTAAATCGGATCCGATATCTAGATGTATT CGCGAGGTACCGAGCTCGAATTCTCTGGCCGTCGTTT TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCT TCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGC CTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTAT TTCACACCGCATATGGTGCACTCTCAGTACAATCTGCT CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGC CAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGC TCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCA CCGAAACGCGCGA
48	pEMBR-1.5A Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT

		TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT

		TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG

		GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCCACAAA AGCGAAGATCAGCTTCGGCGCACGCTGGAAGACGCG

		GAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA AGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGCG AAAACTACGTCATCTCCAGCGGCCACACCCGGCGCCA GCACCTGTTGTCAGCGCCATTGGCGCGCCCGCCCGC CGCGCGCTTCGCTTTTTATAGGGCCGCCGCCGCCGC CGCCTCGCCATAAAAGGAAACTTTCGGAGCGCGCCG CTCTGATTGGCTGCCGCCGCACCTCTCCGCCTCGCC CCGCCCCGCCCCTCGCCCCGCCCCGCCCCGCCTGG CGCGCGCCCCCCCCCCCCCCCCGCCCCCATCGCTGC ACAAAATAATTAAAAAATAAATAAATACAAAATTGGGG GTGGGGAGGGGGGGGAGATGGGGAGAGTGAAGCAG AACGTGGGGCTCACCTCGAGGCCGGCCGAATATCTT CATTTAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCC CAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC ATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCC AGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAT CTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTC CGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCC ATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATG CAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCC AGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTT TTGCAAACGCCGGCGCACCGCGGGCCCGATCCACCG GTACTGTTGGTAAAGCCACCATGTTTTCCGGTGGCGG CGGCCCGCTGTCCCCCGGAGGAAAGTCGGCGGCCA GGGCGGCGTCCGGGTTTTTTGCGCCCGCCGGCCCTC GCGGAGCCAGCCGGGGACCCCCGCCTTGTTTGAGGC AAAACTTTTACAACCCCTACCTCGCCCCAGTCGGGAC GCAACAGAAGCCGACCGGGCCAACCCAGCGCCATAC GTACTATAGCGAATGCGATGAATTTCGATTCATCGCCC CGCGGGTGCTGGACGAGGATGCCCCCCCGGAGAAG CGCGCCGGGGTGCACGACGGTCACCTCAAGCGCGCC CCCAAGGTGTACTGCGGGGGGGACGAGCGCGACGTC CTCCGCGTCGGGTCGGGCGGCTTCTGGCCGCGGCG CTCGCGCCTGTGGGGCGGCGTGGACCACGCCCCGG CGGGGTTCAACCCCACCGTCACCGTCTTTCACGTGTA CGACATCCTGGAGAACGTGGAGCACGCGTACGGCAT GCGCGCGGCCCAGTTCCACGCGCGGTTTATGGACGC CATCACACCGACGGGGACCGTCATCACGCTCCTGGG CCTGACTCCGGAAGGCCACCGGGTGGCCGTTCACGT TTACGGCACGCGGCAGTACTTTTACATGAACAAGGAG GAGGTCGACAGGCACCTACAATGCCGCGCCCCACGA GATCTCTGCGAGCGCATGGCCGCGGCCCTGCGCGAG TCCCCGGGCGCGTCGTTCCGCGGCATCTCCGCGGAC CACTTCGAGGCGGAGGTGGTGGAGCGCACCGACGTG TACTACTACGAGACGCGCCCCGCTCTGTTTTACCGCG TCTACGTCCGAAGCGGGCGCGTGCTGTCGTACCTGT GCGACAACTTCTGCCCGGCCATCAAGAAGTACGAGG

		GTGGGGTCGACGCCACCACCCGGTTCATCCTGGACA ACCCCGGGTTCGTCACCTTCGGCTGGTACCGTCTCAA ACCGGGCCGGAACAACACGCTAGCCCAGCCGCGGGC CCCGATGGCCTTCGGGACATCCAGCGACGTCGAGTTT AACTGTACGGCGGACAACCTGGCCATCGAGGGGGGC ATGAGCGACCTACCGGCATACAAGCTCATGTGCTTCG ATATCGAATGCAAGGCGGGGGGGGAGGACGAGCTGG CCTTTCCGGTGGCCGGGCACCCGGAGGACCTGGTCA TCCAGATATCCTGTCTGCTCTACGACCTGTCCACCAC CGCCCTGGAGCACGTCCTCCTGTTTTCGCTCGGTTCC TGCGACCTCCCCGAATCCCACCTGAACGAGCTGGCG GCCAGGGGCCTGCCCACGCCCGTGGTTCTGGAATTC GACAGCGAATTCGAGATGCTGTTGGCCTTCATGACCC TTGTGAAACAGTACGGCCCCGAGTTCGTGACCGGGTA CAACATCATCAACTTCGACTGGCCCTTCTTGCTGGCC AAGCTGACGGACATTTACAAGGTCCCCCTGGACGGGT ACGGCCGCATGAACGGCCGGGGCGTGTTTCGCGTGT GGGACATAGGCCAGAGCCACTTCCAGAAGCGCAGCA AGATAAAGGTGAACGGCATGGTGAACATCGACATGTA CGGGATTATAACCGACAAGATCAAGCTCTCGAGCTAC AAGCTCAACGCCGTGGCCGAAGCCGTCCTGAAGGAC AAGAAGAAGGACCTGAGCTATCGCGACATCCCCGCCT ACTACGCCGCCGGGCCCGCGCAACGCGGGGTGATC GGCGAGTACTGCATACAGGATTCCCTGCTGGTGGGC CAGCTGTTTTTTAAGTTTTTGCCCCATCTGGAGCTCTC GGCCGTCGCGCGCTTGGCGGGTATTAACATCACCCG CACCATCTACGACGGCCAGCAGATCCGCGTCTTTACG TGCCTGCTGCGCCTGGCCGACCAGAAGGGCTTTATTC TGCCGGACACCCAGGGGCGATTTAGGGGCGCCGGG GGGGAGGCGCCCAAGCGTCCGGCCGCAGCCCGGGA GGACGAGGAGCGGCCAGAGGAGGAGGGGGAGGACG AGGACGAACGCGAGGAGGGCGGGGGCGAGCGGGAG CCGGAGGGCGCGCGGGAGACCGCCGGCAGGCACGT GGGGTACCAGGGGGCCAGGGTCCTTGACCCCACTTC CGGGTTTCACGTGAACCCCGTGGTGGTGTTCGACTTT GCCAGCCTGTACCCCAGCATCATCCAGGCCCACAACC TGTGCTTCAGCACGCTCTCCCTGAGGGCCGACGCAG TGGCGCACCTGGAGGCGGGCAAGGACTACCTGGAGA TCGAGGTGGGGGGGCGACGGCTGTTCTTCGTCAAGG CTCACGTGCGAGAGAGCCTCCTCAGCATCCTCCTGCG GGACTGGCTCGCCATGCGAAAGCAGATCCGCTCGCG GATTCCCCAGAGCAGCCCCGAGGAGGCCGTGCTCCT GGACAAGCAGCAGGCCGCCATCAAGGTCGTGTGTAA CTCGGTGTACGGGTTCACGGGAGTGCAGCACGGACT CCTGCCGTGCCTGCACGTTGCCGCGACGGTGACGAC CATCGGCCGCGAGATGCTGCTCGCGACCCGCGAGTA CGTCCACGCGCGCTGGGCGGCCTTCGAACAGCTCCT

		GGCCGATTTCCCGGAGGCGGCCGACATGCGCGCCCC CGGGCCCTATTCCATGCGCATCATCTACGGGGACACG GACTCCATCTTTGTGCTGTGCCGCGGCCTCACGGCC GCCGGGCTGACGGCCGTGGGCGACAAGATGGCGAG CCACATCTCGCGCGCGCTGTTTCTGCCCCCCATCAAA CTCGAGTGCGAAAAGACGTTCACCAAGCTGCTGCTGA TCGCCAAGAAAAAGTACATCGGCGTCATCTACGGGGG TAAGATGCTCATCAAGGGCGTGGATCTGGTGCGCAAA AACAACTGCGCGTTTATCAACCGCACCTCCAGGGCCC TGGTCGACCTGCTGTTTTACGACGATACCGTCTCCGG AGCCGCCGCGGCGTTAGCCGAGCGCCCCGCGGAGG AGTGGCTGGCGCGACCCCTGCCCGAGGGACTGCAGG CGTTCGGGGCCGTCCTCGTAGACGCCCATCGGCGCA TCACCGACCCGGAGAGGGACATCCAGGACTTTGTCCT CACCGCCGAACTGAGCAGACACCCGCGCGCGTACAC CAACAAGCGCCTGGCCCACCTGACGGTGTATTACAAG CTCATGGCCCGCCGCGCGCAGGTCCCGTCCATCAAG GACCGGATCCCGTACGTGATCGTGGCCCAGACCCGC GAGGTAGAGGAGACGGTCGCGCGGCTGGCCGCCCT CCGCGAGCTAGACGCCGCCGCCCCAGGGGACGAGC CCGCCCCCCCCGCGGCCCTGCCCTCCCCGGCCAAGC GCCCCCGGGAGACGCCGTCGCCTGCCGACCCCCCG GGAGGCGCGTCCAAGCCCCGCAAGCTGCTGGTGTCC GAGCTGGCCGAGGATCCCGCATACGCCATTGCCCAC GGCGTCGCCCTGAACACGGACTATTACTTCTCCCACC TGTTGGGGGCGGCGTGCGTGACATTCAAGGCCCTGT TTGGGAATAACGCCAAGATCACCGAGAGTCTGTTAAA AAGGTTTATTCCCGAAGTGTGGCACCCCCCGGACGAC GTGGCCGCGCGGCTCCGGACCGCAGGGTTCGGGGC GGTGGGTGCCGGCGCTACGGCGGAGGAAACTCGTCG AATGTTGCATAGAGCCTTTGATACTCTAGCAGAATTCG GCAGTGGAGCAACAAACTTCTCTCTGCTGAAACAAGC CGGAGATGTCGAAGAGAATCCTGGACCGACGGATTC CCCTGGCGGTGTGGCCCCCGCCTCCCCCGTGGAGGA CGCGTCGGACGCGTCCCTCGGGCAGCCGGAGGAGG GGGCGCCCTGCCAGGTGGTCCTGCAGGGCGCCGAA CTTAATGGAATCCTACAGGCGTTTGCCCCGCTGCGCA CGAGCCTTCTGGACTCGCTTCTGGTTATGGGCGACCG GGGCATCCTTATCCATAACACGATCTTTGGGGAGCAG GTGTTCCTGCCCCTGGAACACTCGCAATTCAGTCGGT ATCGCTGGCGCGGACCCACGGCGGCGTTCCTGTCTC TCGTGGACCAGAAGCGCTCCCTCCTGAGCGTGTTTCG CGCCAACCAGTACCCGGACCTACGTCGGGTGGAGTT GGCGATCACGGGCCAGGCCCCGTTTCGCACGCTGGT TCAGCGCATATGGACGACGACGTCCGACGGCGAGGC CGTTGAGCTAGCCAGCGAGACGCTGATGAAGCGCGA ACTGACGAGCTTTGTGGTGCTGGTTCCCCAGGGAACC

		CCCGACGTTCAGTTGCGCCTGACGAGGCCGCAGCTC ACCAAGGTCCTTAACGCGACCGGGGCCGATAGTGCC ACGCCCACCACGTTCGAGCTCGGGGTTAACGGCAAAT TTTCCGTGTTCACCACGAGTACCTGCGTCACCTTTGCT GCCCGCGAGGAGGGCGTGTCGTCCAGCACCAGCACC CAGGTCCAGATCCTGTCCAACGCGCTCACCAAGGCG GGCCAGGCCGCCGCGAACGCCAAGACGGTGTACGG GGAAAATACCCATCGCACCTTCTCTGTGGTCGTCGAC GATTGCAGCATGCGGGCGGTGCTCCGGCGACTGCAG GTCGGCGGGGGCACCCTCAAGTTCTTCCTCACGACC CCCGTCCCCAGTCTGTGCGTCACCGCCACCGGTCCC AACGCGGTATCGGCGGTATTTCTCCTGAAACCCCAGA AGATTTGCCTGGACTGGCTGGGTCATAGCCAGGGGT CTCCTTCAGCCGGGAGCTCGGCCTCCCGGGCCTCTG GGAGCGAGCCAACAGACAGCCAGGACTCCGCGTCGG ACGCGGTCAGCCACGGCGATCCGGAAGACCTCGATG GCGCTGCCCGGGCGGGAGAGGCGGGGGCCTTGCAT GCCTGTCCGATGCCGTCGTCGACCACGCGGGTCACT CCCACGACCAAGCGGGGGCGCTCGGGGGGCGAGGA TGCGCGCGCGGACACGGCCCTAAAGAAACCTAAGAC GGGGTCGCCCACCGCACCCCCGCCCGCAGATCCAGT CCCCCTGGACACGGAGGACGACTCCGATGCGGCGGA CGGGACGGCGGCCCGTCCCGCCGCTCCAGACGCCC GGAGCGGAAGCCGTTACGCGTGTTACTTTCGCGACCT CCCGACCGGAGAAGCAAGCCCCGGCGCCTTCTCCGC CTTCCGGGGGGGCCCCCAAACCCCGTATGGTTTTGG ATTCCCCTGATAGGATCCGACTGCAGGTAGCTGTGCC TTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCC GTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTG TCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGT CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATA GCAGGCATGCTGGGGATGCGGTGGGCTCTATGGGTT TAAACATCGATGCGGCCGCCGTTTGTGTTATGTTTCAA CGTGTTTATTTTTCAATTGCAGAAAATTTCAAGTCATTT TTCATTCAGTAGTATAGCCCCACCACCACATAGCTTAT ACAGATCACCGTACCTTAATCAAACTCACAGAACCCTA GTATTCAACCTGCCACCTCCCTCCCAACACACAGAGT ACACAGTCCTTTCTCCCCGGCTGGCCTTAAAAAGCAT CATATCATGGGTAACAGACATATTCTTAGGTGTTATAT TCCACACGGTTTCCTGTCGAGCCAAACGCTCATCAGT GATATTAATAAACTCCCCGGGCAGCTCACTTAAGTTCA TGTCGCTGTCCAGCTGCTGAGCCACAGGCTGCTGTCC AACTTGCGGTTGCTTAACGGGCGGCGAAGGAGAAGT CCACGCCTACATGGGGGTAGAGTCATAATCGTGCATC AGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGAATA AACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAATAC

		AACATGGCAGTGGTCTCCTCAGCGATGATTCGCACCG CCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACAGC AGCGCACCCTGATCTCACTTAAATCAGCACAGTAACT GCAGCACAGCACCACAATATTGTTCAAAATCCCACAG TGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGACC ACAGAACCCACGTGGCCATCATACCACAAGCGCAGGT AGATTAAGTGGCGACCCCTCATAAACACGCTGGACAT AAACATTACCTCTTTTGGCATGTTGTAATTCACCACCT CCCGGTACCATATAAACCTCTGATTAAACATGGCGCC ATCCACCACCATCCTAAACCAGCTGGCCAAAACCTGC CCGCCGGCTATACACTGCAGGGAACCGGGACTGGAA CAATGACAGTGGAGAGCCCAGGACTCGTAACCATGGA TCATCATGCTCGTCATGATATCAATGTTGGCACAACAC AGGCACACGTGCATACACTTCCTCAGGATTACAAGCT CCTCCCGCGTTAGAACCATATCCCAGGGAACAACCCA TTCCTGAATCAGCGTAAATCCCACACTGCAGGGAAGA CCTCGCACGTAACTCACGTTGTGCATTGTCAAAGTGTT ACATTCGGGCAGCAGCGGATGATCCTCCAGTATGGTA GCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGATCCC TACTGTACGGAGTGCGCCGAGACAACCGAGATCGTGT TGGTCGTAGTGTCATGCCAAATGGAACGCCGGACGTA GTCATATTTCCTGAAGCAAAACCAGGTGCGGGCGTGA CAAACAGATCTGCGTCTCCGGTCTCGCCGCTTAGATC GCTCTGTGTAGTAGTTGTAGTATATCCACTCTCTCAAA GCATCCAGGCGCCCCCTGGCTTCGGGTTCTATGTAAA CTCCTTCATGCGCCGCTGCCCTGATAACATCCACCAC CGCAGAATAAGCCACACCCAGCCAACCTACACATTCG TTCTGCGAGTCACACACGGGAGGAGCGGGAAGAGCT GGAAGAACCATGTTTTTTTTTTTATTCCAAAAGATTATC CAAAACCTCAAAATGAAGATCTATTAAGTGAACGCGCT CCCCTCCGGTGGCGTGGTCAAACTCTACAGCCAAAGA ACAGATAATGGCATTTGTAAGATGTTGCACAATGGCTT CCAAAAGGCAAACGGCCCTCACGTCCAAGTGGACGTA AAGGCTAAACCCTTCAGGGTGAATCTCCTCTATAAACA TTCCAGCACCTTCAACCATGCCCAAATAATTCTCATCT CGCCACCTTCTCAATATATCTCTAAGCAAATCCCGAAT ATTAAGTCCGGCCATTGTAAAAATCTGCTCCAGAGCG CCCTCCACCTTCAGCCTCAAGCAGCGAATCATGATTG CAAAAATTCAGGTTCCTCACAGACCTGTATAAGATTCA AAAGCGGAACATTAACAAAAATACCGCGATCCCGTAG GTCCCTTCGCAGGGCCAGCTGAACATAATCGTGCAGG TCTGCACGGACCAGCGCGGCCACTTCCCCGCCAGGA ACCATGACAAAAGAACCCACACTGATTATGACACGCA TACTCGGAGCTATGCTAACCAGCGTAGCCCCGATGTA AGCTTGTTGCATGGGCGGCGATATAAAATGCAAGGTG CTGCTCAAAAAATCAGGCAAAGCCTCGCGCAAAAAAG AAAGCACATCGTAGTCATGCTCATGCAGATAAAGGCA

		GGTAAGCTCCGGAACCACCACAGAAAAAGACACCATT TTTCTCTCAAACATGTCTGCGGGTTTCTGCATAAACAC AAAATAAAATAACAAAAAAACATTTAAACATTAGAAGCC TGTCTTACAACAGGAAAAACAACCCTTATAAGCATAAG ACGGACTACGGCCATGCCGGCGTGACCGTAAAAAAA CTGGTCACCGTGATTAAAAAGCACCACCGACAGCTCC TCGGTCATGTCCGGAGTCATAATGTAAGACTCGGTAA ACACATCAGGTTGATTCACATCGGTCAGTGCTAAAAA GCGACCGAAATAGCCCGGGGGAATACATACCCGCAG GCGTAGAGACAACATTACAGCCCCCATAGGAGGTATA ACAAAATTAATAGGAGAGAAAAACACATAAACACCTGA AAAACCCTCCTGCCTAGGCAAAATAGCACCCTCCCGC TCCAGAACAACATACAGCGCTTCCACAGCGGCAGCCA TAACAGTCAGCCTTACCAGTAAAAAAGAAAACCTATTA AAAAAACACCACTCGACACGGCACCAGCTCAATCAGT CACAGTGTAAAAAAGGGCCAAGTGCAGAGCGAGTATA TATAGGACTAAAAAATGACGTAACGGTTAAAGTCCACA AAAAACACCCAGAAAACCGCACGCGAACCTACGCCCA GAAACGAAAGCCAAAAAACCCACAACTTCCTCAAATC GTCACTTCCGTTTTCCCACGTTACGTCACTTCCCATTT TAAGAAAACTACAATTCCCAACACATACAAGTTACTCC GCCCTTAATTAAATCGGATCCGATATCTAGATGTATTC GCGAGGTACCGAGCTCGAATTCTCTGGCCGTCGTTTT ACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAA CTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCT GGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTT CCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCC TGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATT TCACACCGCATATGGTGCACTCTCAGTACAATCTGCT CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGC CAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGC TCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCA CCGAAACGCGCGA
49	pEMBR-1.7 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT

		GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCTACGTAATCCGT AGATGTACCTGGACATCCAGGTGATGCCGGCGGCGG TGGTGGAGGCGCGCGGAAAGTCGCGGACGCGGTTCC AGATGTTGCGCAGCGGCAAAAAGTGCTCCATGGTCG

		GGACGCTCTGGCCGGTGAGGCGTGCGCAGTCGTTGA CGCTCTAGACCGTGCAAAAGGAGAGCCTGTAAGCGG GCACTCTTCCGTGGTCTGGTGGATAAATTCGCAAGGG TATCATGGCGGACGACCGGGGTTCGAACCCCGGATC CGGCCGTCCGCCGTGATCCATGCGGTTACCGCCCGC GTGTCGAACCCAGGTGTGCGACGTCAGACAACGGGG GAGCGCTCCTTTTGGCTTCCTTCCAGGCGCGGCGGC TGCTGCGCTAGCTTTTTTGGCCACTGGCCGCGCGCG GCGTAAGCGGTTAGGCTGGAAAGCGAAAGCATTAAGT GGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAAGG GTTGAGTCGCAGGACCCCCGGTTCGAGTCTCGGGCC GGCCGGACTGCGGCGAACGGGGGTTTGCCTCCCCGT CATGCAAGACCCCGCTTGCAAATTCCTCCGGAAACAG GGACGAGCCCCTTTTTTGCTTTTCCCAGATGCATCCG GTGCTGCGGCAGATGCGCCCCCCTCCTCAGCAGCGG CAAGAGCAAGAGCAGCGGCAGACATGCAGGGCACCC TCCCCTTCTCCTACCGCGTCAGGAGGGGCAACATCGA TCCAGACATGATAAGATACATTGATGAGTTTGGACAAA CCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGT GAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATA AGCTGCAATAAACAAGTTTGTACACTCTCGGGTGATTA TTTACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATC AAAGGGGTTCTGCCGCGCATCGCTATGCGCCACTGG CAGGGACACGTTGCGATACTGGTGTTTAGTGCTCCAC TTAAACTCAGGCACAACCATCCGCGGCAGCTCGGTGA AGTTTTCACTCCACAGGCTGCGCACCATCACCAACGC GTTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAG TTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATAC ACAGGGTTGCAGCACTGGAACACTATCAGCGCCGGG TGGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATC AGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCG AACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGG GCGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTA GTGGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGT TAGGATACAGCGCCTGCATAAAAGCCTTGATCTGCTT AAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAAC ATGCCGCAAGACTTGCCGGAAAACTGATTGGCCGGAC AGGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGT TGGAGATCTGCACCACATTTCGGCCCCACCGGTTCTT CACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCG CGCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCAC GTGCTCCTTATTTATCATAATGCTTCCGTGTAGACACT TAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC

		AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATGGTGGCTTCCTTCTCCTATAGGCAGGGCGCGCC CGCCCGCCGCGCGCTTCGCTTTTTATAGGGCCGCCG CCGCCGCCGCCTCGCCATAAAAGGAAACTTTCGGAG CGCGCCGCTCTGATTGGCTGCCGCCGCACCTCTCCG CCTCGCCCCGCCCCGCCCCTCGCCCCGCCCCGCCCC GCCTGGCGCGCGCCCCCCCCCCCCCCCGCCCCCAT CGCTGCACAAAATAATTAAAAAATAAATAAATACAAAAT TGGGGGTGGGGAGGGGGGGGAGATGGGGAGAGTGA AGCAGAACGTGGGGCTCACCTCGAGGCCGGCCGAAT ATCTTCATTTAAATGTTTAAACATCGATGCGGCCGCAA CTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCA ATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCAC TGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA TCTTAGCTTAACGGGCGGCGAAGGAGAAGTCCACGC CTACATGGGGGTAGAGTCATAATCGTGCATCAGGATA GGGCGGTGGTGCTGCAGCAGCGCGCGAATAAACTGC TGCCGCCGCCGCTCCGTCCTGCAGGAATACAACATG GCAGTGGTCTCCTCAGCGATGATTCGCACCGCCCGC AGCATAAGGCGCCTTGTCCTCCGGGCACAGCAGCGC ACCCTGATCTCACTTAAATCAGCACAGTAACTGCAGCA CAGCACCACAATATTGTTCAAAATCCCACAGTGCAAG GCGCTGTATCCAAAGCTCATGGCGGGGACCACAGAA CCCACGTGGCCATCATACCACAAGCGCAGGTAGATTA AGTGGCGACCCCTCATAAACACGCTGGACATAAACAT TACCTCTTTTGGCATGTTGTAATTCACCACCTCCCGGT ACCATATAAACCTCTGATTAAACATGGCGCCATCCACC ACCATCCTAAACCAGCTGGCCAAAACCTGCCCGCCGG CTATACACTGCAGGGAACCGGGACTGGAACAATGACA GTGGAGAGCCCAGGACTCGTAACCATGGATCATCATG

		CTCGTCATGATATCAATGTTGGCACAACACAGGCACA CGTGCATACACTTCCTCAGGATTACAAGCTCCTCCCG CGTTAGAACCATATCCCAGGGAACAACCCATTCCTGA ATCAGCGTAAATCCCACACTGCAGGGAAGACCTCGCA CGTAACTCACGTTGTGCATTGTCAAAGTGTTACATTCG GGCAGCAGCGGATGATCCTCCAGTATGGTAGCGCGG GTTTCTGTCTCAAAAGGAGGTAGACGATCCCTACTGT ACGGAGTGCGCCGAGACAACCGAGATCGTGTTGGTC GTAGTGTCATGCCAAATGGAACGCCGGACGTAGTCAT ATTTCCTGAAGCAAAACCAGGTGCGGGCGTGACAAAC AGATCTGCGTCTCCGGTCTCGCCGCTTAGATCGCTCT GTGTAGTAGTTGTAGTATATCCACTCTCTCAAAGCATC CAGGCGCCCCCTGGCTTCGGGTTCTATGTAAACTCCT TCATGCGCCGCTGCCCTGATAACATCCACCACCGCAG AATAAGCCACACCCAGCCAACCTACACATTCGTTCTG CGAGTCACACACGGGAGGAGCGGGAAGAGCTGGAAG AACCATGGTGGCATTTGCAAAAGCCTAGGCCTCCAAA AAAGCCTCCTCACTACTTCTGGAATAGCTCAGAGGCC GAGGCGGCCTCGGCCTCTGCATAAATAAAAAAAATTA GTCAGCCATGGGGCGGAGAATGGGCGGAACTGGGC GGAGTTAGGGGCGGGATGGGCGGAGTTAGGGGCGG GACTATGGTTGCTGACTAATTGAGATGCATGCTTTGCA TACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACA CCTGGTTGCTGACTAATTGAGATGCATGCTTTGCATAC TTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCC TAACTGACACACACGTTACGTCACTTCCCATTTTAAGA AAACTACAATTCCCAACACATACAAGTTACTCCGCCCT TAATTAACATATGGTGCACTCTCAGTACAATCTGCTCT GATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCA ACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTC CCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCG GGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACC GAAACGCGCGA
50	SV40 polyA signal	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG CAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTC ACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATG TATCTTA
51	SV40 promoter upstream of UL30	GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCC CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTA GTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCC AGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAG TCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCC CGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCC CCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCG AGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTG AGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAA

52	Bovine growth hormone PolyA after UL42	CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC CTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACT CCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGT GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGA AGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTC TATGG
53	E4orf1 (3' to 5')	TTAAACATTAGAAGCCTGTCTTACAACAGGAAAAACAA CCCTTATAAGCATAAGACGGACTACGGCCATGCCGGC GTGACCGTAAAAAAACTGGTCACCGTGATTAAAAAGC ACCACCGACAGCTCCTCGGTCATGTCCGGAGTCATAA TGTAAGACTCGGTAAACACATCAGGTTGATTCACATCG GTCAGTGCTAAAAAGCGACCGAAATAGCCCGGGGGA ATACATACCCGCAGGCGTAGAGACAACATTACAGCCC CCATAGGAGGTATAACAAAATTAATAGGAGAGAAAAAC ACATAAACACCTGAAAAACCCTCCTGCCTAGGCAAAAT AGCACCCTCCCGCTCCAGAACAACATACAGCGCTTCC ACAGCGGCAGCCAT
54	E4orf1	ATGGCTGCCGCTGTGGAAGCGCTGTATGTTGTTCTGG AGCGGGAGGGTGCTATTTTGCCTAGGCAGGAGGGTT TTTCAGGTGTTTATGTGTTTTTCTCTCCTATTAATTTTG TTATACCTCCTATGGGGGCTGTAATGTTGTCTCTACGC CTGCGGGTATGTATTCCCCCGGGCTATTTCGGTCGCT TTTTAGCACTGACCGATGTGAATCAACCTGATGTGTTT ACCGAGTCTTACATTATGACTCCGGACATGACCGAGG AGCTGTCGGTGGTGCTTTTTAATCACGGTGACCAGTT TTTTTACGGTCACGCCGGCATGGCCGTAGTCCGTCTT ATGCTTATAAGGGTTGTTTTTCCTGTTGTAAGACAGGC TTCTAATGTTTAA
55	E4orf1	MAAAVEALYVVLEREGAILPRQEGFSGVYVFFSPINFVIP PMGAVMLSLRLRVCIPPGYFGRFLALTDVNQPDVFTESY IMTPDMTEELSVVLFNHGDQFFYGHAGMAVVRLMLIRVV FPVVRQASNV*
56	E4orf2 (3' to 5')	TCATGATTGCAAAAATTCAGGTTCCTCACAGACCTGTA TAAGATTCAAAAGCGGAACATTAACAAAAATACCGCGA TCCCGTAGGTCCCTTCGCAGGGCCAGCTGAACATAAT CGTGCAGGTCTGCACGGACCAGCGCGGCCACTTCCC CGCCAGGAACCATGACAAAAGAACCCACACTGATTAT GACACGCATACTCGGAGCTATGCTAACCAGCGTAGCC CCGATGTAAGCTTGTTGCATGGGCGGCGATATAAAAT GCAAGGTGCTGCTCAAAAAATCAGGCAAAGCCTCGCG CAAAAAAGAAAGCACATCGTAGTCATGCTCATGCAGA TAAAGGCAGGTAAGCTCCGGAACCACCACAGAAAAAG ACACCATTTTTCTCTCAAACATGTCTGCGGGTTTCTGC AT
57	E4orf2	ATGCAGAAACCCGCAGACATGTTTGAGAGAAAAATGG TGTCTTTTTCTGTGGTGGTTCCGGAGCTTACCTGCCTT

		TATCTGCATGAGCATGACTACGATGTGCTTTCTTTTTT GCGCGAGGCTTTGCCTGATTTTTTGAGCAGCACCTTG CATTTTATATCGCCGCCCATGCAACAAGCTTACATCGG GGCTACGCTGGTTAGCATAGCTCCGAGTATGCGTGTC ATAATCAGTGTGGGTTCTTTTGTCATGGTTCCTGGCG GGGAAGTGGCCGCGCTGGTCCGTGCAGACCTGCACG ATTATGTTCAGCTGGCCCTGCGAAGGGACCTACGGGA TCGCGGTATTTTTGTTAATGTTCCGCTTTTGAATCTTAT ACAGGTCTGTGAGGAACCTGAATTTTTGCAATCATGA
58	E4orf2	MQKPADMFERKMVSFSVVVPELTCLYLHEHDYDVLSFL REALPDFLSSTLHFISPPMQQAYIGATLVSIAPSMRVIISV GSFVMVPGGEVAALVRADLHDYVQLALRRDLRDRGIFV NVPLLNLIQVCEEPEFLQS*
59	E4orf3 (3' to 5')	TTATTCCAAAAGATTATCCAAAACCTCAAAATGAAGAT CTATTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTC AAACTCTACAGCCAAAGAACAGATAATGGCATTTGTAA GATGTTGCACAATGGCTTCCAAAAGGCAAACGGCCCT CACGTCCAAGTGGACGTAAAGGCTAAACCCTTCAGGG TGAATCTCCTCTATAAACATTCCAGCACCTTCAACCAT GCCCAAATAATTCTCATCTCGCCACCTTCTCAATATAT CTCTAAGCAAATCCCGAATATTAAGTCCGGCCATTGTA AAAATCTGCTCCAGAGCGCCCTCCACCTTCAGCCTCA AGCAGCGAATCAT
60	E4orf3	ATGATTCGCTGCTTGAGGCTGAAGGTGGAGGGCGCT CTGGAGCAGATTTTTACAATGGCCGGACTTAATATTCG GGATTTGCTTAGAGATATATTGAGAAGGTGGCGAGAT GAGAATTATTTGGGCATGGTTGAAGGTGCTGGAATGT TTATAGAGGAGATTCACCCTGAAGGGTTTAGCCTTTAC GTCCACTTGGACGTGAGGGCCGTTTGCCTTTTGGAAG CCATTGTGCAACATCTTACAAATGCCATTATCTGTTCT TTGGCTGTAGAGTTTGACCACGCCACCGGAGGGGAG CGCGTTCACTTAATAGATCTTCATTTTGAGGTTTTGGA TAATCTTTTGGAATAA
61	E4orf3	MIRCLRLKVEGALEQIFTMAGLNIRDLLRDILRRWRDENY LGMVEGAGMFIEEIHPEGFSLYVHLDVRAVCLLEAIVQHL TNAIICSLAVEFDHATGGERVHLIDLHFEVLDNLLE*
62	E4orf4 (3' to 5')	CTACTGTACGGAGTGCGCCGAGACAACCGAGATCGT GTTGGTCGTAGTGTCATGCCAAATGGAACGCCGGACG TAGTCATATTTCCTGAAGCAAAACCAGGTGCGGGCGT GACAAACAGATCTGCGTCTCCGGTCTCGCCGCTTAGA TCGCTCTGTGTAGTAGTTGTAGTATATCCACTCTCTCA AAGCATCCAGGCGCCCCCTGGCTTCGGGTTCTATGTA AACTCCTTCATGCGCCGCTGCCCTGATAACATCCACC ACCGCAGAATAAGCCACACCCAGCCAACCTACACATT CGTTCTGCGAGTCACACACGGGAGGAGCGGGAAGAG CTGGAAGAACCAT

63	E4orf4	ATGGTTCTTCCAGCTCTTCCCGCTCCTCCCGTGTGTG ACTCGCAGAACGAATGTGTAGGTTGGCTGGGTGTGG CTTATTCTGCGGTGGTGGATGTTATCAGGGCAGCGGC GCATGAAGGAGTTTACATAGAACCCGAAGCCAGGGG GCGCCTGGATGCTTTGAGAGAGTGGATATACTACAAC TACTACACAGAGCGATCTAAGCGGCGAGACCGGAGA CGCAGATCTGTTTGTCACGCCCGCACCTGGTTTTGCT TCAGGAAATATGACTACGTCCGGCGTTCCATTTGGCA TGACACTACGACCAACACGATCTCGGTTGTCTCGGCG CACTCCGTACAGTAG
64	E4orf4	MVLPALPAPPVCDSQNECVGWLGVAYSAVVDVIRAAAH EGVYIEPEARGRLDALREWIYYNYYTERSKRRDRRRRS VCHARTWFCFRKYDYVRRSIWHDTTTNTISVVSAHSVQ*
65	E4orf6 (3' to 5')	CTACATGGGGGTAGAGTCATAATCGTGCATCAGGATA GGGCGGTGGTGCTGCAGCAGCGCGCGAATAAACTGC TGCCGCCGCCGCTCCGTCCTGCAGGAATACAACATG GCAGTGGTCTCCTCAGCGATGATTCGCACCGCCCGC AGCATAAGGCGCCTTGTCCTCCGGGCACAGCAGCGC ACCCTGATCTCACTTAAATCAGCACAGTAACTGCAGCA CAGCACCACAATATTGTTCAAAATCCCACAGTGCAAG GCGCTGTATCCAAAGCTCATGGCGGGGACCACAGAA CCCACGTGGCCATCATACCACAAGCGCAGGTAGATTA AGTGGCGACCCCTCATAAACACGCTGGACATAAACAT TACCTCTTTTGGCATGTTGTAATTCACCACCTCCCGGT ACCATATAAACCTCTGATTAAACATGGCGCCATCCACC ACCATCCTAAACCAGCTGGCCAAAACCTGCCCGCCGG CTATACACTGCAGGGAACCGGGACTGGAACAATGACA GTGGAGAGCCCAGGACTCGTAACCATGGATCATCATG CTCGTCATGATATCAATGTTGGCACAACACAGGCACA CGTGCATACACTTCCTCAGGATTACAAGCTCCTCCCG CGTTAGAACCATATCCCAGGGAACAACCCATTCCTGA ATCAGCGTAAATCCCACACTGCAGGGAAGACCTCGCA CGTAACTCACGTTGTGCATTGTCAAAGTGTTACATTCG GGCAGCAGCGGATGATCCTCCAGTATGGTAGCGCGG GTTTCTGTCTCAAAAGGAGGTAGACGATCCCTACTGT ACGGAGTGCGCCGAGACAACCGAGATCGTGTTGGTC GTAGTGTCATGCCAAATGGAACGCCGGACGTAGTCAT
66	E4orf6	ATGACTACGTCCGGCGTTCCATTTGGCATGACACTAC GACCAACACGATCTCGGTTGTCTCGGCGCACTCCGTA CAGTAGGGATCGTCTACCTCCTTTTGAGACAGAAACC CGCGCTACCATACTGGAGGATCATCCGCTGCTGCCC GAATGTAACACTTTGACAATGCACAACGTGAGTTACGT GCGAGGTCTTCCCTGCAGTGTGGGATTTACGCTGATT CAGGAATGGGTTGTTCCCTGGGATATGGTTCTAACGC GGGAGGAGCTTGTAATCCTGAGGAAGTGTATGCACGT GTGCCTGTGTTGTGCCAACATTGATATCATGACGAGC ATGATGATCCATGGTTACGAGTCCTGGGCTCTCCACT

		GTCATTGTTCCAGTCCCGGTTCCCTGCAGTGTATAGC CGGCGGGCAGGTTTTGGCCAGCTGGTTTAGGATGGT GGTGGATGGCGCCATGTTTAATCAGAGGTTTATATGG TACCGGGAGGTGGTGAATTACAACATGCCAAAAGAGG TAATGTTTATGTCCAGCGTGTTTATGAGGGGTCGCCA CTTAATCTACCTGCGCTTGTGGTATGATGGCCACGTG GGTTCTGTGGTCCCCGCCATGAGCTTTGGATACAGCG CCTTGCACTGTGGGATTTTGAACAATATTGTGGTGCTG TGCTGCAGTTACTGTGCTGATTTAAGTGAGATCAGGG TGCGCTGCTGTGCCCGGAGGACAAGGCGCCTTATGC TGCGGGCGGTGCGAATCATCGCTGAGGAGACCACTG CCATGTTGTATTCCTGCAGGACGGAGCGGCGGCGGC AGCAGTTTATTCGCGCGCTGCTGCAGCACCACCGCCC TATCCTGATGCACGATTATGACTCTACCCCCATGTAG
67	E4orf6	MTTSGVPFGMTLRPTRSRLSRRTPYSRDRLPPFETETR ATILEDHPLLPECNTLTMHNVSYVRGLPCSVGFTLIQEW VVPWDMVLTREELVILRKCMHVCLCCANIDIMTSMMIHG YESWALHCHCSSPGSLQCIAGGQVLASWFRMVVDGAM FNQRFIWYREVVNYNMPKEVMFMSSVFMRGRHLIYLRL WYDGHVGSVVPAMSFGYSALHCGILNNIVVLCCSYCAD LSEIRVRCCARRTRRLMLRAVRIIAEETTAMLYSCRTERR RQQFIRALLQHHRPILMHDYDSTPM*
68	E4orf7	ATGACTACGTCCGGCGTTCCATTTGGCATGACACTAC GACCAACACGATCTCGGTTGTCTCGGCGCACTCCGTA CAGTAGGGATCGTCTACCTCCTTTTGAGACAGAAACC CGCGCTACCATACTGGAGGATCATCCGCTGCTGCCC GAATGTAACACTTTGACAATGCACAACGTGAGTTACGT GCGAGGTCTTCCCTGCAGTGTGGGATTTACGCTGATT CAGGAATGGGTTGTTCCCTGGGATATGGTTCTAACGC GGGAGGAGCTTGTAATCCTGAGGAAGTGTATGCACGT GTGCCTGTGTTGTGCCAACATTGATATCATGACGAGC ATGATGATCCATGGTTACGAGTCCTGGGCTCTCCACT GTCATTGTTCCAGTCCCGGTTCCCTGCAGTGTATAGC CGGCGGGCAGGTTTTGGCCAGCTGGTTTAGGATGGT GGTGGATGGCGCCATGTTTAATCAGAGGTTTATATGG TACCGGGAGGTGGTGAATTACAACATGCCAAAAGAGG TAATGTTTATGTCCAGCGTGTTTATGAGGGGTCGCCA CTTAATCTACCTGCGCTTGTGGTATGATGGCCACGTG GGTTCTGTGGTCCCCGCCATGAGCTTTGGATACAGCG CCTTGCACTGTGGGATTTTGAACAATATTGTGGTGCTG TGCTGCAGTTACTGTGCTGATTTAAGTGAGATCAGGG TGCGCTGCTGTGCCCGGAGGACAAGGCGCCTTATGC TGCGGGCGGTGCGAATCATCGCTGAGGAGACCACTG CCATGTTGTATTCCTGCAGGACGGAGCGGCGGCGGC AGCAGTTTATTCGCGCGCTGCTGCAGCACCACCGCCC TATCCTGATGCACGATTATGACTCTACCCCCATGTAGG CGTGGACTTCTCCTTCGCCGCCCGTTAAGCAACCGCA

		AGTTGGACAGCAGCCTGTGGCTCAGCAGCTGGACAG CGACATGAACTTAAGTGAGCTGCCCGGGGAGTTTATT AATATCACTGATGAGCGTTTGGCTCGACAGGAAACCG TGTGGAATATAACACCTAAGAATATGTCTGTTACCCAT GATATGATGCTTTTTAAGGCCAGCCGGGGAGAAAGGA CTGTGTACTCTGTGTGTTGGGAGGGAGGTGGCAGGTT GAATACTAGGGTTCTGTGA
69	E4orf7	MTTSGVPFGMTLRPTRSRLSRRTPYSRDRLPPFETETR ATILEDHPLLPECNTLTMHNAWTSPSPPVKQPQVGQQP VAQQLDSDMNLSELPGEFINITDERLARQETVWNITPKN MSVTHDMMLFKASRGERTVYSVCWEGGGRLNTRVL
70	pEMBR-1.2B2 Entire sequence	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT

		AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTATCCG TAGATGTACCTGGACATCCAGGTGATGCCGGCGGCG GTGGTGGAGGCGCGCGGAAAGTCGCGGACGCGGTT CCAGATGTTGCGCAGCGGCAAAAAGTGCTCCATGGTC GGGACGCTCTGGCCGGTGAGGCGTGCGCAGTCGTTG ACGCTCTAGACCGTGCAAAAGGAGAGCCTGTAAGCG GGCACTCTTCCGTGGTCTGGTGGATAAATTCGCAAGG GTATCATGGCGGACGACCGGGGTTCGAACCCCGGAT CCGGCCGTCCGCCGTGATCCATGCGGTTACCGCCCG CGTGTCGAACCCAGGTGTGCGACGTCAGACAACGGG GGAGCGCTCCTTTTGGCTTCCTTCCAGGCGCGGCGG CTGCTGCGCTAGCTTTTTTGGCCACTGGCCGCGCGC GGCGTAAGCGGTTAGGCTGGAAAGCGAAAGCATTAA GTGGCTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAA GGGTTGAGTCGCAGGACCCCCGGTTCGAGTCTCGGG CCGGCCGGACTGCGGCGAACGGGGGTTTGCCTCCCC GTCATGCAAGACCCCGCTTGCAAATTCCTCCGGAAAC AGGGACGAGCCCCTTTTTTGCTTTTCCCAGATGCATC CGGTGCTGCGGCAGATGCGCCCCCCTCCTCAGCAGC GGCAAGAGCAAGAGCAGCGGCAGACATGCAGGGCAC CCTCCCCTTCTCCTACCGCGTCAGGAGGGGCAACATC GATCCAGACATGATAAGATACATTGATGAGTTTGGACA AACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTT GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA TAAGCTGCAATAAACAAGTTTGTACACTCTCGGGTGAT TATTTACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAA TCAAAGGGGTTCTGCCGCGCATCGCTATGCGCCACTG

		GCAGGGACACGTTGCGATACTGGTGTTTAGTGCTCCA CTTAAACTCAGGCACAACCATCCGCGGCAGCTCGGTG AAGTTTTCACTCCACAGGCTGCGCACCATCACCAACG CGTTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCA GTTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATA CACAGGGTTGCAGCACTGGAACACTATCAGCGCCGG GTGGTGCACGCTGGCCAGCACGCTCTTGTCGGAGAT CAGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGC GAACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAG GGCGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGT AGTGGCATCAAAAGGTGACCGTGCCCGGTCTGGGCG TTAGGATACAGCGCCTGCATAAAAGCCTTGATCTGCTT AAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAAC ATGCCGCAAGACTTGCCGGAAAACTGATTGGCCGGAC AGGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGT TGGAGATCTGCACCACATTTCGGCCCCACCGGTTCTT CACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCG CGCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCAC GTGCTCCTTATTTATCATAATGCTTCCGTGTAGACACT TAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCATGGAG TCAGTCGAGAAGAAGGACAGCCTAACCGCCCCCTCTG AGTTCGCCACCACCGCCTCCACCGATGCCGCCAACG CGCCTACCACCTTCCCCGTCGAGGCACCCCCGCTTG AGGAGGAGGAAGTGATTATCGAGCAGGACCCAGGTTT

		TGTAAGCGAAGACGACGAGGACCGCTCAGTACCAACA GAGGATAAAAAGCAAGACCAGGACAACGCAGAGGCA AACGAGGAACAAGTCGGGCGGGGGGACGAAAGGCAT GGCGACTACCTAGATGTGGGAGACGACGTGCTGTTG AAGCATCTGCAGCGCCAGTGCGCCATTATCTGCGACG CGTTGCAAGAGCGCAGCGATGTGCCCCTCGCCATAG CGGATGTCAGCCTTGCCTACGAACGCCACCTATTCTC ACCGCGCGTACCCCCCAAACGCCAAGAAAACGGCAC ATGCGAGCCCAACCCGCGCCTCAACTTCTACCCCGTA TTTGCCGTGCCAGAGGTGCTTGCCACCTATCACATCT TTTTCCAAAACTGCAAGATACCCCTATCCTGCCGTGCC AACCGCAGCCGAGCGGACAAGCAGCTGGCCTTGCGG CAGGGCGCTGTCATACCTGATATCGCCTCGCTCAACG AAGTGCCAAAAATCTTTGAGGGTCTTGGACGCGACGA GAAGCGCGCGGCAAACGCTCTGCAACAGGAAAACAG CGAAAATGAAAGTCACTCTGGAGTGTTGGTGGAACTC GAGGGTGACAACGCGCGCCTAGCCGTACTAAAACGC AGCATCGAGGTCACCCACTTTGCCTACCCGGCACTTA ACCTACCCCCCCAAGGTCATGAGCACAGTCATGAGTGA GCTGATCGTGCGCCGTGCGCAGCCCCTGGAGAGGGA TGCAAATTTGCAAGAACAAACAGAGGAGGGCCTACCC GCAGTTGGCGACGAGCAGCTAGCGCGCTGGCTTCAA ACGCGCGAGCCTGCCGACTTGGAGGAGCGACGCAAA CTAATGATGGCCGCAGTGCTCGTTACCGTGGAGCTTG AGTGCATGCAGCGGTTCTTTGCTGACCCGGAGATGCA GCGCAAGCTAGAGGAAACATTGCACTACACCTTTCGA CAGGGCTACGTACGCCAGGCCTGCAAGATCTCCAAC GTGGAGCTCTGCAACCTGGTCTCCTACCTTGGAATTT TGCACGAAAACCGCCTTGGGCAAAACGTGCTTCATTC CACGCTCAAGGGCGAGGCGCGCCGCGACTACGTCCG CGACTGCGTTTACTTATTTCTATGCTACACCTGGCAGA CGGCCATGGGCGTTTGGCAGCAGTGCTTGGAGGAGT GCAACCTCAAGGAGCTGCAGAAACTGCTAAAGCAAAA CTTGAAGGACCTATGGACGGCCTTCAACGAGCGCTCC GTGGCCGCGCACCTGGCGGACATCATTTTCCCCGAA CGCCTGCTTAAAACCCTGCAACAGGGTCTGCCAGACT TCACCAGTCAAAGCATGTTGCAGAACTTTAGGAACTTT ATCCTAGAGCGCTCAGGAATCTTGCCCGCCACCTGCT GTGCACTTCCTAGCGACTTTGTGCCCATTAAGTACCG CGAATGCCCTCCGCCGCTTTGGGGCCACTGCTACCTT CTGCAGCTAGCCAACTACCTTGCCTACCACTCTGACA TAATGGAAGACGTGAGCGGTGACGGTCTACTGGAGT GTCACTGTCGCTGCAACCTATGCACCCCGCACCGCTC CCTGGTTTGCAATTCGCAGCTGCTTAACGAAAGTCAA ATTATCGGTACCTTTGAGCTGCAGGGTCCCTCGCCTG ACGAAAAGTCCGCGGCTCCGGGGTTGAAACTCACTCC GGGGCTGTGGACGTCGGCTTACCTTCGCAAATTTGTA

		CCTGAGGACTACCACGCCCACGAGATTAGGTTCTACG AAGACCAATCCCGCCCGCCTAATGCGGAGCTTACCGC CTGCGTCATTACCCAGGGCCACATTCTTGGCCAATTG CAAGCCATCAACAAAGCCCGCCAAGAGTTTCTGCTAC GAAAGGGACGGGGGGTTTACTTGGACCCCCAGTCCG GCGAGGAGCTCAACCCAATCCCCCCGCCGCCGCAGC CCTATCAGCAGCAGCCGCGGGCCCTTGCTTCCCAGG ATGGCACCCAAAAAGAAGCTGCAGCTGCCGCCGCCA CCCACGGACGAGGAGGAATACTGGGACAGTCAGGCA GAGGAGGTTTTGGACGAGGAGGAGGAGGACATGATG GAAGACTGGGAGAGCCTAGACGAGGAAGCTTCCGAG GTCGAAGAGGTGTCAGACGAAACACCGTCACCCTCG GTCGCATTCCCCTCGCCGGCGCCCCAGAAATCGGCA ACCGGTTCCAGCATGGCTACAACCTCCGCTCCTCAGG CGCCGCCGGCACTGCCCGTTCGCCGACCCAACCGTA GATGGGACACCACTGGAACCAGGGCCGGTAAGTCCA AGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGC GCCAAGGCTACCGCTCATGGCGCGGGCACAAGAACG CCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACAT CTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGC GTGGCCTTCCCCCGTAACATCCTGCATTACTACCGTC ATCTCTACAGCCCATACTGCACCGGCGGCAGCGGCA GCAACAGCAGCGGCCACACAGAAGCAAAGGCGACCG GATAGCAAGACTCTGACAAAGCCCAAGAAATCCACAG CGGCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTG GCGCCCAACGAACCCGTATCGACCCGCGAGCTTAGA AACAGGATTTTTCCCACTCTGTATGCTATATTTCAACA GAGCAGGGGCCAAGAACAAGAGCTGAAAATAAAAAAC AGGTCTCTGCGATCCCTCACCCGCAGCTGCCTGTATC ACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAAG ACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCTGAC TCTTAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAAG CGCGAAAACTACGTCATCTCCAGCGGCCACACCCGG CGCCAGCACCTGTTGTCAGCGCCATTATGAGCAAGGA AATTCCCACGCCCTACATGTGGAGTTACCAGCCACAA ATGGGACTTGCGGCTGGAGCTGCCCAAGACTACTCAA CCCGAATAAACTACATGAGCGCGGGGCGGCCGCCGT TTGTGTTATGTTTCAACGTGTTTATTTTTCAATTGCAGA AAATTTCAAGTCATTTTTCATTCAGTAGTATAGCCCCA CCACCACATAGCTTATACAGATCACCGTACCTTAATCA AACTCACAGAACCCTAGTATTCAACCTGCCACCTCCCT CCCAACACACAGAGTACACAGTCCTTTCTCCCCGGCT GGCCTTAAAAAGCATCATATCATGGGTAACAGACATAT TCTTAGGTGTTATATTCCACACGGTTTCCTGTCGAGCC AAACGCTCATCAGTGATATTAATAAACTCCCCGGGCA GCTCACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGC CACAGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGC

		GGCGAAGGAGAAGTCCACGCCTACATGGGGGTAGAG TCATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCA GCAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCG TCCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGC GATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTC CTCCGGGCACAGCAGCGCACCCTGATCTCACTTAAAT CAGCACAGTAACTGCAGCACAGCACCACAATATTGTT CAAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTC ATGGCGGGGACCACAGAACCCACGTGGCCATCATAC CACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAA ACACGCTGGACATAAACATTACCTCTTTTGGCATGTTG TAATTCACCACCTCCCGGTACCATATAAACCTCTGATT AAACATGGCGCCATCCACCACCATCCTAAACCAGCTG GCCAAAACCTGCCCGCCGGCTATACACTGCAGGGAA CCGGGACTGGAACAATGACAGTGGAGAGCCCAGGAC TCGTAACCATGGATCATCATGCTCGTCATGATATCAAT GTTGGCACAACACAGGCACACGTGCATACACTTCCTC AGGATTACAAGCTCCTCCCGCGTTAGAACCATATCCC AGGGAACAACCCATTCCTGAATCAGCGTAAATCCCAC ACTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGC ATTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGAT CCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGG AGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGA CAACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAAT GGAACGCCGGACGTAGTCATATTTCCTGAAGCAAAAC CAGGTGCGGGCGTGACAAACAGATCTGCGTCTCCGG TCTCGCCGCTTAGATCGCTCTGTGTAGTAGTTGTAGTA TATCCACTCTCTCAAAGCATCCAGGCGCCCCCTGGCT TCGGGTTCTATGTAAACTCCTTCATGCGCCGCTGCCC TGATAACATCCACCACCGCAGAATAAGCCACACCCAG CCAACCTACACATTCGTTCTGCGAGTCACACACGGGA GGAGCGGGAAGAGCTGGAAGAACCATGTTTTTTTTTT TATTCCAAAAGATTATCCAAAACCTCAAAATGAAGATC TATTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCA AACTCTACAGCCAAAGAACAGATAATGGCATTTGTAAG ATGTTGCACAATGGCTTCCAAAAGGCAAACGGCCCTC ACGTCCAAGTGGACGTAAAGGCTAAACCCTTCAGGGT GAATCTCCTCTATAAACATTCCAGCACCTTCAACCATG CCCAAATAATTCTCATCTCGCCACCTTCTCAATATATC TCTAAGCAAATCCCGAATATTAAGTCCGGCCATTGTAA AAATCTGCTCCAGAGCGCCCTCCACCTTCAGCCTCAA GCAGCGAATCATGATTGCAAAAATTCAGGTTCCTCACA GACCTGTATAAGATTCAAAAGCGGAACATTAACAAAAA TACCGCGATCCCGTAGGTCCCTTCGCAGGGCCAGCT GAACATAATCGTGCAGGTCTGCACGGACCAGCGCGG CCACTTCCCCGCCAGGAACCATGACAAAAGAACCCAC ACTGATTATGACACGCATACTCGGAGCTATGCTAACC

		AGCGTAGCCCCGATGTAAGCTTGTTGCATGGGCGGC GATATAAAATGCAAGGTGCTGCTCAAAAAATCAGGCA AAGCCTCGCGCAAAAAAGAAAGCACATCGTAGTCATG CTCATGCAGATAAAGGCAGGTAAGCTCCGGAACCACC ACAGAAAAAGACACCATTTTTCTCTCAAACATGTCTGC GGGTTTCTGCATAAACACAAAATAAAATAACAAAAAAA CATTTAAACATTAGAAGCCTGTCTTACAACAGGAAAAA CAACCCTTATAAGCATAAGACGGACTACGGCCATGCC GGCGTGACCGTAAAAAAACTGGTCACCGTGATTAAAA AGCACCACCGACAGCTCCTCGGTCATGTCCGGAGTCA TAATGTAAGACTCGGTAAACACATCAGGTTGATTCACA TCGGTCAGTGCTAAAAAGCGACCGAAATAGCCCGGG GGAATACATACCCGCAGGCGTAGAGACAACATTACAG CCCCCATAGGAGGTATAACAAAATTAATAGGAGAGAA AAACACATAAACACCTGAAAAACCCTCCTGCCTAGGC AAAATAGCACCCTCCCGCTCCAGAACAACATACAGCG CTTCCACAGCGGCAGCCATAACAGTCAGCCTTACCAG TAAAAAAGAAAACCTATTAAAAAAACACCACTCGACAC GGCACCAGCTCAATCAGTCACAGTGTAAAAAAGGGCC AAGTGCAGAGCGAGTATATATAGGACTAAAAAATGAC GTAACGGTTAAAGTCCACAAAAAACACCCAGAAAACC GCACGCGAACCTACGCCCAGAAACGAAAGCCAAAAAA CCCACAACTTCCTCAAATCGTCACTTCCGTTTTCCCAC GTTACGTCACTTCCCATTTTAAGAAAACTACAATTCCC AACACATACAAGTTACTCCGCCCTTAATTAAATCGGAT CCGATATCTAGATGTATTCGCGAGGTACCGAGCTCGA ATTCTCTGGCCGTCGTTTTACAACGTCGTGACTGGGA AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGG CCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCT GAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTT ACGCATCTGTGCGGTATTTCACACCGCATATGGTGCA CTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGAC AAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGA GGTTTTCACCGTCATCACCGAAACGCGCGA
71	pHelper Entire sequence	GGTACCCAACTCCATGCTTAACAGTCCCCAGGTACAG CCCACCCTGCGTCGCAACCAGGAACAGCTCTACAGCT TCCTGGAGCGCCACTCGCCCTACTTCCGCAGCCACA GTGCGCAGATTAGGAGCGCCACTTCTTTTTGTCACTT GAAAAACATGTAAAAATAATGTACTAGGAGACACTTTC AATAAAGGCAAATGTTTTTATTTGTACACTCTCGGGTG ATTATTTACCCCCCACCCTTGCCGTCTGCGCCGTTTAA AAATCAAAGGGGTTCTGCCGCGCATCGCTATGCGCCA CTGGCAGGGACACGTTGCGATACTGGTGTTTAGTGCT CCACTTAAACTCAGGCACAACCATCCGCGGCAGCTCG

		GTGAAGTTTTCACTCCACAGGCTGCGCACCATCACCA ACGCGTTTAGCAGGTCGGGCGCCGATATCTTGAAGTC GCAGTTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCG ATACACAGGGTTGCAGCACTGGAACACTATCAGCGCC GGGTGGTGCACGCTGGCCAGCACGCTCTTGTCGGAG ATCAGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGG GCGAACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAA AGGGTGCATGCCCAGGCTTTGAGTTGCACTCGCACC GTAGTGGCATCAGAAGGTGACCGTGCCCGGTCTGGG CGTTAGGATACAGCGCCTGCATGAAAGCCTTGATCTG CTTAAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAG AACATGCCGCAAGACTTGCCGGAAAACTGATTGGCCG GACAGGCCGCGTCATGCACGCAGCACCTTGCGTCGG TGTTGGAGATCTGCACCACATTTCGGCCCCACCGGTT CTTCACGATCTTGGCCTTGCTAGACTGCTCCTTCAGC GCGCGCTGCCCGTTTTCGCTCGTCACATCCATTTCAA TCACGTGCTCCTTATTTATCATAATGCTCCCGTGTAGA CACTTAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCA GCCACAACGCGCAGCCCGTGGGCTCGTGGTGCTTGT AGGTTACCTCTGCAAACGACTGCAGGTACGCCTGCAG GAATCGCCCCATCATCGTCACAAAGGTCTTGTTGCTG GTGAAGGTCAGCTGCAACCCGCGGTGCTCCTCGTTTA GCCAGGTCTTGCATACGGCCGCCAGAGCTTCCACTTG GTCAGGCAGTAGCTTGAAGTTTGCCTTTAGATCGTTAT CCACGTGGTACTTGTCCATCAACGCGCGCGCAGCCTC CATGCCCTTCTCCCACGCAGACACGATCGGCAGGCTC AGCGGGTTTATCACCGTGCTTTCACTTTCCGCTTCACT GGACTCTTCCTTTTCCTCTTGCGTCCGCATACCCCGC GCCACTGGGTCGTCTTCATTCAGCCGCCGCACCGTG CGCTTACCTCCCTTGCCGTGCTTGATTAGCACCGGTG GGTTGCTGAAACCCACCATTTGTAGCGCCACATCTTC TCTTTCTTCCTCGCTGTCCACGATCACCTCTGGGGAT GGCGGGCGCTCGGGCTTGGGAGAGGGGCGCTTCTTT TTCTTTTTGGACGCAATGGCCAAATCCGCCGTCGAGG TCGATGGCCGCGGGCTGGGTGTGCGCGGCACCAGC GCATCTTGTGACGAGTCTTCTTCGTCCTCGGACTCGA GACGCCGCCTCAGCCGCTTTTTTGGGGGCGCGCGGG GAGGCGGCGGCGACGGCGACGGGGACGACACGTCC TCCATGGTTGGTGGACGTCGCGCCGCACCGCGTCCG CGCTCGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGA CTGGCCATTTCCTTCTCCTATAGGCAGAAAAAGATCAT GGAGTCAGTCGAGAAGGAGGACAGCCTAACCGCCCC CTTTGAGTTCGCCACCACCGCCTCCACCGATGCCGCC AACGCGCCTACCACCTTCCCCGTCGAGGCACCCCCG CTTGAGGAGGAGGAAGTGATTATCGAGCAGGACCCA GGTTTTGTAAGCGAAGACGACGAGGATCGCTCAGTAC CAACAGAGGATAAAAAGCAAGACCAGGACGACGCAG

		AGGCAAACGAGGAACAAGTCGGGCGGGGGGACCAAA GGCATGGCGACTACCTAGATGTGGGAGACGACGTGC TGTTGAAGCATCTGCAGCGCCAGTGCGCCATTATCTG CGACGCGTTGCAAGAGCGCAGCGATGTGCCCCTCGC CATAGCGGATGTCAGCCTTGCCTACGAACGCCACCTG TTCTCACCGCGCGTACCCCCCAAACGCCAAGAAAACG GCACATGCGAGCCCAACCCGCGCCTCAACTTCTACCC CGTATTTGCCGTGCCAGAGGTGCTTGCCACCTATCAC ATCTTTTTCCAAAACTGCAAGATACCCCTATCCTGCCG TGCCAACCGCAGCCGAGCGGACAAGCAGCTGGCCTT GCGGCAGGGCGCTGTCATACCTGATATCGCCTCGCT CGACGAAGTGCCAAAAATCTTTGAGGGTCTTGGACGC GACGAGAAACGCGCGGCAAACGCTCTGCAACAAGAA AACAGCGAAAATGAAAGTCACTGTGGAGTGCTGGTGG AACTTGAGGGTGACAACGCGCGCCTAGCCGTGCTGA AACGCAGCATCGAGGTCACCCACTTTGCCTACCCGGC ACTTAACCTACCCCCCAAGGTTATGAGCACAGTCATG AGCGAGCTGATCGTGCGCCGTGCACGACCCCTGGAG AGGGATGCAAACTTGCAAGAACAAACCGAGGAGGGC CTACCCGCAGTTGGCGATGAGCAGCTGGCGCGCTGG CTTGAGACGCGCGAGCCTGCCGACTTGGAGGAGCGA CGCAAGCTAATGATGGCCGCAGTGCTTGTTACCGTGG AGCTTGAGTGCATGCAGCGGTTCTTTGCTGACCCGGA GATGCAGCGCAAGCTAGAGGAAACGTTGCACTACACC TTTCGCCAGGGCTACGTGCGCCAGGCCTGCAAAATTT CCAACGTGGAGCTCTGCAACCTGGTCTCCTACCTTGG AATTTTGCACGAAAACCGCCTCGGGCAAAACGTGCTT CATTCCACGCTCAAGGGCGAGGCGCGCCGCGACTAC GTCCGCGACTGCGTTTACTTATTTCTGTGCTACACCTG GCAAACGGCCATGGGCGTGTGGCAGCAATGCCTGGA GGAGCGCAACCTAAAGGAGCTGCAGAAGCTGCTAAA GCAAAACTTGAAGGACCTATGGACGGCCTTCAACGAG CGCTCCGTGGCCGCGCACCTGGCGGACATTATCTTC CCCGAACGCCTGCTTAAAACCCTGCAACAGGGTCTGC CAGACTTCACCAGTCAAAGCATGTTGCAAAACTTTAGG AACTTTATCCTAGAGCGTTCAGGAATTCTGCCCGCCA CCTGCTGTGCGCTTCCTAGCGACTTTGTGCCCATTAA GTACCGTGAATGCCCTCCGCCGCTTTGGGGTCACTGC TACCTTCTGCAGCTAGCCAACTACCTTGCCTACCACTC CGACATCATGGAAGACGTGAGCGGTGACGGCCTACT GGAGTGTCACTGTCGCTGCAACCTATGCACCCCGCAC CGCTCCCTGGTCTGCAATTCGCAACTGCTTAGCGAAA GTCAAATTATCGGTACCTTTGAGCTGCAGGGTCCCTC GCCTGACGAAAAGTCCGCGGCTCCGGGGTTGAAACT CACTCCGGGGCTGTGGACGTCGGCTTACCTTCGCAAA TTTGTACCTGAGGACTACCACGCCCACGAGATTAGGT TCTACGAAGACCAATCCCGCCCGCCAAATGCGGAGCT

		TACCGCCTGCGTCATTACCCAGGGCCACATCCTTGGC CAATTGCAAGCCATCAACAAAGCCCGCCAAGAGTTTC TGCTACGAAAGGGACGGGGGGTTTACCTGGACCCCC AGTCCGGCGAGGAGCTCAACCCAATCCCCCCGCCGC CGCAGCCCTATCAGCAGCCGCGGGCCCTTGCTTCCC AGGATGGCACCCAAAAAGAAGCTGCAGCTGCCGCCG CCGCCACCCACGGACGAGGAGGAATACTGGGACAGT CAGGCAGAGGAGGTTTTGGACGAGGAGGAGGAGATG ATGGAAGACTGGGACAGCCTAGACGAAGCTTCCGAG GCCGAAGAGGTGTCAGACGAAACACCGTCACCCTCG GTCGCATTCCCCTCGCCGGCGCCCCAGAAATTGGCA ACCGTTCCCAGCATCGCTACAACCTCCGCTCCTCAGG CGCCGCCGGCACTGCCTGTTCGCCGACCCAACCGTA GATGGGACACCACTGGAACCAGGGCCGGTAAGTCTA AGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGC GCCAAGGCTACCGCTCGTGGCGCGGGCACAAGAACG CCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACAT CTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGC GTGGCCTTCCCCCGTAACATCCTGCATTACTACCGTC ATCTCTACAGCCCCTACTGCACCGGCGGCAGCGGCA GCGGCAGCAACAGCAGCGGTCACACAGAAGCAAAGG CGACCGGATAGCAAGACTCTGACAAAGCCCAAGAAAT CCACAGCGGCGGCAGCAGCAGGAGGAGGAGCGCTG CGTCTGGCGCCCAACGAACCCGTATCGACCCGCGAG CTTAGAAATAGGATTTTTCCCACTCTGTATGCTATATTT CAACAAAGCAGGGGCCAAGAACAAGAGCTGAAAATAA AAAACAGGTCTCTGCGCTCCCTCACCCGCAGCTGCCT GTATCACAAAAGCGAAGATCAGCTTCGGCGCACGCTG GAAGACGCGGAGGCTCTCTTCAGCAAATACTGCGCG CTGACTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAAT TTAAGCGCGAAAACTACGTCATCTCCAGCGGCCACAC CCGGCGCCAGCACCTGTCGTCAGCGCCATTATGAGC AAGGAAATTCCCACGCCCTACATGTGGAGTTACCAGC CACAAATGGGACTTGCGGCTGGAGCTGCCCAAGACTA CTCAACCCGAATAAACTACATGAGCGCGGGACCCCAC ATGATATCCCGGGTCAACGGAATCCGCGCCCACCGAA ACCGAATTCTCCTCGAACAGGCGGCTATTACCACCAC ACCTCGTAATAACCTTAATCCCCGTAGTTGGCCCGCT GCCCTGGTGTACCAGGAAAGTCCCGCTCCCACCACT GTGGTACTTCCCAGAGACGCCCAGGCCGAAGTTCAG ATGACTAACTCAGGGGCGCAGCTTGCGGGCGGCTTT CGTCACAGGGTGCGGTCGCCCGGGCGTTTTAGGGCG GAGTAACTTGCATGTATTGGGAATTGTAGTTTTTTTAA AATGGGAAGTGACGTATCGTGGGAAAACGGAAGTGAA GATTTGAGGAAGTTGTGGGTTTTTTGGCTTTCGTTTCT GGGCGTAGGTTCGCGTGCGGTTTTCTGGGTGTTTTTT GTGGACTTTAACCGTTACGTCATTTTTTAGTCCTATATA

		TACTCGCTCTGTACTTGGCCCTTTTTACACTGTGACTG ATTGAGCTGGTGCCGTGTCGAGTGGTGTTTTTTAATA GGTTTTTTTACTGGTAAGGCTGACTGTTATGGCTGCC GCTGTGGAAGCGCTGTATGTTGTTCTGGAGCGGGAG GGTGCTATTTTGCCTAGGCAGGAGGGTTTTTCAGGTG TTTATGTGTTTTTCTCTCCTATTAATTTTGTTATACCTC CTATGGGGGCTGTAATGTTGTCTCTACGCCTGCGGGT ATGTATTCCCCCGGGCTATTTCGGTCGCTTTTTAGCAC TGACCGATGTTAACCAACCTGATGTGTTTACCGAGTCT TACATTATGACTCCGGACATGACCGAGGAACTGTCGG TGGTGCTTTTTAATCACGGTGACCAGTTTTTTTACGGT CACGCCGGCATGGCCGTAGTCCGTCTTATGCTTATAA GGGTTGTTTTTCCTGTTGTAAGACAGGCTTCTAATGTT TAAATGTTTTTTTTTTTGTTATTTTATTTTGTGTTTAATG CAGGAACCCGCAGACATGTTTGAGAGAAAAATGGTGT CTTTTTCTGTGGTGGTTCCGGAACTTACCTGCCTTTAT CTGCATGAGCATGACTACGATGTGCTTGCTTTTTTGCG CGAGGCTTTGCCTGATTTTTTGAGCAGCACCTTGCATT TTATATCGCCGCCCATGCAACAAGCTTACATAGGGGC TACGCTGGTTAGCATAGCTCCGAGTATGCGTGTCATA ATCAGTGTGGGTTCTTTTGTCATGGTTCCTGGCGGGG AAGTGGCCGCGCTGGTCCGTGCAGACCTGCACGATT ATGTTCAGCTGGCCCTGCGAAGGGACCTACGGGATC GCGGTATTTTTGTTAATGTTCCGCTTTTGAATCTTATAC AGGTCTGTGAGGAACCTGAATTTTTGCAATCATGATTC GCTGCTTGAGGCTGAAGGTGGAGGGCGCTCTGGAGC AGATTTTTACAATGGCCGGACTTAATATTCGGGATTTG CTTAGAGACATATTGATAAGGTGGCGAGATGAAAATTA TTTGGGCATGGTTGAAGGTGCTGGAATGTTTATAGAG GAGATTCACCCTGAAGGGTTTAGCCTTTACGTCCACTT GGACGTGAGGGCAGTTTGCCTTTTGGAAGCCATTGTG CAACATCTTACAAATGCCATTATCTGTTCTTTGGCTGT AGAGTTTGACCACGCCACCGGAGGGGAGCGCGTTCA CTTAATAGATCTTCATTTTGAGGTTTTGGATAATCTTTT GGAATAAAAAAAAAAAAACATGGTTCTTCCAGCTCTTC CCGCTCCTCCCGTGTGTGACTCGCAGAACGAATGTGT AGGTTGGCTGGGTGTGGCTTATTCTGCGGTGGTGGAT GTTATCAGGGCAGCGGCGCATGAAGGAGTTTACATAG AACCCGAAGCCAGGGGGCGCCTGGATGCTTTGAGAG AGTGGATATACTACAACTACTACACAGAGCGAGCTAA GCGACGAGACCGGAGACGCAGATCTGTTTGTCACGC CCGCACCTGGTTTTGCTTCAGGAAATATGACTACGTC CGGCGTTCCATTTGGCATGACACTACGACCAACACGA TCTCGGTTGTCTCGGCGCACTCCGTACAGTAGGGATC GCCTACCTCCTTTTGAGACAGAGACCCGCGCTACCAT ACTGGAGGATCATCCGCTGCTGCCCGAATGTAACACT TTGACAATGCACAACGTGAGTTACGTGCGAGGTCTTC

		CCTGCAGTGTGGGATTTACGCTGATTCAGGAATGGGT TGTTCCCTGGGATATGGTTCTGACGCGGGAGGAGCTT GTAATCCTGAGGAAGTGTATGCACGTGTGCCTGTGTT GTGCCAACATTGATATCATGACGAGCATGATGATCCAT GGTTACGAGTCCTGGGCTCTCCACTGTCATTGTTCCA GTCCCGGTTCCCTGCAGTGCATAGCCGGCGGGCAGG TTTTGGCCAGCTGGTTTAGGATGGTGGTGGATGGCGC CATGTTTAATCAGAGGTTTATATGGTACCGGGAGGTG GTGAATTACAACATGCCAAAAGAGGTAATGTTTATGTC CAGCGTGTTTATGAGGGGTCGCCACTTAATCTACCTG CGCTTGTGGTATGATGGCCACGTGGGTTCTGTGGTCC CCGCCATGAGCTTTGGATACAGCGCCTTGCACTGTGG GATTTTGAACAATATTGTGGTGCTGTGCTGCAGTTACT GTGCTGATTTAAGTGAGATCAGGGTGCGCTGCTGTGC CCGGAGGACAAGGCGTCTCATGCTGCGGGCGGTGCG AATCATCGCTGAGGAGACCACTGCCATGTTGTATTCC TGCAGGACGGAGCGGCGGCGGCAGCAGTTTATTCGC GCGCTGCTGCAGCACCACCGCCCTATCCTGATGCAC GATTATGACTCTACCCCCATGTAGGCGTGGACTTCCC CTTCGCCGCCCGTTGAGCAACCGCAAGTTGGACAGC AGCCTGTGGCTCAGCAGCTGGACAGCGACATGAACTT AAGCGAGCTGCCCGGGGAGTTTATTAATATCACTGAT GAGCGTTTGGCTCGACAGGAAACCGTGTGGAATATAA CACCTAAGAATATGTCTGTTACCCATGATATGATGCTT TTTAAGGCCAGCCGGGGAGAAAGGACTGTGTACTCTG TGTGTTGGGAGGGAGGTGGCAGGTTGAATACTAGGG TTCTGTGAGTTTGATTAAGGTACGGTGATCAATATAAG CTATGTGGTGGTGGGGCTATACTACTGAATGAAAAAT GACTTGAAATTTTCTGCAATTGAAAAATAAACACGTTG AAACATAACATGCAACAGGTTCACGATTCTCTAGTGAA TCCACAGAAACTAGCGAGGTAAGCACTTACTCTATGT CTTTTACATGGTCCTGGGAAAGTGGAAAATACACCACT GAAACTTTTGCTACCAACTCTTACACCTTCTCCTACAT TGCCCAGGAATAAAATCGATGTAGGATGTTGCCCCTC CTGACGCGGTAGGAGAAGGGGAGGGTGCCCTGCATG TCTGCCGCTGCTCTTGCTCTTGCCGCTGCTGAGGAGG GGGGCGCATCTGCCGCAGCACCGGATGCATCTGGGA AAAGCAAAAAAGGGGCTCGTCCCTGTTTCCGGAGGAA TTTGCAAGCGGGGTCTTGCATGACGGGGAGGCAAAC CCCCGTTCGCCGCAGTCCGGCCGGCCCGAGACTCGA ACCGGGGGTCCTGCGACTCAACCCTTGGAAAATAACC CTCCGGCTACAGGGAGCGAGCCACTTAATGCTTTCGC TTTCCAGCCTAACCGCTTACGCCGCGCGCGGCCAGT GGCCAAAAAAGCTAGCGCAGCAGCCGCCGCGCCTGG AAGGAAGCCAAAAGGAGCGCTCCCCCGTTGTCTGAC GTCGCACACCTGGGTTCGACACGCGGGCGGTAACCG CATGGATCACGGCGGACGGCCGGATCCGGGGTTCGA

		ACCCCGGTCGTCCGCCATGATACCCTTGCGAATTTAT CCACCAGACCACGGAAGAGTGCCCGCTTACAGGCTC TCCTTTTGCACGGTCTAGAGCGTCAACGACTGCGCAC GCCTCACCGGCCAGAGCGTCCCGACCATGGAGCACT TTTTGCCGCTGCGCAACATCTGGAACCGCGTCCGCGA CTTTCCGCGCGCCTCCACCACCGCCGCCGGCATCAC CTGGATGTCCAGGTACATCTACGGATTACGTCGACGT TTAAACCATATGATCAGCTCACTCAAAGGCGGTAATAC GGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT CACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGG ATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAG CGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTC CCCGTCGTGTAGATAACTACGATACGGGAGGGCTTAC CATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG CCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGT CACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGG TTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG TTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACT CATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA

		CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA CCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATT GGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCT TACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG TGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCA GCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC CGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGA ATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGA ATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA CATTTCCCCGAAAAGTGCCACCTAAATTGTAAGCGTTA ATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCA GCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATC CCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGA GTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAG AACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCT ATCAGGGCGATGGCCCACTACGTGAACCATCACCCTA ATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA AATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTT GACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAA GGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCT GGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCAC ACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGAT GGATCC
72	pXX6-80 Entire sequence	TCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGG TCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAA GGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG

		ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTA AACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATA GTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGA AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGG CATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCA TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACAT GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCA GTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGT ATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATAC GGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTC TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGAT ACATATTTGAATGTATTTAGAAAAATAAACAAATAGGG GTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACG TCTAAGAAACCATTATTATCATGACATTAACCTATAAAA ATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTT CGGTGATGACGGTGAAAACCTCTGACACATGCAGCTC CCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCC GGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGG TGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGC ATCAGAGCAGATTGTACTGAGAGTGCACCATAAAATT GTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTT TGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAAT CGGCAAAATCCCTTATAAATCAAAAGAATAGCCCGAG ATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTC CACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCG AAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAA CCATCACCCAAATCAAGTTTTTTGGGGTCGAGGTGCC GTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGC GAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGC TAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGT

		AACCACCACACCCGCCGCGCTTAATGCGCCGCTACA GGGCGCGTACTATGGTTGCTTTGACGTATGCGGTGTG AAATACCGCACAGATGCGTAAGGAGAAAATACCGCAT CAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGG GAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGC CAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTA AGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT GTAAAACGACGGCCAGTGCCAAGCTTAAGGTGCACG GCCCACGTGGCCACTAGTACTTCTCGACAGAAGCACC ATGTCCTTGGGTCCGGCCTGCTGAATGCGCAGGCGG TCGGCCATGCCCCAGGCTTCGTTTTGACATCGGCGCA GGTCTTTGTAGTAGTCTTGCATGAGCCTTTCTACCGG CACTTCTTCTTCTCCTTCCTCTTGTCCTGCATCTCTTG CATCTATCGCTGCGGCGGCGGCGGAGTTTGGCCGTA GGTGGCGCCCTCTTCCTCCCATGCGTGTGACCCCGA AGCCCCTCATCGGCTGAAGCAGGGCTAGGTCGGCGA CAACGCGCTCGGCTAATATGGCCTGCTGCACCTGCGT GAGGGTAGACTGGAAGTCATCCATGTCCACAAAGCGG TGGTATGCGCCCGTGTTGATGGTGTAAGTGCAGTTGG CCATAACGGACCAGTTAACGGTCTGGTGACCCGGCTG CGAGAGCTCGGTGTACCTGAGACGCGAGTAAGCCCT CGAGTCAAATACGTAGTCGTTGCAAGTCCGCACCAGG TACTGGTATCCCACCAAAAAGTGCGGCGGCGGCTGG CGGTAGAGGGGCCAGCGTAGGGTGGCCGGGGCTCC GGGGGCGAGATCTTCCAACATAAGGCGATGATATCCG TAGATGTACCTGGACATCCAGGTGATGCCGGCGGCG GTGGTGGAGGCGCGCGGAAAGTCGCGGACGCGGTT CCAGATGTTGCGCAGCGGCAAAAAGTGCTCCATGGTC GGGACGCTCTGGCCGGTCAGGCGCGCGCAATCGTTG ACGCTCTACCGTGCAAAAGGAGAGCCTGTAAGCGGG CACTCTTCCGTGGTCTGGTGGATAAATTCGCAAGGGT ATCATGGCGGACGACCGGGGTTCGAGCCCCGTATCC GGCCGTCCGCCGTGATCCATGCGGTTACCGCCCGCG TGTCGAACCCAGGTGTGCGACGTCAGACAACGGGGG AGTGCTCCTTTTGGCTTCCTTCCAGGCGCGGCGGCTG CTGCGCTAGCTTTTTTGGCCACTGGCCGCGCGCAGC GTAAGCGGTTAGGCTGGAAAGCGAAAGCATTAAGTGG CTCGCTCCCTGTAGCCGGAGGGTTATTTTCCAAGGGT TGAGTCGCGGGACCCCCGGTTCGAGTCTCGGACCGG CCGGACTGCGGCGAACGGGGGTTTGCCTCCCCGTCA TGCAAGACCCCGCTTGCAAATTCCTCCGGAAACAGGG ACGAGCCCCTTTTTTGCTTTTCCCAGATGCATCCGGT GCTGCGGCAGATGCGCCCCCCTCCTCAGCAGCGGCA AGAGCAAGAGCAGCGGCAGACATGCAGGGCACCCTC CCCTCCTCCTACCGCGTCAGGAGGGGCGACATCCGC GGTTGACGCGGCAGCAGATGGTGATTACGAACCCCC GCGGCGCCGGGCCCGGCACTACCTGGACTTGGAGGA

		GGGCGAGGGCCTGGCGCGGCTAGGAGCGCCCTCTC CTGAGCGGTACCCAAGGGTGCAGCTGAAGCGTGATA CGCGTGAGGCGTACGTGCCGCGGCAGAACCTGTTTC GCGACCGCGAGGGAGAGGAGCCCGAGGAGATGCGG GATCGAAAGTTCCACGCAGGGCGCGAGCTGCGGCAT GGCCTGAATCGCGAGCGGTTGCTGCGCGAGGAGGAC TTTGAGCCCGACGCGCGAACCGGGATTAGTCCCGCG CGCGCACACGTGGCGGCCGCCGACCTGGTAACCGCA TACGAGCAGACGGTGAACCAGGAGATTAACTTTCAAA AAAGCTTTAACAACCACGTGCGTACGCTTGTGGCGCG CGAGGAGGTGGCTATAGGACTGATGCATCTGTGGGA CTTTGTAAGCGCGCTGGAGCAAAACCCAAATAGCAAG CCGCTCATGGCGCAGCTGTTCCTTATAGTGCAGCACA GCAGGGACAACGAGGCATTCAGGGATGCGCTGCTAA ACATAGTAGAGCCCGAGGGCCGCTGGCTGCTCGATTT GATAAACATCCTGCAGAGCATAGTGGTGCAGGAGCGC AGCTTGAGCCTGGCTGACAAGGTGGCCGCCATCAACT ATTCCATGCTTAGCCTGGGCAAGTTTTACGCCCGCAA GATATACCATACCCCTTACGTTCCCATAGACAAGGAG GTAAAGATCGAGGGGTTCTACATGCGCATGGCGCTGA AGGTGCTTACCTTGAGCGACGACCTGGGCGTTTATCG CAACGAGCGCATCCACAAGGCCGTGAGCGTGAGCCG GCGGCGCGAGCTCAGCGACCGCGAGCTGATGCACAG CCTGCAAAGGGCCCTGGCTGGCACGGGCAGCGGCG ATAGAGAGGCCGAGTCCTACTTTGACGCGGGCGCTG ACCTGCGCTGGGCCCCAAGCCGACGCGCCCTGGAGG CAGCTGGGGCCGGACCTGGGCTGGCGGTGGCACCC GCGCGCGCTGGCAACGTCGGCGGCGTGGAGGAATAT GACGAGGACGATGAGTACGAGCCAGAGGACGGCGAG TACTAAGCGGTGATGTTTCTGATCAGATGATGCAAGA CGCAACGGACCCGGCGGTGCGGGCGGCGCTGCAGA GCCAGCCGTCCGGCCTTAACTCCACGGACGACTGGC GCCAGGTCATGGACCGCATCATGTCGCTGACTGCGC GCAATCCTGACGCGTTCCGGCAGCAGCCGCAGGCCA ACCGGCTCTCCGCAATTCTGGAAGCGGTGGTCCCGG CGCGCGCAAACCCCACGCACGAGAAGGTGCTGGCGA TCGTAAACGCGCTGGCCGAAAACAGGGCCATCCGGC CCGACGAGGCCGGCCTGGTCTACGACGCGCTGCTTC AGCGCGTGGCTCGTTACAACAGCGGCAACGTGCAGA CCAACCTGGACCGGCTGGTGGGGGATGTGCGCGAGG CCGTGGCGCAGCGTGAGCGCGCGCAGCAGCAGGGC AACCTGGGCTCCATGGTTGCACTAAACGCCTTCCTGA GTACACAGCCCGCCAACGTGCCGCGGGGACAGGAGG ACTACACCAACTTTGTGAGCGCACTGCGGCTAATGGT GACTGAGACACCGCAAAGTGAGGTGTACCAGTCTGG GCCAGACTATTTTTTCCAGACCAGTAGACAAGGCCTG CAGACCGTAAACCTGAGCCAGGCTTTCAAAAACTTGC

		AGGGGCTGTGGGGGGTGCGGGCTCCCACAGGCGAC CGCGCGACCGTGTCTAGCTTGCTGACGCCCAACTCG CGCCTGTTGCTGCTGCTAATAGCGCCCTTCACGGACA GTGGCAGCGTGTCCCGGGACACATACCTAGGTCACTT GCTGACACTGTACCGCGAGGCCATAGGTCAGGCGCA TGTGGACGAGCATACTTTCCAGGAGATTACAAGTGTC AGCCGCGCGCTGGGGCAGGAGGACACGGGCAGCCT GGAGGCAACCCTAAACTACCTGCTGACCAACCGGCG GCAGAAGATCCCCTCGTTGCACAGTTTCGCACCCTTT GGCGCATCCCATTCTCCAGTAACTTTATGTCCATGGG CGCACTCACAGACCTGGGCCAAAACCTTCTCTACGCC AACTCCGCCCACGCGCTAGACATGACTTTTGAGGTGG ATCCCATGGACGAGCCCACCCTTCTTTATGTTTTGTTT GAAGTCTTTGACGTGGTCCGTGTGCACCGGCCGCAC CGCGGCGTCATCGAAACCGTGTACCTGCGCACGCCC TTCTCGGCCGGCAACGCCACAACATAAAGAAGCAAGC AACATCAACAACAGCTGCCGCCATGGGCTCCAGTGAG CAGGAACTGAAAGCCATTGTCAAAGATCTTGGTTGTG GGCCATATTTTTTGGGCACCTATGACAAGCGCTTTCCA GGCTTTGTTTCTCCACACAAGCTCGCCTGCGCCATAG TCAATACGGCCGGTCGCGAGACTGGGGGCGTACACT GGATGGCCTTTGCCTGGAACCCGCACTCAAAAACATG CTACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGAC TCAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCT GCGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGT ATAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT

		GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATTTCCTTCTCCTATAGGCAGAAAAAGATCATGGAG TCAGTCGAGAAGAAGGACAGCCTAACCGCCCCCTCTG AGTTCGCCACCACCGCCTCCACCGATGCCGCCAACG CGCCTACCACCTTCCCCGTCGAGGCACCCCCGCTTG AGGAGGAGGAAGTGATTATCGAGCAGGACCCAGGTTT TGTAAGCGAAGACGACGAGGACCGCTCAGTACCAACA GAGGATAAAAAGCAAGACCAGGACAACGCAGAGGCA AACGAGGAACAAGTCGGGCGGGGGGACGAAAGGCAT GGCGACTACCTAGATGTGGGAGACGACGTGCTGTTG AAGCATCTGCAGCGCCAGTGCGCCATTATCTGCGACG CGTTGCAAGAGCGCAGCGATGTGCCCCTCGCCATAG CGGATGTCAGCCTTGCCTACGAACGCCACCTATTCTC ACCGCGCGTACCCCCCAAACGCCAAGAAAACGGCAC ATGCGAGCCCAACCCGCGCCTCAACTTCTACCCCGTA TTTGCCGTGCCAGAGGTGCTTGCCACCTATCACATCT

		TTTTCCAAAACTGCAAGATACCCCTATCCTGCCGTGCC AACCGCAGCCGAGCGGACAAGCAGCTGGCCTTGCGG CAGGGCGCTGTCATACCTGATATCGCCTCGCTCAACG AAGTGCCAAAAATCTTTGAGGGTCTTGGACGCGACGA GAAGCGCGCGGCAAACGCTCTGCAACAGGAAAACAG CGAAAATGAAAGTCACTCTGGAGTGTTGGTGGAACTC GAGGGTGACAACGCGCGCCTAGCCGTACTAAAACGC AGCATCGAGGTCACCCACTTTGCCTACCCGGCACTTA ACCTACCCCCCAAGGTCATGAGCACAGTCATGAGTGA GCTGATCGTGCGCCGTGCGCAGCCCCTGGAGAGGGA TGCAAATTTGCAAGAACAAACAGAGGAGGGCCTACCC GCAGTTGGCGACGAGCAGCTAGCGCGCTGGCTTCAA ACGCGCGAGCCTGCCGACTTGGAGGAGCGACGCAAA CTAATGATGGCCGCAGTGCTCGTTACCGTGGAGCTTG AGTGCATGCAGCGGTTCTTTGCTGACCCGGAGATGCA GCGCAAGCTAGAGGAAACATTGCACTACACCTTTCGA CAGGGCTACGTACGCCAGGCCTGCAAGATCTCCAAC GTGGAGCTCTGCAACCTGGTCTCCTACCTTGGAATTT TGCACGAAAACCGCCTTGGGCAAAACGTGCTTCATTC CACGCTCAAGGGCGAGGCGCGCCGCGACTACGTCCG CGACTGCGTTTACTTATTTCTATGCTACACCTGGCAGA CGGCCATGGGCGTTTGGCAGCAGTGCTTGGAGGAGT GCAACCTCAAGGAGCTGCAGAAACTGCTAAAGCAAAA CTTGAAGGACCTATGGACGGCCTTCAACGAGCGCTCC GTGGCCGCGCACCTGGCGGACATCATTTTCCCCGAA CGCCTGCTTAAAACCCTGCAACAGGGTCTGCCAGACT TCACCAGTCAAAGCATGTTGCAGAACTTTAGGAACTTT ATCCTAGAGCGCTCAGGAATCTTGCCCGCCACCTGCT GTGCACTTCCTAGCGACTTTGTGCCCATTAAGTACCG CGAATGCCCTCCGCCGCTTTGGGGCCACTGCTACCTT CTGCAGCTAGCCAACTACCTTGCCTACCACTCTGACA TAATGGAAGACGTGAGCGGTGACGGTCTACTGGAGT GTCACTGTCGCTGCAACCTATGCACCCCGCACCGCTC CCTGGTTTGCAATTCGCAGCTGCTTAACGAAAGTCAA ATTATCGGTACCTTTGAGCTGCAGGGTCCCTCGCCTG ACGAAAAGTCCGCGGCTCCGGGGTTGAAACTCACTCC GGGGCTGTGGACGTCGGCTTACCTTCGCAAATTTGTA CCTGAGGACTACCACGCCCACGAGATTAGGTTCTACG AAGACCAATCCCGCCCGCCAAATGCGGAGCTTACCG CCTGCGTCATTACCCAGGGCCACATTCTTGGCCAATT GCAAGCCATCAACAAAGCCCGCCAAGAGTTTCTGCTA CGAAAGGGACGGGGGGTTTACTTGGACCCCCAGTCC GGCGAGGAGCTCAACCCAATCCCCCCGCCGCCGCAG CCCTATCAGCAGCAGCCGCGGGCCCTTGCTTCCCAG GATGGCACCCAAAAAGAAGCTGCAGCTGCCGCCGCC ACCCACGGACGAGGAGGAATACTGGGACAGTCAGGC AGAGGAGGTTTTGGACGAGGAGGAGGAGGACATGAT

		GGAAGACTGGGAGAGCCTAGACGAGGAAGCTTCCGA GGTCGAAGAGGTGTCAGACGAAACACCGTCACCCTC GGTCGCATTCCCCTCGCCGGCGCCCCAGAAATCGGC AACCGGTTCCAGCATGGCTACAACCTCCGCTCCTCAG GCGCCGCCGGCACTGCCCGTTCGCCGACCCAACCGT AGATGGGACACCACTGGAACCAGGGCCGGTAAGTCC AAGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAG CGCCAAGGCTACCGCTCATGGCGCGGGCACAAGAAC GCCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACA TCTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGG CGTGGCCTTCCCCCGTAACATCCTGCATTACTACCGT CATCTCTACAGCCCATACTGCACCGGCGGCAGCGGC AGCGGCAGCAACAGCAGCGGCCACACAGAAGCAAAG GCGACCGGATAGCAAGACTCTGACAAAGCCCAAGAAA TCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCGCT GCGTCTGGCGCCCAACGAACCCGTATCGACCCGCGA GCTTAGAAACAGGATTTTTCCCACTCTGTATGCTATAT TTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAAAT AAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCTGC CTGTATCACAAAAGCGAAGATCAGCTTCGGCGCACGC TGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGCGC GCTGACTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAA TTTAAGCGCGAAAACTACGTCATCTCCAGCGGCCACA CCCGGCGCCAGCACCTGTCGTCAGCGCCATTATGAG CAAGGAAATTCCCACGCCCTACATGTGGAGTTACCAG CCACAAATGGGACTTGCGGCTGGAGCTGCCCAAGAC TACTCAACCCGAATAAACTACATGAGCGCGGGACCCC ACATGATATCCCGGGTCAACGGAATCCGCGCCCACC GAAACCGAATTCTCTTGGAACAGGCGGCTATTACCAC CACACCTCGTAATAACCTTAATCCCCGTAGTTGGCCC GCTGCCCTGGTGTACCAGGAAAGTCCCGCTCCCACC ACTGTGGTACTTCCCAGAGACGCCCAGGCCGAAGTTC AGATGACTAACTCAGGGGCGCAGCTTGCGGGCGGCT TTCGTCACAGGGTGCGGTCGCCCGGGCAGGGTATAA CTCACCTGACAATCAGAGGGCGAGGTATTCAGCTCAA CGACGAGTCGGTGAGCTCCTCGCTTGGTCTCCGTCC GGACGGGACATTTCAGATCGGCGGCGCCGGCCGTCC TTCATTCACGCCTCGTCAGGCAATCCTAACTCTGCAG ACCTCGTCCTCTGAGCCGCGCTCTGGAGGCATTGGAA CTCTGCAATTTATTGAGGAGTTTGTGCCATCGGTCTAC TTTAACCCCTTCTCGGGACCTCCCGGCCACTATCCGG ATCAATTTATTCCTAACTTTGACGCGGTAAAGGACTCG GCGGACGGCTACGACTGAATGTTAAGTGGAGAGGCA GAGCAACTGCGCCTGAAACACCTGGTCCACTGTCGCC GCCACAAGTGCTTTGCCCGCGACTCCGGTGAGTTTTG CTACTTTGAATTGCCCGAGGATCATATCGAGGGCCCG GCGCACGGCGTCCGGCTTACCGCCCAGGGAGAGCTT

		GCCCGTAGCCTGATTCGGGAGTTTACCCAGCGCCCC CTGCTAGTTGAGCGGGACAGGGGACCCTGTGTTCTCA CTGTGATTTGCAACTGTCCTAACCTTGGATTACATCAA GATCCTCTAGTTAATTAACTAGAGTACCCGGGGATCTT ATTCCCTTTAACTAATAAAAAAAAATAATAAAGCATCAC TTACTTAAAATCAGTTAGCAAATTTCTGTCCAGTTTATT CAGCAGCACCTCCTTGCCCTCCTCCCAGCTCTGGTAT TGCAGCTTCCTCCTGGCTGCAAACTTTCTCCACAATCT AAATGGAATGTCAGTTTCCTCCTGTTCCTGTCCATCCG CACCCACTATCTTCATGTTGTTGCAGATGAAGCGCGC AAGACCGTCTGAAGATACCTTCAACCCCGTGTATCCA TATGACACGGAAACCGGTCCTCCAACTGTGCCTTTTC TTACTCCTCCCTTTGTATCCCCCAATGGGTTTCAAGAG AGTCCCCCTGGGGTACTCTCTTTGCGCCTATCCGAAC CTCTAGTTACCTCCAATGGCATGCTTGCGCTCAAAATG GGCAACGGCCTCTCTCTGGACGAGGCCGGCAACCTT ACCTCCCAAAATGTAACCACTGTGAGCCCACCTCTCA AAAAAACCAAGTCAAACATAAACCTGGAAATATCTGCA CCCCTCACAGTTACCTCAGAAGCCCTAACTGTGGCTG CCGCCGCACCTCTAATGGTCGCGGGCAACACACTCA CCATGCAATCACAGGCCCCGCTAACCGTGCACGACTC CAAACTTAGCATTGCCACCCAAGGACCCCTCACAGTG TCAGAAGGAAAGCTAGCCCTGCAAACATCAGGCCCCC TCACCACCACCGATAGCAGTACCCTTACTATCACTGC CTCACCCCCTCTAACTACTGCCACTGGTAGCTTGGGC ATTGACTTGAAAGAGCCCATTTATACACAAAATGGAAA ACTAGGACTAAAGTACGGGGCTCCTTTGCATGTAACA GACGACCTAAACACTTTGACCGTAGCAACTGGTCCAG GTGTGACTATTAATAATACTTCCTTGCAAACTAAAGTTA CTGGAGCCTTGGGTTTTGATTCACAAGGCAATATGCA ACTTAATGTAGCAGGAGGACTAAGGATTGATTCTCAAA ACAGACGCCTTATACTTGATGTTAGTTATCCGTTTGAT GCTCAAAACCAACTAAATCTAAGACTAGGACAGGGCC CTCTTTTTATAAACTCAGCCCACAACTTGGATATTAACT ACAACAAAGGCCTTTACTTGTTTACAGCTTCAAACAAT TCCAAAAAGCTTGAGGTTAACCTAAGCACTGCCAAGG GGTTGATGTTTGACGCTACAGCCATAGCCATTAATGC AGGAGATGGGCTTGAATTTGGTTCACCTAATGCACCA AACACAAATCCCCTCAAAACAAAAATTGGCCATGGCCT AGAATTTGATTCAAACAAGGCTATGGTTCCTAAACTAG GAACTGGCCTTAGTTTTGACAGCACAGGTGCCATTAC AGTAGGAAACAAAAATAATGATAAGCTAACTTTGTGGA CCACACCAGCTCCATCTCCTAACTGTAGACTAAATGCA GAGAAAGATGCTAAACTCACTTTGGTCTTAACAAAATG TGGCAGTCAAATACTTGCTACAGTTTCAGTTTTGGCTG TTAAAGGCAGTTTGGCTCCAATATCTGGAACAGTTCAA AGTGCTCATCTTATTATAAGATTTGACGAAAATGGAGT

		GCTACTAAACAATTCCTTCCTGGACCCAGAATATTGGA ACTTTAGAAATGGAGATCTTACTGAAGGCACAGCCTAT ACAAACGCTGTTGGATTTATGCCTAACCTATCAGCTTA TCCAAAATCTCACGGTAAAACTGCCAAAAGTAACATTG TCAGTCAAGTTTACTTAAACGGAGACAAAACTAAACCT GTAACACTAACCATTACACTAAACGGTACACAGGAAAC AGGAGACACAACTCCAAGTGCATACTCTATGTCATTTT CATGGGACTGGTCTGGCCACAACTACATTAATGAAAT ATTTGCCACATCCTCTTACACTTTTTCATACATTGCCCA AGAATAAAGAATCGTTTGTGTTATGTTTCAACGTGTTT ATTTTTCAATTGCAGAAAATTTCAAGTCATTTTTCATTC AGTAGTATAGCCCCACCACCACATAGCTTATACAGATC ACCGTACCTTAATCAAACTCACAGAACCCTAGTATTCA ACCTGCCACCTCCCTCCCAACACACAGAGTACACAGT CCTTTCTCCCCGGCTGGCCTTAAAAAGCATCATATCAT GGGTAACAGACATATTCTTAGGTGTTATATTCCACACG GTTTCCTGTCGAGCCAAACGCTCATCAGTGATATTAAT AAACTCCCCGGGCAGCTCACTTAAGTTCATGTCGCTG TCCAGCTGCTGAGCCACAGGCTGCTGTCCAACTTGCG GTTGCTTAACGGGCGGCGAAGGAGAAGTCCACGCCT ACATGGGGGTAGAGTCATAATCGTGCATCAGGATAGG GCGGTGGTGCTGCAGCAGCGCGCGAATAAACTGCTG CCGCCGCCGCTCCGTCCTGCAGGAATACAACATGGC AGTGGTCTCCTCAGCGATGATTCGCACCGCCCGCAG CATAAGGCGCCTTGTCCTCCGGGCACAGCAGCGCAC CCTGATCTCACTTAAATCAGCACAGTAACTGCAGCACA GCACCACAATATTGTTCAAAATCCCACAGTGCAAGGC GCTGTATCCAAAGCTCATGGCGGGGACCACAGAACC CACGTGGCCATCATACCACAAGCGCAGGTAGATTAAG TGGCGACCCCTCATAAACACGCTGGACATAAACATTA CCTCTTTTGGCATGTTGTAATTCACCACCTCCCGGTAC CATATAAACCTCTGATTAAACATGGCGCCATCCACCAC CATCCTAAACCAGCTGGCCAAAACCTGCCCGCCGGCT ATACACTGCAGGGAACCGGGACTGGAACAATGACAGT GGAGAGCCCAGGACTCGTAACCATGGATCATCATGCT CGTCATGATATCAATGTTGGCACAACACAGGCACACG TGCATACACTTCCTCAGGATTACAAGCTCCTCCCGCG TTAGAACCATATCCCAGGGAACAACCCATTCCTGAATC AGCGTAAATCCCACACTGCAGGGAAGACCTCGCACGT AACTCACGTTGTGCATTGTCAAAGTGTTACATTCGGGC AGCAGCGGATGATCCTCCAGTATGGTAGCGCGGGTTT CTGTCTCAAAAGGAGGTAGACGATCCCTACTGTACGG AGTGCGCCGAGACAACCGAGATCGTGTTGGTCGTAGT GTCATGCCAAATGGAACGCCGGACGTAGTCATATTTC CTGAAGCAAAACCAGGTGCGGGCGTGACAAACAGAT CTGCGTCTCCGGTCTCGCCGCTTAGATCGCTCTGTGT AGTAGTTGTAGTATATCCACTCTCTCAAAGCATCCAGG

		CGCCCCCTGGCTTCGGGTTCTATGTAAACTCCTTCAT GCGCCGCTGCCCTGATAACATCCACCACCGCAGAATA AGCCACACCCAGCCAACCTACACATTCGTTCTGCGAG TCACACACGGGAGGAGCGGGAAGAGCTGGAAGAACC ATGTTTTTTTTTTTATTCCAAAAGATTATCCAAAACCTC AAAATGAAGATCTATTAAGTGAACGCGCTCCCCTCCG GTGGCGTGGTCAAACTCTACAGCCAAAGAACAGATAA TGGCATTTGTAAGATGTTGCACAATGGCTTCCAAAAG GCAAACGGCCCTCACGTCCAAGTGGACGTAAAGGCTA AACCCTTCAGGGTGAATCTCCTCTATAAACATTCCAGC ACCTTCAACCATGCCCAAATAATTCTCATCTCGCCACC TTCTCAATATATCTCTAAGCAAATCCCGAATATTAAGTC CGGCCATTGTAAAAATCTGCTCCAGAGCGCCCTCCAC CTTCAGCCTCAAGCAGCGAATCATGATTGCAAAAATTC AGGTTCCTCACAGACCTGTATAAGATTCAAAAGCGGA ACATTAACAAAAATACCGCGATCCCGTAGGTCCCTTC GCAGGGCCAGCTGAACATAATCGTGCAGGTCTGCAC GGACCAGCGCGGCCACTTCCCCGCCAGGAACCTTGA CAAAAGAACCCACACTGATTATGACACGCATACTCGG AGCTATGCTAACCAGCGTAGCCCCGATGTAAGCTTTG TTGCATGGGCGGCGATATAAAATGCAAGGTGCTGCTC AAAAAATCAGGCAAAGCCTCGCGCAAAAAAGAAAGCA CATCGTAGTCATGCTCATGCAGATAAAGGCAGGTAAG CTCCGGAACCACCACAGAAAAAGACACCATTTTTCTCT CAAACATGTCTGCGGGTTTCTGCATAAACACAAAATAA AATAACAAAAAAACATTTAAACATTAGAAGCCTGTCTTA CAACAGGAAAAACAACCCTTATAAGCATAAGACGGAC TACGGCCATGCCGGCGTGACCGTAAAAAAACTGGTCA CCGTGATTAAAAAGCACCACCGACAGCTCCTCGGTCA TGTCCGGAGTCATAATGTAAGACTCGGTAAACACATC AGGTTGATTCATCGGTCAGTGCTAAAAAGCGACCGAA ATAGCCCGGGGGAATACATACCCGCAGGCGTAGAGA CAACATTACAGCCCCCATAGGAGGTATAACAAAATTAA TAGGAGAGAAAAACACATAAACACCTGAAAAACCCTC CTGCCTAGGCAAAATAGCACCCTCCCGCTCCAGAACA ACATACAGCGCTTCCACAGCGGCAGCCATAACAGTCA GCCTTACCAGTAAAAAAGAAAACCTATTAAAAAAACAC CACTCGACACGGCACCAGCTCAATCAGTCACAGTGTA AAAAAGGGCCAAGTGCAGAGCGAGTATATATAGGACT AAAAAATGACGTAACGGTTAAAGTCCACAAAAAACACC CAGAAAACCGCACGCGAACCTACGCCCAGAAACGAAA GCCAAAAAACCCACAACTTCCTCAAATCGTCACTTCCG TTTTCCCACGTTACGTAACTTCCCATTTTAAGAAAACTA CAATTCCCAACACATACAAGTTACTCCGCCCTAAAACC TACGTCACCCGCCCCGTTCCCACGCCCCGCGCCACG TCACAAACTCCACCCCCTCATTATCATATTGGCTTCAA TCCAAAATAAGGTATATTATTGATGATTTATTTTGGATT

		GAAGCCAATATGATAATGAGGGGGTGGAGTTTGTGAC GTGGCGCGGGGCGTGGGAACGGGGCGGGTGACGTA GTAGTGTGGCGGAAGTGTGATGTTGCAAGTGTGGCG GAACACATGTAAGCGACGGATGTGGCAAAAGTGACGT TTTTGGTGTGCGCCGGATCCACAGGACGGGTGTGGT CGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGC GAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTC GGACAGTGCTCCGAGAACGGGTGCGCATAGAAATTG CATCAACGCATATAGCGCTAGCAGCACGCCATAGTGA CTGGCGATGCTGTCGGAATGGACGATATCCCGCAAGA GGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTA CAGCATCCAGGGTGACGGTGCCGAGGATGACGATGA GCGCATTGTTAGATTTCATACACGGTGCCTGACTGCG TTAGCAATTTAACTGTGATAAACTACCGCATTAAAGCT TATCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTG TGAAATTGTTATCCGCTCACAATTCCACACAACATACG AGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAA TGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT GCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGG TTTGCGTATTGGGCGC
80	pEMBR-1.2B2-GEP	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcga

		acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttcc aggcgcggcggctgctgcgctagctttttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggaccccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgccccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccctcctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC

		ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATcttgaaggacctatggacggccttcaacgagcgctccgtggccgcgc acctggcggacatcattttccccgaacgcctgcttaaaaccctgcaacagggtc tgccagacttcaccagtcaaagcatgttgcagaactttaggaactttatcctaga gcgctcaggaatcttgcccgccacctgctgtgcacttcctagcgactttgtgccca ttaagtaccgcgaatgccctccgccgctttggggccactgctaccttctgcagcta gccaactaccttgcctaccactctgacataatggaagacgtgagcggtgacggt ctactggagtgtcactgtcgctgcaacctatgcaccccgcaccgctccctggtttg caattcgcagctgcttaacgaaagtcaaattatcggtacctttgagctgcagggt ccctcgcctgacgaaaagtccgcggctccggggttgaaactcactccggggct gtggacgtcggcttaccttcgcaaatttgtacctgaggactaccacgcccacga gattaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgcctg cgtcattacccagggccacattcttggccaattgcaagccatcaacaaagccc gccaagagtttctgctacgaaagggacgggggtttacttggacccccagtcc ggcgaggagctcaacccaatccccccgccgccgcagccctatcagcagcag ccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgccg ccgccacccacggacgaggaggaatactgggacagtcaggcagaggaggt tttggacgaggaggaggaggacatgatggaagactgggagagcctagacga ggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcggtc gcattcccctcgccggcgccccagaaatcggcaaccggttccagcatggctac aacctccgctcctcaggcgccgccggcactgcccgttcgccgacccaaccgta gATGGGACACCACTGGAACCAGGGCCGGTAAGTCCAA

		GCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGCG CCAAGGCTACCGCTCATGGCGCGGGCACAAGAACGC CATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACATC TCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGCG TGGCCTTCCCCCGTAACATCCTGCATTACTACCGTCAT CTCTACAGCCCATACTGCACCGGCGGCAGCGGCAGC AACAGCAGCGGCCACACAGAAGCAAAGGCGACCGGA TAGCAAGACTCTGACAAAGCCCAAGAAATCCACAGCG GCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTGGC GCCCAACGAACCCGTATCGACCCGCGAGCTTAGAAAC AGGATTTTTCCCACTCTGTATGCTATATTTCAACAGAG CAGGGGCCAAGAACAAGAGCTGAAAATAAAAAACAGG TCTCTGCGATCCCTCACCCGCAGCTGCCTGTATCACA AAAGCGAAGATCAGCTTCGGCGCACGCTGGAAGACG CGGAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCT TAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGC GAAAACTACGTCATCTCCAGCGGCCACACCCGGCGC CAGCACCTGTTGTCAGCGCCATTATGAGCAAGGAAAT TCCCACGCCCTACATGTGGAGTTACCAGCCACAAATG GGACTTGCGGCTGGAGCTGCCCAAGACTACTCAACC CGAATAAACTACATGAGCGCGGGGCggccgccgtttgtgttatg tttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattcagtagtata gccccaccaccacatagcttatacagatcaccgtaccttaatcaaacTCACA GAACCCTAGTATTCAACCTGCCACCTCCCTCCCAACA CACAGAGTACACAGTCCTTTCTCCCCGGCTGGCCTTA AAAAGCATCATATCATGGGTAACAGACATATTCTTAGG TGTTATATTCCACACGGTTTCCTGTCGAGCCAAACGCT CATCAGTGATATTAATAAACTCCCCGGGCAGCTCACTT AAGTTCATGTCGCTGTCCAGCTGCTGAGCCACAGGCT GCTGTCCAACTTGCGGTTGCTTAACGGGCGGCGAAG GAGAAGTCCACGCCTACATGGGGGTAGAGTCATAATC GTGCATCAGGATAGGGCGGTGGTGCTGCAGCAGCGC GCGAATAAACTGCTGCCGCCGCCGCTCCGTCCTGCA GGAATACAACATGGCAGTGGTCTCCTCAGCGATGATT CGCACCGCCCGCAGCATAAGGCGCCTTGTCCTCCGG GCACAGCAGCGCACCCTGATCTCACTTAAATCAGCAC AGTAACTGCAGCACAGCACCACAATATTGTTCAAAATC CCACAGTGCAAGGCGCTGTATCCAAAGCTCATGGCG GGGACCACAGAACCCACGTGGCCATCATACCACAAG CGCAGGTAGATTAAGTGGCGACCCCTCATAAACACGC TGGACATAAACATTACCTCTTTTGGCATGTTGTAATTC ACCACCTCCCGGTACCATATAAACCTCTGATTAAACAT GGCGCCATCCACCACCATCCTAAACCAGCTGGCCAAA ACCTGCCCGCCGGCTATACACTGCAGGGAACCGGGA CTGGAACAATGACAGTGGAGAGCCCAGGACTCGTAAC CATGGATCATCATGCTCGTCATGATATCAATGTTGGCA CAACACAGGCACACGTGCATACACTTCCTCAGGATTA

		CAAGCTCCTCCCGCGTTAGAACCATATCCCAGGGAAC AACCCATTCCTGAATCAGCGTAAATCCCACACTGCAG GGAAGACCTCGCACGTAACTCACGTTGTGCATTGTCA AAGTGTTACATTCGGGCAGCAGCGGATGATCCTCCAG TATGGTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGA CGATCCCTACTGTACGGAGTGCGCCGAGACAACCGA GATCGTGTTGGTCGTAGTGTCATGCCAAATGGAACGC CGGACGTAGTCATatttcctgaagcaaaaccaggtgcgggcgtgaca aacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagtagttgtagta tatccactctctcaaagcatccaggcgccccctggcttcgggttctatgtaaactc cttcatgcgccgctgccctgataacatccaccaccgcagaataagccacaccc agccaacctacacattcgttctgcgagtcacacacgggaggagcgggaaga gctggaagaaccatgttttttttttttattccaaaagattatccaaaacctcaaaatga agatctattaagtgaacgcgctcccctccggtggcgtggtcaaactctacagcc aaagaacagataatggcatttgtaagatgttgcacaatggcttccaaaaggca aacggccctcacgtccaagtggacgtaaaggctaaacccttcagggtgaatct cctctataaacattccagcaccttcaaccatgcccaaataattctcatctcgccac cttctcaatatatctctaagcaaatcccgaatattaagtccggccattgtaaaaat ctgctccagagcgccctccaccttcagcctcaagcagcgaatcatgattgcaaa aattcaggttcctcacagacctgtataagattcaaaagcggaacattaacaaaa ataccgcgatcccgtaggtcccttcgcagggccagctgaacataatcgtgcag gtctgcacggaccagcgcggccacttccccgccaggaaccatgacaaaaga acccacactgattatgacacgcatactcggagctatgctaaccagcgtagccc cgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaaa aatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtcatgctcatg cagataaaggcaggtaagctccggaaccaccacagaaaaagacaccatttttt ctctcaaacatgtctgcgggtttctgcataaacacaaaataaaataacaaaaaa acatttaaacattagaagcctgtcttacaacaggaaaaacaacccttataagca taagacggactacggccatgccggcgtgaccgtaaaaaaactggtcaccgtg attaaaaagcaccaccgacagctcctcggtcatgtccggagtcataatgtaag actcggtaaacacatcaggttgattcacatcggtcagtgctaaaaagcgaccg aaatagcccgggggaatacatacccgcaggcgtagagacaacattacagcc cccataggaggtataacaaaattaataggagagaaaaacacataaacacct gaaaaaccctcctgcctaggcaaaatagcaccctcccgctccagaacaacat acagcgcttccacagcggcagccataacagtcagccttaccagtaaaaaaga aaacctattaaaaaaacaccactcgacacggcaccagctcaatcagtcacag tgtaaaaaagggccaagtgcagagcgagtatatataggactaaaaaatgacg taacggttaaagtccacaaaaaacacccagaaaaccgcacgcgaacctacg cccagaaacgaaagccaaaaaacccacaacttcctcaaatcgtcacttccgtt ttcccacgttacgtcacttcccattttaagaaaactacaattcccaacacatacaa gttactccgcccttaattaaatcggatccgatatctagatgtattcgcgaggtacc gagctcgaattctctggccgtcgttttacaacgtcgtgactgggaaaaccctggc gttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatag cgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggc gaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgc atatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagcccc gacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggc

		atccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttt tcaccgtcatcaccgaaacgcgcga
81	pEMBR- 1.2B2- GEP+1kbp	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc

		aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC

		TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcatggagtcagtcgagaagaa ggacagcctaaccgccccctctgagttcgccaccaccgcctccaccgatgccg ccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggagga agtgattatcgagcaggacccaggttttgtaagcgaagacgacgaggaccgct cagtaccaacagaggataaaaagcaagaccaggacaacgcagaggcaaa cgaggaacaagtcgggcgggggacgaaaggcatggcgactacctagatgt gggagacgacgtgctgttgaagcatctgcagcgccagtgcgccattatctgcg acgcgttgcaagagcgcagcgatgtgcccctcgccatagcggatgtcagcctt gcctacgaacgccacctattctcaccgcgcgtacccccccaaacgccaagaaa acggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgccgtgc cagaggtgcttgccacctatcacatctttttccaaaactgcaagataccctatcc tgccgtgccaaccgcagccgagcggacaagcagctggccttgcggcagggc gctgtcatacctgatatcgcctcgctcaacgaagtgccaaaaatctttgagggtct tggacgcgacgagaagcgcgcggcaaacgctctgcaacaggaaaacagc gaaaatgaaagtcactctggagtgttggtggaactcgagggtgacaacgcgc gcctagccgtactaaaacgcagcatcgaggtcacccactttgcctacccggca cttaacctacccccccaaggtcatgagcacagtcatgagtgagctgatcgtgcgc cgtgcgcagcccctggagagggatgcaaatttgcaagaacaaacagaggag ggcctacccgcagttggcgacgagcagctagcgcgctggcttcaaacgcgcg agcctgcccttgaaggacctatggacggccttcaacgagcgctccgtggccgc gcacctggcggacatcattttccccgaacgcctgcttaaaaccctgcaacagg gtctgccagacttcaccagtcaaagcatgttgcagaactttaggaactttatccta gagcgctcaggaatcttgcccgccacctgctgtgcacttcctagcgactttgtgcc cattaagtaccgcgaatgccctccgccgctttggggccactgctaccttctgcag ctagccaactaccttgcctaccactctgacataatggaagacgtgagcggtgac ggtctactggagtgtcactgtcgctgcaacctatgcaccccgcaccgctccctgg tttgcaattcgcagctgcttaacgaaagtcaaattatcggtacctttgagctgcag ggtccctcgcctgacgaaaagtccgcggctccggggttgaaactcactccggg gctgtggacgtcggcttaccttcgcaaatttgtacctgaggactaccacgcccac gagattaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgc ctgcgtcattacccagggccacattcttggccaattgcaagccatcaacaaagc ccgccaagagtttctgctacgaaagggacgggggtttacttggacccccagtc cggcgaggagctcaacccaatccccccgccgccgcagccctatcagcagca gccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgcc gccgccacccacggacgaggaggaatactgggacagtcaggcagaggag gttttggacgaggaggaggaggacatgatggaagactgggagagcctagac gaggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcg gtcgcattcccctcgccggcgccccagaaatcggcaaccggttccagcatggc tacaacctccgctcctcaggcgccgccggcactgcccgttcgccgacccaacc gtagATGGGACACCACTGGAACCAGGGCCGGTAAGTC

		CAAGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACA GCGCCAAGGCTACCGCTCATGGCGCGGGCACAAGAA CGCCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAAC ATCTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACG GCGTGGCCTTCCCCCGTAACATCCTGCATTACTACCG TCATCTCTACAGCCCATACTGCACCGGCGGCAGCGG CAGCAACAGCAGCGGCCACACAGAAGCAAAGGCGAC CGGATAGCAAGACTCTGACAAAGCCCAAGAAATCCAC AGCGGCGGCAGCAGCAGGAGGAGGAGCGCTGCGTC TGGCGCCCAACGAACCCGTATCGACCCGCGAGCTTA GAAACAGGATTTTTCCCACTCTGTATGCTATATTTCAA CAGAGCAGGGGCCAAGAACAAGAGCTGAAAATAAAAA ACAGGTCTCTGCGATCCCTCACCCGCAGCTGCCTGTA TCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAA GACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCTGA CTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAA GCGCGAAAACTACGTCATCTCCAGCGGCCACACCCG GCGCCAGCACCTGTTGTCAGCGCCATTATGAGCAAGG AAATTCCCACGCCCTACATGTGGAGTTACCAGCCACA AATGGGACTTGCGGCTGGAGCTGCCCAAGACTACTCA ACCCGAATAAACTACATGAGCGCGGGGCggccgccgtttgt gttatgtttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattcagta gtatagccccaccaccacatagcttatacagatcaccgtaccttaatcaaacT CACAGAACCCTAGTATTCAACCTGCCACCTCCCTCCC AACACACAGAGTACACAGTCCTTTCTCCCCGGCTGGC CTTAAAAAGCATCATATCATGGGTAACAGACATATTCT TAGGTGTTATATTCCACACGGTTTCCTGTCGAGCCAAA CGCTCATCAGTGATATTAATAAACTCCCCGGGCAGCT CACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGCCAC AGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGCGG CGAAGGAGAAGTCCACGCCTACATGGGGGTAGAGTC ATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCAG CAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCGT CCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCG ATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTCC TCCGGGCACAGCAGCGCACCCTGATCTCACTTAAATC AGCACAGTAACTGCAGCACAGCACCACAATATTGTTC AAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTCA TGGCGGGGACCACAGAACCCACGTGGCCATCATACC ACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAAA CACGCTGGACATAAACATTACCTCTTTTGGCATGTTGT AATTCACCACCTCCCGGTACCATATAAACCTCTGATTA AACATGGCGCCATCCACCACCATCCTAAACCAGCTGG CCAAAACCTGCCCGCCGGCTATACACTGCAGGGAAC CGGGACTGGAACAATGACAGTGGAGAGCCCAGGACT CGTAACCATGGATCATCATGCTCGTCATGATATCAATG TTGGCACAACACAGGCACACGTGCATACACTTCCTCA

		GGATTACAAGCTCCTCCCGCGTTAGAACCATATCCCA GGGAACAACCCATTCCTGAATCAGCGTAAATCCCACA CTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCA TTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGATC CTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGA GGTAGACGATCCCTACTGTACGGAGTGCGCCGAGAC AACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAATG GAACGCCGGACGTAGTCATatttcctgaagcaaaaccaggtgcg ggcgtgacaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagt agttgtagtatatccactctctcaaagcatccaggcgccccctggcttcgggttct atgtaaactccttcatgcgccgctgccctgataacatccaccaccgcagaataa gccacacccagccaacctacacattcgttctgcgagtcacacacgggaggag cgggaagagctggaagaaccatgttttttttttttattccaaaagattatccaaaacc tcaaaatgaagatctattaagtgaacgcgctcccctccggtggcgtggtcaaac tctacagccaaagaacagataatggcatttgtaagatgttgcacaatggcttcca aaaggcaaacggccctcacgtccaagtggacgtaaaggctaaacccttcagg gtgaatctcctctataaacattccagcaccttcaaccatgcccaaataattctcat ctcgccaccttctcaatatatctctaagcaaatcccgaatattaagtccggccatt gtaaaaatctgctccagagcgccctccaccttcagcctcaagcagcgaatcat gattgcaaaaattcaggttcctcacagacctgtataagattcaaaagcggaaca ttaacaaaaataccgcgatcccgtaggtcccttcgcagggccagctgaacata atcgtgcaggtctgcacggaccagcgcggccacttccccgccaggaaccatg acaaaagaacccacactgattatgacacgcatactcggagctatgctaaccag cgtagccccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgct gctcaaaaaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtc atgctcatgcagataaaggcaggtaagctccggaaccaccacagaaaaaga caccatttttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaata acaaaaaaacatttaaacattagaagcctgtcttacaacaggaaaaacaacc cttataagcataagacggactacggccatgccggcgtgaccgtaaaaaaact ggtcaccgtgattaaaaagcaccaccgacagctcctcggtcatgtccggagtc ataatgtaagactcggtaaacacatcaggttgattcacatcggtcagtgctaaaa agcgaccgaaatagcccgggggaatacatacccgcaggcgtagagacaac attacagcccccataggaggtataacaaaattaataggagagaaaaacacat aaacacctgaaaaaccctcctgcctaggcaaaatagcaccctcccgctccag aacaacatacagcgcttccacagcggcagccataacagtcagccttaccagt aaaaaagaaaacctattaaaaaaacaccactcgacacggcaccagctcaat cagtcacagtgtaaaaaagggccaagtgcagagcgagtatatataggactaa aaaatgacgtaacggttaaagtccacaaaaaacacccagaaaaccgcacg cgaacctacgcccagaaacgaaagccaaaaaacccacaacttcctcaaatc gtcacttccgttttcccacgttacgtcacttcccattttaagaaaactacaattccca acacatacaagttactccgcccttaattaaatcggatccgatatctagatgtattc gcgaggtaccgagctcgaattctctggccgtcgttttacaacgtcgtgactggga aaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagc tggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgca gcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggt atttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagtta agccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtc

		tgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgt gtcagaggttttcaccgtcatcaccgaaacgcgcga
82	pEMBR-1.2B2-GEP+50bp	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc

		aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC

		TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcatggagtcagtcgagaagaa ggcttgaaggacctatggacggccttcaacgagcgctccgtggccgcgcacct ggcggacatcattttccccgaacgcctgcttaaaaccctgcaacagggtctgcc agacttcaccagtcaaagcatgttgcagaactttaggaactttatcctagagcgc tcaggaatcttgcccgccacctgctgtgcacttcctagcgactttgtgcccattaa gtaccgcgaatgccctccgccgctttggggccactgctaccttctgcagctagcc aactaccttgcctaccactctgacataatggaagacgtgagcggtgacggtcta ctggagtgtcactgtcgctgcaacctatgcaccccgcaccgctccctggtttgca attcgcagctgcttaacgaaagtcaaattatcggtacctttgagctgcagggtcc ctcgcctgacgaaaagtccgcggctccggggttgaaactcactccggggctgt ggacgtcggcttaccttcgcaaatttgtacctgaggactaccacgcccacgaga ttaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgcctgcg tcattacccagggccacattcttggccaattgcaagccatcaacaaagcccgcc aagagtttctgctacgaaagggacgggggtttacttggacccccagtccggc gaggagctcaacccaatccccccgccgccgcagccctatcagcagcagccg cgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgccgccg ccacccacggacgaggaggaatactgggacagtcaggcagaggaggttttg gacgaggaggaggaggacatgatggaagactgggagagcctagacgagg aagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgc attcccctcgccggcgccccagaaatcggcaaccggttccagcatggctacaa cctccgctcctcaggcgccgccggcactgcccgttcgccgacccaaccgtag ATGGGACACCACTGGAACCAGGGCCGGTAAGTCCAA GCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGCG CCAAGGCTACCGCTCATGGCGCGGGCACAAGAACGC CATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACATC TCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGCG TGGCCTTCCCCCGTAACATCCTGCATTACTACCGTCAT CTCTACAGCCCATACTGCACCGGCGGCAGCGGCAGC AACAGCAGCGGCCACACAGAAGCAAAGGCGACCGGA TAGCAAGACTCTGACAAAGCCCAAGAAATCCACAGCG GCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTGGC GCCCAACGAACCCGTATCGACCCGCGAGCTTAGAAAC AGGATTTTTCCCACTCTGTATGCTATATTTCAACAGAG CAGGGGCCAAGAACAAGAGCTGAAAATAAAAAACAGG TCTCTGCGATCCCTCACCCGCAGCTGCCTGTATCACA AAAGCGAAGATCAGCTTCGGCGCACGCTGGAAGACG CGGAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCT TAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGCGC GAAAACTACGTCATCTCCAGCGGCCACACCCGGCGC CAGCACCTGTTGTCAGCGCCATTATGAGCAAGGAAAT

		TCCCACGCCCTACATGTGGAGTTACCAGCCACAAATG GGACTTGCGGCTGGAGCTGCCCAAGACTACTCAACC CGAATAAACTACATGAGCGCGGGGCggccgccgtttgtgttatg tttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattcagtagtata gccccaccaccacatagcttatacagatcaccgtaccttaatcaaacTCACA GAACCCTAGTATTCAACCTGCCACCTCCCTCCCAACA CACAGAGTACACAGTCCTTTCTCCCCGGCTGGCCTTA AAAAGCATCATATCATGGGTAACAGACATATTCTTAGG TGTTATATTCCACACGGTTTCCTGTCGAGCCAAACGCT CATCAGTGATATTAATAAACTCCCCGGGCAGCTCACTT AAGTTCATGTCGCTGTCCAGCTGCTGAGCCACAGGCT GCTGTCCAACTTGCGGTTGCTTAACGGGCGGCGAAG GAGAAGTCCACGCCTACATGGGGGTAGAGTCATAATC GTGCATCAGGATAGGGCGGTGGTGCTGCAGCAGCGC GCGAATAAACTGCTGCCGCCGCCGCTCCGTCCTGCA GGAATACAACATGGCAGTGGTCTCCTCAGCGATGATT CGCACCGCCCGCAGCATAAGGCGCCTTGTCCTCCGG GCACAGCAGCGCACCCTGATCTCACTTAAATCAGCAC AGTAACTGCAGCACAGCACCACAATATTGTTCAAAATC CCACAGTGCAAGGCGCTGTATCCAAAGCTCATGGCG GGGACCACAGAACCCACGTGGCCATCATACCACAAG CGCAGGTAGATTAAGTGGCGACCCCTCATAAACACGC TGGACATAAACATTACCTCTTTTGGCATGTTGTAATTC ACCACCTCCCGGTACCATATAAACCTCTGATTAAACAT GGCGCCATCCACCACCATCCTAAACCAGCTGGCCAAA ACCTGCCCGCCGGCTATACACTGCAGGGAACCGGGA CTGGAACAATGACAGTGGAGAGCCCAGGACTCGTAAC CATGGATCATCATGCTCGTCATGATATCAATGTTGGCA CAACACAGGCACACGTGCATACACTTCCTCAGGATTA CAAGCTCCTCCCGCGTTAGAACCATATCCCAGGGAAC AACCCATTCCTGAATCAGCGTAAATCCCACACTGCAG GGAAGACCTCGCACGTAACTCACGTTGTGCATTGTCA AAGTGTTACATTCGGGCAGCAGCGGATGATCCTCCAG TATGGTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGA CGATCCCTACTGTACGGAGTGCGCCGAGACAACCGA GATCGTGTTGGTCGTAGTGTCATGCCAAATGGAACGC CGGACGTAGTCATatttcctgaagcaaaaccaggtgcgggcgtgaca aacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagtagttgtagta tatccactctctcaaagcatccaggcgccccctggcttcgggttctatgtaaactc cttcatgcgccgctgccctgataacatccaccaccgcagaataagccacaccc agccaacctacacattcgttctgcgagtcacacacgggaggagcgggaaga gctggaagaaccatgttttttttttttattccaaaagattatccaaaacctcaaaatga agatctattaagtgaacgcgctcccctccggtggcgtggtcaaactctacagcc aaagaacagataatggcatttgtaagatgttgcacaatggcttccaaaaggca aacggccctcacgtccaagtggacgtaaaggctaaacccttcagggtgaatct cctctataaacattccagcaccttcaaccatgcccaaataattctcatctcgccac cttctcaatatatctctaagcaaatcccgaatattaagtccggccattgtaaaaat

		ctgctccagagcgccctccaccttcagcctcaagcagcgaatcatgattgcaaa aattcaggttcctcacagacctgtataagattcaaaagcggaacattaacaaaa ataccgcgatcccgtaggtcccttcgcagggccagctgaacataatcgtgcag gtctgcacggaccagcgcggccacttccccgccaggaaccatgacaaaaga acccacactgattatgacacgcatactcggagctatgctaaccagcgtagccc cgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaaa aatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtcatgctcatg cagataaaggcaggtaagctccggaaccaccacagaaaaagacaccatttttt ctctcaaacatgtctgcgggtttctgcataaacacaaaataaaataacaaaaaa acatttaaacattagaagcctgtcttacaacaggaaaaacaacccttataagca taagacggactacggccatgccggcgtgaccgtaaaaaaactggtcaccgtg attaaaaagcaccaccgacagctcctcggtcatgtccggagtcataatgtaag actcggtaaacacatcaggttgattcacatcggtcagtgctaaaaagcgaccg aaatagcccgggggaatacatacccgcaggcgtagagacaacattacagcc cccataggaggtataacaaaattaataggagagaaaaacacataaacacct gaaaaaccctcctgcctaggcaaaatagcaccctcccgctccagaacaacat acagcgcttccacagcggcagccataacagtcagccttaccagtaaaaaaga aaacctattaaaaaaacaccactcgacacggcaccagctcaatcagtcacag tgtaaaaaagggccaagtgcagagcgagtatatataggactaaaaaatgacg taacggttaaagtccacaaaaaacacccagaaaaccgcacgcgaacctacg cccagaaacgaaagccaaaaaacccacaacttcctcaaatcgtcacttccgtt ttcccacgttacgtcacttcccattttaagaaaactacaattcccaacacatacaa gttactccgcccttaattaaatcggatccgatatctagatgtattcgcgaggtacc gagctcgaattctctggccgtcgttttacaacgtcgtgactgggaaaaccctggc gttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatag cgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggc gaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgc atatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagcccc gacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggc atccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttt tcaccgtcatcaccgaaacgcgcga
83	pEMBR-1.2B2-GEP+50bp+ Kozak	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc

		atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggaccccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT

		GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATggtggcttccttctcctataggcagaaaaagatcatggagtcagtcgag aagaaggcttgaaggacctatggacggccttcaacgagcgctccgtggccgc gcacctggcggacatcattttccccgaacgcctgcttaaaaccctgcaacagg gtctgccagacttcaccagtcaaagcatgttgcagaactttaggaactttatccta gagcgctcaggaatcttgcccgccacctgctgtgcacttcctagcgactttgtgcc cattaagtaccgcgaatgccctccgccgctttggggccactgctaccttctgcag ctagccaactaccttgcctaccactctgacataatggaagacgtgagcggtgac ggtctactggagtgtcactgtcgctgcaacctatgcaccccgcaccgctccctgg tttgcaattcgcagctgcttaacgaaagtcaaattatcggtacctttgagctgcag ggtccctcgcctgacgaaaagtccgcggctccggggttgaaactcactccggg gctgtggacgtcggcttaccttcgcaaatttgtacctgaggactaccacgcccac

		gagattaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgc ctgcgtcattacccagggccacattcttggccaattgcaagccatcaacaaagc ccgccaagagtttctgctacgaaagggacgggggtttacttggacccccagtc cggcgaggagctcaacccaatccccccgccgccgcagccctatcagcagca gccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgcc gccgccacccacggacgaggaggaatactgggacagtcaggcagaggag gttttggacgaggaggaggaggacatgatggaagactgggagagcctagac gaggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcg gtcgcattcccctcgccggcgccccagaaatcggcaaccggttccagcatggc tacaacctccgctcctcaggcgccgccggcactgcccgttcgccgacccaacc gtagATGGGACACCACTGGAACCAGGGCCGGTAAGTC CAAGCAGCCGCCGCCGTTAGCCCAAGAGCAACAACA GCGCCAAGGCTACCGCTCATGGCGCGGGCACAAGAA CGCCATAGTTGCTTGCTTGCAAGACTGTGGGGGCAAC ATCTCCTTCGCCCGCCGCTTTCTTCTCTACCATCACG GCGTGGCCTTCCCCCGTAACATCCTGCATTACTACCG TCATCTCTACAGCCCATACTGCACCGGCGGCAGCGG CAGCAACAGCAGCGGCCACACAGAAGCAAAGGCGAC CGGATAGCAAGACTCTGACAAAGCCCAAGAAATCCAC AGCGGCGGCAGCAGCAGGAGGAGGAGCGCTGCGTC TGGCGCCCAACGAACCCGTATCGACCCGCGAGCTTA GAAACAGGATTTTTCCCACTCTGTATGCTATATTTCAA CAGAGCAGGGGCCAAGAACAAGAGCTGAAAATAAAAA ACAGGTCTCTGCGATCCCTCACCCGCAGCTGCCTGTA TCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAA GACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCTGA CTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAA GCGCGAAAACTACGTCATCTCCAGCGGCCACACCCG GCGCCAGCACCTGTTGTCAGCGCCATTATGAGCAAGG AAATTCCCACGCCCTACATGTGGAGTTACCAGCCACA AATGGGACTTGCGGCTGGAGCTGCCCAAGACTACTCA ACCCGAATAAACTACATGAGCGCGGGGCggccgccgtttgt gttatgtttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattcagta gtatagccccaccaccacatagcttatacagatcaccgtaccttaatcaaacT CACAGAACCCTAGTATTCAACCTGCCACCTCCCTCCC AACACACAGAGTACACAGTCCTTTCTCCCCGGCTGGC CTTAAAAAGCATCATATCATGGGTAACAGACATATTCT TAGGTGTTATATTCCACACGGTTTCCTGTCGAGCCAAA CGCTCATCAGTGATATTAATAAACTCCCCGGGCAGCT CACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGCCAC AGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGCGG CGAAGGAGAAGTCCACGCCTACATGGGGGTAGAGTC ATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCAG CAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCGT CCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCG ATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTCC TCCGGGCACAGCAGCGCACCCTGATCTCACTTAAATC

		AGCACAGTAACTGCAGCACAGCACCACAATATTGTTC AAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTCA TGGCGGGGACCACAGAACCCACGTGGCCATCATACC ACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAAA CACGCTGGACATAAACATTACCTCTTTTGGCATGTTGT AATTCACCACCTCCCGGTACCATATAAACCTCTGATTA AACATGGCGCCATCCACCACCATCCTAAACCAGCTGG CCAAAACCTGCCCGCCGGCTATACACTGCAGGGAAC CGGGACTGGAACAATGACAGTGGAGAGCCCAGGACT CGTAACCATGGATCATCATGCTCGTCATGATATCAATG TTGGCACAACACAGGCACACGTGCATACACTTCCTCA GGATTACAAGCTCCTCCCGCGTTAGAACCATATCCCA GGGAACAACCCATTCCTGAATCAGCGTAAATCCCACA CTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCA TTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGATC CTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGA GGTAGACGATCCCTACTGTACGGAGTGCGCCGAGAC AACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAATG GAACGCCGGACGTAGTCATatttcctgaagcaaaaccaggtgcg ggcgtgacaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagt agttgtagtatatccactctctcaaagcatccaggcgccccctggcttcgggttct atgtaaactccttcatgcgccgctgccctgataacatccaccaccgcagaataa gccacacccagccaacctacacattcgttctgcgagtcacacacgggaggag cgggaagagctggaagaaccatgttttttttttttattccaaaagattatccaaaacc tcaaaatgaagatctattaagtgaacgcgctcccctccggtggcgtggtcaaac tctacagccaaagaacagataatggcatttgtaagatgttgcacaatggcttcca aaaggcaaacggccctcacgtccaagtggacgtaaaggctaaacccttcagg gtgaatctcctctataaacattccagcaccttcaaccatgcccaaataattctcat ctcgccaccttctcaatatatctctaagcaaatcccgaatattaagtccggccatt gtaaaaatctgctccagagcgccctccaccttcagcctcaagcagcgaatcat gattgcaaaaattcaggttcctcacagacctgtataagattcaaaagcggaaca ttaacaaaaataccgcgatcccgtaggtcccttcgcagggccagctgaacata atcgtgcaggtctgcacggaccagcgcggccacttccccgccaggaaccatg acaaaagaacccacactgattatgacacgcatactcggagctatgctaaccag cgtagccccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgct gctcaaaaaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtc atgctcatgcagataaaggcaggtaagctccggaaccaccacagaaaaaga caccatttttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaata acaaaaaaacatttaaacattagaagcctgtcttacaacaggaaaaacaacc cttataagcataagacggactacggccatgccggcgtgaccgtaaaaaaact ggtcaccgtgattaaaaagcaccaccgacagctcctcggtcatgtccggagtc ataatgtaagactcggtaaacacatcaggttgattcacatcggtcagtgctaaaa agcgaccgaaatagcccgggggaatacatacccgcaggcgtagagacaac attacagcccccataggaggtataacaaaattaataggagagaaaaacacat aaacacctgaaaaaccctcctgcctaggcaaaatagcaccctcccgctccag aacaacatacagcgcttccacagcggcagccataacagtcagccttaccagt aaaaaagaaaacctattaaaaaaacaccactcgacacggcaccagctcaat

		cagtcacagtgtaaaaaagggccaagtgcagagcgagtatatataggactaa aaaatgacgtaacggttaaagtccacaaaaaacacccagaaaaccgcacg cgaacctacgcccagaaacgaaagccaaaaaacccacaacttcctcaaatc gtcacttccgttttcccacgttacgtcacttcccattttaagaaaactacaattccca acacatacaagttactccgcccttaattaaatcggatccgatatctagatgtattc gcgaggtaccgagctcgaattctctggccgtcgttttacaacgtcgtgactggga aaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagc tggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgca gcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggt atttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagtta agccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtc tgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgt gtcagaggttttcaccgtcatcaccgaaacgcgcga
84	pEMBR-1.2B2-HexDeltaC32stop	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca

		aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC

		AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGAGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcatggagtcagtcgagaagaa ggacagcctaaccgccccctctgagttcgccaccaccgcctccaccgatgccg ccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggagga agtgattatcgagcaggacccaggttttgtaagcgaagacgacgaggaccgct cagtaccaacagaggataaaaagcaagaccaggacaacgcagaggcaaa cgaggaacaagtcgggcgggggacgaaaggcatggcgactacctagatgt gggagacgacgtgctgttgaagcatctgcagcgccagtgcgccattatctgcg acgcgttgcaagagcgcagcgatgtgcccctcgccatagcggatgtcagcctt gcctacgaacgccacctattctcaccgcgcgtacccccccaaacgccaagaaa acggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgccgtgc cagaggtgcttgccacctatcacatctttttccaaaactgcaagataccctatcc tgccgtgccaaccgcagccgagcggacaagcagctggccttgcggcagggc gctgtcatacctgatatcgcctcgctcaacgaagtgccaaaaatctttgagggtct tggacgcgacgagaagcgcgcggcaaacgctctgcaacaggaaaacagc gaaaatgaaagtcactctggagtgttggtggaactcgagggtgacaacgcgc gcctagccgtactaaaacgcagcatcgaggtcacccactttgcctacccggca cttaacctacccccccaaggtcatgagcacagtcatgagtgagctgatcgtgcgc cgtgcgcagcccctggagagggatgcaaatttgcaagaacaaacagaggag ggcctacccgcagttggcgacgagcagctagcgcgctggcttcaaacgcgcg agcctgccgacttggaggagcgacgcaaactaatgatggccgcagtgctcgtt accgtggagcttgagtgcatgcagcggttctttgctgacccggagatgcagcgc aagctagaggaaacattgcactacacctttcgacagggctacgtacgccaggc ctgcaagatctccaacgtggagctctgcaacctggtctcctaccttggaattttgc acgaaaaccgccttgggcaaaacgtgcttcattccacgctcaagggcgaggc gcgccgcgactacgtccgcgactgcgtttacttatttctatgctacacctggcaga cggccatgggcgtttggcagcagtgcttggaggagtgcaacctcaaggagctg cagaaactgctaaagcaaaacttgaaggacctatggacggccttcaacgagc gctccgtggccgcgcacctggcggacatcattttccccgaacgcctgcttaaaa ccctgcaacagggtctgccagacttcaccagtcaaagcatgttgcagaacttta

		ggaactttatcctagagcgctcaggaatcttgcccgccacctgctgtgcacttcct agcgactttgtgcccattaagtaccgcgaatgccctccgccgctttggggccact gctaccttctgcagctagccaactaccttgcctaccactctgacataatggaaga cgtgagcggtgacggtctactggagtgtcactgtcgctgcaacctatgcacccc gcaccgctccctggtttgcaattcgcagctgcttaacgaaagtcaaattatcggta cctttgagctgcagggtccctcgcctgacgaaaagtccgcggctccggggttga aactcactccggggctgtggacgtcggcttaccttcgcaaatttgtacctgagga ctaccacgcccacgagattaggttctacgaagaccaatcccgcccgcctaatg cggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaag ccatcaacaaagcccgccaagagtttctgctacgaaagggacgggggttta cttggacccccagtccggcgaggagctcaacccaatccccccgccgccgcag ccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaag aagctgcagctgccgccgccacccacggacgaggaggaatactgggacagt caggcagaggaggttttggacgaggaggaggaggacatgatggaagactgg gagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaaca ccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccgg ttccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcg ccgacccaaccgtagATGGGACACCACTGGAACCAGGGCC GGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAG CAACAACAGCGCCAAGGCTACCGCTCATGGCGCGGG CACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTG GGGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTA CCATCACGGCGTGGCCTTCCCCCGTAACATCCTGCAT TACTACCGTCATCTCTACAGCCCATACTGCACCGGCG GCAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAA AGGCGACCGGATAGCAAGACTCTGACAAAGCCCAAG AAATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGTTTCGCGCCCTTTCTC AAATTTAAGCGCGAAAACTACGTCATCTCCAGCGGCC ACACCCGGCGCCAGCACCTGTTGTCAGCGCCATTATG AGCAAGGAAATTCCCACGCCCTACATGTGGAGTTACC AGCCACAAATGGGACTTGCGGCTGGAGCTGCCCAAG ACTACTCAACCCGAATAAACTACATGAGCGCGGGGCg gccgccgtttgtgttatgtttcaacgtgtttatttttcaattgcagaaaatttcaagtcat ttttcattcagtagtatagccccaccaccacatagcttatacagatcaccgtacctt aatcaaacTCACAGAACCCTAGTATTCAACCTGCCACCTC CCTCCCAACACACAGAGTACACAGTCCTTTCTCCCCG GCTGGCCTTAAAAAGCATCATATCATGGGTAACAGAC ATATTCTTAGGTGTTATATTCCACACGGTTTCCTGTCG AGCCAAACGCTCATCAGTGATATTAATAAACTCCCCG GGCAGCTCACTTAAGTTCATGTCGCTGTCCAGCTGCT

		GAGCCACAGGCTGCTGTCCAACTTGCGGTTGCTTAAC GGGCGGCGAAGGAGAAGTCCACGCCTACATGGGGGT AGAGTCATAATCGTGCATCAGGATAGGGCGGTGGTGC TGCAGCAGCGCGCGAATAAACTGCTGCCGCCGCCGC TCCGTCCTGCAGGAATACAACATGGCAGTGGTCTCCT CAGCGATGATTCGCACCGCCCGCAGCATAAGGCGCC TTGTCCTCCGGGCACAGCAGCGCACCCTGATCTCACT TAAATCAGCACAGTAACTGCAGCACAGCACCACAATA TTGTTCAAAATCCCACAGTGCAAGGCGCTGTATCCAA AGCTCATGGCGGGGACCACAGAACCCACGTGGCCAT CATACCACAAGCGCAGGTAGATTAAGTGGCGACCCCT CATAAACACGCTGGACATAAACATTACCTCTTTTGGCA TGTTGTAATTCACCACCTCCCGGTACCATATAAACCTC TGATTAAACATGGCGCCATCCACCACCATCCTAAACC AGCTGGCCAAAACCTGCCCGCCGGCTATACACTGCA GGGAACCGGGACTGGAACAATGACAGTGGAGAGCCC AGGACTCGTAACCATGGATCATCATGCTCGTCATGAT ATCAATGTTGGCACAACACAGGCACACGTGCATACAC TTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAACCA TATCCCAGGGAACAACCCATTCCTGAATCAGCGTAAA TCCCACACTGCAGGGAAGACCTCGCACGTAACTCACG TTGTGCATTGTCAAAGTGTTACATTCGGGCAGCAGCG GATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTCTC AAAAGGAGGTAGACGATCCCTACTGTACGGAGTGCGC CGAGACAACCGAGATCGTGTTGGTCGTAGTGTCATGC CAAATGGAACGCCGGACGTAGTCATatttcctgaagcaaaac caggtgcgggcgtgacaaacagatctgcgtctccggtctcgccgcttagatcgc tctgtgtagtagttgtagtatatccactctctcaaagcatccaggcgccccctggct tcgggttctatgtaaactccttcatgcgccgctgccctgataacatccaccaccgc agaataagccacacccagccaacctacacattcgttctgcgagtcacacacg ggaggagcgggaagagctggaagaaccatgttttttttttttattccaaaagattat ccaaaacctcaaaatgaagatctattaagtgaacgcgctcccctccggtggcgt ggtcaaactctacagccaaagaacagataatggcatttgtaagatgttgcacaa tggcttccaaaaggcaaacggccctcacgtccaagtggacgtaaaggctaaa cccttcagggtgaatctcctctataaacattccagcaccttcaaccatgcccaaat aattctcatctcgccaccttctcaatatatctctaagcaaatcccgaatattaagtc cggccattgtaaaaatctgctccagagcgccctccaccttcagcctcaagcagc gaatcatgattgcaaaaattcaggttcctcacagacctgtataagattcaaaagc ggaacattaacaaaaataccgcgatcccgtaggtcccttcgcagggccagctg aacataatcgtgcaggtctgcacggaccagcgcggccacttccccgccagga accatgacaaaagaacccacactgattatgacacgcatactcggagctatgct aaccagcgtagccccgatgtaagcttgttgcatgggcggcgatataaaatgca aggtgctgctcaaaaaatcaggcaaagcctcgcgcaaaaaagaaagcacat cgtagtcatgctcatgcagataaaggcaggtaagctccggaaccaccacaga aaaagacaccatttttctctcaaacatgtctgcgggtttctgcataaacacaaaat aaaataacaaaaaaacatttaaacattagaagcctgtcttacaacaggaaaa acaacccttataagcataagacggactacggccatgccggcgtgaccgtaaa

		aaaactggtcaccgtgattaaaaagcaccaccgacagctcctcggtcatgtcc ggagtcataatgtaagactcggtaaacacatcaggttgattcacatcggtcagtg ctaaaaagcgaccgaaatagcccgggggaatacatacccgcaggcgtaga gacaacattacagcccccataggaggtataacaaaattaataggagagaaa aacacataaacacctgaaaaaccctcctgcctaggcaaaatagcaccctccc gctccagaacaacatacagcgcttccacagcggcagccataacagtcagcct taccagtaaaaaagaaaacctattaaaaaaacaccactcgacacggcacca gctcaatcagtcacagtgtaaaaaagggccaagtgcagagcgagtatatata ggactaaaaaatgacgtaacggttaaagtccacaaaaaacacccagaaaac cgcacgcgaacctacgcccagaaacgaaagccaaaaaacccacaacttcct caaatcgtcacttccgttttcccacgttacgtcacttcccattttaagaaaactaca attcccaacacatacaagttactccgcccttaattaaatcggatccgatatctaga tgtattcgcgaggtaccgagctcgaattctctggccgtcgttttacaacgtcgtgac tgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcg ccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagtt gcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgt gcggtatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgca tagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgg gcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagct gcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcga
85	pEMBR- 1.3+L4	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT

		TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC

		AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC

		GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcCACAAAAGCGAAGAT CAGCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTC TTCAGTAAATACTGCGCGCTGACTCTTAAGGACTAGTT TCGCGCCCTTTCTCAAATTTAAGCGCGAAAACTACGTC ATCTCCAGCGGCCACACCCGGCGCCAGCACCTGTTG TCAGCGCCATTGGCGCGCCGGCCGGCCGAAtatcTTCA TTTAAATGTTTAAACATCGATGCGGCCGCgattaggttctacg aagaccaatcccgcccgcctaatgcggagcttaccgcctgcgtcattacccag ggccacattcttggccaattgcaagccatcaacaaagcccgccaagagtttctg ctacgaaagggacgggggtttacttggacccccagtccggcgaggagctca acccaatccccccgccgccgcagccctatcagcagcagccgcgggcccttgc ttcccaggatggcacccaaaaagaagctgcagctgccgccgccacccacgg acgaggaggaatactgggacagtcaggcagaggaggttttggacgaggagg aggaggacatgatggaagactgggagagcctagacgaggaagcttccgag gtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgcc ggcgccccagaaatcggcaaccggttccagcatggctacaacctccgctcctc aggcgccgccggcactgcccgttcgccgacccaaccgtagATGGGACA CCACTGGAACCAGGGCCGGTAAGTCCAAGCAGCCGC CGCCGTTAGCCCAAGAGCAACAACAGCGCCAAGGCT ACCGCTCATGGCGCGGGCACAAGAACGCCATAGTTG CTTGCTTGCAAGACTGTGGGGGCAACATCTCCTTCGC CCGCCGCTTTCTTCTCTACCATCACGGCGTGGCCTTC

		CCCCGTAACATCCTGCATTACTACCGTCATCTCTACAG CCCATACTGCACCGGCGGCAGCGGCAGCAACAGCAG CGGCCACACAGAAGCAAAGGCGACCGGATAGCAAGA CTCTGACAAAGCCCAAGAAATCCACAGCGGCGGCAG CAGCAGGAGGAGGAGCGCTGCGTCTGGCGCCCAACG AACCCGTATCGACCCGCGAGCTTAGAAACAGGATTTT TCCCACTCTGTATGCTATATTTCAACAGAGCAGGGGC CAAGAACAAGAGCTGAAAATAAAAAACAGGTCTCTGC GATCCCTCACCCGCAGCTGCCTGTATCACAAAAGCGA AGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGGC TCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGACT AGaataaaagatctttattttcattagatctgtgtgttggttttttgtgtgGCGGCC GCCGTTTGTGTTATGTTTCAACGTGTTTATTTTTCAATT GCAGAAAATTTCAAGTCATTTTTCATTCAGTAGTATAG CCCCACCACCACATAGCTTATACAGATCACCGTACCTT AATCAAACTCACAGAACCCTAGTATTCAACCTGCCACC TCCCTCCCAACACACAGAGTACACAGTCCTTTCTCCC CGGCTGGCCTTAAAAAGCATCATATCATGGGTAACAG ACATATTCTTAGGTGTTATATTCCACACGGTTTCCTGT CGAGCCAAACGCTCATCAGTGATATTAATAAACTCCCC GGGCAGCTCACTTAAGTTCATGTCGCTGTCCAGCTGC TGAGCCACAGGCTGCTGTCCAACTTGCGGTTGCTTAA CGGGCGGCGAAGGAGAAGTCCACGCCTACATGGGGG TAGAGTCATAATCGTGCATCAGGATAGGGCGGTGGTG CTGCAGCAGCGCGCGAATAAACTGCTGCCGCCGCCG CTCCGTCCTGCAGGAATACAACATGGCAGTGGTCTCC TCAGCGATGATTCGCACCGCCCGCAGCATAAGGCGC CTTGTCCTCCGGGCACAGCAGCGCACCCTGATCTCAC TTAAATCAGCACAGTAACTGCAGCACAGCACCACAAT ATTGTTCAAAATCCCACAGTGCAAGGCGCTGTATCCA AAGCTCATGGCGGGGACCACAGAACCCACGTGGCCA TCATACCACAAGCGCAGGTAGATTAAGTGGCGACCCC TCATAAACACGCTGGACATAAACATTACCTCTTTTGGC ATGTTGTAATTCACCACCTCCCGGTACCATATAAACCT CTGATTAAACATGGCGCCATCCACCACCATCCTAAAC CAGCTGGCCAAAACCTGCCCGCCGGCTATACACTGCA GGGAACCGGGACTGGAACAATGACAGTGGAGAGCCC AGGACTCGTAACCATGGATCATCATGCTCGTCATGAT ATCAATGTTGGCACAACACAGGCACACGTGCATACAC TTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAACCA TATCCCAGGGAACAACCCATTCCTGAATCAGCGTAAA TCCCACACTGCAGGGAAGACCTCGCACGTAACTCACG TTGTGCATTGTCAAAGTGTTACATTCGGGCAGCAGCG GATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTCTC AAAAGGAGGTAGACGATCCCTACTGTACGGAGTGCGC CGAGACAACCGAGATCGTGTTGGTCGTAGTGTCATGC CAAATGGAACGCCGGACGTAGTCATATTTCCTGAAGC

		AAAACCAGGTGCGGGCGTGACAAACAGATCTGCGTCT CCGGTCTCGCCGCTTAGATCGCTCTGTGTAGTAGTTG TAGTATATCCACTCTCTCAAAGCATCCAGGCGCCCCC TGGCTTCGGGTTCTATGTAAACTCCTTCATGCGCCGC TGCCCTGATAACATCCACCACCGCAGAATAAGCCACA CCCAGCCAACCTACACATTCGTTCTGCGAGTCACACA CGGGAGGAGCGGGAAGAGCTGGAAGAACCATGTTTT TTTTTTTATTCCAAAAGATTATCCAAAACCTCAAAATGA AGATCTATTAAGTGAACGCGCTCCCCTCCGGTGGCGT GGTCAAACTCTACAGCCAAAGAACAGATAATGGCATTT GTAAGATGTTGCACAATGGCTTCCAAAAGGCAAACGG CCCTCACGTCCAAGTGGACGTAAAGGCTAAACCCTTC AGGGTGAATCTCCTCTATAAACATTCCAGCACCTTCAA CCATGCCCAAATAATTCTCATCTCGCCACCTTCTCAAT ATATCTCTAAGCAAATCCCGAATATTAAGTCCGGCCAT TGTAAAAATCTGCTCCAGAGCGCCCTCCACCTTCAGC CTCAAGCAGCGAATCATGATTGCAAAAATTCAGGTTCC TCACAGACCTGTATAAGATTCAAAAGCGGAACATTAAC AAAAATACCGCGATCCCGTAGGTCCCTTCGCAGGGCC AGCTGAACATAATCGTGCAGGTCTGCACGGACCAGCG CGGCCACTTCCCCGCCAGGAACCATGACAAAAGAACC CACACTGATTATGACACGCATACTCGGAGCTATGCTA ACCAGCGTAGCCCCGATGTAAGCTTGTTGCATGGGCG GCGATATAAAATGCAAGGTGCTGCTCAAAAAATCAGG CAAAGCCTCGCGCAAAAAAGAAAGCACATCGTAGTCA TGCTCATGCAGATAAAGGCAGGTAAGCTCCGGAACCA CCACAGAAAAAGACACCATTTTTCTCTCAAACATGTCT GCGGGTTTCTGCATAAACACAAAATAAAATAACAAAAA AACATTTAAACATTAGAAGCCTGTCTTACAACAGGAAA AACAACCCTTATAAGCATAAGACGGACTACGGCCATG CCGGCGTGACCGTAAAAAAACTGGTCACCGTGATTAA AAAGCACCACCGACAGCTCCTCGGTCATGTCCGGAGT CATAATGTAAGACTCGGTAAACACATCAGGTTGATTCA CATCGGTCAGTGCTAAAAAGCGACCGAAATAGCCCGG GGGAATACATACCCGCAGGCGTAGAGACAACATTACA GCCCCCATAGGAGGTATAACAAAATTAATAGGAGAGA AAAACACATAAACACCTGAAAAACCCTCCTGCCTAGG CAAAATAGCACCCTCCCGCTCCAGAACAACATACAGC GCTTCCACAGCGGCAGCCATAACAGTCAGCCTTACCA GTAAAAAAGAAAACCTATTAAAAAAACACCACTCGACA CGGCACCAGCTCAATCAGTCACAGTGTAAAAAAGGGC CAAGTGCAGAGCGAGTATATATAGGACTAAAAAATGA CGTAACGGTTAAAGTCCACAAAAAACACCCAGAAAAC CGCACGCGAACCTACGCCCAGAAACGAAAGCCAAAAA ACCCACAACTTCCTCAAATCGTCACTTCCGTTTTCCCA CGTTACGTCACTTCCCATTTTAAGAAAACTACAATTCC CAACACATACAAGTTACTCCGCCCTTAATTAAATCGGA

		TCCGATATCTAGATGTATTCGCGAGGTACCGAGCTCG AATTCTCTGGCCGTCGTTTTACAACGTCGTGACTGGG AAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGC ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAG GCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCC TTACGCATCTGTGCGGTATTTCACACCGCATATGGTG CACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA GCCAGCCCCGACACCCGCCAACACCCGCTGACGCGC CCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAG ACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAG AGGTTTTCACCGTCATCACCGAAACGCGCGA
86	pEMBR-1.4+L4	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt

		ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctctgcagtcgaccagaagcaccatgtccttgggtccgg cctgctgaatgcgcaggcggtcggccatgccccaggcttcgttttgacatcggc gcaggtctttgtagtagtcttgcatgagcctttctaccggcacttcttcttctccttcct cttgtcctgcatctcttgcatctatcgctgcggcggcggcggagtttggccgtaggt ggcgccctcttcctcccatgcgtgtgaccccgaagcccctcatcggctgaagca gggctaggtcggcgacaacgcgctcggctaatatggcctgctgcacctgcgtg agggtagactggaagtcatccatgtccacaaagcggtggtatgcgcccgtgttg atggtgtaagtgcagttggccataacggaccagttaacggtctggtgacccggc tgcgagagctcggtgtacctgagacgcgagtaagccctcgagtcaaatacgta gtcgttgcaagtccgcaccaggtactggtatcccaccaaaaagtgcggcggcg gctggcggtagaggggccagcgtagggtggccggggctccgggggcgagat cttccaacataaggcgatgatatccgtagatgtacctggacatccaggtgatgc cggcggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgtt gcgcagcggcaaaaagtgctccatggtcgggacgctctggccggtcaggcgc gcgcaatcgttgacgctctagcgtgcaaaaggagagcctgtaagcgggcact cttccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggtt cgagccccgtatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtc gaacccaggtgtgcgacgtcagacaacgggggagtgctccttttggcttccttcc aggcgcggcggctgctgcgctagctttttttggccactggccgcgcgcagcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcgggacccccggttcgagtctcggaccgag actgggggcgtacactggatggcctttgcctggaacccgcactcaaaaacatg ctacctctttgagccctttggcttttctgaccagcgactcaagcaggtttaccagttt gagtacgagtcactcctgcgccgtagcgccattgcttcttcccccgaccgctgtat aacgctggaaaagtccacccaaagcgtacaggggcccaactcggccgcctg tggactattctgctgcatgtttctccacgcctttgccaactggccccaaactcccat ggatcacaaccccaccatgaaccttattaccggggtacccaactccatgctcaa cagtccccaggtacagcccaccctgcgtcgcaaccaggaacagctctacagc ttcctggagcgccactcgccctacttccgcagccacagtgcgcagattaggagc gccacttctttttgtcacttgaaaaacatgtaaaaataatgtactagagacactttc aataaaggcaaatgcttttatttgtacactctcgggtgattatttacccccacccttg ccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgcca ctggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggcac aaccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatcac caacgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctcc gccctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatca gcgccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcg tccaggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgccttc ccaaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatc aaaaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaag ccttgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgcc gcaagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgcag caccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttcac gatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcac

		atccatttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagct cgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctc gtgatgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcg ccccatcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcg gtgctcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtca ggcagtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagc gcgcgcgcagcctccatgcccttctcccacgcagacacgatcggcacactca gcgggttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcg tccgcataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgctt acctcctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtag cgccacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgc tcgggcttgggagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgc cgccgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtg atgagtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcg cccggggaggcggcggcgacggggacggggacgacacgtcctccatggttg ggggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcct cttcccgactggccatttccttctcctataggcagaaaaagatccacaaaagcg aagatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatac tgcgcgctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaa actacgtcatctccagcggccacacccggcgccagcacctgttgtcagcgcca ttggcgcgcccgcccgccgcgcgcttcgctttttatagggccgccgccgccgcc gcctcgccataaaaggaaactttcggagcgcgccgctctgattggctgccgcc gcacctctccgcctcgccccgccccgccctcgccccgccccgccccgcctgg cgcgcgccccccccccccccccgccccatcgctgcacaaaataattaaaaa ataaataaatacaaaattgggggtggggaggggggggagatggggagagtg aagcagaacgtggggctcacctcgaggccggccgaatatcttcatttaaatgttt aaacatcgatgcGGCCGCgattaggttctacgaagaccaatcccgcccgc ctaatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaatt gcaagccatcaacaaagcccgccaagagtttctgctacgaaagggacgggg ggtttacttggacccccagtccggcgaggagctcaacccaatccccccgccgc cgcagccctatcagcagcagccgcgggcccttgcttcccaggatggcaccca aaaagaagctgcagctgccgccgccacccacggacgaggaggaatactgg gacagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaa gactgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagac gaaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcgg caaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactg cccgttcgccgacccaaccgtagATGGGACACCACTGGAACCA GGGCCGGTAAGTCCAAGCAGCCGCCGCCGTTAGCCC AAGAGCAACAACAGCGCCAAGGCTACCGCTCATGGC GCGGGCACAAGAACGCCATAGTTGCTTGCTTGCAAGA CTGTGGGGGCAACATCTCCTTCGCCCGCCGCTTTCTT CTCTACCATCACGGCGTGGCCTTCCCCCGTAACATCC TGCATTACTACCGTCATCTCTACAGCCCATACTGCACC GGCGGCAGCGGCAGCAACAGCAGCGGCCACACAGAA GCAAAGGCGACCGGATAGCAAGACTCTGACAAAGCC CAAGAAATCCACAGCGGCGGCAGCAGCAGGAGGAGG AGCGCTGCGTCTGGCGCCCAACGAACCCGTATCGAC

		CCGCGAGCTTAGAAACAGGATTTTTCCCACTCTGTAT GCTATATTTCAACAGAGCAGGGGCCAAGAACAAGAGC TGAAAATAAAAAAACAGGTCTCTGCGATCCCTCACCCG CAGCTGCCTGTATCACAAAAGCGAAGATCAGCTTCGG CGCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAAT ACTGCGCGCTGACTCTTAAGGACTAGaataaaagatctttattt tcattagatctgtgtgttggttttttgtgtgGCggccgccgtttgtgttatgtttcaacgt gtttatttttcaattgcagaaaatttcaagtcatttttcattcagtagtatagccccacc accacatagcttatacagatcaccgtaccttaatcaaactcacagaaccctagt attcaacctgccacctccctcccaacacacagagtacacagtcctttctccccgg ctggccttaaaaagcatcatatcatgggtaacagacatattcttaggtgttatattc cacacggtttcctgtcgagccaaacgctcatcagtgatattaataaactccccgg gcagctcacttaagttcatgtcgctgtccagctgctgagccacaggctgctgtcc aacttgcggttgcttaacgggcggcgaaggagaagtccacgcctacatgggg gtagagtcataatcgtgcatcaggatagggcggtggtgctgcagcagcgcgcg aataaactgctgccgccgccgctccgtcctgcaggaatacaacatggcagtgg tctcctcagcgatgattcgcaccgcccgcagcataaggcgccttgtcctccggg cacagcagcgcaccctgatctcacttaaatcagcacagtaactgcagcacag caccacaatattgttcaaaatcccacagtgcaaggcgctgtatccaaagctcat ggcggggaccacagaacccacgtggccatcataccacaagcgcaggtagat taagtggcgacccctataaacacgctggacataaacattacctcttttggcatgt tgtaattcaccacctcccggtaccatataaacctctgattaaacatggcgccatc caccaccatcctaaaccagctggccaaaacctgcccgccggctatacactgc agggaaccgggactggaacaatgacagtggagagcccaggactcgtaacc atggatcatcatgctcgtcatgatatcaatgttggcacaacacaggcacacgtg catacacttcctcaggattacaagctcctcccgcgttagaaccatatcccaggg aacaacccattcctgaatcagcgtaaatcccacactgcagggaagacctcgc acgtaactcacgttgtgcattgtcaaagtgttacattcgggcagcagcggatgat cctccagtatggtagcgcgggtttctgtctcaaaaggaggtagacgatccctact gtacggagtgcgccgagacaaccgagatcgtgttggtcgtagtgtcatgccaa atggaacgccggacgtagtcatatttcctgaagcaaaaccaggtgcgggcgtg acaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagtagttgta gtatatccactctctcaaagcatccaggcgccccctggcttcgggttctatgtaaa ctccttcatgcgccgctgccctgataacatccaccaccgcagaataagccaca cccagccaacctacacattcgttctgcgagtcacacacgggaggagcgggaa gagctggaagaaccatgttttttttttttattccaaaagattatccaaaacctcaaaat gaagatctattaagtgaacgcgctcccctccggtggcgtggtcaaactctacag ccaaagaacagataatggcatttgtaagatgttgcacaatggcttccaaaaggc aaacggccctcacgtccaagtggacgtaaaggctaaacccttcagggtgaatc tcctctataaacattccagcaccttcaaccatgcccaaataattctcatctcgcca ccttctcaatatatctctaagcaaatcccgaatattaagtccggccattgtaaaaa tctgctccagagcgccctccaccttcagcctcaagcagcgaatcatgattgcaa aaattcaggttcctcacagacctgtataagattcaaaagcggaacattaacaaa aataccgcgatcccgtaggtcccttcgcagggccagctgaacataatcgtgca ggtctgcacggaccagcgcggccacttccccgccaggaaccatgacaaaag aacccacactgattatgacacgcatactcggagctatgctaaccagcgtagcc ccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaa

		aaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtcatgctcat gcagataaaggcaggtaagctccggaaccaccacagaaaaagacaccattt ttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaataacaaaaa aacatttaaacattagaagcctgtcttacaacaggaaaaacaacccttataagc ataagacggactacggccatgccggcgtgaccgtaaaaaaactggtcaccgt gattaaaaagcaccaccgacagctcctcggtcatgtccggagtcataatgtaa gactcggtaaacacatcaggttgattcacatcggtcagtgctaaaaagcgacc gaaatagcccgggggaatacatacccgcaggcgtagagacaacattacagc ccccataggaggtataacaaaattaataggagagaaaaacacataaacacct gaaaaaccctcctgcctaggcaaaatagcaccctcccgctccagaacaacat acagcgcttccacagcggcagccataacagtcagccttaccagtaaaaaaga aaacctattaaaaaaacaccactcgacacggcaccagctcaatcagtcacag tgtaaaaaagggccaagtgcagagcgagtatatataggactaaaaaatgacg taacggttaaagtccacaaaaaacacccagaaaaccgcacgcgaacctacg cccagaaacgaaagccaaaaaacccacaacttcctcaaatcgtcacttccgtt ttcccacgttacgtcacttcccattttaagaaaactacaattcccaacacatacaa gttactccgcccttaattaaatcggatccgatatctagatgtattcgcgaggtacc gagctcgaattctctggccgtcgttttacaacgtcgtgactgggaaaaccctggc gttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatag cgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggc gaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgc atatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagcccc gacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggc atccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttt tcaccgtcatcaccgaaacgcgcga
87	pEMBR- 1.5A+L4	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT

		TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT

		GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA

		AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcCACAAAAGCGAAGAT CAGCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTC TTCAGTAAATACTGCGCGCTGACTCTTAAGGACTAGTT TCGCGCCCTTTCTCAAATTTAAGCGCGAAAACTACGTC ATCTCCAGCGGCCACACCCGGCGCCAGCACCTGTTG TCAGCGCCATTGGCGCGCCcgcccgccgcgcgcttcgcttttttata gggccgccgccgccgccgcctcgccataaaaggaaactttcggagcgcgcc gctctgattggctgccgccgcacctctccgcctcgccccgccccgcccctcgcc ccgccccgccccgcctggcgcgcgcccccccccccccccgcccccatcgct gcacaaaataattaaaaaataaataaatacaaaattgggggtggggagggg ggggagatggggagagtgaagcagaacgtggggctcacctcgaGGCCG GCCGAAtatcTTCATTTAAATgtgtgtcagttagggtgtggaaagtccc caggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagca accaggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagc atgcatctcaattagtcagcaaccatagtcccgcccctaactccgcccatcccg cccctaactccgcccagttccgcccattctccgccccatggctgactaatttttttta tttatgcagaggccgaggccgcctcggcctctgagctattccagaagtagtgag

		gaggctttttggaggcctaggcttttgcaaaCGCCGGCGcaccgcgggc ccgatccaccggtactgttggtaaagccaccatgttttccggtggcggcggccc gctgtcccccggaggaaagtcggcggccagggcggcgtccgggtttttgcgc ccgccggccctcgcggagccagccggggaccccccgccttgtttgaggcaaaa cttttacaacccctacctcgccccagtcgggacgcaacagaagccgaccggg ccaacccagcgccatacgtactatagcgaatgcgatgaatttcgattcatcgcc ccgcgggtgctggacgaggatgcccccccggagaagcgcgccggggtgca cgacggtcacctcaagcgcgcccccaaggtgtactgcgggggggacgagcg cgacgtcctccgcgtcgggtcgggcggcttctggccgcggcgctcgcgcctgtg gggcggcgtggaccacgccccggcggggttcaaccccaccgtcaccgtctttc acgtgtacgacatcctggagaacgtggagcacgcgtacggcatgcgcgcggc ccagttccacgcgcggtttatggacgccatcacaccgacggggaccgtcatca cgctcctgggcctgactccggaaggccaccgggtggccgttcacgtttacggc acgcggcagtacttttacatgaacaaggaggaggtcgacaggcacctacaat gccgcgccccacgagatctctgcgagcgcatggccgcggccctgcgcgagtc cccgggcgcgtcgttccgcggcatctccgcggaccacttcgaggcggaggtg gtggagcgcaccgacgtgtactactacgagacgcgccccgctctgttttaccgc gtctacgtccgaagcgggcgcgtgctgtcgtacctgtgcgacaacttctgcccg gccatcaagaagtacgagggtggggtcgacgccaccacccggttcatcctgg acaaccccgggttcgtcaccttcggctggtaccgtctcaaaccgggccggaac aacacgctagcccagccgcgggccccgatggccttcgggacatccagcgac gtcgagtttaactgtacggcggacaacctggccatcgaggggggcatgagcg acctaccggcatacaagctcatgtgcttcgatatcgaatgcaaggcgggggggg gaggacgagctggcctttccggtggccgggcacccggaggacctggtcatcc agatatcctgtctgctctacgacctgtccaccaccgccctggagcacgtcctcct gttttcgctcggttcctgcgacctccccgaatcccacctgaacgagctggcggcc aggggcctgcccacgcccgtggttctggaattcgacagcgaattcgagatgctg ttggccttcatgacccttgtgaaacagtacggccccgagttcgtgaccgggtaca acatcatcaacttcgactggcccttcttgctggccaagctgacggacatttacaa ggtccccctggacgggtacggccgcatgaacggccggggcgtgtttcgcgtgt gggacataggccagagccacttccagaagcgcagcaagataaaggtgaac ggcatggtgaacatcgacatgtacgggattataaccgacaagatcaagctctc gagctacaagctcaacgccgtggccgaagccgtcctgaaggacaagaaga aggacctgagctatcgcgacatccccgcctactacgccgccgggcccgcgca acgcggggtgatcggcgagtactgcatacaggattccctgctggtgggccagc tgtttttaagtttttgccccatctggagctctcggccgtcgcgcgcttggcgggtatt aacatcacccgcaccatctacgacggccagcagatccgcgtctttacgtgcctg ctgcgcctggccgaccagaagggctttattctgccggacacccaggggcgattt aggggcgccgggggggaggcgcccaagcgtccggccgcagcccgggagg acgaggagcggccagaggaggaggggggaggacgaggacgaacgcgag gagggcgggggcgagcgggagccggagggcgcgcgggagaccgccggc aggcacgtggggtaccaggggccagggtccttgaccccacttccgggtttca cgtgaaccccgtggtggtgttcgactttgccagcctgtaccccagcatcatccag gcccacaacctgtgcttcagcacgctctccctgagggccgacgcagtggcgca cctggaggcgggcaaggactacctggagatcgaggtgggggggcgacggct gttcttcgtcaaggctcacgtgcgagagagcctcctcagcatcctcctgcggga ctggctcgccatgcgaaagcagatccgctcgcggattccccagagcagcccc

		gaggaggccgtgctcctggacaagcagcaggccgccatcaaggtcgtgtgta actcggtgtacgggttcacgggagtgcagcacggactcctgccgtgcctgcac gttgccgcgacggtgacgaccatcggccgcgagatgctgctcgcgacccgcg agtacgtccacgcgcgctgggcggccttcgaacagctcctggccgatttcccgg aggcggccgacatgcgcgcccccgggccctattccatgcgcatcatctacggg gacacggactccatctttgtgctgtgccgcggcctcacggccgccgggctgacg gccgtgggcgacaagatggcgagccacatctcgcgcgcgctgtttctgccccc catcaaactcgagtgcgaaaagacgttcaccaagctgctgctgatcgccaaga aaaagtacatcggcgtcatctacgggggtaagatgctcatcaagggcgtggat ctggtgcgcaaaaacaactgcgcgtttatcaaccgcacctccagggccctggt cgacctgctgttttacgacgataccgtctccggagcCgccgcGgcgttagccg agcgccccgcggaggagtggctggcgcgacccctgcccgagggactgcag gcgttcggggccgtcctcgtagacgcccatcggcgcatcaccgacccggaga gggacatccaggactttgtcctcaccgccgaactgagcagacacccgcgcgc gtacaccaacaagcgcctggcccacctgacggtgtattacaagctcatggccc gccgcgcgcaggtcccgtccatcaaggaccggatcccgtacgtgatcgtggcc cagacccgcgaggtagaggagacggtcgcgcggctggccgccctccgcga gctagacgccgccgccccaggggacgagcccgccccccccgcggccctgc cctccccggccaagcgcccccgggagacgccgtcgcctgccgacccccccgg gaggcgcgtccaagccccgcaagctgctggtgtccgagctggccgaggatcc cgcatacgccattgcccacggcgtcgccctgaacacggactattacttctccca cctgttgggggcggcgtgcgtgacattcaaggccctgtttgggaataacgccaa gatcaccgagagtctgttaaaaaggtttattcccgaagtgtggcacccccggga cgacgtggccgcgcggctccggaccgcagggttcggggcggtgggtgccgg cgctacggcggaggaaactcgtcgaatgttgcatagagcctttgatactctagc agaattcggcagtggagcaacaaacttctctctgctgaaacaagccggagatg tcgaagagaatcctggaccgacggattcccctggcggtgtggcccccgcctcc cccgtggaggacgcgtcggacgcgtccctcgggcagccggaggagggggc gccctgccaggtggtcctgcagggcgccgaacttaatggaatcctacaggcgtt tgccccgctgcgcacgagccttctggactcgcttctggttatgggcgaccgggg catccttatccataacacgatctttggggagcaggtgttcctgcccctggaacact cgcaattcagtcggtatcgctggcgcggacccacggcggcgttcctgtctctcgt ggaccagaagcgctccctcctgagcgtgtttcgcgccaaccagtacccggacc tacgtcgggtggagttggcgatcacgggccaggccccgtttcgcacgctggttc agcgcatatggacgacgacgtccgacggcgaggccgttgagctagccagcg agacgctgatgaagcgcgaactgacgagctttgtggtgctggttccccaggga acccccgacgttcagttgcgcctgacgaggccgcagctcaccaaggtccttaa cgcgaccggggccgatagtgccacgcccaccacgttcgagctcggggttaac ggcaaattttccgtgttcaccacgagtacctgcgtcacctttgctgcccgcgagg agggcgtgtcgtccagcaccagcacccaggtccagatcctgtccaacgcgctc accaaggcgggccaggcCgccgcGaacgccaagacggtgtacggggaa aatacccatcgcaccttctctgtggtcgtcgacgattgcagcatgcgggcggtgc tccggcgactgcaggtcggcgggggcaccctcaagttcttcctcacgacccccc gtccccagtctgtgcgtcaccgccaccggtcccaacgcggtatcggcggtatttc tcctgaaaccccagaagatttgcctggactggctgggtcatagccaggggtctc cttcagccgggagctcggcctcccgggcctctgggagcgagccaacagaca gccaggactccgcgtcggacgcggtcagccacggcgatccggaagacctcg

		atggcgctgcccgggcgggagaggcgggggccttgcatgcctgtccgatgcc gtcgtcgaccacgcgggtcactcccacgaccaagcgggggcgctcgggggg cgaggatgcgcgcgcggacacggccctaaagaaacctaagacggggtcgc ccaccgcacccccgcccgcagatccagtcccctggacacggaggacgact ccgatgcggcggacgggacggcggcccgtcccgccgctccagacgcccgg agcggaagccgttacgcgtgttactttcgcgacctcccgaccggagaagcaag ccccggcgccttctccgccttccgggggggccccccaaaccccgtatggttttgg attccccTGAtaggatccgactgcaggtagctgtgccttctagttgccagccatc tgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgt cctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctatt ctgggggtggggtggggcaggacagcaagggggaggattgggaagacaa tagcaggcatgctggggatgcggtgggctctatggGTTTAAACATCGA TGCGGCCGCgattaggttctacgaagaccaatcccgcccgcctaatgcg gagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaagcc atcaacaaagcccgccaagagtttctgctacgaaagggacgggggtttactt ggacccccagtccggcgaggagctcaacccaatccccccgccgccgcagcc ctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaagaa gctgcagctgccgccgccacccacggacgaggaggaatactgggacagtca ggcagaggaggttttggacgaggaggaggaggacatgatggaagactggg agagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaacac cgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccggt tccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcgc cgacccaaccgtagATGGGACACCACTGGAACCAGGGCCG GTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAGC AACAACAGCGCCAAGGCTACCGCTCATGGCGCGGGC ACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTGG GGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTAC CATCACGGCGTGGCCTTCCCCCGTAACATCCTGCATT ACTACCGTCATCTCTACAGCCCATACTGCACCGGCGG CAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAAA GGCGACCGGATAGCAAGACTCTGACAAAGCCCAAGA AATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGaataaaagatctttattttcattaga tctgtgtgttggttttttgtgtgGCGGCCGCCGTTTGTGTTATGTTT CAACGTGTTTATTTTTCAATTGCAGAAAATTTCAAGTCA TTTTTCATTCAGTAGTATAGCCCCACCACCACATAGCT TATACAGATCACCGTACCTTAATCAAACTCACAGAACC CTAGTATTCAACCTGCCACCTCCCTCCCAACACACAG AGTACACAGTCCTTTCTCCCCGGCTGGCCTTAAAAAG CATCATATCATGGGTAACAGACATATTCTTAGGTGTTA TATTCCACACGGTTTCCTGTCGAGCCAAACGCTCATC

		AGTGATATTAATAAACTCCCCGGGCAGCTCACTTAAGT TCATGTCGCTGTCCAGCTGCTGAGCCACAGGCTGCTG TCCAACTTGCGGTTGCTTAACGGGCGGCGAAGGAGA AGTCCACGCCTACATGGGGGTAGAGTCATAATCGTGC ATCAGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGA ATAAACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAAT ACAACATGGCAGTGGTCTCCTCAGCGATGATTCGCAC CGCCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACA GCAGCGCACCCTGATCTCACTTAAATCAGCACAGTAA CTGCAGCACAGCACCACAATATTGTTCAAAATCCCACA GTGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGAC CACAGAACCCACGTGGCCATCATACCACAAGCGCAG GTAGATTAAGTGGCGACCCCTCATAAACACGCTGGAC ATAAACATTACCTCTTTTGGCATGTTGTAATTCACCAC CTCCCGGTACCATATAAACCTCTGATTAAACATGGCG CCATCCACCACCATCCTAAACCAGCTGGCCAAAACCT GCCCGCCGGCTATACACTGCAGGGAACCGGGACTGG AACAATGACAGTGGAGAGCCCAGGACTCGTAACCATG GATCATCATGCTCGTCATGATATCAATGTTGGCACAAC ACAGGCACACGTGCATACACTTCCTCAGGATTACAAG CTCCTCCCGCGTTAGAACCATATCCCAGGGAACAACC CATTCCTGAATCAGCGTAAATCCCACACTGCAGGGAA GACCTCGCACGTAACTCACGTTGTGCATTGTCAAAGT GTTACATTCGGGCAGCAGCGGATGATCCTCCAGTATG GTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGAT CCCTACTGTACGGAGTGCGCCGAGACAACCGAGATC GTGTTGGTCGTAGTGTCATGCCAAATGGAACGCCGGA CGTAGTCATATTTCCTGAAGCAAAACCAGGTGCGGGC GTGACAAACAGATCTGCGTCTCCGGTCTCGCCGCTTA GATCGCTCTGTGTAGTAGTTGTAGTATATCCACTCTCT CAAAGCATCCAGGCGCCCCCTGGCTTCGGGTTCTATG TAAACTCCTTCATGCGCCGCTGCCCTGATAACATCCA CCACCGCAGAATAAGCCACACCCAGCCAACCTACACA TTCGTTCTGCGAGTCACACACGGGAGGAGCGGGAAG AGCTGGAAGAACCATGTTTTTTTTTTTATTCCAAAAGAT TATCCAAAACCTCAAAATGAAGATCTATTAAGTGAACG CGCTCCCCTCCGGTGGCGTGGTCAAACTCTACAGCCA AAGAACAGATAATGGCATTTGTAAGATGTTGCACAATG GCTTCCAAAAGGCAAACGGCCCTCACGTCCAAGTGGA CGTAAAGGCTAAACCCTTCAGGGTGAATCTCCTCTATA AACATTCCAGCACCTTCAACCATGCCCAAATAATTCTC ATCTCGCCACCTTCTCAATATATCTCTAAGCAAATCCC GAATATTAAGTCCGGCCATTGTAAAAATCTGCTCCAGA GCGCCCTCCACCTTCAGCCTCAAGCAGCGAATCATGA TTGCAAAAATTCAGGTTCCTCACAGACCTGTATAAGAT TCAAAAGCGGAACATTAACAAAAATACCGCGATCCCG TAGGTCCCTTCGCAGGGCCAGCTGAACATAATCGTGC

		AGGTCTGCACGGACCAGCGCGGCCACTTCCCCGCCA GGAACCATGACAAAAGAACCCACACTGATTATGACAC GCATACTCGGAGCTATGCTAACCAGCGTAGCCCCGAT GTAAGCTTGTTGCATGGGCGGCGATATAAAATGCAAG GTGCTGCTCAAAAAATCAGGCAAAGCCTCGCGCAAAA AAGAAAGCACATCGTAGTCATGCTCATGCAGATAAAG GCAGGTAAGCTCCGGAACCACCACAGAAAAAGACACC ATTTTTCTCTCAAACATGTCTGCGGGTTTCTGCATAAA CACAAAATAAAATAACAAAAAAACATTTAAACATTAGAA GCCTGTCTTACAACAGGAAAAACAACCCTTATAAGCAT AAGACGGACTACGGCCATGCCGGCGTGACCGTAAAA AAACTGGTCACCGTGATTAAAAAGCACCACCGACAGC TCCTCGGTCATGTCCGGAGTCATAATGTAAGACTCGG TAAACACATCAGGTTGATTCACATCGGTCAGTGCTAAA AAGCGACCGAAATAGCCCGGGGGAATACATACCCGC AGGCGTAGAGACAACATTACAGCCCCCATAGGAGGTA TAACAAAATTAATAGGAGAGAAAAACACATAAACACCT GAAAAACCCTCCTGCCTAGGCAAAATAGCACCCTCCC GCTCCAGAACAACATACAGCGCTTCCACAGCGGCAGC CATAACAGTCAGCCTTACCAGTAAAAAAGAAAACCTAT TAAAAAAACACCACTCGACACGGCACCAGCTCAATCA GTCACAGTGTAAAAAAGGGCCAAGTGCAGAGCGAGTA TATATAGGACTAAAAAATGACGTAACGGTTAAAGTCCA CAAAAAACACCCAGAAAACCGCACGCGAACCTACGCC CAGAAACGAAAGCCAAAAAACCCACAACTTCCTCAAAT CGTCACTTCCGTTTTCCCACGTTACGTCACTTCCCATT TTAAGAAAACTACAATTCCCAACACATACAAGTTACTC CGCCCTTAATTAAATCGGATCCGATATCTAGATGTATT CGCGAGGTACCGAGCTCGAATTCTCTGGCCGTCGTTT TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCT TCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGC CTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTAT TTCACACCGCATATGGTGCACTCTCAGTACAATCTGCT CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGC CAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGC TCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCA CCGAAACGCGCGA
88	pEMBR-1.7B2	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat

		gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgccccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccctctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat

		tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcatggagtcagtcgagaagaa ggacagcctaaccgccccctctgagttcgccaccaccgcctccaccgatgccg

		ccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggagga agtgattatcgagcaggacccaggttttgtaagcgaagacgacgaggaccgct cagtaccaacagaggataaaaagcaagaccaggacaacgcagaggcaaa cgaggaacaagtcgggcgggggacgaaaggcatggcgactacctagatgt gggagacgacgtgctgttgaagcatctgcagcgccagtgcgccattatctgcg acgcgttgcaagagcgcagcgatgtgcccctcgccatagcggatgtcagcctt gcctacgaacgccacctattctcaccgcgcgtacccccccaaacgccaagaaa acggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgccgtgc cagaggtgcttgccacctatcacatctttttccaaaactgcaagataccctatcc tgccgtgccaaccgcagccgagcggacaagcagctggccttgcggcagggc gctgtcatacctgatatcgcctcgctcaacgaagtgccaaaaatctttgagggtct tggacgcgacgagaagcgcgcggcaaacgctctgcaacaggaaaacagc gaaaatgaaagtcactctggagtgttggtggaactcgagggtgacaacgcgc gcctagccgtactaaaacgcagcatcgaggtcacccactttgcctacccggca cttaacctacccccccaaggtcatgagcacagtcatgagtgagctgatcgtgcgc cgtgcgcagcccctggagagggatgcaaatttgcaagaacaaacagaggag ggcctacccgcagttggcgacgagcagctagcgcgctggcttcaaacgcgcg agcctgccgacttggaggagcgacgcaaactaatgatggccgcagtgctcgtt accgtggagcttgagtgcatgcagcggttctttgctgacccggagatgcagcgc aagctagaggaaacattgcactacacctttcgacagggctacgtacgccaggc ctgcaagatctccaacgtggagctctgcaacctggtctcctaccttggaattttgc acgaaaaccgccttgggcaaaacgtgcttcattccacgctcaagggcgaggc gcgccgcgactacgtccgcgactgcgtttacttatttctatgctacacctggcaga cggccatgggcgtttggcagcagtgcttggaggagtgcaacctcaaggagctg cagaaactgctaaagcaaaacttgaaggacctatggacggccttcaacgagc gctccgtggccgcgcacctggcggacatcattttccccgaacgcctgcttaaaa ccctgcaacagggtctgccagacttcaccagtcaaagcatgttgcagaacttta ggaactttatcctagagcgctcaggaatcttgcccgccacctgctgtgcacttcct agcgactttgtgcccattaagtaccgcgaatgccctccgccgctttggggccact gctaccttctgcagctagccaactaccttgcctaccactctgacataatggaaga cgtgagcggtgacggtctactggagtgtcactgtcgctgcaacctatgcacccc gcaccgctccctggtttgcaattcgcagctgcttaacgaaagtcaaattatcggta cctttgagctgcagggtccctcgcctgacgaaaagtccgcggctccggggttga aactcactccggggctgtggacgtcggcttaccttcgcaaatttgtacctgagga ctaccacgcccacgagattaggttctacgaagaccaatcccgcccgcctaatg cggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaag ccatcaacaaagcccgccaagagtttctgctacgaaagggacgggggttta cttggacccccagtccggcgaggagctcaacccaatccccccgccgccgcag ccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaag aagctgcagctgccgccgccacccacggacgaggaggaatactgggacagt caggcagaggaggttttggacgaggaggaggaggacatgatggaagactgg gagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaaca ccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccgg ttccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcg ccgacccaaccgtagATGGGACACCACTGGAACCAGGGCC GGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAG CAACAACAGCGCCAAGGCTACCGCTCATGGCGCGGG

		CACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTG GGGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTA CCATCACGGCGTGGCCTTCCCCCGTAACATCCTGCAT TACTACCGTCATCTCTACAGCCCATACTGCACCGGCG GCAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAA AGGCGACCGGATAGCAAGACTCTGACAAAGCCCAAG AAATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGTTTCGCGCCCTTTCTC AAATTTAAGCGCGAAAACTACGTCATCTCCAGCGGCC ACACCCGGCGCCAGCACCTGTTGTCAGCGCCATTATG AGCAAGGAAATTCCCACGCCCTACATGTGGAGTTACC AGCCACAAATGGGACTTGCGGCTGGAGCTGCCCAAG ACTACTCAACCCGAATAAACTACATGAGCGCGGGGCG GCCGCAACTTGTTTATTGCAGCTTATAATGGTTACAAA TAAAGCAATAGCATCACAAATTTCACAAATAAAGCATT TTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT CAATGTATCTTAGCTTAACGGGCGGCGAAGGAGAAGT CCACGCCTACATGGGGGTAGAGTCATAATCGTGCATC AGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGAATA AACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAATAC AACATGGCAGTGGTCTCCTCAGCGATGATTCGCACCG CCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACAGC AGCGCACCCTGATCTCACTTAAATCAGCACAGTAACT GCAGCACAGCACCACAATATTGTTCAAAATCCCACAG TGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGACC ACAGAACCCACGTGGCCATCATACCACAAGCGCAGGT AGATTAAGTGGCGACCCCTCATAAACACGCTGGACAT AAACATTACCTCTTTTGGCATGTTGTAATTCACCACCT CCCGGTACCATATAAACCTCTGATTAAACATGGCGCC ATCCACCACCATCCTAAACCAGCTGGCCAAAACCTGC CCGCCGGCTATACACTGCAGGGAACCGGGACTGGAA CAATGACAGTGGAGAGCCCAGGACTCGTAACCATGGA TCATCATGCTCGTCATGATATCAATGTTGGCACAACAC AGGCACACGTGCATACACTTCCTCAGGATTACAAGCT CCTCCCGCGTTAGAACCATATCCCAGGGAACAACCCA TTCCTGAATCAGCGTAAATCCCACACTGCAGGGAAGA CCTCGCACGTAACTCACGTTGTGCATTGTCAAAGTGTT ACATTCGGGCAGCAGCGGATGATCCTCCAGTATGGTA GCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGATCCC TACTGTACGGAGTGCGCCGAGACAACCGAGATCGTGT TGGTCGTAGTGTCATGCCAAATGGAACGCCGGACGTA

		GTCATATTTCCTGAAGCAAAACCAGGTGCGGGCGTGA CAAACAGATCTGCGTCTCCGGTCTCGCCGCTTAGATC GCTCTGTGTAGTAGTTGTAGTATATCCACTCTCTCAAA GCATCCAGGCGCCCCCTGGCTTCGGGTTCTATGTAAA CTCCTTCATGCGCCGCTGCCCTGATAACATCCACCAC CGCAGAATAAGCCACACCCAGCCAACCTACACATTCG TTCTGCGAGTCACACACGGGAGGAGCGGGAAGAGCT GGAAGAACCATGGTGGCATTTGCAAAAGCCTAGGCCT CCAAAAAAGCCTCCTCACTACTTCTGGAATAGCTCAGA GGCCGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAA ATTAGTCAGCCATGGGGCGGAGAATGGGCGGAACTG GGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGG CGGGACTATGGTTGCTGACTAATTGAGATGCATGCTT TGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTC CACACCTGGTTGCTGACTAATTGAGATGCATGCTTTG CATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCA CACCCTAACTGACACACACGTTACGTCACTTCCCATTT TAAGAAAACTACAATTCCCAACACATACAAGTTACTCC GCCCTTAATtaaatcggatccgatatctagatgtattcgcgaggtaccgag ctcgaattctctggccgtcgttttacaacgtcgtgactgggaaaaccctggcgtta cccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcga agaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaa tggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatat ggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgac acccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatcc gcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcac cgtcatcaccgaaacgcgcga
89	pEMBR-1.8+L4	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG

		TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA

		TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTgtacactctcgggtgattatttacccccaccctt gccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgcc actggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggca caaccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatca ccaacgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctc cgccctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatc agcgccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgc gtccaggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgcctt cccaaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcat caaaaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaa gccttgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgc cgcaagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgca gcaccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttca

		cgatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtca catccatttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagc tcgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctc gtgatgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcg ccccatcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcg gtgctcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtca ggcagtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagc gcgcgcgcagcctccatgcccttctcccacgcagacacgatcggcacactca gcgggttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcg tccgcataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgctt acctcctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtag cgccacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgc tcgggcttgggagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgc cgccgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtg atgagtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcg cccggggaggcggcggcgacggggacggggacgacacgtcctccatggttg ggggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcct cttcccgactggccatggtggcttccttctcctataggcagaaaaagggcgcgc catccgcccgccgcgcgcttcgcttttatagggccgccgccgccgccgcctcg ccataaaaggaaactttcggagcgcgccgctctgattggctgccgccgcacct ctccgcctcgccccgccccgcccctcgccccgccccgccccgcctggcgcgc gccccccccccccccccgcccccatcgctgcacaaaataattaaaaaataaat aaatacaaaattgggggtggggaggggggggagatggggagagtgaagca gaacgtggggctcacctcgaccatggtaatagcgatgactaatacgtagatgta ctgccaagtaggaaagtcccgtaaggtcatgtactgggcataatgccaggcgg gccatttaccgtcattgacgtcaataggggggcggacttggcatatgatacacttg atgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtcaatgg aaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcg ggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgcg gaactccatatatgggctatgaactaatgaccccgtaattgattactattaataact agtcaataatcaatgccggccggccgaatatcttcatttaaatgtttaaacatcga tgcGGCCGCgattaggttctacgaagaccaatcccgcccgcctaatgcgg agcttaccgcctgcgtcattacccagggccacattcttggccaattgcaagccat caacaaagcccgccaagagtttctgctacgaaagggacgggggggtttacttgg acccccagtccggcgaggagctcaacccaatccccccgccgccgcagccct atcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaagaag ctgcagctgccgccgccacccacggacgaggaggaatactgggacagtcag gcagaggaggttttggacgaggaggaggaggacatgatggaagactggga gagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaacacc gtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccggtt ccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcgc cgacccaaccgtagATGGGACACCACTGGAACCAGGGCCG GTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAGC AACAACAGCGCCAAGGCTACCGCTCATGGCGCGGGC ACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTGG GGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTAC CATCACGGCGTGGCCTTCCCCCGTAACATCCTGCATT

		ACTACCGTCATCTCTACAGCCCATACTGCACCGGCGG CAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAAA GGCGACCGGATAGCAAGACTCTGACAAAGCCCAAGA AATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGaataaaagatctttattttcattaga tctgtgtgttggttttttgtgtgGCGGCCGCCGTTTGTGTTATGTTT CAACGTGTTTATTTTTCAATTGCAGAAAATTTCAAGTCA TTTTTCATTCAGTAGTATAGCCCCACCACCACATAGCT TATACAGATCACCGTACCTTAATCAAACTCACAGAACC CTAGTATTCAACCTGCCACCTCCCTCCCAACACACAG AGTACACAGTCCTTTCTCCCCGGCTGGCCTTAAAAAG CATCATATCATGGGTAACAGACATATTCTTAGGTGTTA TATTCCACACGGTTTCCTGTCGAGCCAAACGCTCATC AGTGATATTAATAAACTCCCCGGGCAGCTCACTTAAGT TCATGTCGCTGTCCAGCTGCTGAGCCACAGGCTGCTG TCCAACTTGCGGTTGCTTAACGGGCGGCGAAGGAGA AGTCCACGCCTACATGGGGGTAGAGTCATAATCGTGC ATCAGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGA ATAAACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAAT ACAACATGGCAGTGGTCTCCTCAGCGATGATTCGCAC CGCCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACA GCAGCGCACCCTGATCTCACTTAAATCAGCACAGTAA CTGCAGCACAGCACCACAATATTGTTCAAAATCCCACA GTGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGAC CACAGAACCCACGTGGCCATCATACCACAAGCGCAG GTAGATTAAGTGGCGACCCCTCATAAACACGCTGGAC ATAAACATTACCTCTTTTGGCATGTTGTAATTCACCAC CTCCCGGTACCATATAAACCTCTGATTAAACATGGCG CCATCCACCACCATCCTAAACCAGCTGGCCAAAACCT GCCCGCCGGCTATACACTGCAGGGAACCGGGACTGG AACAATGACAGTGGAGAGCCCAGGACTCGTAACCATG GATCATCATGCTCGTCATGATATCAATGTTGGCACAAC ACAGGCACACGTGCATACACTTCCTCAGGATTACAAG CTCCTCCCGCGTTAGAACCATATCCCAGGGAACAACC CATTCCTGAATCAGCGTAAATCCCACACTGCAGGGAA GACCTCGCACGTAACTCACGTTGTGCATTGTCAAAGT GTTACATTCGGGCAGCAGCGGATGATCCTCCAGTATG GTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGAT CCCTACTGTACGGAGTGCGCCGAGACAACCGAGATC GTGTTGGTCGTAGTGTCATGCCAAATGGAACGCCGGA CGTAGTCATATTTCCTGAAGCAAAACCAGGTGCGGGC

		GTGACAAACAGATCTGCGTCTCCGGTCTCGCCGCTTA GATCGCTCTGTGTAGTAGTTGTAGTATATCCACTCTCT CAAAGCATCCAGGCGCCCCCTGGCTTCGGGTTCTATG TAAACTCCTTCATGCGCCGCTGCCCTGATAACATCCA CCACCGCAGAATAAGCCACACCCAGCCAACCTACACA TTCGTTCTGCGAGTCACACACGGGAGGAGCGGGAAG AGCTGGAAGAACCATGTTTTTTTTTTTATTCCAAAAGAT TATCCAAAACCTCAAAATGAAGATCTATTAAGTGAACG CGCTCCCCTCCGGTGGCGTGGTCAAACTCTACAGCCA AAGAACAGATAATGGCATTTGTAAGATGTTGCACAATG GCTTCCAAAAGGCAAACGGCCCTCACGTCCAAGTGGA CGTAAAGGCTAAACCCTTCAGGGTGAATCTCCTCTATA AACATTCCAGCACCTTCAACCATGCCCAAATAATTCTC ATCTCGCCACCTTCTCAATATATCTCTAAGCAAATCCC GAATATTAAGTCCGGCCATTGTAAAAATCTGCTCCAGA GCGCCCTCCACCTTCAGCCTCAAGCAGCGAATCATGA TTGCAAAAATTCAGGTTCCTCACAGACCTGTATAAGAT TCAAAAGCGGAACATTAACAAAAATACCGCGATCCCG TAGGTCCCTTCGCAGGGCCAGCTGAACATAATCGTGC AGGTCTGCACGGACCAGCGCGGCCACTTCCCCGCCA GGAACCATGACAAAAGAACCCACACTGATTATGACAC GCATACTCGGAGCTATGCTAACCAGCGTAGCCCCGAT GTAAGCTTGTTGCATGGGCGGCGATATAAAATGCAAG GTGCTGCTCAAAAAATCAGGCAAAGCCTCGCGCAAAA AAGAAAGCACATCGTAGTCATGCTCATGCAGATAAAG GCAGGTAAGCTCCGGAACCACCACAGAAAAAGACACC ATTTTTCTCTCAAACATGTCTGCGGGTTTCTGCATAAA CACAAAATAAAATAACAAAAAAACATTTAAACATTAGAA GCCTGTCTTACAACAGGAAAAACAACCCTTATAAGCAT AAGACGGACTACGGCCATGCCGGCGTGACCGTAAAA AAACTGGTCACCGTGATTAAAAAGCACCACCGACAGC TCCTCGGTCATGTCCGGAGTCATAATGTAAGACTCGG TAAACACATCAGGTTGATTCACATCGGTCAGTGCTAAA AAGCGACCGAAATAGCCCGGGGGAATACATACCCGC AGGCGTAGAGACAACATTACAGCCCCCATAGGAGGTA TAACAAAATTAATAGGAGAGAAAAACACATAAACACCT GAAAAACCCTCCTGCCTAGGCAAAATAGCACCCTCCC GCTCCAGAACAACATACAGCGCTTCCACAGCGGCAGC CATAACAGTCAGCCTTACCAGTAAAAAAGAAAACCTAT TAAAAAAACACCACTCGACACGGCACCAGCTCAATCA GTCACAGTGTAAAAAAGGGCCAAGTGCAGAGCGAGTA TATATAGGACTAAAAAATGACGTAACGGTTAAAGTCCA CAAAAAACACCCAGAAAACCGCACGCGAACCTACGCC CAGAAACGAAAGCCAAAAAACCCACAACTTCCTCAAAT CGTCACTTCCGTTTTCCCACGTTACGTCACTTCCCATT TTAAGAAAACTACAATTCCCAACACATACAAGTTACTC CGCCCTTAATTAAATCGGATCCGATATCTAGATGTATT

		CGCGAGGTACCGAGCTCGAATTCTCTGGCCGTCGTTT TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCT TCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGC CTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTAT TTCACACCGCATATGGTGCACTCTCAGTACAATCTGCT CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGC CAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGC TCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCA CCGAAACGCGCGA
90	pEMBR-1.9+L4	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT

		GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC

		GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTgtacactctcgggtgattatttacccccaccctt gccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgcc actggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggca caaccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatca ccaacgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctc cgccctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatc agcgccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgc gtccaggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgcctt cccaaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcat caaaaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaa gccttgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgc cgcaagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgca gcaccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttca cgatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtca catccatttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagc tcgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctc gtgatgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcg ccccatcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcg gtgctcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtca ggcagtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagc gcgcgcgcagcctccatgcccttctcccacgcagacacgatcggcacactca gcgggttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcg tccgcataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgctt acctcctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtag cgccacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgc tcgggcttgggagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgc cgccgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtg atgagtcttcctcgtcctcggactcgatacgccgcctcatccgcttttttgggggcg

		cccggggaggcggcggcgacggggacggggacgacacgtcctccatggttg ggggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcct cttcccgactggccatggtggcttccttctcctataggcagaaaaagggcgcgc cctgtggagagaaaggcaaagtggatgtcagtaagaccaataggtgcctatc agaaacgcaagagtcttctctgtctcgacaagcccagtttctattggtctccttaaa cctgtcttgtaaccttgatacttacatccgcccgccgcgcgcttcgcttttttataggg ccgccgccgccgccgcctcgccataaaaggaaactttcggagcgcgccgctc tgattggctgccgccgcacctctccgcctcgccccgccccgcccctcgccccgc cccgccccgcctggcgcgcgccccccccccccccccgccccatcgctgcac aaaataattaaaaaataaataaatacaaaattgggggtggggaggggggggg agatggggagagtgaagcagaacgtggggctcacctcgaccatggtaatagc gatgactaatacgtagatgtactgccaagtaggaaagtcccgtaaggtcatgta ctgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggcg gacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatac tccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacg tcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccat ttaccgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccc cgtaattgattactattaataactagtcaataatcaatgccggccggccgaatatc ttcatttaaatgtttaaacatcgatgcGGCCGCgattaggttctacgaagacc aatcccgcccgcctaatgcggagcttaccgcctgcgtcattacccagggccac attcttggccaattgcaagccatcaacaaagcccgccaagagtttctgctacga aagggacgggggggtttacttggacccccagtccggcgaggagctcaacccaa tccccccgccgccgcagccctatcagcagcagccgcgggcccttgcttcccag gatggcacccaaaaagaagctgcagctgccgccgccacccacggacgagg aggaatactgggacagtcaggcagaggaggttttggacgaggaggaggagg acatgatggaagactgggagagcctagacgaggaagcttccgaggtcgaag aggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgccggcgccc cagaaatcggcaaccggttccagcatggctacaacctccgctcctcaggcgcc gccggcactgcccgttcgccgacccaaccgtagATGGGACACCACT GGAACCAGGGCCGGTAAGTCCAAGCAGCCGCCGCCG TTAGCCCAAGAGCAACAACAGCGCCAAGGCTACCGCT CATGGCGCGGGCACAAGAACGCCATAGTTGCTTGCTT GCAAGACTGTGGGGGCAACATCTCCTTCGCCCGCCG CTTTCTTCTCTACCATCACGGCGTGGCCTTCCCCCGT AACATCCTGCATTACTACCGTCATCTCTACAGCCCATA CTGCACCGGCGGCAGCGGCAGCAACAGCAGCGGCC ACACAGAAGCAAAGGCGACCGGATAGCAAGACTCTGA CAAAGCCCAAGAAATCCACAGCGGCGGCAGCAGCAG GAGGAGGAGCGCTGCGTCTGGCGCCCAACGAACCCG TATCGACCCGCGAGCTTAGAAACAGGATTTTTCCCAC TCTGTATGCTATATTTCAACAGAGCAGGGGCCAAGAA CAAGAGCTGAAAATAAAAAACAGGTCTCTGCGATCCC TCACCCGCAGCTGCCTGTATCACAAAAGCGAAGATCA GCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTCTT CAGTAAATACTGCGCGCTGACTCTTAAGGACTAGaataa aagatctttattttcattagatctgtgtgttggtttttgtgtgGCGGCCGCCGT TTGTGTTATGTTTCAACGTGTTTATTTTTCAATTGCAGA

		AAATTTCAAGTCATTTTTCATTCAGTAGTATAGCCCCA CCACCACATAGCTTATACAGATCACCGTACCTTAATCA AACTCACAGAACCCTAGTATTCAACCTGCCACCTCCCT CCCAACACACAGAGTACACAGTCCTTTCTCCCCGGCT GGCCTTAAAAAGCATCATATCATGGGTAACAGACATAT TCTTAGGTGTTATATTCCACACGGTTTCCTGTCGAGCC AAACGCTCATCAGTGATATTAATAAACTCCCCGGGCA GCTCACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGC CACAGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGC GGCGAAGGAGAAGTCCACGCCTACATGGGGGTAGAG TCATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCA GCAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCG TCCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGC GATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTC CTCCGGGCACAGCAGCGCACCCTGATCTCACTTAAAT CAGCACAGTAACTGCAGCACAGCACCACAATATTGTT CAAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTC ATGGCGGGGACCACAGAACCCACGTGGCCATCATAC CACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAA ACACGCTGGACATAAACATTACCTCTTTTGGCATGTTG TAATTCACCACCTCCCGGTACCATATAAACCTCTGATT AAACATGGCGCCATCCACCACCATCCTAAACCAGCTG GCCAAAACCTGCCCGCCGGCTATACACTGCAGGGAA CCGGGACTGGAACAATGACAGTGGAGAGCCCAGGAC TCGTAACCATGGATCATCATGCTCGTCATGATATCAAT GTTGGCACAACACAGGCACACGTGCATACACTTCCTC AGGATTACAAGCTCCTCCCGCGTTAGAACCATATCCC AGGGAACAACCCATTCCTGAATCAGCGTAAATCCCAC ACTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGC ATTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGAT CCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGG AGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGA CAACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAAT GGAACGCCGGACGTAGTCATATTTCCTGAAGCAAAAC CAGGTGCGGGCGTGACAAACAGATCTGCGTCTCCGG TCTCGCCGCTTAGATCGCTCTGTGTAGTAGTTGTAGTA TATCCACTCTCTCAAAGCATCCAGGCGCCCCCTGGCT TCGGGTTCTATGTAAACTCCTTCATGCGCCGCTGCCC TGATAACATCCACCACCGCAGAATAAGCCACACCCAG CCAACCTACACATTCGTTCTGCGAGTCACACACGGGA GGAGCGGGAAGAGCTGGAAGAACCATGTTTTTTTTTT TATTCCAAAAGATTATCCAAAACCTCAAAATGAAGATC TATTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCA AACTCTACAGCCAAAGAACAGATAATGGCATTTGTAAG ATGTTGCACAATGGCTTCCAAAAGGCAAACGGCCCTC ACGTCCAAGTGGACGTAAAGGCTAAACCCTTCAGGGT GAATCTCCTCTATAAACATTCCAGCACCTTCAACCATG

		CCCAAATAATTCTCATCTCGCCACCTTCTCAATATATC TCTAAGCAAATCCCGAATATTAAGTCCGGCCATTGTAA AAATCTGCTCCAGAGCGCCCTCCACCTTCAGCCTCAA GCAGCGAATCATGATTGCAAAAATTCAGGTTCCTCACA GACCTGTATAAGATTCAAAAGCGGAACATTAACAAAAA TACCGCGATCCCGTAGGTCCCTTCGCAGGGCCAGCT GAACATAATCGTGCAGGTCTGCACGGACCAGCGCGG CCACTTCCCCGCCAGGAACCATGACAAAAGAACCCAC ACTGATTATGACACGCATACTCGGAGCTATGCTAACC AGCGTAGCCCCGATGTAAGCTTGTTGCATGGGCGGC GATATAAAATGCAAGGTGCTGCTCAAAAAATCAGGCA AAGCCTCGCGCAAAAAAGAAAGCACATCGTAGTCATG CTCATGCAGATAAAGGCAGGTAAGCTCCGGAACCACC ACAGAAAAAGACACCATTTTTCTCTCAAACATGTCTGC GGGTTTCTGCATAAACACAAAATAAAATAACAAAAAAA CATTTAAACATTAGAAGCCTGTCTTACAACAGGAAAAA CAACCCTTATAAGCATAAGACGGACTACGGCCATGCC GGCGTGACCGTAAAAAAACTGGTCACCGTGATTAAAA AGCACCACCGACAGCTCCTCGGTCATGTCCGGAGTCA TAATGTAAGACTCGGTAAACACATCAGGTTGATTCACA TCGGTCAGTGCTAAAAAGCGACCGAAATAGCCCGGG GGAATACATACCCGCAGGCGTAGAGACAACATTACAG CCCCCATAGGAGGTATAACAAAATTAATAGGAGAGAA AAACACATAAACACCTGAAAAACCCTCCTGCCTAGGC AAAATAGCACCCTCCCGCTCCAGAACAACATACAGCG CTTCCACAGCGGCAGCCATAACAGTCAGCCTTACCAG TAAAAAAGAAAACCTATTAAAAAAACACCACTCGACAC GGCACCAGCTCAATCAGTCACAGTGTAAAAAAGGGCC AAGTGCAGAGCGAGTATATATAGGACTAAAAAATGAC GTAACGGTTAAAGTCCACAAAAAACACCCAGAAAACC GCACGCGAACCTACGCCCAGAAACGAAAGCCAAAAAA CCCACAACTTCCTCAAATCGTCACTTCCGTTTTCCCAC GTTACGTCACTTCCCATTTTAAGAAAACTACAATTCCC AACACATACAAGTTACTCCGCCCTTAATTAAATCGGAT CCGATATCTAGATGTATTCGCGAGGTACCGAGCTCGA ATTCTCTGGCCGTCGTTTTACAACGTCGTGACTGGGA AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGG CCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCT GAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTT ACGCATCTGTGCGGTATTTCACACCGCATATGGTGCA CTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGAC AAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGA GGTTTTCACCGTCATCACCGAAACGCGCGA

91	pEMBR-2.0	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag

		ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC

		TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcatggagtcagtcgagaagaa gggattaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgc ctgcgtcattacccagggccacattcttggccaattgcaagccatcaacaaagc ccgccaagagtttctgctacgaaagggacgggggtttacttggacccccagtc cggcgaggagctcaacccaatccccccgccgccgcagccctatcagcagca gccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgcc gccgccacccacggacgaggaggaatactgggacagtcaggcagaggag gttttggacgaggaggaggaggacatgatggaagactgggagagcctagac gaggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcg gtcgcattcccctcgccggcgccccagaaatcggcaaccggttccagcatggc tacaacctccgctcctcaggcgccgccggcactgcccgttcgccgacccaacc gtagatgggacaccactggaaccagggccggtaagtccaagcagccgccgc cgttagcccaagagcaacaacagcgccaaggctaccgctcatggcgcgggc acaagaacgccatagttgcttgcttgcaagactgtgggggcaacatctccttcgc ccgccgctttcttctctaccatcacggcgtggccttcccccgtaacatcctgcatta ctaccgtcatctctacagcccatactgcaccggcggcagcggcagcaacagc agcggccacacagaagcaaaggcgaccggatagcaagactctgacaaag cccaagaaatccacagcggcggcagcagcaggaggaggagcgctgcgtct ggcgcccaacgaacccgtatcgacccgcgagcttagaaacaggatttttccca ctctgtatgctatatttcaacagagcaggggccaagaacaagagctgaaaata aaaaacaggtctctgcgatccctcacccgcagctgcctgtatcacaaaagcga agatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatact gcgcgctgactcttaaggactagTTTCGCGCCCTTTCTCAAATTT AAGCGCGAAAACTACGTCATCTCCAGCGGCCACACCC GGCGCCAGCACCTGTTGTCAGCGCCATTATGAGCAAG GAAATTCCCACGCCCTACATGTGGAGTTACCAGCCAC AAATGGGACTTGCGGCTGGAGCTGCCCAAGACTACTC AACCCGAATAAACTACATGAGCGCGGGGCggccgccgtttg tgttatgtttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattcagt agtatagccccaccaccacatagcttatacagatcaccgtaccttaatcaaacT CACAGAACCCTAGTATTCAACCTGCCACCTCCCTCCC AACACACAGAGTACACAGTCCTTTCTCCCCGGCTGGC CTTAAAAAGCATCATATCATGGGTAACAGACATATTCT TAGGTGTTATATTCCACACGGTTTCCTGTCGAGCCAAA CGCTCATCAGTGATATTAATAAACTCCCCGGGCAGCT CACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGCCAC AGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGCGG CGAAGGAGAAGTCCACGCCTACATGGGGGTAGAGTC ATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCAG CAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCGT CCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCG ATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTCC

		TCCGGGCACAGCAGCGCACCCTGATCTCACTTAAATC AGCACAGTAACTGCAGCACAGCACCACAATATTGTTC AAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTCA TGGCGGGGACCACAGAACCCACGTGGCCATCATACC ACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAAA CACGCTGGACATAAACATTACCTCTTTTGGCATGTTGT AATTCACCACCTCCCGGTACCATATAAACCTCTGATTA AACATGGCGCCATCCACCACCATCCTAAACCAGCTGG CCAAAACCTGCCCGCCGGCTATACACTGCAGGGAAC CGGGACTGGAACAATGACAGTGGAGAGCCCAGGACT CGTAACCATGGATCATCATGCTCGTCATGATATCAATG TTGGCACAACACAGGCACACGTGCATACACTTCCTCA GGATTACAAGCTCCTCCCGCGTTAGAACCATATCCCA GGGAACAACCCATTCCTGAATCAGCGTAAATCCCACA CTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCA TTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGATC CTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGA GGTAGACGATCCCTACTGTACGGAGTGCGCCGAGAC AACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAATG GAACGCCGGACGTAGTCATatttcctgaagcaaaaccaggtgcg ggcgtgacaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagt agttgtagtatatccactctctcaaagcatccaggcgccccctggcttcgggttct atgtaaactccttcatgcgccgctgccctgataacatccaccaccgcagaataa gccacacccagccaacctacacattcgttctgcgagtcacacacgggaggag cgggaagagctggaagaaccatgttttttttttattccaaaagattatccaaaacc tcaaaatgaagatctattaagtgaacgcgctcccctccggtggcgtggtcaaac tctacagccaaagaacagataatggcatttgtaagatgttgcacaatggcttcca aaaggcaaacggccctcacgtccaagtggacgtaaaggctaaacccttcagg gtgaatctcctctataaacattccagcaccttcaaccatgcccaaataattctcat ctcgccaccttctcaatatatctctaagcaaatcccgaatattaagtccggccatt gtaaaaatctgctccagagcgccctccaccttcagcctcaagcagcgaatcat gattgcaaaaattcaggttcctcacagacctgtataagattcaaaagcggaaca ttaacaaaaataccgcgatcccgtaggtcccttcgcagggccagctgaacata atcgtgcaggtctgcacggaccagcgcggccacttccccgccaggaaccatg acaaaagaacccacactgattatgacacgcatactcggagctatgctaaccag cgtagccccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgct gctcaaaaaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtc atgctcatgcagataaaggcaggtaagctccggaaccaccacagaaaaaga caccatttttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaata acaaaaaaacatttaaacattagaagcctgtcttacaacaggaaaaacaacc cttataagcataagacggactacggccatgccggcgtgaccgtaaaaaaact ggtcaccgtgattaaaaagcaccaccgacagctcctcggtcatgtccggagtc ataatgtaagactcggtaaacacatcaggttgattcacatcggtcagtgctaaaa agcgaccgaaatagcccgggggaatacatacccgcaggcgtagagacaac attacagcccccataggaggtataacaaaattaataggagagaaaaacacat aaacacctgaaaaaccctcctgcctaggcaaaatagcaccctcccgctccag aacaacatacagcgcttccacagcggcagccataacagtcagccttaccagt

		aaaaaagaaaacctattaaaaaaacaccactcgacacggcaccagctcaat cagtcacagtgtaaaaaagggccaagtgcagagcgagtatatataggactaa aaaatgacgtaacggttaaagtccacaaaaaacacccagaaaaccgcacg cgaacctacgcccagaaacgaaagccaaaaaacccacaacttcctcaaatc gtcacttccgttttcccacgttacgtcacttcccattttaagaaaactacaattccca acacatacaagttactccgcccttaattaaatcggatccgatatctagatgtattc gcgaggtaccgagctcgaattctctggccgtcgttttacaacgtcgtgactggga aaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagc tggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgca gcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggt atttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagtta agccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtc tgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgt gtcagaggttttcaccgtcatcaccgaaacgcgcga
92	pEMBR-2.0A	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc

		gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccctctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA

		AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcatggagtcagtcgagaagaa gggattaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgg ctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaag ttgggggagggtcggcaattgaaccggtgcctagagaaggtggcgcggg gtaaactgggaaagtgatgtcgtgtactggctccgccttttcccgagggtgggg gagaaccgtatataagtgcagtagtcgccgtgaacgttcttttcgcaacgggttt gccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctcttt acgggttatggcccttgcgtgccttgaattacttccacgcccctggctgcagtacg tgattcttgatcccgagcttcgggttggaagtgggtgggagagttcgaggccttgc gcttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggg gccgccgcgtgcgaatctggtggcaccttcgcgcctgtctcgctgctttcgataa gtctctagccatttaaaatttttgatgacctgctgcgacgctttttttctggcaagata gtcttgtaaatgcgggccaagatctgcacactggtatttcggtttttggggccgcg ggcggcgacggggcccgtgcgtcccagcgcacatgttcggcgaggcggggc ctgcgagcgcggccaccgagaatcggacgggggtagtctcaagctggccgg cctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcggca aggctggcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccg gccctgctgcagggagctcaaaatggaggacgcggcgctcgggagagcgg gcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagccgtcg cttcatgtgactccacggagtaccgggcgccgtccaggcacctcgattagttctc gagcttttggagtacgtcgtctttaggttgggggggagggggttttatgcgatggagttt ccccacactgagtgggtggagactgaagttaggccagcttggcacttgatgtaa ttctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagaca gtggttcaaagtttttttcttccatttcaggtgtcgtgaccgccgcagccctatcagc agcagccgcgggcccttgcttcccagggccaccatggcacccaaaaagaag ctgcagctgccgccgccacccacggacgaggaggaatactgggacagtcag gcagaggaggttttggacgaggaggaggaggacatgatggaagactggga gagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaacacc

		gtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccggtt ccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcgc cgacccaaccgtagATGGGACACCACTGGAACCAGGGCCG GTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAGC AACAACAGCGCCAAGGCTACCGCTCATGGCGCGGGC ACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTGG GGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTAC CATCACGGCGTGGCCTTCCCCCGTAACATCCTGCATT ACTACCGTCATCTCTACAGCCCATACTGCACCGGCGG CAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAAA GGCGACCGGATAGCAAGACTCTGACAAAGCCCAAGA AATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGaataaaagatctttattttcattaga tctgtgtgttggttttttgtgtgGCGGCCGCcgtttgtgttatgtttcaacgtgtttat ttttcaattgcagaaaatttcaagtcatttttcattcagtagtatagccccaccacca catagcttatacagatcaccgtaccttaatcaaacTCACAGAACCCTAG TATTCAACCTGCCACCTCCCTCCCAACACACAGAGTA CACAGTCCTTTCTCCCCGGCTGGCCTTAAAAAGCATC ATATCATGGGTAACAGACATATTCTTAGGTGTTATATT CCACACGGTTTCCTGTCGAGCCAAACGCTCATCAGTG ATATTAATAAACTCCCCGGGCAGCTCACTTAAGTTCAT GTCGCTGTCCAGCTGCTGAGCCACAGGCTGCTGTCC AACTTGCGGTTGCTTAACGGGCGGCGAAGGAGAAGT CCACGCCTACATGGGGGTAGAGTCATAATCGTGCATC AGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGAATA AACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAATAC AACATGGCAGTGGTCTCCTCAGCGATGATTCGCACCG CCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACAGC AGCGCACCCTGATCTCACTTAAATCAGCACAGTAACT GCAGCACAGCACCACAATATTGTTCAAAATCCCACAG TGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGACC ACAGAACCCACGTGGCCATCATACCACAAGCGCAGGT AGATTAAGTGGCGACCCCTCATAAACACGCTGGACAT AAACATTACCTCTTTTGGCATGTTGTAATTCACCACCT CCCGGTACCATATAAACCTCTGATTAAACATGGCGCC ATCCACCACCATCCTAAACCAGCTGGCCAAAACCTGC CCGCCGGCTATACACTGCAGGGAACCGGGACTGGAA CAATGACAGTGGAGAGCCCAGGACTCGTAACCATGGA TCATCATGCTCGTCATGATATCAATGTTGGCACAACAC AGGCACACGTGCATACACTTCCTCAGGATTACAAGCT CCTCCCGCGTTAGAACCATATCCCAGGGAACAACCCA

		TTCCTGAATCAGCGTAAATCCCACACTGCAGGGAAGA CCTCGCACGTAACTCACGTTGTGCATTGTCAAAGTGTT ACATTCGGGCAGCAGCGGATGATCCTCCAGTATGGTA GCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGATCCC TACTGTACGGAGTGCGCCGAGACAACCGAGATCGTGT TGGTCGTAGTGTCATGCCAAATGGAACGCCGGACGTA GTCATatttcctgaagcaaaaccaggtgcgggcgtgacaaacagatctgcg tctccggtctcgccgcttagatcgctctgtgtagtagttgtagtatatccactctcta aagcatccaggcgccccctggcttcgggttctatgtaaactccttcatgcgccgct gccctgataacatccaccaccgcagaataagccacacccagccaacctaca cattcgttctgcgagtcacacacgggaggagcgggaagagctggaagaacc atgtttttttttttattccaaaagattatccaaaacctcaaaatgaagatctattaagt gaacgcgctcccctccggtggcgtggtcaaactctacagccaaagaacagat aatggcatttgtaagatgttgcacaatggcttccaaaaggcaaacggccctcac gtccaagtggacgtaaaggctaaacccttcagggtgaatctcctctataaacatt ccagcaccttcaaccatgcccaaataattctcatctcgccaccttctcaatatatct ctaagcaaatcccgaatattaagtccggccattgtaaaaatctgctccagagcg ccctccaccttcagcctcaagcagcgaatcatgattgcaaaaattcaggttcctc acagacctgtataagattcaaaagcggaacattaacaaaaataccgcgatcc cgtaggtcccttcgcagggccagctgaacataatcgtgcaggtctgcacggac cagcgcggccacttccccgccaggaaccatgacaaaagaacccacactgatt atgacacgcatactcggagctatgctaaccagcgtagccccgatgtaagcttgtt gcatgggcggcgatataaaatgcaaggtgctgctcaaaaaatcaggcaaagc ctcgcgcaaaaaagaaagcacatcgtagtcatgctcatgcagataaaggcag gtaagctccggaaccaccacagaaaaagacaccatttttctctcaaacatgtct gcgggtttctgcataaacacaaaataaaataacaaaaaaacatttaaacatta gaagcctgtcttacaacaggaaaaacaacccttataagcataagacggacta cggccatgccggcgtgaccgtaaaaaaactggtcaccgtgattaaaaagcac caccgacagctcctcggtcatgtccggagtcataatgtaagactcggtaaacac atcaggttgattcacatcggtcagtgctaaaaagcgaccgaaatagcccgggg gaatacatacccgcaggcgtagagacaacattacagcccccataggaggtat aacaaaattaataggagagaaaaacacataaacacctgaaaaaccctcctg cctaggcaaaatagcaccctcccgctccagaacaacatacagcgcttccaca gcggcagccataacagtcagccttaccagtaaaaaagaaaacctattaaaaa aacaccactcgacacggcaccagctcaatcagtcacagtgtaaaaaagggc caagtgcagagcgagtatatataggactaaaaaatgacgtaacggttaaagtc cacaaaaaacacccagaaaaccgcacgcgaacctacgcccagaaacgaa agccaaaaaacccacaacttcctcaaatcgtcacttccgttttcccacgttacgtc acttcccattttaagaaaactacaattcccaacacatacaagttactccgccctta attaaatcggatccgatatctagatgtattcgcgaggtaccgagctcgaattctct ggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaat cgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccg caccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcctga tgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatggtgcactctc agtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaac acccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagaca

		agctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccg aaacgcgcga
93	pEMBR-2.1	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc

		aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccctttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACAgattaggttctacgaagaccaatcc cgcccgcctaatgcggagcttaccgcctgcgtcattacccagggccacattcttg gccaattgcaagccatcaacaaagcccgccaagagtttctgctacgaaaggg acgggggtttacttggacccccagtccggcgaggagctcaacccaatccccc cgccgccgcagccctatcagcagcagccgcgggcccttgcttcccaggatgg cacccaaaaagaagctgcagctgccgccgccacccacggacgaggagga atactgggacagtcaggcagaggaggttttggacgaggaggaggaggacat gatggaagactgggagagcctagacgaggaagcttccgaggtcgaagaggt gtcagacgaaacaccgtcaccctcggtcgcattcccctcgccggcgccccag aaatcggcaaccggttccagcatggctacaacctccgctcctcaggcgccgcc ggcactgcccgttcgccgacccaaccgtagatgggacaccactggaaccag ggccggtaagtccaagcagccgccgccgttagcccaagagcaacaacagc gccaaggctaccgctcatggcgcgggcacaagaacgccatagttgcttgcttg caagactgtgggggcaacatctccttcgcccgccgctttcttctctaccatcacgg cgtggccttcccccgtaacatcctgcattactaccgtcatctctacagcccatact gcaccggcggcagcggcagcaacagcagcggccacacagaagcaaagg cgaccggatagcaagactctgacaaagcccaagaaatccacagcggcggc agcagcaggaggaggagcgctgcgtctggcgcccaacgaacccgtatcgac ccgcgagcttagaaacaggatttttcccactctgtatgctatatttcaacagagca ggggccaagaacaagagctgaaaataaaaaacaggtctctgcgatccctca cccgcagctgcctgtatcacaaaagcgaagatcagcttcggcgcacgctgga agacgcggaggctctcttcagtaaatactgcgcgctgactcttaaggactagTT TCGCGCCCTTTCTCAAATTTAAGCGCGAAAACTACGTC ATCTCCAGCGGCCACACCCGGCGCCAGCACCTGTTG TCAGCGCCATTATGAGCAAGGAAATTCCCACGCCCTA CATGTGGAGTTACCAGCCACAAATGGGACTTGCGGCT GGAGCTGCCCAAGACTACTCAACCCGAATAAACTACA TGAGCGCGGGGCggccgccgtttgtgttatgtttcaacgtgtttatttttcaa ttgcagaaaatttcaagtcatttttcattcagtagtatagccccaccaccacatagc ttatacagatcaccgtaccttaatcaaacTCACAGAACCCTAGTATT CAACCTGCCACCTCCCTCCCAACACACAGAGTACACA GTCCTTTCTCCCCGGCTGGCCTTAAAAAGCATCATATC ATGGGTAACAGACATATTCTTAGGTGTTATATTCCACA CGGTTTCCTGTCGAGCCAAACGCTCATCAGTGATATT AATAAACTCCCCGGGCAGCTCACTTAAGTTCATGTCG CTGTCCAGCTGCTGAGCCACAGGCTGCTGTCCAACTT GCGGTTGCTTAACGGGCGGCGAAGGAGAAGTCCACG CCTACATGGGGGTAGAGTCATAATCGTGCATCAGGAT

		AGGGCGGTGGTGCTGCAGCAGCGCGCGAATAAACTG CTGCCGCCGCCGCTCCGTCCTGCAGGAATACAACAT GGCAGTGGTCTCCTCAGCGATGATTCGCACCGCCCG CAGCATAAGGCGCCTTGTCCTCCGGGCACAGCAGCG CACCCTGATCTCACTTAAATCAGCACAGTAACTGCAG CACAGCACCACAATATTGTTCAAAATCCCACAGTGCAA GGCGCTGTATCCAAAGCTCATGGCGGGGACCACAGA ACCCACGTGGCCATCATACCACAAGCGCAGGTAGATT AAGTGGCGACCCCTCATAAACACGCTGGACATAAACA TTACCTCTTTTGGCATGTTGTAATTCACCACCTCCCGG TACCATATAAACCTCTGATTAAACATGGCGCCATCCAC CACCATCCTAAACCAGCTGGCCAAAACCTGCCCGCCG GCTATACACTGCAGGGAACCGGGACTGGAACAATGAC AGTGGAGAGCCCAGGACTCGTAACCATGGATCATCAT GCTCGTCATGATATCAATGTTGGCACAACACAGGCAC ACGTGCATACACTTCCTCAGGATTACAAGCTCCTCCC GCGTTAGAACCATATCCCAGGGAACAACCCATTCCTG AATCAGCGTAAATCCCACACTGCAGGGAAGACCTCGC ACGTAACTCACGTTGTGCATTGTCAAAGTGTTACATTC GGGCAGCAGCGGATGATCCTCCAGTATGGTAGCGCG GGTTTCTGTCTCAAAAGGAGGTAGACGATCCCTACTG TACGGAGTGCGCCGAGACAACCGAGATCGTGTTGGT CGTAGTGTCATGCCAAATGGAACGCCGGACGTAGTCA Tatttcctgaagcaaaaccaggtgcgggcgtgacaaacagatctgcgtctccg gtctcgccgcttagatcgctctgtgtagtagttgtagtatatccactctctcaaagca tccaggcgccccctggcttcgggttctatgtaaactccttcatgcgccgctgccct gataacatccaccaccgcagaataagccacacccagccaacctacacattcg ttctgcgagtcacacacgggaggagcgggaagagctggaagaaccatgttttttt tttttattccaaaagattatccaaaacctcaaaatgaagatctattaagtgaacgc gctcccctccggtggcgtggtcaaactctacagccaaagaacagataatggca tttgtaagatgttgcacaatggcttccaaaaggcaaacggccctcacgtccaag tggacgtaaaggctaaacccttcagggtgaatctcctctataaacattccagcac cttcaaccatgcccaaataattctcatctcgccaccttctcaatatatctctaagca aatcccgaatattaagtccggccattgtaaaaatctgctccagagcgccctcca ccttcagcctcaagcagcgaatcatgattgcaaaaattcaggttcctcacagac ctgtataagattcaaaagcggaacattaacaaaaataccgcgatcccgtaggt cccttcgcagggccagctgaacataatcgtgcaggtctgcacggaccagcgc ggccacttccccgccaggaaccatgacaaaagaacccacactgattatgaca cgcatactcggagctatgctaaccagcgtagccccgatgtaagcttgttgcatgg gcggcgatataaaatgcaaggtgctgctcaaaaaatcaggcaaagcctcgcg caaaaaagaaagcacatcgtagtcatgctcatgcagataaaggcaggtaagc tccggaaccaccacagaaaaagacaccatttttctctcaaacatgtctgcgggtt tctgcataaacacaaaataaaataacaaaaaaacatttaaacattagaagcct gtcttacaacaggaaaaacaacccttataagcataagacggactacggccat gccggcgtgaccgtaaaaaaactggtcaccgtgattaaaaagcaccaccgac agctcctcggtcatgtccggagtcataatgtaagactcggtaaacacatcaggtt gattcacatcggtcagtgctaaaaagcgaccgaaatagcccgggggaataca

		tacccgcaggcgtagagacaacattacagcccccataggaggtataacaaa attaataggagagaaaaacacataaacacctgaaaaaccctcctgcctaggc aaaatagcaccctcccgctccagaacaacatacagcgcttccacagcggcag ccataacagtcagccttaccagtaaaaaagaaaacctattaaaaaaacacca ctcgacacggcaccagctcaatcagtcacagtgtaaaaaagggccaagtgca gagcgagtatatataggactaaaaaatgacgtaacggttaaagtccacaaaa aacacccagaaaaccgcacgcgaacctacgcccagaaacgaaagccaaa aaacccacaacttcctcaaatcgtcacttccgttttcccacgttacgtcacttccca ttttaagaaaactacaattcccaacacatacaagttactccgcccttaattaaatc ggatccgatatctagatgtattcgcgaggtaccgagctcgaattctctggccgtc gttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgc agcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatc gcccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattt tctccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagtacaat ctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctg acgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtga ccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcg cga
94	pEMBR-2.1A	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta

		tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACAgattaggttctacgaagaccaatcc cgcccgcctaatgcggagcttaccggctccggtgcccgtcagtgggcagagc gcacatcgcccacagtccccgagaagttgggggggaggggtcggcaattgaa ccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgta ctggctccgccttttcccgagggtgggggagaaccgtatataagtgcagtagtc gccgtgaacgttcttttcgcaacgggtttgccgccagaacacaggtaagtgccg tgtgtggttcccgcgggcctggcctctttacgggttatggcccttgcgtgccttgaat tacttccacgcccctggctgcagtacgtgattcttgatcccgagcttcgggttgga agtgggtgggagagttcgaggccttgcgcttaaggagccccttcgcctcgtgctt gagttgaggcctggcctgggcgctggggccgccgcgtgcgaatctggtggcac cttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacct gctgcgacgctttttttctggcaagatagtcttgtaaatgcgggccaagatctgca cactggtatttcggtttttggggccgcgggcggcgacggggcccgtgcgtccca gcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcgg acgggggtagtctcaagctggccggcctgctctggtgcctggcctcgcgccgcc gtgtatcgccccgccctgggcggcaaggctggcccggtcggcaccagttgcgt gagcggaaagatggccgcttcccggccctgctgcagggagctcaaaatgga ggacgcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaa aagggcctttccgtcctcagccgtcgcttcatgtgactccacggagtaccgggc gccgtccaggcacctcgattagttctcgagcttttggagtacgtcgtctttaggttg gggggagggttttatgcgatggagtttccccacactgagtgggtggagactga agttaggccagcttggcacttgatgtaattctccttggaatttgccctttttgagtttgg

		atcttggttcattctcaagcctcagacagtggttcaaagttttttcttccatttcaggt gtcgtgaccgccgcagccctatcagcagcagccgcgggcccttgcttcccagg gccaccatggcacccaaaaagaagctgcagctgccgccgccacccacgga cgaggaggaatactgggacagtcaggcagaggaggttttggacgaggagga ggaggacatgatggaagactgggagagcctagacgaggaagcttccgaggt cgaagaggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgccgg cgccccagaaatcggcaaccggttccagcatggctacaacctccgctcctcag gcgccgccggcactgcccgttcgccgacccaaccgtagATGGGACACC ACTGGAACCAGGGCCGGTAAGTCCAAGCAGCCGCCG CCGTTAGCCCAAGAGCAACAACAGCGCCAAGGCTAC CGCTCATGGCGCGGGCACAAGAACGCCATAGTTGCTT GCTTGCAAGACTGTGGGGGCAACATCTCCTTCGCCCG CCGCTTTCTTCTCTACCATCACGGCGTGGCCTTCCCC CGTAACATCCTGCATTACTACCGTCATCTCTACAGCCC ATACTGCACCGGCGGCAGCGGCAGCAACAGCAGCGG CCACACAGAAGCAAAGGCGACCGGATAGCAAGACTCT GACAAAGCCCAAGAAATCCACAGCGGCGGCAGCAGC AGGAGGAGGAGCGCTGCGTCTGGCGCCCAACGAACC CGTATCGACCCGCGAGCTTAGAAACAGGATTTTTCCC ACTCTGTATGCTATATTTCAACAGAGCAGGGGCCAAG AACAAGAGCTGAAAATAAAAAACAGGTCTCTGCGATC CCTCACCCGCAGCTGCCTGTATCACAAAAGCGAAGAT CAGCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTC TTCAGTAAATACTGCGCGCTGACTCTTAAGGACTAGaa taaaagatctttattttcattagatctgtgtgttggttttttgtgtgGCGGCCGCcg tttgtgttatgtttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattc agtagtatagccccaccaccacatagcttatacagatcaccgtaccttaatcaa acTCACAGAACCCTAGTATTCAACCTGCCACCTCCCTC CCAACACACAGAGTACACAGTCCTTTCTCCCCGGCTG GCCTTAAAAAGCATCATATCATGGGTAACAGACATATT CTTAGGTGTTATATTCCACACGGTTTCCTGTCGAGCCA AACGCTCATCAGTGATATTAATAAACTCCCCGGGCAG CTCACTTAAGTTCATGTCGCTGTCCAGCTGCTGAGCC ACAGGCTGCTGTCCAACTTGCGGTTGCTTAACGGGCG GCGAAGGAGAAGTCCACGCCTACATGGGGGTAGAGT CATAATCGTGCATCAGGATAGGGCGGTGGTGCTGCA GCAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCG TCCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGC GATGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTC CTCCGGGCACAGCAGCGCACCCTGATCTCACTTAAAT CAGCACAGTAACTGCAGCACAGCACCACAATATTGTT CAAAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTC ATGGCGGGGACCACAGAACCCACGTGGCCATCATAC CACAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAA ACACGCTGGACATAAACATTACCTCTTTTGGCATGTTG TAATTCACCACCTCCCGGTACCATATAAACCTCTGATT AAACATGGCGCCATCCACCACCATCCTAAACCAGCTG

		GCCAAAACCTGCCCGCCGGCTATACACTGCAGGGAA CCGGGACTGGAACAATGACAGTGGAGAGCCCAGGAC TCGTAACCATGGATCATCATGCTCGTCATGATATCAAT GTTGGCACAACACAGGCACACGTGCATACACTTCCTC AGGATTACAAGCTCCTCCCGCGTTAGAACCATATCCC AGGGAACAACCCATTCCTGAATCAGCGTAAATCCCAC ACTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGC ATTGTCAAAGTGTTACATTCGGGCAGCAGCGGATGAT CCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGG AGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGA CAACCGAGATCGTGTTGGTCGTAGTGTCATGCCAAAT GGAACGCCGGACGTAGTCATatttcctgaagcaaaaccaggtgc gggcgtgacaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgta gtagttgtagtatatccactctctcaaagcatccaggcgccccctggcttcgggtt ctatgtaaactccttcatgcgccgctgccctgataacatccaccaccgcagaata agccacacccagccaacctacacattcgttctgcgagtcacacacgggagga gcgggaagagctggaagaaccatgttttttttttattccaaaagattatccaaaa cctcaaaatgaagatctattaagtgaacgcgctcccctccggtggcgtggtcaa actctacagccaaagaacagataatggcatttgtaagatgttgcacaatggcttc caaaaggcaaacggccctcacgtccaagtggacgtaaaggctaaacccttca gggtgaatctcctctataaacattccagcaccttcaaccatgcccaaataattctc atctcgccaccttctcaatatatctctaagcaaatcccgaatattaagtccggcca ttgtaaaaatctgctccagagcgccctccaccttcagcctcaagcagcgaatca tgattgcaaaaattcaggttcctcacagacctgtataagattcaaaagcggaac attaacaaaaataccgcgatcccgtaggtcccttcgcagggccagctgaacat aatcgtgcaggtctgcacggaccagcgcggccacttccccgccaggaaccat gacaaaagaacccacactgattatgacacgcatactcggagctatgctaacca gcgtagccccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgc tgctcaaaaaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagt catgctcatgcagataaaggcaggtaagctccggaaccaccacagaaaaag acaccatttttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaat aacaaaaaaacatttaaacattagaagcctgtcttacaacaggaaaaacaac ccttataagcataagacggactacggccatgccggcgtgaccgtaaaaaaac tggtcaccgtgattaaaaagcaccaccgacagctcctcggtcatgtccggagtc ataatgtaagactcggtaaacacatcaggttgattcacatcggtcagtgctaaaa agcgaccgaaatagcccgggggaatacatacccgcaggcgtagagacaac attacagcccccataggaggtataacaaaattaataggagagaaaaacacat aaacacctgaaaaaccctcctgcctaggcaaaatagcaccctcccgctccag aacaacatacagcgcttccacagcggcagccataacagtcagccttaccagt aaaaaagaaaacctattaaaaaaacaccactcgacacggcaccagctcaat cagtcacagtgtaaaaaagggccaagtgcagagcgagtatatataggactaa aaaatgacgtaacggttaaagtccacaaaaaacacccagaaaaccgcacg cgaacctacgcccagaaacgaaagccaaaaaacccacaacttcctcaaatc gtcacttccgttttcccacgttacgtcacttcccattttaagaaaactacaattccca acacatacaagttactccgcccttaattaaatcggatccgatatctagatgtattc gcgaggtaccgagctcgaattctctggccgtcgttttacaacgtcgtgactggga aaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagc

		tggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgca gcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggt atttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagtta agccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtc tgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgt gtcagaggttttcaccgtcatcaccgaaacgcgcga
95	pMinKanRLacZ backbone	ACATTAATTAAggatccGAATTCAGTACTCATATGTTGAC AATTAATCATCGGCATAGTATATCGGCATAGTATAATA CGACTCACTATAGGAGGGCCACCAAGGTGAGGAACTA AACCATGAGCCATATTCAACGGGAAACGTCGAGGCCG CGATTAAATTCCAACATGGATGCTGATTTATATGGGTA TAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCG ACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAG AGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAA TGATGTTACAGATGAGATGGTCAGACTAAACTGGCTG ACGGAATTTATGCCACTTCCGACCATCAAGCATTTTAT CCGTACTCCTGATGATGCATGGTTACTCACCACTGCG ATCCCCGGAAAAACAGCGTTCCAGGTATTAGAAGAAT ATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCA GTGTTCCTGCGCCGGTTGCACTCGATTCCTGTTTGTA ATTGTCCTTTTAACAGCGATCGCGTATTTCGCCTCGCT CAGGCGCAATCACGAATGAATAACGGTTTGGTTGATG CGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGT TGAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCAT TCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCA CTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGG TTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGA TACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTG AGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAAT ATGGTATTGATAATCCTGATATGAATAAATTGCAGTTT CATTTGATGCTCGATGAGTTTTTCTAATAAAAACCCGC TTCGGCGGGTTTTTTTATGCATGTGAGCAAAAGGCCA GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCC TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGC CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC

		CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTT GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAC GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG TCATGtcgacGTTAACAAGCTTTACGTACAGtaactatgcggc atcagagcagattgtactgagagtgcaccatatgcggtgtgaaataccgcaca gatgcgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgc aactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcg aaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttccca gtcacgacgttgtaaaacgacggccagagaattcgagctcggtacctcgcga atacatctagatatcggatcccgggcccgtcgactgcagaggcctgcatgcaa gcttggtgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaatt ccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaat gagtgagctaactcacattaattg
100	pEMBR-1.37+L4	TGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTA CATTGCACAAGATAAAAATATATCATCATGAACAATAA AACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTT ATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGAT TAAATTCCAACATGGATGCTGATTTATATGGGTATAAA TGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA ATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGT TGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG TACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATC CTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGT GTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA ACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCT CACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTT GATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGT TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT TGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACATTATTCAGATTGGGCTTGATTTAAAAC TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA

		AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA ATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCC TGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCG CCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTT TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT GTGAGCGGATAACAATTTCACACAGGAAACAGCTATG ACCATGATTACACCAAGCTTGCATGCAGGCCTCTGCA GTCGACCAGAAGCACCATGTCCTTGGGTCCGGCCTG CTGAATGCGCAGGCGGTCGGCCATGCCCCAGGCTTC GTTTTGACATCGGCGCAGGTCTTTGTAGTAGTCTTGC ATGAGCCTTTCTACCGGCACTTCTTCTTCTCCTTCCTC TTGTCCTGCATCTCTTGCATCTATCGCTGCGGCGGCG GCGGAGTTTGGCCGTAGGTGGCGCCCTCTTCCTCCC ATGCGTGTGACCCCGAAGCCCCTCATCGGCTGAAGC AGGGCTAGGTCGGCGACAACGCGCTCGGCTAATATG GCCTGCTGCACCTGCGTGAGGGTAGACTGGAAGTCA TCCATGTCCACAAAGCGGTGGTATGCGCCCGTGTTGA TGGTGTAAGTGCAGTTGGCCATAACGGACCAGTTAAC GGTCTGGTGACCCGGCTGCGAGAGCTCGGTGTACCT GAGACGCGAGTAAGCCCTCGAGTCAAATACGTAGTCG TTGCAAGTCCGCACCAGGTACTGGTATCCCACCAAAA AGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGT AGGGTGGCCGGGGCTCCGGGGGCGAGATCTTCCAAC ATAAGGCGATGATATCCGTAGATGTACCTGGACATCC AGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGA AAGTCGCGGACGCGGTTCCAGATGTTGCGCAGCGGC

		AAAAAGTGCTCCATGGTCGGGACGCTCTGGCCGGTC AGGCGCGCGCAATCGTTGACGCTCTAGCGTGCAAAA GGAGAGCCTGTAAGCGGGCACTCTTCCGTGGTCTGG TGGATAAATTCGCAAGGGTATCATGGCGGACGACCGG GGTTCGAGCCCCGTATCCGGCCGTCCGCCGTGATCC ATGCGGTTACCGCCCGCGTGTCGAACCCAGGTGTGC GACGTCAGACAACGGGGGAGTGCTCCTTTTGGCTTCC TTCCAGGCGCGGCGGCTGCTGCGCTAGCTTTTTTGGC CACTGGCCGCGCGCAGCGTAAGCGGTTAGGCTGGAA AGCGAAAGCATTAAGTGGCTCGCTCCCTGTAGCCGGA GGGTTATTTTCCAAGGGTTGAGTCGCGGGACCCCCG GTTCGAGTCTCGGACCGAGACTGGGGGCGTACACTG GATGGCCTTTGCCTGGAACCCGCACTCAAAAACATGC TACCTCTTTGAGCCCTTTGGCTTTTCTGACCAGCGACT CAAGCAGGTTTACCAGTTTGAGTACGAGTCACTCCTG CGCCGTAGCGCCATTGCTTCTTCCCCCGACCGCTGTA TAACGCTGGAAAAGTCCACCCAAAGCGTACAGGGGC CCAACTCGGCCGCCTGTGGACTATTCTGCTGCATGTT TCTCCACGCCTTTGCCAACTGGCCCCAAACTCCCATG GATCACAACCCCACCATGAACCTTATTACCGGGGTAC CCAACTCCATGCTCAACAGTCCCCAGGTACAGCCCAC CCTGCGTCGCAACCAGGAACAGCTCTACAGCTTCCTG GAGCGCCACTCGCCCTACTTCCGCAGCCACAGTGCG CAGATTAGGAGCGCCACTTCTTTTTGTCACTTGAAAAA CATGTAAAAATAATGTACTAGAGACACTTTCAATAAAG GCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT

		AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcCACAAAAGCGAAGAT CAGCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTC TTCAGTAAATACTGCGCGCTGACTCTTAAGGACTAGTT TCGCGCCCTTTCTCAAATTTAAGCGCGAAAACTACGTC ATCTCCAGCGGCCACACCCGGCGCCAGCACCTGTTG TCAGCGCCATTGGCGCGCCGGCCGGCCGAAtatcTTCA TTTAAATGTTTAAACATCGATGCGGCCGCgattaggttctacg aagaccaatcccgcccgcctaatgcggagcttaccgcctgcgtcattacccag ggccacattcttggccaattgcaagccatcaacaaagcccgccaagagtttctg ctacgaaagggacgggggtttacttggacccccagtccggcgaggagctca acccaatccccccgccgccgcagccctatcagcagcagccgcgggcccttgc ttcccaggatggcacccaaaaagaagctgcagctgccgccgccacccacgg acgaggaggaatactgggacagtcaggcagaggaggttttggacgaggagg aggaggacatgatggaagactgggagagcctagacgaggaagcttccgag gtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgcc ggcgccccagaaatcggcaaccggttccagcatggctacaacctccgctcctc aggcgccgccggcactgcccgttcgccgacccaaccgtagATGGGACA CCACTGGAACCAGGGCCGGTAAGTCCAAGCAGCCGC CGCCGTTAGCCCAAGAGCAACAACAGCGCCAAGGCT ACCGCTCATGGCGCGGGCACAAGAACGCCATAGTTG CTTGCTTGCAAGACTGTGGGGGCAACATCTCCTTCGC CCGCCGCTTTCTTCTCTACCATCACGGCGTGGCCTTC CCCCGTAACATCCTGCATTACTACCGTCATCTCTACAG CCCATACTGCACCGGCGGCAGCGGCAGCAACAGCAG

		CGGCCACACAGAAGCAAAGGCGACCGGATAGCAAGA CTCTGACAAAGCCCAAGAAATCCACAGCGGCGGCAG CAGCAGGAGGAGGAGCGCTGCGTCTGGCGCCCAACG AACCCGTATCGACCCGCGAGCTTAGAAACAGGATTTT TCCCACTCTGTATGCTATATTTCAACAGAGCAGGGGC CAAGAACAAGAGCTGAAAATAAAAAACAGGTCTCTGC GATCCCTCACCCGCAGCTGCCTGTATCACAAAAGCGA AGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGGC TCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGACT AGaataaaagatctttattttcattagatctgtgtgttggttttttgtgtgGCGGCC GCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAA AGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTT TCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAA TGTATCTTAGCTTAACGGGCGGCGAAGGAGAAGTCCA CGCCTACATGGGGGTAGAGTCATAATCGTGCATCAGG ATAGGGCGGTGGTGCTGCAGCAGCGCGCGAATAAAC TGCTGCCGCCGCCGCTCCGTCCTGCAGGAATACAAC ATGGCAGTGGTCTCCTCAGCGATGATTCGCACCGCCC GCAGCATAAGGCGCCTTGTCCTCCGGGCACAGCAGC GCACCCTGATCTCACTTAAATCAGCACAGTAACTGCA GCACAGCACCACAATATTGTTCAAAATCCCACAGTGC AAGGCGCTGTATCCAAAGCTCATGGCGGGGACCACA GAACCCACGTGGCCATCATACCACAAGCGCAGGTAGA TTAAGTGGCGACCCCTCATAAACACGCTGGACATAAA CATTACCTCTTTTGGCATGTTGTAATTCACCACCTCCC GGTACCATATAAACCTCTGATTAAACATGGCGCCATCC ACCACCATCCTAAACCAGCTGGCCAAAACCTGCCCGC CGGCTATACACTGCAGGGAACCGGGACTGGAACAAT GACAGTGGAGAGCCCAGGACTCGTAACCATGGATCAT CATGCTCGTCATGATATCAATGTTGGCACAACACAGG CACACGTGCATACACTTCCTCAGGATTACAAGCTCCT CCCGCGTTAGAACCATATCCCAGGGAACAACCCATTC CTGAATCAGCGTAAATCCCACACTGCAGGGAAGACCT CGCACGTAACTCACGTTGTGCATTGTCAAAGTGTTACA TTCGGGCAGCAGCGGATGATCCTCCAGTATGGTAGC GCGGGTTTCTGTCTCAAAAGGAGGTAGACGATCCCTA CTGTACGGAGTGCGCCGAGACAACCGAGATCGTGTT GGTCGTAGTGTCATGCCAAATGGAACGCCGGACGTA GTCATATTTCCTGAAGCAAAACCAGGTGCGGGCGTGA CAAACAGATCTGCGTCTCCGGTCTCGCCGCTTAGATC GCTCTGTGTAGTAGTTGTAGTATATCCACTCTCTCAAA GCATCCAGGCGCCCCCTGGCTTCGGGTTCTATGTAAA CTCCTTCATGCGCCGCTGCCCTGATAACATCCACCAC CGCAGAATAAGCCACACCCAGCCAACCTACACATTCG TTCTGCGAGTCACACACGGGAGGAGCGGGAAGAGCT GGAAGAACCATGGTGGCATTTGCAAAAGCCTAGGCCT CCAAAAAAGCCTCCTCACTACTTCTGGAATAGCTCAGA

		GGCCGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAA ATTAGTCAGCCATGGGGCGGAGAATGGGCGGAACTG GGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGG CGGGACTATGGTTGCTGACTAATTGAGATGCATGCTT TGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTC CACACCTGGTTGCTGACTAATTGAGATGCATGCTTTG CATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCA CACCCTAACTGACACACACGTTACGTCACTTCCCATTT TAAGAAAACTACAATTCCCAACACATACAAGTTACTCC GCCCTTAATtaaatcggatccgatatctagatgtattcgcgaggtaccgag ctcgAATTCTCTGGCCGTCGTTTTACAACGTCGTGACTG GGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCA GCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAG AGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA GCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCT CCTTACGCATCTGTGCGGTATTTCACACCGCATATGG TGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTT AAGCCAGCCCCGACACCCGCCAACACCCGCTGACGC GCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTAC AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTC AGAGGTTTTCACCGTCATCACCGAAACGCGCGA
101	pEMBR-1.37+L4-minKanR	AATTCAGTACTCATATGTTGACAATTAATCATCGGCAT AGTATATCGGCATAGTATAATACGACTCACTATAGGAG GGCCACCAAGGTGAGGAACTAAACCATGAGCCATATT CAACGGGAAACGTCGAGGCCGCGATTAAATTCCAACA TGGATGCTGATTTATATGGGTATAAATGGGCTCGCGA TAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTG TATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAAC ATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGA GATGGTCAGACTAAACTGGCTGACGGAATTTATGCCA CTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGA TGCATGGTTACTCACCACTGCGATCCCCGGAAAAACA GCGTTCCAGGTATTAGAAGAATATCCTGATTCAGGTG AAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCG GTTGCACTCGATTCCTGTTTGTAATTGTCCTTTTAACA GCGATCGCGTATTTCGCCTCGCTCAGGCGCAATCACG AATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATG ACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAA AGAAATGCATAAACTTTTGCCATTCTCACCGGATTCAG TCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTT TTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGA CGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCA TCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTA CAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATG AGTTTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTT ATGCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA

		CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAA AACTCACGTTAAGGGATTTTGGTCATGtcgacGTTAACA AGCTTTACCCTCTGCAGTCGACCAGAAGCACCATGTC CTTGGGTCCGGCCTGCTGAATGCGCAGGCGGTCGGC CATGCCCCAGGCTTCGTTTTGACATCGGCGCAGGTCT TTGTAGTAGTCTTGCATGAGCCTTTCTACCGGCACTTC TTCTTCTCCTTCCTCTTGTCCTGCATCTCTTGCATCTAT CGCTGCGGCGGCGGCGGAGTTTGGCCGTAGGTGGC GCCCTCTTCCTCCCATGCGTGTGACCCCGAAGCCCCT CATCGGCTGAAGCAGGGCTAGGTCGGCGACAACGCG CTCGGCTAATATGGCCTGCTGCACCTGCGTGAGGGTA GACTGGAAGTCATCCATGTCCACAAAGCGGTGGTATG CGCCCGTGTTGATGGTGTAAGTGCAGTTGGCCATAAC GGACCAGTTAACGGTCTGGTGACCCGGCTGCGAGAG CTCGGTGTACCTGAGACGCGAGTAAGCCCTCGAGTCA AATACGTAGTCGTTGCAAGTCCGCACCAGGTACTGGT ATCCCACCAAAAAGTGCGGCGGCGGCTGGCGGTAGA GGGGCCAGCGTAGGGTGGCCGGGGCTCCGGGGGCG AGATCTTCCAACATAAGGCGATGATATCCGTAGATGTA CCTGGACATCCAGGTGATGCCGGCGGCGGTGGTGGA GGCGCGCGGAAAGTCGCGGACGCGGTTCCAGATGTT GCGCAGCGGCAAAAAGTGCTCCATGGTCGGGACGCT CTGGCCGGTCAGGCGCGCGCAATCGTTGACGCTCTA GCGTGCAAAAGGAGAGCCTGTAAGCGGGCACTCTTC CGTGGTCTGGTGGATAAATTCGCAAGGGTATCATGGC GGACGACCGGGGTTCGAGCCCCGTATCCGGCCGTCC GCCGTGATCCATGCGGTTACCGCCCGCGTGTCGAAC CCAGGTGTGCGACGTCAGACAACGGGGGAGTGCTCC TTTTGGCTTCCTTCCAGGCGCGGCGGCTGCTGCGCTA GCTTTTTTGGCCACTGGCCGCGCGCAGCGTAAGCGG

		TTAGGCTGGAAAGCGAAAGCATTAAGTGGCTCGCTCC CTGTAGCCGGAGGGTTATTTTCCAAGGGTTGAGTCGC GGGACCCCCGGTTCGAGTCTCGGACCGAGACTGGGG GCGTACACTGGATGGCCTTTGCCTGGAACCCGCACTC AAAAACATGCTACCTCTTTGAGCCCTTTGGCTTTTCTG ACCAGCGACTCAAGCAGGTTTACCAGTTTGAGTACGA GTCACTCCTGCGCCGTAGCGCCATTGCTTCTTCCCCC GACCGCTGTATAACGCTGGAAAAGTCCACCCAAAGCG TACAGGGGCCCAACTCGGCCGCCTGTGGACTATTCTG CTGCATGTTTCTCCACGCCTTTGCCAACTGGCCCCAA ACTCCCATGGATCACAACCCCACCATGAACCTTATTAC CGGGGTACCCAACTCCATGCTCAACAGTCCCCAGGTA CAGCCCACCCTGCGTCGCAACCAGGAACAGCTCTACA GCTTCCTGGAGCGCCACTCGCCCTACTTCCGCAGCCA CAGTGCGCAGATTAGGAGCGCCACTTCTTTTTGTCAC TTGAAAAACATGTAAAAATAATGTACTAGAGACACTTT CAATAAAGGCAAATGCTTTTATTTGTACACTCTCGGGT GATTATTTACCCCCACCCTTGCCGTCTGCGCCGTTTAA AAATCAAAGGGGTTCTGCCGCGCATCGCTATGCGCCA CTGGCAGGGACACGTTGCGATACTGGTGTTTAGTGCT CCACTTAAACTCAGGCACAACCATCCGCGGCAGCTCG GTGAAGTTTTCACTCCACAGGCTGCGCACCATCACCA ACGCGTTTAGCAGGTCGGGCGCCGATATCTTGAAGTC GCAGTTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCG ATACACAGGGTTGCAGCACTGGAACACTATCAGCGCC GGGTGGTGCACGCTGGCCAGCACGCTCTTGTCGGAG ATCAGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGG GCGAACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAA AGGGCGCGTGCCCAGGCTTTGAGTTGCACTCGCACC GTAGTGGCATCAAAAGGTGACCGTGCCCGGTCTGGG CGTTAGGATACAGCGCCTGCATAAAAGCCTTGATCTG CTTAAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAG AACATGCCGCAAGACTTGCCGGAAAACTGATTGGCCG GACAGGCCGCGTCGTGCACGCAGCACCTTGCGTCGG TGTTGGAGATCTGCACCACATTTCGGCCCCACCGGTT CTTCACGATCTTGGCCTTGCTAGACTGCTCCTTCAGC GCGCGCTGCCCGTTTTCGCTCGTCACATCCATTTCAA TCACGTGCTCCTTATTTATCATAATGCTTCCGTGTAGA CACTTAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCA GCCACAACGCGCAGCCCGTGGGCTCGTGATGCTTGT AGGTCACCTCTGCAAACGACTGCAGGTACGCCTGCAG GAATCGCCCCATCATCGTCACAAAGGTCTTGTTGCTG GTGAAGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCA GCCAGGTCTTGCATACGGCCGCCAGAGCTTCCACTTG GTCAGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTAT CCACGTGGTACTTGTCCATCAGCGCGCGCGCAGCCT CCATGCCCTTCTCCCACGCAGACACGATCGGCACACT

		CAGCGGGTTCATCACCGTAATTTCACTTTCCGCTTCG CTGGGCTCTTCCTCTTCCTCTTGCGTCCGCATACCAC GCGCCACTGGGTCGTCTTCATTCAGCCGCCGCACTGT GCGCTTACCTCCTTTGCCATGCTTGATTAGCACCGGT GGGTTGCTGAAACCCACCATTTGTAGCGCCACATCTT CTCTTTCTTCCTCGCTGTCCACGATTACCTCTGGTGAT GGCGGGCGCTCGGGCTTGGGAGAAGGGCGCTTCTTT TTCTTCTTGGGCGCAATGGCCAAATCCGCCGCCGAGG TCGATGGCCGCGGGCTGGGTGTGCGCGGCACCAGC GCGTCTTGTGATGAGTCTTCCTCGTCCTCGGACTCGA TACGCCGCCTCATCCGCTTTTTTGGGGGCGCCCGGG GAGGCGGCGGCGACGGGGACGGGGACGACACGTCC TCCATGGTTGGGGGACGTCGCGCCGCACCGCGTCCG CGCTCGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGA CTGGCCATttccttctcctataggcagaaaaagatcCACAAAAGCG AAGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGG CTCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGAC TAGTTTCGCGCCCTTTCTCAAATTTAAGCGCGAAAACT ACGTCATCTCCAGCGGCCACACCCGGCGCCAGCACC TGTTGTCAGCGCCATTGGCGCGCCGGCCGGCCGAAta tcTTCATTTAAATGTTTAAACATCGATGCGGCCGCgattag gttctacgaagaccaatcccgcccgcctaatgcggagcttaccgcctgcgtcatt acccagggccacattcttggccaattgcaagccatcaacaaagcccgccaag agtttctgctacgaaagggacgggggtttacttggacccccagtccggcgag gagctcaacccaatccccccgccgccgcagccctatcagcagcagccgcgg gcccttgcttcccaggatggcacccaaaaagaagctgcagctgccgccgcca cccacggacgaggaggaatactgggacagtcaggcagaggaggttttggac gaggaggaggaggacatgatggaagactgggagagcctagacgaggaag cttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgcattc ccctcgccggcgccccagaaatcggcaaccggttccagcatggctacaacct ccgctcctcaggcgccgccggcactgcccgttcgccgacccaaccgtagAT GGGACACCACTGGAACCAGGGCCGGTAAGTCCAAGC AGCCGCCGCCGTTAGCCCAAGAGCAACAACAGCGCC AAGGCTACCGCTCATGGCGCGGGCACAAGAACGCCA TAGTTGCTTGCTTGCAAGACTGTGGGGGCAACATCTC CTTCGCCCGCCGCTTTCTTCTCTACCATCACGGCGTG GCCTTCCCCCGTAACATCCTGCATTACTACCGTCATCT CTACAGCCCATACTGCACCGGCGGCAGCGGCAGCAA CAGCAGCGGCCACACAGAAGCAAAGGCGACCGGATA GCAAGACTCTGACAAAGCCCAAGAAATCCACAGCGGC GGCAGCAGCAGGAGGAGGAGCGCTGCGTCTGGCGC CCAACGAACCCGTATCGACCCGCGAGCTTAGAAACAG GATTTTTCCCACTCTGTATGCTATATTTCAACAGAGCA GGGGCCAAGAACAAGAGCTGAAAATAAAAAACAGGTC TCTGCGATCCCTCACCCGCAGCTGCCTGTATCACAAA AGCGAAGATCAGCTTCGGCGCACGCTGGAAGACGCG GAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCTTA

		AGGACTAGaataaaagatctttattttcattagatctgtgtgttggttttttgtgtg GCGGCCGCAACTTGTTTATTGCAGCTTATAATGGTTAC AAATAAAGCAATAGCATCACAAATTTCACAAATAAAGC ATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACT CATCAATGTATCTTAGCTTAACGGGCGGCGAAGGAGA AGTCCACGCCTACATGGGGGTAGAGTCATAATCGTGC ATCAGGATAGGGCGGTGGTGCTGCAGCAGCGCGCGA ATAAACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAAT ACAACATGGCAGTGGTCTCCTCAGCGATGATTCGCAC CGCCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACA GCAGCGCACCCTGATCTCACTTAAATCAGCACAGTAA CTGCAGCACAGCACCACAATATTGTTCAAAATCCCACA GTGCAAGGCGCTGTATCCAAAGCTCATGGCGGGGAC CACAGAACCCACGTGGCCATCATACCACAAGCGCAG GTAGATTAAGTGGCGACCCCTCATAAACACGCTGGAC ATAAACATTACCTCTTTTGGCATGTTGTAATTCACCAC CTCCCGGTACCATATAAACCTCTGATTAAACATGGCG CCATCCACCACCATCCTAAACCAGCTGGCCAAAACCT GCCCGCCGGCTATACACTGCAGGGAACCGGGACTGG AACAATGACAGTGGAGAGCCCAGGACTCGTAACCATG GATCATCATGCTCGTCATGATATCAATGTTGGCACAAC ACAGGCACACGTGCATACACTTCCTCAGGATTACAAG CTCCTCCCGCGTTAGAACCATATCCCAGGGAACAACC CATTCCTGAATCAGCGTAAATCCCACACTGCAGGGAA GACCTCGCACGTAACTCACGTTGTGCATTGTCAAAGT GTTACATTCGGGCAGCAGCGGATGATCCTCCAGTATG GTAGCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGAT CCCTACTGTACGGAGTGCGCCGAGACAACCGAGATC GTGTTGGTCGTAGTGTCATGCCAAATGGAACGCCGGA CGTAGTCATATTTCCTGAAGCAAAACCAGGTGCGGGC GTGACAAACAGATCTGCGTCTCCGGTCTCGCCGCTTA GATCGCTCTGTGTAGTAGTTGTAGTATATCCACTCTCT CAAAGCATCCAGGCGCCCCCTGGCTTCGGGTTCTATG TAAACTCCTTCATGCGCCGCTGCCCTGATAACATCCA CCACCGCAGAATAAGCCACACCCAGCCAACCTACACA TTCGTTCTGCGAGTCACACACGGGAGGAGCGGGAAG AGCTGGAAGAACCATGGTGGCATTTGCAAAAGCCTAG GCCTCCAAAAAAGCCTCCTCACTACTTCTGGAATAGCT CAGAGGCCGAGGCGGCCTCGGCCTCTGCATAAATAA AAAAAATTAGTCAGCCATGGGGCGGAGAATGGGCGG AACTGGGCGGAGTTAGGGGCGGGATGGGCGGAGTTA GGGGCGGGACTATGGTTGCTGACTAATTGAGATGCAT GCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGAC TTTCCACACCTGGTTGCTGACTAATTGAGATGCATGCT TTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTT CCACACCCTAACTGACACACACGTTACGTCACTTCCC ATTTTAAGAAAACTACAATTCCCAACACATACAAGTTA

		CTCCGCCCTTAATtaaatcggatccgatatctagatgtattcgcgaggtaccgagctcg
102	pEMBR-1.37B2+L4	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaatatatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatggatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccggttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgccattctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttcatttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcagattgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacaccaagcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggcggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgcagcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgcagtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcgaaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc

		aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttGTACACTCTCGGGTGATTATTT ACCCCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAA AGGGGTTCTGCCGCGCATCGCTATGCGCCACTGGCA GGGACACGTTGCGATACTGGTGTTTAGTGCTCCACTT AAACTCAGGCACAACCATCCGCGGCAGCTCGGTGAA GTTTTCACTCCACAGGCTGCGCACCATCACCAACGCG TTTAGCAGGTCGGGCGCCGATATCTTGAAGTCGCAGT TGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGATACA CAGGGTTGCAGCACTGGAACACTATCAGCGCCGGGT GGTGCACGCTGGCCAGCACGCTCTTGTCGGAGATCA GATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGGCGA ACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAAGGG CGCGTGCCCAGGCTTTGAGTTGCACTCGCACCGTAGT GGCATCAAAAGGTGACCGTGCCCGGTCTGGGCGTTA GGATACAGCGCCTGCATAAAAGCCTTGATCTGCTTAA AAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGAACAT GCCGCAAGACTTGCCGGAAAACTGATTGGCCGGACA GGCCGCGTCGTGCACGCAGCACCTTGCGTCGGTGTT GGAGATCTGCACCACATTTCGGCCCCACCGGTTCTTC ACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCGCGC GCTGCCCGTTTTCGCTCGTCACATCCATTTCAATCACG TGCTCCTTATTTATCATAATGCTTCCGTGTAGACACTT AAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAGCCA CAACGCGCAGCCCGTGGGCTCGTGATGCTTGTAGGT CACCTCTGCAAACGACTGCAGGTACGCCTGCAGGAAT CGCCCCATCATCGTCACAAAGGTCTTGTTGCTGGTGA AGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCAGCC AGGTCTTGCATACGGCCGCCAGAGCTTCCACTTGGTC AGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCA CGTGGTACTTGTCCATCAGCGCGCGCGCAGCCTCCAT GCCCTTCTCCCACGCAGACACGATCGGCACACTCAGC GGGTTCATCACCGTAATTTCACTTTCCGCTTCGCTGG GCTCTTCCTCTTCCTCTTGCGTCCGCATACCACGCGC CACTGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGC TTACCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTT GCTGAAACCCACCATTTGTAGCGCCACATCTTCTCTTT CTTCCTCGCTGTCCACGATTACCTCTGGTGATGGCGG GCGCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTC

		TTGGGCGCAATGGCCAAATCCGCCGCCGAGGTCGAT GGCCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTC TTGTGATGAGTCTTCCTCGTCCTCGGACTCGATACGC CGCCTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGC GGCGGCGACGGGGACGGGGACGACACGTCCTCCAT GGTTGGGGGACGTCGCGCCGCACCGCGTCCGCGCT CGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGACTGG CCATttccttctcctataggcagaaaaagatcCACAAAAGCGAAGAT CAGCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTC TTCAGTAAATACTGCGCGCTGACTCTTAAGGACTAGTT TCGCGCCCTTTCTCAAATTTAAGCGCGAAAACTACGTC ATCTCCAGCGGCCACACCCGGCGCCAGCACCTGTTG TCAGCGCCATTGGCGCGCCGGCCGGCCGAAtatcTTCA TTTAAATGTTTAAACATCGATGCGGCCGCgattaggttctacg aagaccaatcccgcccgcctaatgcggagcttaccgcctgcgtcattacccag ggccacattcttggccaattgcaagccatcaacaaagcccgccaagagtttctg ctacgaaagggacgggggtttacttggacccccagtccggcgaggagctca acccaatccccccgccgccgcagccctatcagcagcagccgcgggcccttgc ttcccaggatggcacccaaaaagaagctgcagctgccgccgccacccacgg acgaggaggaatactgggacagtcaggcagaggaggttttggacgaggagg aggaggacatgatggaagactgggagagcctagacgaggaagcttccgag gtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgcc ggcgccccagaaatcggcaaccggttccagcatggctacaacctccgctcctc aggcgccgccggcactgcccgttcgccgacccaaccgtagATGGGACA CCACTGGAACCAGGGCCGGTAAGTCCAAGCAGCCGC CGCCGTTAGCCCAAGAGCAACAACAGCGCCAAGGCT ACCGCTCATGGCGCGGGCACAAGAACGCCATAGTTG CTTGCTTGCAAGACTGTGGGGGCAACATCTCCTTCGC CCGCCGCTTTCTTCTCTACCATCACGGCGTGGCCTTC CCCCGTAACATCCTGCATTACTACCGTCATCTCTACAG CCCATACTGCACCGGCGGCAGCGGCAGCAACAGCAG CGGCCACACAGAAGCAAAGGCGACCGGATAGCAAGA CTCTGACAAAGCCCAAGAAATCCACAGCGGCGGCAG CAGCAGGAGGAGGAGCGCTGCGTCTGGCGCCCAACG AACCCGTATCGACCCGCGAGCTTAGAAACAGGATTTT TCCCACTCTGTATGCTATATTTCAACAGAGCAGGGGC CAAGAACAAGAGCTGAAAATAAAAAACAGGTCTCTGC GATCCCTCACCCGCAGCTGCCTGTATCACAAAAGCGA AGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGGC TCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGACT AGaataaaagatctttattttcattagatctgtgtgttggttttttgtgtgGCGGCC GCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAA AGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTT TCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAA TGTATCTTAGCTTAACGGGCGGCGAAGGAGAAGTCCA CGCCTACATGGGGGGTAGAGTCATAATCGTGCATCAGG ATAGGGCGGTGGTGCTGCAGCAGCGCGCGAATAAAC

		TGCTGCCGCCGCCGCTCCGTCCTGCAGGAATACAAC ATGGCAGTGGTCTCCTCAGCGATGATTCGCACCGCCC GCAGCATAAGGCGCCTTGTCCTCCGGGCACAGCAGC GCACCCTGATCTCACTTAAATCAGCACAGTAACTGCA GCACAGCACCACAATATTGTTCAAAATCCCACAGTGC AAGGCGCTGTATCCAAAGCTCATGGCGGGGACCACA GAACCCACGTGGCCATCATACCACAAGCGCAGGTAGA TTAAGTGGCGACCCCTCATAAACACGCTGGACATAAA CATTACCTCTTTTGGCATGTTGTAATTCACCACCTCCC GGTACCATATAAACCTCTGATTAAACATGGCGCCATCC ACCACCATCCTAAACCAGCTGGCCAAAACCTGCCCGC CGGCTATACACTGCAGGGAACCGGGACTGGAACAAT GACAGTGGAGAGCCCAGGACTCGTAACCATGGATCAT CATGCTCGTCATGATATCAATGTTGGCACAACACAGG CACACGTGCATACACTTCCTCAGGATTACAAGCTCCT CCCGCGTTAGAACCATATCCCAGGGAACAACCCATTC CTGAATCAGCGTAAATCCCACACTGCAGGGAAGACCT CGCACGTAACTCACGTTGTGCATTGTCAAAGTGTTACA TTCGGGCAGCAGCGGATGATCCTCCAGTATGGTAGC GCGGGTTTCTGTCTCAAAAGGAGGTAGACGATCCCTA CTGTACGGAGTGCGCCGAGACAACCGAGATCGTGTT GGTCGTAGTGTCATGCCAAATGGAACGCCGGACGTA GTCATATTTCCTGAAGCAAAACCAGGTGCGGGCGTGA CAAACAGATCTGCGTCTCCGGTCTCGCCGCTTAGATC GCTCTGTGTAGTAGTTGTAGTATATCCACTCTCTCAAA GCATCCAGGCGCCCCCTGGCTTCGGGTTCTATGTAAA CTCCTTCATGCGCCGCTGCCCTGATAACATCCACCAC CGCAGAATAAGCCACACCCAGCCAACCTACACATTCG TTCTGCGAGTCACACACGGGAGGAGCGGGAAGAGCT GGAAGAACCATGGTGGCATTTGCAAAAGCCTAGGCCT CCAAAAAAGCCTCCTCACTACTTCTGGAATAGCTCAGA GGCCGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAA ATTAGTCAGCCATGGGGCGGAGAATGGGCGGAACTG GGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGG CGGGACTATGGTTGCTGACTAATTGAGATGCATGCTT TGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTC CACACCTGGTTGCTGACTAATTGAGATGCATGCTTTG CATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCA CACCCTAACTGACACACACGTTACGTCACTTCCCATTT TAAGAAAACTACAATTCCCAACACATACAAGTTACTCC GCCCTTAATtaaatcggatccgatatctagatgtattcgcgaggtaccgag ctcgaattctctggccgtcgttttacaacgtcgtgactgggaaaaccctggcgtta cccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcga agaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaa tggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatat ggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgac acccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatcc

		gcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcga
103	pEMBR-1.37B2+L4-minKanR-v2	AATTCAGTACTCATATGTTGACAATTAATCATCGGCAT AGTATATCGGCATAGTATAATACGACTCACTATAGGAG GGCCACCAAGGTGAGGAACTAAACCATGAGCCATATT CAACGGGAAACGTCGAGGCCGCGATTAAATTCCAACA TGGATGCTGATTTATATGGGTATAAATGGGCTCGCGA TAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTG TATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAAC ATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGA GATGGTCAGACTAAACTGGCTGACGGAATTTATGCCA CTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGA TGCATGGTTACTCACCACTGCGATCCCCGGAAAAACA GCGTTCCAGGTATTAGAAGAATATCCTGATTCAGGTG AAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCG GTTGCACTCGATTCCTGTTTGTAATTGTCCTTTTAACA GCGATCGCGTATTTCGCCTCGCTCAGGCGCAATCACG AATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATG ACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAA AGAAATGCATAAACTTTTGCCATTCTCACCGGATTCAG TCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTT TTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGA CGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCA TCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTA CAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATG AGTTTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTT ATGCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAA AACTCACGTTAAGGGATTTTGGTCATGtcgacGTTAACA

		AGCTTTACcctatccgtagatgtacctggacatccaggtgatgccggcggc ggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgcagc ggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgcagt cgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgt ggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcgaac cccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaac ccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttccagg cgcggcggctgctgcgctagctttttttggccactggccgcgcgcggcgtaagcg gttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggagggtt attttccaagggttgagtcgcaggacccccggttcgagtctcgggccggccgga ctgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaaattc ctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtgctgc ggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcggcag acatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacatcgat ccagacatgataagatacattgatgagtttggacaaaccacaactagaatgca gtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata agctgcaataaacaagtttGTACACTCTCGGGTGATTATTTACC CCCACCCTTGCCGTCTGCGCCGTTTAAAAATCAAAGG GGTTCTGCCGCGCATCGCTATGCGCCACTGGCAGGG ACACGTTGCGATACTGGTGTTTAGTGCTCCACTTAAAC TCAGGCACAACCATCCGCGGCAGCTCGGTGAAGTTTT CACTCCACAGGCTGCGCACCATCACCAACGCGTTTAG CAGGTCGGGCGCCGATATCTTGAAGTCGCAGTTGGG GCCTCCGCCCTGCGCGCGCGAGTTGCGATACACAGG GTTGCAGCACTGGAACACTATCAGCGCCGGGTGGTG CACGCTGGCCAGCACGCTCTTGTCGGAGATCAGATCC GCGTCCAGGTCCTCCGCGTTGCTCAGGGCGAACGGA GTCAACTTTGGTAGCTGCCTTCCCAAAAAGGGCGCGT GCCCAGGCTTTGAGTTGCACTCGCACCGTAGTGGCAT CAAAAGGTGACCGTGCCCGGTCTGGGCGTTAGGATA CAGCGCCTGCATAAAAGCCTTGATCTGCTTAAAAGCC ACCTGAGCCTTTGCGCCTTCAGAGAAGAACATGCCGC AAGACTTGCCGGAAAACTGATTGGCCGGACAGGCCG CGTCGTGCACGCAGCACCTTGCGTCGGTGTTGGAGA TCTGCACCACATTTCGGCCCCACCGGTTCTTCACGAT CTTGGCCTTGCTAGACTGCTCCTTCAGCGCGCGCTGC CCGTTTTCGCTCGTCACATCCATTTCAATCACGTGCTC CTTATTTATCATAATGCTTCCGTGTAGACACTTAAGCT CGCCTTCGATCTCAGCGCAGCGGTGCAGCCACAACG CGCAGCCCGTGGGCTCGTGATGCTTGTAGGTCACCT CTGCAAACGACTGCAGGTACGCCTGCAGGAATCGCC CCATCATCGTCACAAAGGTCTTGTTGCTGGTGAAGGT CAGCTGCAACCCGCGGTGCTCCTCGTTCAGCCAGGT CTTGCATACGGCCGCCAGAGCTTCCACTTGGTCAGGC AGTAGTTTGAAGTTCGCCTTTAGATCGTTATCCACGTG GTACTTGTCCATCAGCGCGCGCGCAGCCTCCATGCC CTTCTCCCACGCAGACACGATCGGCACACTCAGCGG

		GTTCATCACCGTAATTTCACTTTCCGCTTCGCTGGGCT CTTCCTCTTCCTCTTGCGTCCGCATACCACGCGCCAC TGGGTCGTCTTCATTCAGCCGCCGCACTGTGCGCTTA CCTCCTTTGCCATGCTTGATTAGCACCGGTGGGTTGC TGAAACCCACCATTTGTAGCGCCACATCTTCTCTTTCT TCCTCGCTGTCCACGATTACCTCTGGTGATGGCGGGC GCTCGGGCTTGGGAGAAGGGCGCTTCTTTTTCTTCTT GGGCGCAATGGCCAAATCCGCCGCCGAGGTCGATGG CCGCGGGCTGGGTGTGCGCGGCACCAGCGCGTCTTG TGATGAGTCTTCCTCGTCCTCGGACTCGATACGCCGC CTCATCCGCTTTTTTGGGGGCGCCCGGGGAGGCGGC GGCGACGGGGACGGGGACGACACGTCCTCCATGGTT GGGGGACGTCGCGCCGCACCGCGTCCGCGCTCGGG GGTGGTTTCGCGCTGCTCCTCTTCCCGACTGGCCATtt ccttctcctataggcagaaaaagatcCACAAAAGCGAAGATCAGC TTCGGCGCACGCTGGAAGACGCGGAGGCTCTCTTCA GTAAATACTGCGCGCTGACTCTTAAGGACTAGTTTCG CGCCCTTTCTCAAATTTAAGCGCGAAAACTACGTCATC TCCAGCGGCCACACCCGGCGCCAGCACCTGTTGTCA GCGCCATTGGCGCGCCGGCCGGCCGAAtatcTTCATTT AAATgattaggttctacgaagaccaatcccgcccgcctaatgcggagcttac cgcctgcgtcattacccagggccacattcttggccaattgcaagccatcaacaa agcccgccaagagtttctgctacgaaagggacgggggtttacttggaccccc agtccggcgaggagctcaacccaatccccccgccgccgcagccctatcagc agcagccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcag ctgccgccgccacccacggacgaggaggaatactgggacagtcaggcaga ggaggttttggacgaggaggaggaggacatgatggaagactgggagagcct agacgaggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcacc ctcggtcgcattcccctcgccggcgccccagaaatcggcaaccggttccagca tggctacaacctccgctcctcaggcgccgccggcactgcccgttcgccgaccc aaccgtagATGGGACACCACTGGAACCAGGGCCGGTAA GTCCAAGCAGCCGCCGCCGTTAGCCCAAGAGCAACA ACAGCGCCAAGGCTACCGCTCATGGCGCGGGCACAA GAACGCCATAGTTGCTTGCTTGCAAGACTGTGGGGGC AACATCTCCTTCGCCCGCCGCTTTCTTCTCTACCATCA CGGCGTGGCCTTCCCCCGTAACATCCTGCATTACTAC CGTCATCTCTACAGCCCATACTGCACCGGCGGCAGC GGCAGCAACAGCAGCGGCCACACAGAAGCAAAGGCG ACCGGATAGCAAGACTCTGACAAAGCCCAAGAAATCC ACAGCGGCGGCAGCAGCAGGAGGAGGAGCGCTGCG TCTGGCGCCCAACGAACCCGTATCGACCCGCGAGCT TAGAAACAGGATTTTTCCCACTCTGTATGCTATATTTC AACAGAGCAGGGGCCAAGAACAAGAGCTGAAAATAAA AAACAGGTCTCTGCGATCCCTCACCCGCAGCTGCCTG TATCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGG AAGACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCT GACTCTTAAGGACTAGaataaaagatctttattttcattagatctgtgtgt

		tggttttttgtgtgGCGGCCGCAACTTGTTTATTGCAGCTTATA ATGGTTACAAATAAAGCAATAGCATCACAAATTTCACA AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTG TCCAAACTCATCAATGTATCTTAGCTTAACGGGCGGC GAAGGAGAAGTCCACGCCTACATGGGGGTAGAGTCA TAATCGTGCATCAGGATAGGGCGGTGGTGCTGCAGC AGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCGTC CTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCGA TGATTCGCACCGCCCGCAGCATAAGGCGCCTTGTCCT CCGGGCACAGCAGCGCACCCTGATCTCACTTAAATCA GCACAGTAACTGCAGCACAGCACCACAATATTGTTCA AAATCCCACAGTGCAAGGCGCTGTATCCAAAGCTCAT GGCGGGGACCACAGAACCCACGTGGCCATCATACCA CAAGCGCAGGTAGATTAAGTGGCGACCCCTCATAAAC ACGCTGGACATAAACATTACCTCTTTTGGCATGTTGTA ATTCACCACCTCCCGGTACCATATAAACCTCTGATTAA ACATGGCGCCATCCACCACCATCCTAAACCAGCTGGC CAAAACCTGCCCGCCGGCTATACACTGCAGGGAACC GGGACTGGAACAATGACAGTGGAGAGCCCAGGACTC GTAACCATGGATCATCATGCTCGTCATGATATCAATGT TGGCACAACACAGGCACACGTGCATACACTTCCTCAG GATTACAAGCTCCTCCCGCGTTAGAACCATATCCCAG GGAACAACCCATTCCTGAATCAGCGTAAATCCCACAC TGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCAT TGTCAAAGTGTTACATTCGGGCAGCAGCGGATGATCC TCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGAG GTAGACGATCCCTACTGTACGGAGTGCGCCGAGACAA CCGAGATCGTGTTGGTCGTAGTGTCATGCCAAATGGA ACGCCGGACGTAGTCATATTTCCTGAAGCAAAACCAG GTGCGGGCGTGACAAACAGATCTGCGTCTCCGGTCT CGCCGCTTAGATCGCTCTGTGTAGTAGTTGTAGTATAT CCACTCTCTCAAAGCATCCAGGCGCCCCCTGGCTTCG GGTTCTATGTAAACTCCTTCATGCGCCGCTGCCCTGA TAACATCCACCACCGCAGAATAAGCCACACCCAGCCA ACCTACACATTCGTTCTGCGAGTCACACACGGGAGGA GCGGGAAGAGCTGGAAGAACCATGGTGGCATTTGCA AAAGCCTAGGCCTCCAAAAAAGCCTCCTCACTACTTCT GGAATAGCTCAGAGGCCGAGGCGGCCTCGGCCTCTG CATAAATAAAAAAAATTAGTCAGCCATGGGGCGGAGA ATGGGCGGAACTGGGCGGAGTTAGGGGCGGGATGG GCGGAGTTAGGGGCGGGACTATGGTTGCTGACTAATT GAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAG CCTGGGGACTTTCCACACCTGGTTGCTGACTAATTGA GATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCC TGGGGACTTTCCACACCCTAACTGACACACACGTTAC GTCACTTCCCATTTTAAGAAAACTACAATTCCCAACAC

		ATACAAGTTACTCCGCCCTTAATtaaatcggatccgatatctaga tgtattcgcgaggtaccgagctcg
104	pEMBR- 1.47B2+L4	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc

		aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttgtacactctcgggtgattatttacccccacccttt gccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgcc actggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggca caaccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatca ccaacgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctc cgccctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatc agcgccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgc gtccaggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgcctt cccaaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcat caaaaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaa gccttgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgc cgcaagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgca gcaccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttca cgatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtca catccatttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagc tcgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctc gtgatgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcg ccccatcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcg gtgctcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtca ggcagtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagc gcgcgcgcagcctccatgcccttctcccacgcagacacgatcggcacactca gcgggttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcg tccgcataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgctt acctcctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtag cgccacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgc tcgggcttgggagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgc cgccgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtg atgagtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcg cccggggaggcggcggcgacggggacggggacgacacgtcctccatggttg ggggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcct cttcccgactggccatttccttctcctataggcagaaaaagatccacaaaagcg aagatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatac tgcgcgctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaa actacgtcatctccagcggccacacccggcgccagcacctgttgtcagcgcca ttggcgcgcccgcccgccgcgcgcttcgctttttatagggccgccgccgccgcc gcctcgccataaaaggaaactttcggagcgcgccgctctgattggctgccgcc gcacctctccgcctcgccccgccccgcccctcgccccgccccgccccgcctgg cgcgcgccccccccccccccccgcccccatcgctgcacaaaataattaaaaa

		ataaataaatacaaaattgggggtggggaggggggggagatggggagagtg aagcagaacgtggggctcacctcgaggccggccgaatatcttcatttaaatgttt aaacatcgatgcGGCCGCgattaggttctacgaagaccaatcccgcccgc ctaatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaatt gcaagccatcaacaaagcccgccaagagtttctgctacgaaagggacgggg ggtttacttggacccccagtccggcgaggagctcaacccaatccccccgccgc cgcagccctatcagcagcagccgcgggcccttgcttcccaggatggcaccca aaaagaagctgcagctgccgccgccacccacggacgaggaggaatactgg gacagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaa gactgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagac gaaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcgg caaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactg cccgttcgccgacccaaccgtagATGGGACACCACTGGAACCA GGGCCGGTAAGTCCAAGCAGCCGCCGCCGTTAGCCC AAGAGCAACAACAGCGCCAAGGCTACCGCTCATGGC GCGGGCACAAGAACGCCATAGTTGCTTGCTTGCAAGA CTGTGGGGGCAACATCTCCTTCGCCCGCCGCTTTCTT CTCTACCATCACGGCGTGGCCTTCCCCCGTAACATCC TGCATTACTACCGTCATCTCTACAGCCCATACTGCACC GGCGGCAGCGGCAGCAACAGCAGCGGCCACACAGAA GCAAAGGCGACCGGATAGCAAGACTCTGACAAAGCC CAAGAAATCCACAGCGGCGGCAGCAGCAGGAGGAGG AGCGCTGCGTCTGGCGCCCAACGAACCCGTATCGAC CCGCGAGCTTAGAAACAGGATTTTTCCCACTCTGTAT GCTATATTTCAACAGAGCAGGGGCCAAGAACAAGAGC TGAAAATAAAAAACAGGTCTCTGCGATCCCTCACCCG CAGCTGCCTGTATCACAAAAGCGAAGATCAGCTTCGG CGCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAAT ACTGCGCGCTGACTCTTAAGGACTAGaataaaagatctttattt tcattagatctgtgtgttggttttttgtgtgGCGGCCGCAACTTGTTTAT TGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTA GTTGTGGTTTGTCCAAACTCATCAATGTATCTTAGCTT AACGGGCGGCGAAGGAGAAGTCCACGCCTACATGGG GGTAGAGTCATAATCGTGCATCAGGATAGGGCGGTG GTGCTGCAGCAGCGCGCGAATAAACTGCTGCCGCCG CCGCTCCGTCCTGCAGGAATACAACATGGCAGTGGTC TCCTCAGCGATGATTCGCACCGCCCGCAGCATAAGGC GCCTTGTCCTCCGGGCACAGCAGCGCACCCTGATCT CACTTAAATCAGCACAGTAACTGCAGCACAGCACCAC AATATTGTTCAAAATCCCACAGTGCAAGGCGCTGTATC CAAAGCTCATGGCGGGGACCACAGAACCCACGTGGC CATCATACCACAAGCGCAGGTAGATTAAGTGGCGACC CCTCATAAACACGCTGGACATAAACATTACCTCTTTTG GCATGTTGTAATTCACCACCTCCCGGTACCATATAAAC CTCTGATTAAACATGGCGCCATCCACCACCATCCTAAA CCAGCTGGCCAAAACCTGCCCGCCGGCTATACACTG

		CAGGGAACCGGGACTGGAACAATGACAGTGGAGAGC CCAGGACTCGTAACCATGGATCATCATGCTCGTCATG ATATCAATGTTGGCACAACACAGGCACACGTGCATAC ACTTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAAC CATATCCCAGGGAACAACCCATTCCTGAATCAGCGTA AATCCCACACTGCAGGGAAGACCTCGCACGTAACTCA CGTTGTGCATTGTCAAAGTGTTACATTCGGGCAGCAG CGGATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTC TCAAAAGGAGGTAGACGATCCCTACTGTACGGAGTGC GCCGAGACAACCGAGATCGTGTTGGTCGTAGTGTCAT GCCAAATGGAACGCCGGACGTAGTCATATTTCCTGAA GCAAAACCAGGTGCGGGCGTGACAAACAGATCTGCG TCTCCGGTCTCGCCGCTTAGATCGCTCTGTGTAGTAG TTGTAGTATATCCACTCTCTCAAAGCATCCAGGCGCC CCCTGGCTTCGGGTTCTATGTAAACTCCTTCATGCGC CGCTGCCCTGATAACATCCACCACCGCAGAATAAGCC ACACCCAGCCAACCTACACATTCGTTCTGCGAGTCAC ACACGGGAGGAGCGGGAAGAGCTGGAAGAACCATGG TGGCATTTGCAAAAGCCTAGGCCTCCAAAAAAGCCTC CTCACTACTTCTGGAATAGCTCAGAGGCCGAGGCGGC CTCGGCCTCTGCATAAATAAAAAAAATTAGTCAGCCAT GGGGCGGAGAATGGGCGGAACTGGGCGGAGTTAGG GGCGGGATGGGCGGAGTTAGGGGCGGGACTATGGTT GCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCC TGCTGGGGAGCCTGGGGACTTTCCACACCTGGTTGCT GACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGC TGGGGAGCCTGGGGACTTTCCACACCCTAACTGACAC ACACGTTACGTCACTTCCCATTTTAAGAAAACTACAAT TCCCAACACATACAAGTTACTCCGCCCTTAATtaaatcgg atccgatatctagatgtattcgcgaggtaccgagctcgaattctctggccgtcgttt tacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcag cacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgc ccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttct ccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagtacaatctg ctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgac gcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgacc gtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcg a
105	pEMBR-1.47B2+L4-minKanR	AATTCAGTACTCATATGTTGACAATTAATCATCGGCAT AGTATATCGGCATAGTATAATACGACTCACTATAGGAG GGCCACCAAGGTGAGGAACTAAACCATGAGCCATATT CAACGGGAAACGTCGAGGCCGCGATTAAATTCCAACA TGGATGCTGATTTATATGGGTATAAATGGGCTCGCGA TAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTG TATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAAC ATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGA GATGGTCAGACTAAACTGGCTGACGGAATTTATGCCA

		CTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGA TGCATGGTTACTCACCACTGCGATCCCCGGAAAAACA GCGTTCCAGGTATTAGAAGAATATCCTGATTCAGGTG AAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCG GTTGCACTCGATTCCTGTTTGTAATTGTCCTTTTAACA GCGATCGCGTATTTCGCCTCGCTCAGGCGCAATCACG AATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATG ACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAA AGAAATGCATAAACTTTTGCCATTCTCACCGGATTCAG TCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTT TTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGA CGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCA TCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTA CAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATG AGTTTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTT ATGCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAA AACTCACGTTAAGGGATTTTGGTCATGtcgacGTTAACA AGCTTTACcctatccgtagatgtacctggacatccaggtgatgccggcggc ggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgcagc ggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgcagt cgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgt ggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcgaac cccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaac ccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttccagg cgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaagcg gttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggagggtt attttccaagggttgagtcgcaggacccccggttcgagtctcgggccggccgga ctgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaaattc

		ctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtgctgc ggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcggcag acatgcagggcaccctccccttctcctaccgcgtcaggaggggcaacatcgat ccagacatgataagatacattgatgagtttggacaaaccacaactagaatgca gtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata agctgcaataaacaagtttgtacactctcgggtgattatttacccccacccttgcc gtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgccactg gcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggcacaa ccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatcacca acgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgc cctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatcagc gccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcgtc caggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgccttccc aaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatcaa aaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaagcct tgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgccgc aagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgcagca ccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttcacga tcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcacatc catttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagctcgc cttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtg atgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcgccc catcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgc tcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtcaggca gtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagcgcgc gcgcagcctccatgcccttctcccacgcagacacgatcggcacactcagcgg gttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcgtccg cataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgcttacct cctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtagcgc cacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgctcg ggcttgggagaagggcgcttctttttcttcttgggcgcaatggccaaatccgccgc cgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtgatg agtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcgccc ggggaggcggcggcgacggggacggggacgacacgtcctccatggttggg ggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcctctt cccgactggccatttccttctcctataggcagaaaaagatccacaaaagcgaa gatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatactgc gcgctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaaact acgtcatctccagcggccacacccggcgccagcacctgttgtcagcgccattg gcgcgcccgcccgccgcgcgcttcgctttttatagggccgccgccgccgccgc ctcgccataaaaggaaactttcggagcgcgccgctctgattggctgccgccgc acctctccgcctcgccccgccccgcccctcgccccgccccgccccgcctggcg cgcgccccccccccccccccgcccccatcgctgcacaaaataattaaaaaat aaataaatacaaaattgggggtggggaggggggggagatggggagagtga agcagaacgtggggctcacctcgaggccggccgaatatcttcatttaaatgttta aacatcgatgcGGCCGCgattaggttctacgaagaccaatcccgcccgcc taatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaattg

		caagccatcaacaaagcccgccaagagtttctgctacgaaagggacggggg gtttacttggacccccagtccggcgaggagctcaacccaatccccccgccgcc gcagccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaa aaagaagctgcagctgccgccgccacccacggacgaggaggaatactggg acagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaag actgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacg aaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggc aaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactgc ccgttcgccgacccaaccgtagATGGGACACCACTGGAACCAG GGCCGGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCA AGAGCAACAACAGCGCCAAGGCTACCGCTCATGGCG CGGGCACAAGAACGCCATAGTTGCTTGCTTGCAAGAC TGTGGGGGCAACATCTCCTTCGCCCGCCGCTTTCTTC TCTACCATCACGGCGTGGCCTTCCCCCGTAACATCCT GCATTACTACCGTCATCTCTACAGCCCATACTGCACC GGCGGCAGCGGCAGCAACAGCAGCGGCCACACAGAA GCAAAGGCGACCGGATAGCAAGACTCTGACAAAGCC CAAGAAATCCACAGCGGCGGCAGCAGCAGGAGGAGG AGCGCTGCGTCTGGCGCCCAACGAACCCGTATCGAC CCGCGAGCTTAGAAACAGGATTTTTCCCACTCTGTAT GCTATATTTCAACAGAGCAGGGGCCAAGAACAAGAGC TGAAAATAAAAAACAGGTCTCTGCGATCCCTCACCCG CAGCTGCCTGTATCACAAAAGCGAAGATCAGCTTCGG CGCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAAT ACTGCGCGCTGACTCTTAAGGACTAGaataaaagatctttattt tcattagatctgtgtgttggttttttgtgtgGCGGCCGCAACTTGTTTAT TGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTA GTTGTGGTTTGTCCAAACTCATCAATGTATCTTAGCTT AACGGGCGGCGAAGGAGAAGTCCACGCCTACATGGG GGTAGAGTCATAATCGTGCATCAGGATAGGGCGGTG GTGCTGCAGCAGCGCGCGAATAAACTGCTGCCGCCG CCGCTCCGTCCTGCAGGAATACAACATGGCAGTGGTC TCCTCAGCGATGATTCGCACCGCCCGCAGCATAAGGC GCCTTGTCCTCCGGGCACAGCAGCGCACCCTGATCT CACTTAAATCAGCACAGTAACTGCAGCACAGCACCAC AATATTGTTCAAAATCCCACAGTGCAAGGCGCTGTATC CAAAGCTCATGGCGGGGACCACAGAACCCACGTGGC CATCATACCACAAGCGCAGGTAGATTAAGTGGCGACC CCTCATAAACACGCTGGACATAAACATTACCTCTTTTG GCATGTTGTAATTCACCACCTCCCGGTACCATATAAAC CTCTGATTAAACATGGCGCCATCCACCACCATCCTAAA CCAGCTGGCCAAAACCTGCCCGCCGGCTATACACTG CAGGGAACCGGGACTGGAACAATGACAGTGGAGAGC CCAGGACTCGTAACCATGGATCATCATGCTCGTCATG ATATCAATGTTGGCACAACACAGGCACACGTGCATAC ACTTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAAC

		CATATCCCAGGGAACAACCCATTCCTGAATCAGCGTA AATCCCACACTGCAGGGAAGACCTCGCACGTAACTCA CGTTGTGCATTGTCAAAGTGTTACATTCGGGCAGCAG CGGATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTC TCAAAAGGAGGTAGACGATCCCTACTGTACGGAGTGC GCCGAGACAACCGAGATCGTGTTGGTCGTAGTGTCAT GCCAAATGGAACGCCGGACGTAGTCATATTTCCTGAA GCAAAACCAGGTGCGGGCGTGACAAACAGATCTGCG TCTCCGGTCTCGCCGCTTAGATCGCTCTGTGTAGTAG TTGTAGTATATCCACTCTCTCAAAGCATCCAGGCGCC CCCTGGCTTCGGGTTCTATGTAAACTCCTTCATGCGC CGCTGCCCTGATAACATCCACCACCGCAGAATAAGCC ACACCCAGCCAACCTACACATTCGTTCTGCGAGTCAC ACACGGGAGGAGCGGGAAGAGCTGGAAGAACCATGG TGGCATTTGCAAAAGCCTAGGCCTCCAAAAAAGCCTC CTCACTACTTCTGGAATAGCTCAGAGGCCGAGGCGGC CTCGGCCTCTGCATAAATAAAAAAAATTAGTCAGCCAT GGGGCGGAGAATGGGCGGAACTGGGCGGAGTTAGG GGCGGGATGGGCGGAGTTAGGGGCGGGACTATGGTT GCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCC TGCTGGGGAGCCTGGGGACTTTCCACACCTGGTTGCT GACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGC TGGGGAGCCTGGGGACTTTCCACACCCTAACTGACAC ACACGTTACGTCACTTCCCATTTTAAGAAAACTACAAT TCCCAACACATACAAGTTACTCCGCCCTTAATtaaatcgg atccgatatctagatgtattcgcgaggtaccgagctcg
106	E2a region - A	cacacaaaaaaccaacacacagatctaatgaaaataaagatcttttattCTA GTCCTTAAGAGTCAGCGCGCAGTATTTACTGAAGAGA GCCTCCGCGTCTTCCAGCGTGCGCCGAAGCTGATCTT CGCTTTTGTGATACAGGCAGCTGCGGGTGAGGGATC GCAGAGACCTGTTTTTTATTTTCAGCTCTTGTTCTTGG CCCCTGCTCTGTTGAAATATAGCATACAGAGTGGGAA AAATCCTGTTTCTAAGCTCGCGGGTCGATACGGGTTC GTTGGGCGCCAGACGCAGCGCTCCTCCTCCTGCTGC TGCCGCCGCTGTGGATTTCTTGGGCTTTGTCAGAGTC TTGCTATCCGGTCGCCTTTGCTTCTGTGTGGCCGCTG CTGTTGCTGCCGCTGCCGCCGGTGCAGTATGGGCTG TAGAGATGACGGTAGTAATGCAGGATGTTACGGGGGA AGGCCACGCCGTGATGGTAGAGAAGAAAGCGGCGGG CGAAGGAGATGTTGCCCCCACAGTCTTGCAAGCAAGC AACTATGGCGTTCTTGTGCCCGCGCCATGAGCGGTAG CCTTGGCGCTGTTGTTGCTCTTGGGCTAACGGCGGC GGCTGCTTGGACTTACCGGCCCTGGTTCCAGTGGTGT CCCATctacggttgggtcggcgaacgggcagtgccggcggcgcctgagga gcggaggttgtagccatgctggaaccggttgccgatttctggggcgccggcga ggggaatgcgaccgagggtgacggtgtttcgtctgacacctcttcgacctcgga agcttcctcgtctaggctctcccagtcttccatcatgtcctcctcctcctcgtccaaa

		acctcctctgcctgactgtcccagtattcctcctcgtccgtgggtggcggcggcag ctgcagcttctttttgggtgccatcctgggaagcaagggcccgcggctgctgctg atagggctgcggcggcgggggattgggttgagctcctcgccggactgggggt ccaagtaaacccccgtccctttcgtagcagaaactcttggcgggctttgttgatg gcttgcaattggccaagaatgtggccctgggtaatgacgcaggcggtaagctc cgcattaggcgggcgggattggtcttcgtagaacctaatcATTTAAATGA AgataTTCGGCCGGCCGGCGCGCCAATGGCGCTGACA ACAGGTGCTGGCGCCGGGTGTGGCCGCTGGAGATGA CGTAGTTTTCGCGCTTAAATTTGAGAAAGGGCGCGAA ACTAGTCCTTAAGAGTCAGCGCGCAGTATTTACTGAA GAGAGCCTCCGCGTCTTCCAGCGTGCGCCGAAGCTG ATCTTCGCTTTTGTGgatcttttctgcctataggagaaggaaATGG CCAGTCGGGAAGAGGAGCAGCGCGAAACCACCCCCG AGCGCGGACGCGGTGCGGCGCGACGTCCCCCAACC ATGGAGGACGTGTCGTCCCCGTCCCCGTCGCCGCCG CCTCCCCGGGCGCCCCCAAAAAAGCGGATGAGGCGG CGTATCGAGTCCGAGGACGAGGAAGACTCATCACAAG ACGCGCTGGTGCCGCGCACACCCAGCCCGCGGCCAT CGACCTCGGCGGCGGATTTGGCCATTGCGCCCAAGA AGAAAAAGAAGCGCCCTTCTCCCAAGCCCGAGCGCC CGCCATCACCAGAGGTAATCGTGGACAGCGAGGAAG AAAGAGAAGATGTGGCGCTACAAATGGTGGGTTTCAG CAACCCACCGGTGCTAATCAAGCATGGCAAAGGAGGT AAGCGCACAGTGCGGCGGCTGAATGAAGACGACCCA GTGGCGCGTGGTATGCGGACGCAAGAGGAAGAGGAA GAGCCCAGCGAAGCGGAAAGTGAAATTACGGTGATG AACCCGCTGAGTGTGCCGATCGTGTCTGCGTGGGAG AAGGGCATGGAGGCTGCGCGCGCGCTGATGGACAAG TACCACGTGGATAACGATCTAAAGGCGAACTTCAAAC TACTGCCTGACCAAGTGGAAGCTCTGGCGGCCGTATG CAAGACCTGGCTGAACGAGGAGCACCGCGGGTTGCA GCTGACCTTCACCAGCAACAAGACCTTTGTGACGATG ATGGGGCGATTCCTGCAGGCGTACCTGCAGTCGTTTG CAGAGGTGACCTACAAGCATCACGAGCCCACGGGCT GCGCGTTGTGGCTGCACCGCTGCGCTGAGATCGAAG GCGAGCTTAAGTGTCTACACGGAAGCATTATGATAAAT AAGGAGCACGTGATTGAAATGGATGTGACGAGCGAAA ACGGGCAGCGCGCGCTGAAGGAGCAGTCTAGCAAGG CCAAGATCGTGAAGAACCGGTGGGGCCGAAATGTGG TGCAGATCTCCAACACCGACGCAAGGTGCTGCGTGCA CGACGCGGCCTGTCCGGCCAATCAGTTTTCCGGCAA GTCTTGCGGCATGTTCTTCTCTGAAGGCGCAAAGGCT CAGGTGGCTTTTAAGCAGATCAAGGCTTTTATGCAGG CGCTGTATCCTAACGCCCAGACCGGGCACGGTCACC TTTTGATGCCACTACGGTGCGAGTGCAACTCAAAGCC TGGGCACGCGCCCTTTTTGGGAAGGCAGCTACCAAA GTTGACTCCGTTCGCCCTGAGCAACGCGGAGGACCT

		GGACGCGGATCTGATCTCCGACAAGAGCGTGCTGGC CAGCGTGCACCACCCGGCGCTGATAGTGTTCCAGTG CTGCAACCCTGTGTATCGCAACTCGCGCGCGCAGGG CGGAGGCCCCAACTGCGACTTCAAGATATCGGCGCC CGACCTGCTAAACGCGTTGGTGATGGTGCGCAGCCT GTGGAGTGAAAACTTCACCGAGCTGCCGCGGATGGTT GTGCCTGAGTTTAAGTGGAGCACTAAACACCAGTATC GCAACGTGTCCCTGCCAGTGGCGCATAGCGATGCGC GGCAGAACCCCTTTGATTTTTAAACGGCGCAGACGGC AAGGGTGGGGGTAAATAATCA
107	E2a fragment of human mastadeno-virus C serotype 5	Ggatcccatggacgagcccacccttctttatgttttgtttgaagtctttgacgtggtc cgtgtgcaccagccgcaccgcggcgtcatcgaaaccgtgtacctgcgcacgc ccttctcggccggcaacgccacaacataaagaagcaagcaacatcaacaac agctgccgccatgggctccagtgagcaggaactgaaagccattgtcaaagat cttggttgtgggccatatttttgggcacctatgacaagcgctttccaggctttgtttct ccacacaagctcgcctgcgccatagtcaatacggccggtcgcgagactgggg gcgtacactggatggcctttgcctggaacccgcactcaaaaacatgctacctctt tgagccctttggcttttctgaccagcgactcaagcaggtttaccagtttgagtacg agtcactcctgcgccgtagcgccattgcttcttcccccgaccgctgtataacgctg gaaaagtccacccaaagcgtacaggggcccaactcggccgcctgtggactat tctgctgcatgtttctccacgcctttgccaactggccccaaactcccatggatcac aaccccaccatgaaccttattaccggggtacccaactccatgctcaacagtccc caggtacagcccaccctgcgtcgcaaccaggaacagctctacagcttcctgga gcgccactcgccctacttccgcagccacagtgcgcagattaggagcgccactt ctttttgtcacttgaaaaacatgtaaaaataatgtactagagacactttcaataaa ggcaaatgcttttatttgtacactctcgggtgattatttacccccacccttgccgtctg cgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgccactggcag ggacacgttgcgatactggtgtttagtgctccacttaaactcaggcacaaccatc cgcggcagctcggtgaagttttcactccacaggctgcgcaccatcaccaacgc gtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgccctgc gcgcgcgagttgcgatacacagggttgcagcactggaacactatcagcgccg ggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcgtccaggt cctccgcgttgctcagggcgaacggagtcaactttggtagctgccttcccaaaa agggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatcaaaagg tgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaagccttgatct gcttaaaagccacctgagcctttgcgccttcagagaagaacatgccgcaagac ttgccggaaaactgattggccggacaggccgcgtcgtgcacgcagcaccttgc gtcggtgttggagatctgcaccacatttcggccccaccggttcttcacgatcttgg ccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcacatccatttc aatcacgtgctccttatttatcataatgcttccgtgtagacacttaagctcgccttcg atctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtgatgctt gtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcgccccatca tcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgctcctc gttcagccaggtcttgcatacggccgccagagcttccacttggtcaggcagtagt ttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagcgcgcgcgca gcctccatgcccttctcccacgcagacacgatcggcacactcagcgggttcatc accgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcgtccgcatacc

		acgcgccactgggtcgtcttcattcagccgccgcactgtgcgcttacctcctttgc catgcttgattagcaccggtgggttgctgaaacccaccatttgtagcgccacatct tctctttcttcctcgctgtccacgattacctctggtgatggcgggcgctcgggcttgg gagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgccgccgaggt cgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtgatgagtcttc ctcgtcctcggactcgatacgccgcctcatccgctttttgggggcgcccgggga ggcggcggcgacggggacggggacgacacgtcctccatggttgggggacgt cgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcctcttcccga ctggccatttccttctcctataggcagaaaaagatcatggagtcagtcgagaag aaggacagcctaaccgccccctctgagttcgccaccaccgcctccaccgatgc cgccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggag gaagtgattatcgagcaggacccaggttttgtaagcgaagacgacgaggacc gctcagtaccaacagaggataaaaagcaagaccaggacaacgcagaggc aaacgaggaacaagtcgggcgggggacgaaaggcatggcgactacctag atgtgggagacgacgtgctgttgaagcatctgcagcgccagtgcgccattatct gcgacgcgttgcaagagcgcagcgatgtgcccctcgccatagcggatgtcag ccttgcctacgaacgccacctattctcaccgcgcgtacccccaaacgccaag aaaacggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgcc gtgccagaggtgcttgccacctatcacatctttttccaaaactgcaagataccct atcctgccgtgccaaccgcagccgagcggacaagcagctggccttgcggca gggcgctgtcatacctgatatcgcctcgctcaacgaagtgccaaaaatctttgag ggtcttggacgcgacgagaagcgcgcggcaaacgctctgcaacaggaaaa cagcgaaaatgaaagtcactctggagtgttggtggaactcgagggtgacaac gcgcgcctagccgtactaaaacgcagcatcgaggtcacccactttgcctaccc ggcacttaacctacccccccaaggtcatgagcacagtcatgagtgagctgatcgt gcgccgtgcgcagcccctggagagggatgcaaatttgcaagaacaaacaga ggagggcctacccgcagttggcgacgagcagctagcgcgctggcttcaaacg cgcgagcctgccgacttggaggagcgacgcaaactaatgatggccgcagtgc tcgttaccgtggagcttgagtgcatgcagcggttctttgctgacccggagatgca gcgcaagctagaggaaacattgcactacacctttcgacagggctacgtacgcc aggcctgcaagatctccaacgtggagctctgcaacctggtctcctaccttggaat tttgcacgaaaaccgccttgggcaaaacgtgcttcattccacgctcaagggcga ggcgcgccgcgactacgtccgcgactgcgtttacttatttctatgctacacctggc agacggccatgggcgtttggcagcagtgcttggaggagtgcaacctcaagga gctgcagaaactgctaaagcaaaacttgaaggacctatggacggccttcaac gagcgctccgtggccgcgcacctggcggacatcattttccccgaacgcctgctt aaaaccctgcaacagggtctgccagacttcaccagtcaaagcatgttgcagaa ctttaggaactttatcctagagcgctcaggaatcttgcccgccacctgctgtgcac ttcctagcgactttgtgcccattaagtaccgcgaatgccctccgccgctttggggc cactgctaccttctgcagctagccaactaccttgcctaccactctgacataatgga agacgtgagcggtgacggtctactggagtgtcactgtcgctgcaacctatgcac cccgcaccgctccctggtttgcaattcgcagctgcttaacgaaagtcaaattatc ggtacctttgagctgcagggtccctcgcctgacgaaaagtccgcggctccggg gttgaaactcactccggggctgtggacgtcggcttaccttcgcaaatttgtacctg aggactaccacgcccacgagattaggttctacgaagaccaatcccgcccgcct aatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaattgc aagccatcaacaaagcccgccaagagtttctgctacgaaagggacgggggg

		tttacttggacccccagtccggcgaggagctcaacccaatccccccgccgccg cagccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaa aagaagctgcagctgccgccgccacccacggacgaggaggaatactggga cagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaaga ctgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacga aacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggca accggttccagcatggctacaacctccgctcctcaggcgccgccggcactgcc cgttcgccgacccaaccgtagatgggacaccactggaaccagggccggtaa gtccaagcagccgccgccgttagcccaagagcaacaacagcgccaaggct accgctcatggcgcgggcacaagaacgccatagttgcttgcttgcaagactgtg ggggcaacatctccttcgcccgccgctttcttctctaccatcacggcgtggccttc ccccgtaacatcctgcattactaccgtcatctctacagcccatactgcaccggcg gcagcggcagcaacagcagcggccacacagaagcaaaggcgaccggata gcaagactctgacaaagcccaagaaatccacagcggcggcagcagcagga ggaggagcgctgcgtctggcgcccaacgaacccgtatcgacccgcgagctta gaaacaggatttttcccactctgtatgctatatttcaacagagcaggggccaaga acaagagctgaaaataaaaaacaggtctctgcgatccctcacccgcagctgc ctgtatcacaaaagcgaagatcagcttcggcgcacgctggaagacgcggag gctctcttcagtaaatactgcgcgctgactcttaaggactagtttcgcgccctttctc aaatttaagcgcgaaaactacgtcatctccagcggccacacccggcgccagc acctgttgtcagcgccattatgagcaaggaaattcccacgccctacatgtggagt taccagccacaaatgggacttgcggctggagctgcccaagactactcaaccc gaataaactacatgagcgcgggaccccacatgatatcccgggtcaacggaat acgcgcccaccgaaaccgaattctcctggaacaggcggctattaccaccaca cctcgtaataaccttaatccccgtagttggcccgctgccctggtgtaccaggaaa gtcccgctcccaccactgtggtacttcccagagacgcccaggccgaagttcag atgactaactcaggggcgcagcttgcgggcggctttcgtcacagggtgcggtc gcccgggc
108	49bp synthetic polyA tail	AAAtAAatatctttattttCAttacatctGTGTGTTGGTTTTTTGTGT G
109	E4 region Matsushita 1998	gagaatttggttcacctaatgcaccaaacacaaatcccctcaaaacaaaaatt ggccatggcctagaatttgattcaaacaaggctatggttcctaaactaggaactg gccttagttttgacagcacaggtgccattacagtaggaaacaaaaataatgata agctaactttgtggaccacaccagctccatctcctaactgtagactaaatgcaga gaaagatgctaaactcactttggtcttaacaaaatgtggcagtcaaatacttgcta cagtttcagttttggctgttaaaggcagtttggctccaatatctggaacagttcaaa gtgctcatcttattataagatttgacgaaaatggagtgctactaaacaattccttcct ggacccagaatattggaactttagaaatggagatcttactgaaggcacagccta tacaaacgctgttggatttatgcctaacctatcagcttatccaaaatctcacggta aaactgccaaaagtaacattgtcagtcaagtttacttaaacggagacaaaacta aacctgtaacactaaccattacactaaacggtacacaggaaacaggagaca caactccaagtgcatactctatgtcattttcatgggactggtctggccacaactac attaatgaaatatttgccacatcctcttacacttttcatacattgcccaagaataaa gaatcgtttgtgttatgtttcaacgtgtttatttttcaattgcagaaaatttcaagtcattt ttcattcagtagtatagccccaccaccacatagcttatacagatcaccgtacctta atcaaactcacagaaccctagtattcaacctgccacctccctcccaacacaca

		gagtacacagtcctttctccccggctggccttaaaaagcatcatatcatgggtaa cagacatattcttaggtgttatattccacacggtttcctgtcgagccaaacgctcat cagtgatattaataaactccccgggcagctcacttaagttcatgtcgctgtccagc tgctgagccacaggctgctgtccaacttgcggttgcttaacgggcggcgaagg agaagtccacgcctacatgggggtagagtcataatcgtgcatcaggatagggc ggtggtgctgcagcagcgcgcgaataaactgctgccgccgccgctccgtcctg caggaatacaacatggcagtggtctcctcagcgatgattcgcaccgcccgcag cataaggcgccttgtcctccgggcacagcagcgcaccctgatctcacttaaatc agcacagtaactgcagcacagcaccacaatattgttcaaaatcccacagtgca aggcgctgtatccaaagctcatggcggggaccacagaacccacgtggccatc ataccacaagcgcaggtagattaagtggcgacccctcataaacacgctggac ataaacattacctcttttggcatgttgtaattcaccacctcccggtaccatataaac ctctgattaaacatggcgccatccaccaccatcctaaaccagctggccaaaac ctgcccgccggctatacactgcagggaaccgggactggaacaatgacagtgg agagcccaggactcgtaaccatggatcatcatgctcgtcatgatatcaatgttgg cacaacacaggcacacgtgcatacacttcctcaggattacaagctcctcccgc gttagaaccatatcccagggaacaacccattcctgaatcagcgtaaatcccac actgcagggaagacctcgcacgtaactcacgttgtgcattgtcaaagtgttacat tcgggcagcagcggatgatcctccagtatggtagcgcgggtttctgtctcaaaa ggaggtagacgatccctactgtacggagtgcgccgagacaaccgagatcgtg ttggtcgtagtgtcatgccaaatggaacgccggacgtagtcatatttcctgaagc aaaaccaggtgcgggcgtgacaaacagatctgcgtctccggtctcgccgctta gatcgctctgtgtagtagttgtagtatatccactctctcaaagcatccaggcgccc cctggcttcgggttctatgtaaactccttcatgcgccgctgccctgataacatccac caccgcagaataagccacacccagccaacctacacattcgttctgcgagtcac acacgggaggagcgggaagagctggaagaaccatgttttttttttttattccaaaa gattatccaaaacctcaaaatgaagatctattaagtgaacgcgctcccctccggt ggcgtggtcaaactctacagccaaagaacagataatggcatttgtaagatgttg cacaatggcttccaaaaggcaaacggccctcacgtccaagtggacgtaaagg ctaaacccttcagggtgaatctcctctataaacattccagcaccttcaaccatgcc caaataattctcatctcgccaccttctcaatatatctctaagcaaatcccgaatatt aagtccggccattgtaaaaatctgctccagagcgccctccaccttcagcctcaa gcagcgaatcatgattgcaaaaattcaggttcctcacagacctgtataagattca aaagcggaacattaacaaaaataccgcgatcccgtaggtcccttcgcagggc cagctgaacataatcgtgcaggtctgcacggaccagcgcggccacttccccgc caggaaccatgacaaaagaacccacactgattatgacacgcatactcggagc tatgctaaccagcgtagccccgatgtaagcttgttgcatgggcggcgatataaa atgcaaggtgctgctcaaaaaatcaggcaaagcctcgcgcaaaaaagaaag cacatcgtagtcatgctcatgcagataaaggcaggtaagctccggaaccacca cagaaaaagacaccatttttctctcaaacatgtctgcgggtttctgcataaacac aaaataaaataacaaaaaaacatttaaacattagaagcctgtcttacaacagg aaaaacaacccttataagcataagacggactacggccatgccggcgtgaccg taaaaaaactggtcaccgtgattaaaaagcaccaccgacagctcctcggtcat gtccggagtcataatgtaagactcggtaaacacatcaggttgattcacatcggtc agtgctaaaaagcgaccgaaatagcccgggggaatacatacccgcaggcgt agagacaacattacagcccccataggaggtataacaaaattaataggagag aaaaacacataaacacctgaaaaaccctcctgcctaggcaaaatagcaccct

		cccgctccagaacaacatacagcgcttccacagcggcagccataacagtcag ccttaccagtaaaaaagaaaacctattaaaaaaacaccactcgacacggcac cagctcaatcagtcacagtgtaaaaaagggccaagtgcagagcgagtatatat aggactaaaaaatgacgtaacggttaaagtccacaaaaaacacccagaaaa ccgcacgcgaacctacgcccagaaacgaaagccaaaaaacccacaacttc ctcaaatcgtcacttccgttttcccacgttacgtaacttcccattttaagaaaactac aattcccaacacatacaagttactccgccctaaaac
110	E2a region Matsushita 1998	ggatcccatggacgagcccacccttctttatgttttgtttgaagtctttgacgtggtcc gtgtgcaccggccgcaccgcggcgtcatcgaaaccgtgtacctgcgcacgcc cttctcggccggcaacgccacaacataaagaagcaagcaacatcaacaaca gctgccgccatgggctccagtgagcaggaactgaaagccattgtcaaagatct tggttgtgggccatatttttttgggcacctatgacaagcgctttccaggctttgtttctc cacacaagctcgcctgcgccatagtcaatacggccggtcgcgagactggggg cgtacactggatggcctttgcctggaacccgcactcaaaaacatgctacctctttt gagccctttggcttttctgaccagcgactcaagcaggtttaccagtttgagtacga gtcactcctgcgccgtagcgccattgcttcttcccccgaccgctgtataacgctgg aaaagtccacccaaagcgtacaggggcccaactcggccgcctgtggactatt ctgctgcatgtttctccacgcctttgccaactggccccaaactcccatggatcaca accccaccatgaaccttattaccggggtacccaactccatgctcaacagtcccc aggtacagcccaccctgcgtcgcaaccaggaacagctctacagcttcctgga gcgccactcgccctacttccgcagccacagtgcgcagattaggagcgccactt ctttttgtcacttgaaaaacatgtaaaaataatgtactagagacactttcaataaa ggcaaatgcttttatttgtacactctcgggtgattatttacccccacccttgccgtctg cgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgccactggcag ggacacgttgcgatactggtgtttagtgctccacttaaactcaggcacaaccatc cgcggcagctcggtgaagttttcactccacaggctgcgcaccatcaccaacgc gtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgccctgc gcgcgcgagttgcgatacacagggttgcagcactggaacactatcagcgccg ggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcgtccaggt cctccgcgttgctcagggcgaacggagtcaactttggtagctgccttcccaaaa agggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatcaaaagg tgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaagccttgatct gcttaaaagccacctgagcctttgcgccttcagagaagaacatgccgcaagac ttgccggaaaactgattggccggacaggccgcgtcgtgcacgcagcaccttgc gtcggtgttggagatctgcaccacatttcggccccaccggttcttcacgatcttgg ccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcacatccatttc aatcacgtgctccttatttatcataatgcttccgtgtagacacttaagctcgccttcg atctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtgatgctt gtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcgccccatca tcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgctcctc gttcagccaggtcttgcatacggccgccagagcttccacttggtcaggcagtagt ttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagcgcgcgcgca gcctccatgcccttctcccacgcagacacgatcggcacactcagcgggttcatc accgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcgtccgcatacc acgcgccactgggtcgtcttcattcagccgccgcactgtgcgcttacctcctttgc catgcttgattagcaccggtgggttgctgaaacccaccatttgtagcgccacatct tctctttcttcctcgctgtccacgattacctctggtgatggcgggcgctcgggcttgg

		gagaagggcgcttctttttcttcttgggcgcaatggccaaatccgccgccgaggt cgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtgatgagtcttc ctcgtcctcggactcgatacgccgcctcatccgctttttgggggcgcccgggga ggcggcggcgacggggacggggacgacacgtcctccatggttgggggacgt cgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcctcttcccga ctggccatttccttctcctataggcagaaaaagatcatggagtcagtcgagaag aaggacagcctaaccgccccctctgagttcgccaccaccgcctccaccgatgc cgccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggag gaagtgattatcgagcaggacccaggttttgtaagcgaagacgacgaggacc gctcagtaccaacagaggataaaaagcaagaccaggacaacgcagaggc aaacgaggaacaagtcgggcgggggggacgaaaggcatggcgactacctag atgtgggagacgacgtgctgttgaagcatctgcagcgccagtgcgccattatct gcgacgcgttgcaagagcgcagcgatgtgcccctcgccatagcggatgtcag ccttgcctacgaacgccacctattctcaccgcgcgtacccccccaaacgccaag aaaacggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgcc gtgccagaggtgcttgccacctatcacatctttttccaaaactgcaagataccccct atcctgccgtgccaaccgcagccgagcggacaagcagctggccttgcggca gggcgctgtcatacctgatatcgcctcgctcaacgaagtgccaaaaatctttgag ggtcttggacgcgacgagaagcgcgcggcaaacgctctgcaacaggaaaa cagcgaaaatgaaagtcactctggagtgttggtggaactcgagggtgacaac gcgcgcctagccgtactaaaacgcagcatcgaggtcacccactttgcctaccc ggcacttaacctacccccccaaggtcatgagcacagtcatgagtgagctgatcgt gcgccgtgcgcagcccctggagagggatgcaaatttgcaagaacaaacaga ggagggcctacccgcagttggcgacgagcagctagcgcgctggcttcaaacg cgcgagcctgccgacttggaggagcgacgcaaactaatgatggccgcagtgc tcgttaccgtggagcttgagtgcatgcagcggttctttgctgacccggagatgca gcgcaagctagaggaaacattgcactacacctttcgacagggctacgtacgcc aggcctgcaagatctccaacgtggagctctgcaacctggtctcctaccttggaat tttgcacgaaaaccgccttgggcaaaacgtgcttcattccacgctcaagggcga ggcgcgccgcgactacgtccgcgactgcgtttacttatttctatgctacacctggc agacggccatgggcgtttggcagcagtgcttggaggagtgcaacctcaagga gctgcagaaactgctaaagcaaaacttgaaggacctatggacggccttcaac gagcgctccgtggccgcgcacctggcggacatcattttccccgaacgcctgctt aaaaccctgcaacagggtctgccagacttcaccagtcaaagcatgttgcagaa ctttaggaactttatcctagagcgctcaggaatcttgcccgccacctgctgtgcac ttcctagcgactttgtgcccattaagtaccgcgaatgccctccgccgctttggggc cactgctaccttctgcagctagccaactaccttgcctaccactctgacataatgga agacgtgagcggtgacggtctactggagtgtcactgtcgctgcaacctatgcac cccgcaccgctccctggtttgcaattcgcagctgcttaacgaaagtcaaattatc ggtacctttgagctgcagggtccctcgcctgacgaaaagtccgcggctccggg gttgaaactcactccggggctgtggacgtcggcttaccttcgcaaatttgtacctg aggactaccacgcccacgagattaggttctacgaagaccaatcccgcccgcc aaatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaattg caagccatcaacaaagcccgccaagagtttctgctacgaaagggacggggg gtttacttggacccccagtccggcgaggagctcaacccaatcccccccgccgcc gcagccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaa aaagaagctgcagctgccgccgccacccacggacgaggaggaatactggg

		acagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaag actgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacg aaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggc aaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactgc ccgttcgccgacccaaccgtagatgggacaccactggaaccagggccggta agtccaagcagccgccgccgttagcccaagagcaacaacagcgccaaggc taccgctcatggcgcgggcacaagaacgccatagttgcttgcttgcaagactgt gggggcaacatctccttcgcccgccgctttcttctctaccatcacggcgtggcctt cccccgtaacatcctgcattactaccgtcatctctacagcccatactgcaccggc ggcagcggcagcggcagcaacagcagcggccacacagaagcaaaggcg accggatagcaagactctgacaaagcccaagaaatccacagcggcggcag cagcaggaggaggagcgctgcgtctggcgcccaacgaacccgtatcgaccc gcgagcttagaaacaggatttttcccactctgtatgctatatttcaacagagcagg ggccaagaacaagagctgaaaataaaaaacaggtctctgcgatccctcaccc gcagctgcctgtatcacaaaagcgaagatcagcttcggcgcacgctggaaga cgcggaggctctcttcagtaaatactgcgcgctgactcttaaggactagtttcgcg ccctttctcaaatttaagcgcgaaaactacgtcatctccagcggccacacccgg cgccagcacctgtcgtcagcgccattatgagcaaggaaattcccacgccctac atgtggagttaccagccacaaatgggacttgcggctggagctgcccaagacta ctcaacccgaataaactacatgagcgcgggaccccacatgatatcccgggtc aacggaatccgcgcccaccgaaaccgaattctcttggaacaggcggctattac caccacacctcgtaataaccttaatccccgtagttggcccgctgccctggtgtac caggaaagtcccgctcccaccactgtggtacttcccagagacgcccaggccg aagttcagatgactaactcaggggcgcagcttgcgggcggctttcgtcacagg gtgcggtcgcccgggc
111	VA RNA region Matsushita 1998	tcgacagaagcaccatgtccttgggtccggcctgctgaatgcgcaggcggtcg gccatgccccaggcttcgttttgacatcggcgcaggtctttgtagtagtcttgcatg agcctttctaccggcacttcttcttctccttcctcttgtcctgcatctcttgcatctatcg ctgcggcggcggcggagtttggccgtaggtggcgccctcttcctcccatgcgtgt gaccccgaagcccctcatcggctgaagcagggctaggtcggcgacaacgcg ctcggctaatatggcctgctgcacctgcgtgagggtagactggaagtcatccat gtccacaaagcggtggtatgcgcccgtgttgatggtgtaagtgcagttggccata acggaccagttaacggtctggtgacccggctgcgagagctcggtgtacctgag acgcgagtaagccctcgagtcaaatacgtagtcgttgcaagtccgcaccaggt actggtatcccaccaaaaagtgcggcggcggctggcggtagaggggccagc gtagggtggccggggctccgggggcgagatcttccaacataaggcgatgatat ccgtagatgtacctggacatccaggtgatgccggcggcggtggtggaggcgc gcggaaagtcgcggacgcggttccagatgttgcgcagcggcaaaaagtgctc catggtcgggacgctctggccggtcaggcgcgcgcaatcgttgacgctctaga ccgtgcaaaaggagagcctgtaagcgggcactcttccgtggtctggtggataa attcgcaagggtatcatggcggacgaccggggttcgagccccgtatccggccg tccgccgtgatccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgt cagacaacgggggagtgctccttttggcttccttccaggcgcggcggctgctgc gctagcttttttggccactggccgcgcgcagcgtaagcggttaggctggaaagc gaaagcattaagtggctcgctccctgtagccggagggttattttccaagggttga gtcgcgggacccccggttcgagtctcggaccggccggactgcggcgaacgg gggtttgcctccccgtcatgcaagaccccgcttgcaaattcctccggaaacagg

		gacgagcccctttttgcttttcccagatgcatccggtgctgcggcagatgcgccc ccctcctcagcagcggcaagagcaagagcagcggcagacatgcagggcac cctcccctcctcctaccgcgtcaggaggggcgacatccgcggttgacgcggca gcagatggtgattacgaacccccgcggcgccgggcccggcactacctggact tggaggagggcgagggcctggcgcggctaggagcgccctctcctgagcggt acccaagggtgcagctgaagcgtgatacgcgtgaggcgtacgtgccgcggc agaacctgtttcgcgaccgcgagggagaggagcccgaggagatgcgggatc gaaagttccacgcagggcgcgagctgcggcatggcctgaatcgcgagcggtt gctgcgcgaggaggactttgagcccgacgcgcgaaccgggattagtcccgcg cgcgcacacgtggcggccgccgacctggtaaccgcatacgagcagacggtg aaccaggagattaactttcaaaaaagct
112	E2a region – B	ggatcccatggacgagcccacccttctttatgttttgtttgaagtctttgacgtggtcc gtgtgcaccggccgcaccgcggcgtcatcgaaaccgtgtacctgcgcacgcc cttctcggccggcaacgccacaacataaagaagcaagcaacatcaacaaca gctgccgccatgggctccagtgagcaggaactgaaagccattgtcaaagatct tggttgtgggccatatttttgggcacctatgacaagcgctttccaggctttgtttctc cacacaagctcgcctgcgccatagtcaatacggccggtcgcgagactggggg cgtacactggatggcctttgcctggaacccgcactcaaaaacatgctacctctttt gagccctttggcttttctgaccagcgactcaagcaggtttaccagtttgagtacga gtcactcctgcgccgtagcgccattgcttcttcccccgaccgctgtataacgctgg aaaagtccacccaaagcgtacaggggcccaactcggccgcctgtggactatt ctgctgcatgtttctccacgcctttgccaactggccccaaactcccatggatcaca accccaccatgaaccttattaccggggtacccaactccatgctcaacagtcccc aggtacagcccaccctgcgtcgcaaccaggaacagctctacagcttcctgga gcgccactcgccctacttccgcagccacagtgcgcagattaggagcgccactt ctttttgtcacttgaaaaacatgtaaaaataatgtactagagacactttcaataaa ggcaaatgcttttatttgtacactctcgggtgattatttacccccacccttgccgtctg cgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgccactggcag ggacacgttgcgatactggtgtttagtgctccacttaaactcaggcacaaccatc cgcggcagctcggtgaagttttcactccacaggctgcgcaccatcaccaacgc gtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgccctgc gcgcgcgagttgcgatacacagggttgcagcactggaacactatcagcgccg ggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcgtccaggt cctccgcgttgctcagggcgaacggagtcaactttggtagctgccttcccaaaa agggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatcaaaagg tgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaagccttgatct gcttaaaagccacctgagcctttgcgccttcagagaagaacatgccgcaagac ttgccggaaaactgattggccggacaggccgcgtcgtgcacgcagcaccttgc gtcggtgttggagatctgcaccacatttcggccccaccggttcttcacgatcttgg ccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcacatccatttc aatcacgtgctccttatttatcataatgcttccgtgtagacacttaagctcgccttcg atctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtgatgctt gtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcgccccatca tcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgctcctc gttcagccaggtcttgcatacggccgccagagcttccacttggtcaggcagtagt ttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagcgcgcgcgca gcctccatgcccttctcccacgcagacacgatcggcacactcagcgggttcatc

		accgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcgtccgcatacc acgcgccactgggtcgtcttcattcagccgccgcactgtgcgcttacctcctttgc catgcttgattagcaccggtgggttgctgaaacccaccatttgtagcgccacatct tctctttcttcctcgctgtccacgattacctctggtgatggcgggcgctcgggcttgg gagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgccgccgaggt cgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtgatgagtcttc ctcgtcctcggactcgatacgccgcctcatccgctttttttgggggcgcccgggga ggcggcggcgacggggacggggacgacacgtcctccatggttgggggacgt cgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcctcttcccga ctggccatttccttctcctataggcagaaaaagatcatggagtcagtcgagaag aaggacagcctaaccgccccctctgagttcgccaccaccgcctccaccgatgc cgccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggag gaagtgattatcgagcaggacccaggttttgtaagcgaagacgacgaggacc gctcagtaccaacagaggataaaaagcaagaccaggacaacgcagaggc aaacgaggaacaagtcgggcgggggggacgaaaggcatggcgactacctag atgtgggagacgacgtgctgttgaagcatctgcagcgccagtgcgccattatct gcgacgcgttgcaagagcgcagcgatgtgcccctcgccatagcggatgtcag ccttgcctacgaacgccacctattctcaccgcgcgtacccccccaaacgccaag aaaacggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgcc gtgccagaggtgcttgccacctatcacatctttttccaaaactgcaagataccct atcctgccgtgccaaccgcagccgagcggacaagcagctggccttgcggca gggcgctgtcatacctgatatcgcctcgctcaacgaagtgccaaaaatctttgag ggtcttggacgcgacgagaagcgcgcggcaaacgctctgcaacaggaaaa cagcgaaaatgaaagtcactctggagtgttggtggaactcgagggtgacaac gcgcgcctagccgtactaaaacgcagcatcgaggtcacccactttgcctaccc ggcacttaacctacccccccaaggtcatgagcacagtcatgagtgagctgatcgt gcgccgtgcgcagcccctggagagggatgcaaatttgcaagaacaaacaga ggagggcctacccgcagttggcgacgagcagctagcgcgctggcttcaaacg cgcgagcctgccgacttggaggagcgacgcaaactaatgatggccgcagtgc tcgttaccgtggagcttgagtgcatgcagcggttctttgctgacccggagatgca gcgcaagctagaggaaacattgcactacacctttcgacagggctacgtacgcc aggcctgcaagatctccaacgtggagctctgcaacctggtctcctaccttggaat tttgcacgaaaaccgccttgggcaaaacgtgcttcattccacgctcaagggcga ggcgcgccgcgactacgtccgcgactgcgtttacttatttctatgctacacctggc agacggccatgggcgtttggcagcagtgcttggaggagtgcaacctcaagga gctgcagaaactgctaaagcaaaacttgaaggacctatggacggccttcaac gagcgctccgtggccgcgcacctggcggacatcattttccccgaacgcctgctt aaaaccctgcaacagggtctgccagacttcaccagtcaaagcatgttgcagaa ctttaggaactttatcctagagcgctcaggaatcttgcccgccacctgctgtgcac ttcctagcgactttgtgcccattaagtaccgcgaatgccctccgccgctttggggc cactgctaccttctgcagctagccaactaccttgcctaccactctgacataatgga agacgtgagcggtgacggtctactggagtgtcactgtcgctgcaacctatgcac cccgcaccgctccctggtttgcaattcgcagctgcttaacgaaagtcaaattatc ggtacctttgagctgcagggtccctcgcctgacgaaaagtccgcggctccggg gttgaaactcactccggggctgtggacgtcggcttaccttcgcaaatttgtacctg aggactaccacgcccacgagattaggttctacgaagaccaatcccgcccgcc aaatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaattg

		caagccatcaacaaagcccgccaagagtttctgctacgaaagggacggggg gtttacttggacccccagtccggcgaggagctcaacccaatccccccgccgcc gcagccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaa aaagaagctgcagctgccgccgccacccacggacgaggaggaatactggg acagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaag actgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacg aaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggc aaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactgc ccgttcgccgacccaaccgtagatgggacaccactggaaccagggccggta agtccaagcagccgccgccgttagcccaagagcaacaacagcgccaaggc taccgctcatggcgcgggcacaagaacgccatagttgcttgcttgcaagactgt gggggcaacatctccttcgcccgccgctttcttctctaccatcacggcgtggcctt cccccgtaacatcctgcattactaccgtcatctctacagcccatactgcaccggc ggcagcggcagcggcagcaacagcagcggccacacagaagcaaaggcg accggatagcaagactctgacaaagcccaagaaatccacagcggcggcag cagcaggaggaggagcgctgcgtctggcgcccaacgaacccgtatcgaccc gcgagcttagaaacaggatttttcccactctgtatgctatatttcaacagagcagg ggccaagaacaagagctgaaaataaaaaacaggtctctgcgatccctcaccc gcagctgcctgtatcacaaaagcgaagatcagcttcggcgcacgctggaaga cgcggaggctctcttcagtaaatactgcgcgctgactcttaaggactagtttcgcg ccctttctcaaatttaagcgcgaaaactacgtcatctccagcggccacacccgg cgccagcacctgtcgtcagcgccattatgagcaaggaaattcccacgccctac atgtggagttaccagccacaaatgggacttgcggctggagctgcccaagacta ctcaacccgaataaactacatgagcgcgggaccccacatgatatcccgggtc aacggaatccgcgcccaccgaaaccgaattctcttggaacaggcggctattac caccacacctcgtaataaccttaatccccgtagttggcccgctgccctggtgtac caggaaagtcccgctcccaccactgtggtacttcccagagacgcccaggccg aagttcagatgactaactcaggggcgcagcttgcgggcggctttcgtcacagg gtgcggtcgcccgggc
113	pEMBR-1.4B2+ native L4 (BAC)	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt

		gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcttttggcttccttcc aggcgcggcggctgctgcgctagcttttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccctctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttgtacactctcgggtgattatttacccccacccttt gccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgcc actggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggca caaccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatca ccaacgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctc cgccctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatc agcgccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgc gtccaggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgcctt cccaaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcat caaaaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaa gccttgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgc

		cgcaagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgca gcaccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttca cgatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtca catccatttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagc tcgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctc gtgatgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcg ccccatcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcg gtgctcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtca ggcagtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagc gcgcgcgcagcctccatgcccttctcccacgcagacacgatcggcacactca gcgggttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcg tccgcataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgctt acctcctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtag cgccacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgc tcgggcttgggagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgc cgccgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtg atgagtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcg cccggggaggcggcggcgacggggacggggacgacacgtcctccatggttg ggggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcct cttcccgactggccatttccttctcctataggcagaaaaagatccacaaaagcg aagatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatac tgcgcgctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaa actacgtcatctccagcggccacacccggcgccagcacctgttgtcagcgcca ttggcgcgcccgcccgccgcgcgcttcgctttttatagggccgccgccgccgcc gcctcgccataaaaggaaactttcggagcgcgccgctctgattggctgccgcc gcacctctccgcctcgccccgccccgcccctcgccccgccccgccccgcctgg cgcgcgccccccccccccccccgcccccatcgctgcacaaaataattaaaaa ataaataaatacaaaattgggggtggggaggggggggagatggggagagtg aagcagaacgtggggctcacctcgaggccggccgaatatcttcatttcacgcc cacgagattaggttctacgaagaccaatcccgcccgcctaatgcggagcttac cgcctgcgtcattacccagggccacattcttggccaattgcaagccatcaacaa agcccgccaagagtttctgctacgaaagggacgggggtttacttggaccccc agtccggcgaggagctcaacccaatccccccgccgccgcagccctatcagc agcagccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcag ctgccgccgccacccacggacgaggaggaatactgggacagtcaggcaga ggaggttttggacgaggaggaggaggacatgatggaagactgggagagcct agacgaggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcacc ctcggtcgcattcccctcgccggcgccccagaaatcggcaaccggttccagca tggctacaacctccgctcctcaggcgccgccggcactgcccgttcgccgaccc aaccgtagATGGGACACCACTGGAACCAGGGCCGGTAA GTCCAAGCAGCCGCCGCCGTTAGCCCAAGAGCAACA ACAGCGCCAAGGCTACCGCTCATGGCGCGGGCACAA GAACGCCATAGTTGCTTGCTTGCAAGACTGTGGGGGC AACATCTCCTTCGCCCGCCGCTTTCTTCTCTACCATCA CGGCGTGGCCTTCCCCCGTAACATCCTGCATTACTAC CGTCATCTCTACAGCCCATACTGCACCGGCGGCAGC GGCAGCAACAGCAGCGGCCACACAGAAGCAAAGGCG

		ACCGGATAGCAAGACTCTGACAAAGCCCAAGAAATCC ACAGCGGCGGCAGCAGCAGGAGGAGGAGCGCTGCG TCTGGCGCCCAACGAACCCGTATCGACCCGCGAGCT TAGAAACAGGATTTTTCCCACTCTGTATGCTATATTTC AACAGAGCAGGGGCCAAGAACAAGAGCTGAAAATAAA AAACAGGTCTCTGCGATCCCTCACCCGCAGCTGCCTG TATCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGG AAGACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCT GACTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAATTT AAGCGCGAAAACTACGTCATCTCCAGCGGCCACACCC GGCGCCAGCACCTGTTGTCAGCGCCATTATGAGCAAG GAAATTCCCACGCCCTACATGTGGAGTTACCAGCCAC AAATGGGACTTGCGGCTGGAGCTGCCCAAGACTACTC AACCCGAATAAACTACATGAGCGCGGGGCaaatgtttaaac atcgatgcggccgccgtttgtgttatgtttcaacgtgtttatttttcaattgcagaaaat ttcaagtcatttttcattcagtagtatagccccaccaccacatagcttatacagatc accgtaccttaatcaaactcacagaaccctagtattcaacctgccacctccctcc caacacacagagtacacagtcctttctccccggctggccttaaaaagcatcata tcatgggtaacagacatattcttaggtgttatattccacacggtttcctgtcgagcc aaacgctcatcagtgatattaataaactccccgggcagctcacttaagttcatgtc gctgtccagctgctgagccacaggctgctgtccaacttgcggttgcttaacgggc ggcgaaggagaagtccacgcctacatgggggtagagtcataatcgtgcatca ggatagggcggtggtgctgcagcagcgcgcgaataaactgctgccgccgccg ctccgtcctgcaggaatacaacatggcagtggtctcctcagcgatgattcgcac cgcccgcagcataaggcgccttgtcctccgggcacagcagcgcaccctgatct cacttaaatcagcacagtaactgcagcacagcaccacaatattgttcaaaatcc cacagtgcaaggcgctgtatccaaagctcatggcggggaccacagaaccca cgtggccatcataccacaagcgcaggtagattaagtggcgacccctcataaac acgctggacataaacattacctcttttggcatgttgtaattcaccacctcccggtac catataaacctctgattaaacatggcgccatccaccaccatcctaaaccagctg gccaaaacctgcccgccggctatacactgcagggaaccgggactggaacaa tgacagtggagagcccaggactcgtaaccatggatcatcatgctcgtcatgata tcaatgttggcacaacacaggcacacgtgcatacacttcctcaggattacaagc tcctcccgcgttagaaccatatcccagggaacaacccattcctgaatcagcgta aatcccacactgcagggaagacctcgcacgtaactcacgttgtgcattgtcaaa gtgttacattcgggcagcagcggatgatcctccagtatggtagcgcgggtttctgt ctcaaaaggaggtagacgatccctactgtacggagtgcgccgagacaaccg agatcgtgttggtcgtagtgtcatgccaaatggaacgccggacgtagtcatatttc ctgaagcaaaaccaggtgcgggcgtgacaaacagatctgcgtctccggtctcg ccgcttagatcgctctgtgtagtagttgtagtatatccactctctcaaagcatccag gcgccccctggcttcgggttctatgtaaactccttcatgcgccgctgccctgataa catccaccaccgcagaataagccacacccagccaacctacacattcgttctgc gagtcacacacgggaggagcgggaagagctggaagaaccatgttttttttttttat tccaaaagattatccaaaacctcaaaatgaagatctattaagtgaacgcgctcc cctccggtggcgtggtcaaactctacagccaaagaacagataatggcatttgta agatgttgcacaatggcttccaaaaggcaaacggccctcacgtccaagtggac gtaaaggctaaaccccttcagggtgaatctcctctataaacattccagcaccttca

		accatgcccaaataattctcatctcgccaccttctcaatatatctctaagcaaatcc cgaatattaagtccggccattgtaaaaatctgctccagagcgccctccaccttca gcctcaagcagcgaatcatgattgcaaaaattcaggttcctcacagacctgtat aagattcaaaagcggaacattaacaaaaataccgcgatcccgtaggtcccttc gcagggccagctgaacataatcgtgcaggtctgcacggaccagcgcggcca cttccccgccaggaaccatgacaaaagaacccacactgattatgacacgcat actcggagctatgctaaccagcgtagccccgatgtaagcttgttgcatgggcgg cgatataaaatgcaaggtgctgctcaaaaaatcaggcaaagcctcgcgcaaa aaagaaagcacatcgtagtcatgctcatgcagataaaggcaggtaagctccg gaaccaccacagaaaaagacaccatttttctctcaaacatgtctgcgggtttctg cataaacacaaaataaaataacaaaaaaacatttaaacattagaagcctgtct tacaacaggaaaaacaacccttataagcataagacggactacggccatgcc ggcgtgaccgtaaaaaaactggtcaccgtgattaaaaagcaccaccgacag ctcctcggtcatgtccggagtcataatgtaagactcggtaaacacatcaggttga ttcacatcggtcagtgctaaaaagcgaccgaaatagcccgggggaatacata cccgcaggcgtagagacaacattacagcccccataggaggtataacaaaatt aataggagagaaaaacacataaacacctgaaaaaccctcctgcctaggcaa aatagcaccctcccgctccagaacaacatacagcgcttccacagcggcagcc ataacagtcagccttaccagtaaaaaagaaaacctattaaaaaaacaccact cgacacggcaccagctcaatcagtcacagtgtaaaaaagggccaagtgcag agcgagtatatataggactaaaaaatgacgtaacggttaaagtccacaaaaa acacccagaaaaccgcacgcgaacctacgcccagaaacgaaagccaaaa aacccacaacttcctcaaatcgtcacttccgttttcccacgttacgtcacttcccatt ttaagaaaactacaattcccaacacatacaagttactccgcccttaattaaatcg gatccgatatctagatgtattcgcgaggtaccgagctcgaattctctggccgtcgt tttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgca gcacatcccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcg cccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtatttt ctccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagtacaatc tgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctga cgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgac cgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgc ga
114	pEMBR-1.4B2+L4 (U5579HI060-9)	tgcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaat atatcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggt gttatgagccatattcaacgggaaacgtcgaggccgcgattaaattccaacatg gatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtg cgacaatctatcgcttgtatgggaagcccgatgcgccagagttgtttctgaaaca tggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactg gctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgat gcatggttactcaccactgcgatccccggaaaaacagcattccaggtattagaa gaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccg gttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgct caggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacg agcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgcc attctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttga cgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccg

		ataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacag aaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttc atttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacattattcag attgggcttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatca agagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtag caccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcga acgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcca cgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcg gaacaggagagcgcacgagggagcttccagggggaaacgcctggtatcttta tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag gggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctg gccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggata accgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc gagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgca aaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggt ttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctc actcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaa ttgtgagcggataacaatttcacacaggaaacagctatgaccatgattacacca agcttgcatgcaggcctatccgtagatgtacctggacatccaggtgatgccggc ggcggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgc agcggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgc agtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactctt ccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcg aaccccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcg aacccaggtgtgcgacgtcagacaacgggggagcgctcctttggcttccttcc aggcgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaa gcggttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggag ggttattttccaagggttgagtcgcaggacccccggttcgagtctcgggccggcc ggactgcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaa attcctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtg ctgcggcagatgcgccccccctcctcagcagcggcaagagcaagagcagcg gcagacatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacat cgatccagacatgataagatacattgatgagtttggacaaaccacaactagaat gcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccat tataagctgcaataaacaagtttgtacactctcgggtgattatttacccccaccctt gccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgcc actggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggca caaccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatca ccaacgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctc cgccctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatc agcgccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgc

		gtccaggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgcctt cccaaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcat caaaaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaa gccttgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgc cgcaagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgca gcaccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttca cgatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtca catccatttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagc tcgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctc gtgatgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcg ccccatcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcg gtgctcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtca ggcagtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagc gcgcgcgcagcctccatgcccttctcccacgcagacacgatcggcacactca gcgggttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcg tccgcataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgctt acctcctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtag cgccacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgc tcgggcttgggagaagggcgcttcttttttcttcttgggcgcaatggccaaatccgc cgccgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtg atgagtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcg cccggggaggcggcggcgacggggacggggacgacacgtcctccatggttg ggggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcct cttcccgactggccatttccttctcctataggcagaaaaagatccacaaaagcg aagatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatac tgcgcgctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaa actacgtcatctccagcggccacacccggcgccagcacctgttgtcagcgcca ttggcgcgcccgcccgccgcgcgcttcgctttttatagggccgccgccgccgcc gcctcgccataaaaggaaactttcggagcgcgccgctctgattggctgccgcc gcacctctccgcctcgccccgccccgccctcgccccgccccgccccgcctgg cgcgcgccccccccccccccccgcccccatcgctgcacaaaataattaaaaa ataaataaatacaaaattgggggtggggaggggggggagatggggagagtg aagcagaacgtggggctcacctcgaggccggccgaatatcttcatttaaatgttt aaacatcgatgcGGCCGCgattaggttctacgaagaccaatcccgcccgc ctaatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaatt gcaagccatcaacaaagcccgccaagagtttctgctacgaaagggacgggg ggtttacttggacccccagtccggcgaggagctcaacccaatccccccgccgc cgcagccctatcagcagcagccgcgggcccttgcttcccaggatggcaccca aaaagaagctgcagctgccgccgccacccacggacgaggaggaatactgg gacagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaa gactgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagac gaaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcgg caaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactg cccgttcgccgacccaaccgtagATGGGACACCACTGGAACCA GGGCCGGTAAGTCCAAGCAGCCGCCGCCGTTAGCCC AAGAGCAACAACAGCGCCAAGGCTACCGCTCATGGC GCGGGCACAAGAACGCCATAGTTGCTTGCTTGCAAGA

		CTGTGGGGGCAACATCTCCTTCGCCCGCCGCTTTCTT CTCTACCATCACGGCGTGGCCTTCCCCCGTAACATCC TGCATTACTACCGTCATCTCTACAGCCCATACTGCACC GGCGGCAGCGGCAGCAACAGCAGCGGCCACACAGAA GCAAAGGCGACCGGATAGCAAGACTCTGACAAAGCC CAAGAAATCCACAGCGGCGGCAGCAGCAGGAGGAGG AGCGCTGCGTCTGGCGCCCAACGAACCCGTATCGAC CCGCGAGCTTAGAAACAGGATTTTTCCCACTCTGTAT GCTATATTTCAACAGAGCAGGGGCCAAGAACAAGAGC TGAAAATAAAAAACAGGTCTCTGCGATCCCTCACCCG CAGCTGCCTGTATCACAAAAGCGAAGATCAGCTTCGG CGCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAAT ACTGCGCGCTGACTCTTAAGGACTAGaataaaagatctttattt tcattagatctgtgtgttggttttttgtgtgGCggccgccgtttgtgttatgtttcaacgt gtttatttttcaattgcagaaaatttcaagtcatttttcattcagtagtatagccccacc accacatagcttatacagatcaccgtaccttaatcaaactcacagaaccctagt attcaacctgccacctccctcccaacacacagagtacacagtcctttctccccgg ctggccttaaaaagcatcatatcatgggtaacagacatattcttaggtgttatattc cacacggtttcctgtcgagccaaacgctcatcagtgatattaataaactccccgg gcagctcacttaagttcatgtcgctgtccagctgctgagccacaggctgctgtcc aacttgcggttgcttaacgggcggcgaaggagaagtccacgcctacatgggg gtagagtcataatcgtgcatcaggatagggcggtggtgctgcagcagcgcgcg aataaactgctgccgccgccgctccgtcctgcaggaatacaacatggcagtgg tctcctcagcgatgattcgcaccgcccgcagcataaggcgccttgtcctccggg cacagcagcgcaccctgatctcacttaaatcagcacagtaactgcagcacag caccacaatattgttcaaaatcccacagtgcaaggcgctgtatccaaagctcat ggcggggaccacagaacccacgtggccatcataccacaagcgcaggtagat taagtggcgacccctcataaacacgctggacataaacattacctcttttggcatgt tgtaattcaccacctcccggtaccatataaacctctgattaaacatggcgccatc caccaccatcctaaaccagctggccaaaacctgcccgccggctatacactgc agggaaccgggactggaacaatgacagtggagagcccaggactcgtaacc atggatcatcatgctcgtcatgatatcaatgttggcacaacacaggcacacgtg catacacttcctcaggattacaagctcctcccgcgttagaaccatatcccaggg aacaacccattcctgaatcagcgtaaatcccacactgcagggaagacctcgc acgtaactcacgttgtgcattgtcaaagtgttacattcgggcagcagcggatgat cctccagtatggtagcgcgggtttctgtctcaaaaggaggtagacgatccctact gtacggagtgcgccgagacaaccgagatcgtgttggtcgtagtgtcatgccaa atggaacgccggacgtagtcatatttcctgaagcaaaaccaggtgcgggcgtg acaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagtagttgta gtatatccactctctcaaagcatccaggcgccccctggcttcgggttctatgtaaa ctccttcatgcgccgctgccctgataacatccaccaccgcagaataagccaca cccagccaacctacacattcgttctgcgagtcacacacgggaggagcgggaa gagctggaagaaccatgttttttttttttattccaaaagattatccaaaacctcaaaat gaagatctattaagtgaacgcgctcccctccggtggcgtggtcaaactctacag ccaaagaacagataatggcatttgtaagatgttgcacaatggcttccaaaaggc aaacggccctcacgtccaagtggacgtaaaggctaaacccttcagggtgaatc tcctctataaacattccagcaccttcaaccatgcccaaataattctcatctcgcca

		ccttctcaatatatctctaagcaaatcccgaatattaagtccggccattgtaaaaa tctgctccagagcgccctccaccttcagcctcaagcagcgaatcatgattgcaa aaattcaggttcctcacagacctgtataagattcaaaagcggaacattaacaaa aataccgcgatcccgtaggtcccttcgcagggccagctgaacataatcgtgca ggtctgcacggaccagcgcggccacttccccgccaggaaccatgacaaaag aacccacactgattatgacacgcatactcggagctatgctaaccagcgtagcc ccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaa aaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtcatgctcat gcagataaaggcaggtaagctccggaaccaccacagaaaaagacaccattt ttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaataacaaaaa aacatttaaacattagaagcctgtcttacaacaggaaaaacaacccttataagc ataagacggactacggccatgccggcgtgaccgtaaaaaaactggtcaccgt gattaaaaagcaccaccgacagctcctcggtcatgtccggagtcataatgtaa gactcggtaaacacatcaggttgattcacatcggtcagtgctaaaaagcgacc gaaatagcccgggggaatacatacccgcaggcgtagagacaacattacagc ccccataggaggtataacaaaattaataggagagaaaaacacataaacacct gaaaaaccctcctgcctaggcaaaatagcaccctcccgctccagaacaacat acagcgcttccacagcggcagccataacagtcagccttaccagtaaaaaaga aaacctattaaaaaaacaccactcgacacggcaccagctcaatcagtcacag tgtaaaaaagggccaagtgcagagcgagtatatataggactaaaaaatgacg taacggttaaagtccacaaaaaacacccagaaaaccgcacgcgaacctacg cccagaaacgaaagccaaaaaacccacaacttcctcaaatcgtcacttccgtt ttcccacgttacgtcacttcccattttaagaaaactacaattcccaacacatacaa gttactccgcccttaattaaatcggatccgatatctagatgtattcgcgaggtacc gagctcgaattctctggccgtcgttttacaacgtcgtgactgggaaaaccctggc gttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatag cgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggc gaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgc atatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagcccc gacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggc atccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttt tcaccgtcatcaccgaaacgcgcga

115	L4 nucleotide sequence	atggcacccaaaaagaagctgcagctgccgccgccacccacggacgagga ggaatactgggacagtcaggcagaggaggttttggacgaggaggaggagga catgatggaagactgggagagcctagacgaggaagcttccgaggtcgaaga ggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgccggcgcccc agaaatcggcaaccggttccagcatggctacaacctccgctcctcaggcgccg ccggcactgcccgttcgccgacccaaccgtagatgggacaccactggaacca gggccggtaagtccaagcagccgccgccgttagcccaagagcaacaacag cgccaaggctaccgctcatggcgcgggcacaagaacgccatagttgcttgctt gcaagactgtgggggcaacatctccttcgcccgccgctttcttctctaccatcacg gcgtggccttcccccgtaacatcctgcattactaccgtcatctctacagcccatac tgcaccggcggcagcggcagcaacagcagcggccacacagaagcaaagg cgaccggatagcaagactctgacaaagcccaagaaatccacagcggcggc agcagcaggaggaggagcgctgcgtctggcgcccaacgaacccgtatcgac ccgcgagcttagaaacaggatttttcccactctgtatgctatatttcaacagagca ggggccaagaacaagagctgaaaataaaaaacaggtctctgcgatccctca cccgcagctgcctgtatcacaaaagcgaagatcagcttcggcgcacgctgga agacgcggaggctctcttcagtaaatactgcgcgctgactcttaaggactag

116	L4 Protein (33K native)	MAPKKKLQLPPPPTDEEEYWDSQAEEVLDEEEEDMME DWESLDEEASEVEEVSDETPSPSVAFPSPAPQKSATGS SMATTSAPQAPPALPVRRPNRRWDTTGTRAAHTAPAAA AATAAATQKQRRPDSKTLTKPKKSTAAAAAGGGALRLA PNEPVSTRELRNRIFPTLYAIFQQSRGQEQELKIKNRSLR SLTRSCLYHKSEDQLRRTLEDAEALFSKYCALTLKD*
	L4 Protein (22K native)	MAPKKKLQLPPPPTDEEEYWDSQAEEVLDEEEEDMME DWESLDEEASEVEEVSDETPSPSVAFPSPAPQKSATGS SMATTSAPQAPPALPVRRPNRRWDTTGTRAGKSKQPP PLAQEQQQRQGYRSWRGHKNAIVACLQDCGGNISFAR RFLLYHHGVAFPRNILHYYRHLYSPYCTGGSGSNSSGHT EAKATG*

117	L4-33K (E1)	atggcacccaaaaagaagctgcagctgccgccgccacccacggacgagga ggaatactgggacagtcaggcagaggaggttttggacgaggaggaggagga catgatggaagactgggagagcctagacgaggaagcttccgaggtcgaaga ggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgccggcgcccc agaaatcggcaaccggttccagcatggctacaacctccgctcctcaggcgccg ccggcactgcccgttcgccgacccaaccgtagatgggacaccactggaacca gggccg

118	L4 33K (E1) Protein	MAPKKKLQLPPPPTDEEEYWDSQAEEVLDEEEEDMME DWESLDEEASEVEEVSDETPSPSVAFPSPAPQKSATGS SMATTSAPQAPPALPVRRPNRRWDTTGTRA

119	L4 33K optimized	atggcccccaagaagaagctgcagctgccccccccccccaccgacgagga ggaatactgggacagccaggccgaggaggtgctggacgaggaggaggag gacatgatggaagactgggagagcctggacgaggaagccagcgaggtgga ggaggtgagcgacgagacccccagcccctccgtggccttccccagccccgcc ccccagaagagcgccaccggctccagcatggccaccacctccgcccccag gccccccccgccctgcccgtgagaagacccaacaggagATGGGACAC CACCGGCACCAGGGCCGCCCACACCGCCCCCGCCG CAGCTGCAGCAACAGCAGCTGCCACCCAGAAGCAGA GGAGACCTGATAGCAAGACCCTGACCAAGCCCAAGAA AAGCACCGCCGCCGCTGCCGCAGGGGGCGGAGCCC TGAGACTGGCCCCCAACGAGCCCGTGAGCACCAGGG AGCTGAGAAACAGGATTTTTCCCACCCTGTACGCCAT CTTTCAGCAGAGCAGGGGCCAGGAGCAGGAGCTGAA GATCAAGAACAGGAGCCTGAGAAGCCTGACCAGAAG CTGCCTGTACCACAAGAGCGAGGACCAGCTGAGAAG GACACTGGAAGATGCCGAGGCTCTGTTTAGCAAGTAT TGCGCCCTGACCCTGAAGGACTA

120	L4 33K	MAPKKKLQLPPPPTDEEEYWDSQAEEVLDEEEEDMME DWESLDEEASEVEEVSDETPSPSVAFPSPAPQKSATGS SMATTSAPQAPPALPVRRPNRRWDTTGTRAAHTAPAAA AATAAATQKQRRPDSKTLTKPKKSTAAAAAGGGALRLA PNEPVSTRELRNRIFPTLYAIFQQSRGQEQELKIKNRSLR SLTRSCLYHKSEDQLRRTLEDAEALFSKYCALTLKD*

121	L4 Insert_pUC57	tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggaga cggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcaggg cgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcaga gcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgt aaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgtt gggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagg gggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacg acgttgtaaaacgacggccagtgaattcgagctcggtacctcgcgaatgcatct agatatttaaatgattaggttctacgaagaccaatcccgcccgcctaatgcgga gcttaccgcctgcgtcattacccagggccacattcttggccaattgcaagccatc aacaaagcccgccaagagtttctgctacgaaagggacgggggtttacttgga cccccagtccggcgaggagctcaacccaatccccccgccgccgcagccctat cagcagcagccgcgggcccttgcttcccaggatggcacccaaaaagaagct gcagctgccgccgccacccacggacgaggaggaatactgggacagtcagg cagaggaggttttggacgaggaggaggaggacatgatggaagactgggag agcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaacaccg tcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccggttcc agcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcgccg acccaaccgtagatgggacaccactggaaccagggccggtaagtccaagca gccgccgccgttagcccaagagcaacaacagcgccaaggctaccgctcatg gcgcgggcacaagaacgccatagttgcttgcttgcaagactgtgggggcaac atctccttcgcccgccgctttcttctctaccatcacggcgtggccttcccccgtaac atcctgcattactaccgtcatctctacagcccatactgcaccggcggcagcggc agcaacagcagcggccacacagaagcaaaggcgaccggatagcaagact ctgacaaagcccaagaaatccacagcggcggcagcagcaggaggaggag cgctgcgtctggcgcccaacgaacccgtatcgacccgcgagcttagaaacag gatttttcccactctgtatgctatatttcaacagagcaggggccaagaacaagag ctgaaaataaaaaacaggtctctgcgatccctcacccgcagctgcctgtatcac aaaagcgaagatcagcttcggcgcacgctggaagacgcggaggctctcttca gtaaatactgcgcgctgactcttaaggactagaataaaagatctttattttcattag atctgtgtgttggtttttgtgtggcggccgcatcggatcccgggcccgtcgactgc agaggcctgcatgcaagcttggcgtaatcatggtcatagctgtttcctgtgtgaaa ttgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaa agcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactg cccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaa cgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactg actcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaagg cggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtg agcaaaaggccagcaaaaggccaggaaccgtaaaaggccgcgttgctgg cgtttttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaa gtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccct ggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtc cgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatct cagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgtt cagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaa gacacgacttatcgccactggcagcagccactggtaacaggattagcagagc gaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac

		actagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaa aaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtt tttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatc ctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaaggg attttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaat gaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaat gcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgc ctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccc cagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcag caataaaccagccagccggaagggccgagcgcagaagtggtcctgcaacttt atccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgc cagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgct cgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacat gatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtc agaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattc tcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaa gtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaata cgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaa acgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcg atgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgttt ctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataag ggcgacacggaaatgttgaatactcatactcttccttttcaatattattgaagcattt atcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaac aaatagggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaa accattattatcatgacattaacctataaaaataggcgtatcacgaggcccttcg tc

122	33K opt insert_pUC57	tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggaga cggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcaggg cgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcaga gcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgt aaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgtt gggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagg gggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacg acgttgtaaaacgacggccagtgaattcgagctcggtacctcgcgaatgcatct agatATTTAAATgattaggttctacgaagaccaatcccgcccgcctaatgc ggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaagc catcaacaaagcccgccaagagtttctgctacgaaagggacgggggtttact tggacccccagtccggcgaggagctcaacccaatccccccgccgccgcagc cctatcagcagcagccgcgggcccttgcttcccaggatggcccccaagaaga agctgcagctgccccccccccccaccgacgaggaggaatactgggacagcc aggccgaggaggtgctggacgaggaggaggaggacatgatggaagactgg gagagcctggacgaggaagccagcgaggtggaggaggtgagcgacgaga cccccagcccctccgtggccttccccagccccgccccccagaagagcgccac cggctccagcatggccaccacctccgccccccaggccccccccgccctgccc gtgagaagacccaacaggagATGGGACACCACCGGCACCAG GGCCGCCCACACCGCCCCCGCCGCAGCTGCAGCAAC AGCAGCTGCCACCCAGAAGCAGAGGAGACCTGATAG CAAGACCCTGACCAAGCCCAAGAAAAGCACCGCCGC CGCTGCCGCAGGGGGCGGAGCCCTGAGACTGGCCC CCAACGAGCCCGTGAGCACCAGGGAGCTGAGAAACA GGATTTTTCCCACCCTGTACGCCATCTTTCAGCAGAG CAGGGGCCAGGAGCAGGAGCTGAAGATCAAGAACAG GAGCCTGAGAAGCCTGACCAGAAGCTGCCTGTACCA CAAGAGCGAGGACCAGCTGAGAAGGACACTGGAAGA TGCCGAGGCTCTGTTTAGCAAGTATTGCGCCCTGACC CTGAAGGACTAGaataaaagatctttattttcattagatctgtgtgttggttttt tgtgtgGCGGCCGCatcggatcccgggcccgtcgactgcagaggcctgc atgcaagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgct cacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggt gcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttcca gtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcgggg agaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcg ctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacg gttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaagg ccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccata ggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtg gcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctcc ctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcc cttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgt aggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgacc gctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgactt atcgccactggcagcagccactggtaacaggattagcagagcgaggtatgta ggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaga

		acagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttgg tagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaa gcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttct acggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatg agattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaa tcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagt gaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactcccc gtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgca atgataccgcgagacccacgctcaccggctccagatttatcagcaataaacca gccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctcc atccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatag tttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggt atggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccca tgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaag ttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtca tgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgag aatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataata ccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcgg ggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccac tcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagca aaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacgg aaatgttgaatactcatactcttccttttcaatattattgaagcatttatcagggttatt gtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggtt ccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatc atgacattaacctataaaaataggcgtatcacgaggccctttcgtc

123	22K opt insert_pUC57	tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggaga cggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcaggg cgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcaga gcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgt aaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgtt gggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagg gggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacg acgttgtaaaacgacggccagtgaattcgagctcggtacctcgcgaatgcatct agatATTTAAATgattaggttctacgaagaccaatcccgcccgcctaatgc ggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaagc catcaacaaagcccgccaagagtttctgctacgaaagggacgggggtttact tggacccccagtccggcgaggagctcaacccaatccccccgccgccgcagc cctatcagcagcagccgcgggcccttgcttcccaggatggcccccaagaaga agctgcagctgccgcccccacccaccgacgaggaggaatactgggacagcc aggccgaggaggtgctggacgaggaggaggaggacatgatggaagactgg gagagcctggacgaggaagcctccgaggtggaggaggtgagcgacgaaac acccagccccagcgtggccttccccagccccgccccccagaagagcgccac cggcagcagcatggccaccaccagcgccccccaggccccccccgccctgcc cgtgagaaggcccaacagaagATGGGACACCACAGGCACCAG GGCCGGCAAGAGCAAGCAGCCCCCTCCACTGGCCCA GGAGCAACAGCAGAGACAGGGCTACAGATCCTGGAG AGGACACAAGAACGCCATTGTGGCCTGCCTGCAGGA CTGCGGCGGCAACATTAGCTTCGCCAGAAGGTTTCTG CTCTATCACCACGGCGTGGCCTTCCCCAGAAACATCC TGCATTACTACAGACATCTCTACAGCCCTTACTGCACC GGCGGAAGCGGCAGCAACAGCAGCGGCCACACCGA GGCCAAGGCCACCGGCTGAaataaaagatctttattttcattagatct gtgtgttggtttttgtgtgGCGGCCGCatcggatcccgggcccgtcgactgc agaggcctgcatgcaagcttggcgtaatcatggtcatagctgtttcctgtgtgaaa ttgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaa agcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactg cccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaa cgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactg actcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaagg cggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtg agcaaaaggccagcaaaaggccaggaaccgtaaaaggccgcgttgctgg cgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaa gtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccct ggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtc cgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatct cagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgtt cagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaa gacacgacttatcgccactggcagcagccactggtaacaggattagcagagc gaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac actagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaa aaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtt tttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatc

		ctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaaggg attttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaat gaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaat gcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgc ctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccc cagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcag caataaaccagccagccggaagggccgagcgcagaagtggtcctgcaacttt atccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgc cagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgct cgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacat gatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtc agaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattc tcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaa gtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaata cgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaa acgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcg atgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgttt ctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataag ggcgacacggaaatgttgaatactcatactcttccttttcaatattattgaagcattt atcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaac aaatagggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaa accattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcg tc

124	shortsynL4insert_pUC57	tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggaga cggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcaggg cgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcaga gcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgt aaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgtt gggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagg gggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacg acgttgtaaaacgacggccagtgaattcgagctcggtacctcgcgaatgcatct agatATTTAAATgattaggttctacgaagaccaatcccgcccgcctaatgc ggagcttaccggctccggtgcccgtcagtgggcagagcgcacatcgcccaca gtccccgagaagttgggggggaggggtcggcaattgaaccggtgcctagagaa ggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttcc cgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttcttttt cgcaacgggtttgccgccagaacacagccgccgcagccctatcagcagcag ccgcgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgccg ccgccacccacggacgaggaggaatactgggacagtcaggcagaggaggt tttggacgaggaggaggaggacatgatggaagactgggagagcctagacga ggaagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcggtc gcattcccctcgccggcgccccagaaatcggcaaccggttccagcatggctac aacctccgctcctcaggcgccgccggcactgcccgttcgccgacccaaccgta gATGGGACACCACTGGAACCAGGGCCGGTAAGTCCAA GCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGCG CCAAGGCTACCGCTCATGGCGCGGGCACAAGAACGC CATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACATC TCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGCG TGGCCTTCCCCCGTAACATCCTGCATTACTACCGTCAT CTCTACAGCCCATACTGCACCGGCGGCAGCGGCAGC AACAGCAGCGGCCACACAGAAGCAAAGGCGACCGGA TAGCAAGACTCTGACAAAGCCCAAGAAATCCACAGCG GCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTGGC GCCCAACGAACCCGTATCGACCCGCGAGCTTAGAAAC AGGATTTTTCCCACTCTGTATGCTATATTTCAACAGAG CAGGGGCCAAGAACAAGAGCTGAAAATAAAAAACAGG TCTCTGCGATCCCTCACCCGCAGCTGCCTGTATCACA AAAGCGAAGATCAGCTTCGGCGCACGCTGGAAGACG CGGAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCT TAAGGACTAGaataaaagatctttattttcattagatctgtgtgttggtttttgt gtgGCGGCCGCatcggatcccgggcccgtcgactgcagaggcctgcatg caagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcac aattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcc taatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcg ggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagag gcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggt cgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatc cacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagc aaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcga

		aacccgacaggactataaagataccaggcgtttccccctggaagctccctcgt gcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcg ggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggt cgttcgctccaagctgggctgtgtgcacgaacccccccgttcagcccgaccgctg cgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcg ccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg gtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacag tatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagct cttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcag cagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacgg ggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagatt atcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaat ctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgagg cacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtg tagataactacgatacgggagggcttaccatctggccccagtgctgcaatgata ccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccag ccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccag tctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgc aacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggctt cattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtg caaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggcc gcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccat ccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagt gtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgc cacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaa aactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgca cccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaaca ggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttg aatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatg agcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgc acatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacatt aacctataaaaataggcgtatcacgaggccctttcgtc

125	pEMBR1.4C MVe-33F	TAAggatccGAATTCAGTACTCATATGTTGACAATTAATC ATCGGCATAGTATATCGGCATAGTATAATACGACTCAC TATAGGAGGGCCACCAAGGTGAGGAACTAAACCATGA GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAA TTCCAACATGGATGCTGATTTATATGGGTATAAATGGG CTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTA TCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTT CTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTA CAGATGAGATGGTCAGACTAAACTGGCTGACGGAATT TATGCCACTTCCGACCATCAAGCATTTTATCCGTACTC CTGATGATGCATGGTTACTCACCACTGCGATCCCCGG AAAAACAGCGTTCCAGGTATTAGAAGAATATCCTGATT CAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCT GCGCCGGTTGCACTCGATTCCTGTTTGTAATTGTCCTT TTAACAGCGATCGCGTATTTCGCCTCGCTCAGGCGCA ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATT TTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGT CTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGG ATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAAC CTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGAT CTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTC CTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATT GATAATCCTGATATGAATAAATTGCAGTTTCATTTGAT GCTCGATGAGTTTTTCTAATAAAAACCCGCTTCGGCG GGTTTTTTTATGCATGTGAGCAAAAGGCCAGCAAAAG GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGtcga actgacattgattattgactagttattaatagtaatcaattacggggtcattagttcat agcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggct

		gaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatag taacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaact gcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacg tcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatggg actttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcg gttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaa gtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacggga ctttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcg tgtacggtgggaggtctatataagcagagctggcgaataacgcgttctttttcgca acgggtttgccgccagaacacagccgccgcagccctatcagcagcagccgc gggcccttgcttcccaggatggcccccaagaagaagctgcagctgcccccccc cccccaccgacgaggaggaatactgggacagccaggccgaggaggtgctg gacgaggaggaggaggacatgatggaagactgggagagcctggacgagg aagccagcgaggtggaggaggtgagcgacgagacccccagcccctccgtg gccttccccagccccgccccccagaagagcgccaccggctccagcatggcc accacctccgccccccaggccccccccgccctgcccgtgagaagacccaac aggagATGGGACACCACCGGCACCAGGGCCGCCCACA CCGCCCCCGCCGCAGCTGCAGCAACAGCAGCTGCCA CCCAGAAGCAGAGGAGACCTGATAGCAAGACCCTGA CCAAGCCCAAGAAAAGCACCGCCGCCGCTGCCGCAG GGGGCGGAGCCCTGAGACTGGCCCCCAACGAGCCC GTGAGCACCAGGGAGCTGAGAAACAGGATTTTTCCCA CCCTGTACGCCATCTTTCAGCAGAGCAGGGGCCAGG AGCAGGAGCTGAAGATCAAGAACAGGAGCCTGAGAA GCCTGACCAGAAGCTGCCTGTACCACAAGAGCGAGG ACCAGCTGAGAAGGACACTGGAAGATGCCGAGGCTC TGTTTAGCAAGTATTGCGCCCTGACCCTGAAGGACTA GaataaaagatctttattttcattagatctgtgtgttggtttttgtgtgGCGGCCt cgacGTTcctatccgtagatgtacctggacatccaggtgatgccggcggcgg tggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgcagcgg caaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgcagtcgt tgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgtgg tctggtggataaattcgcaagggtatcatggcggacgaccggggttcgaaccc cggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaaccc aggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttccaggcg cggcggctgctgcgctagctttttttggccactggccgcgcgcggcgtaagcggtt aggctggaaagcgaaagcattaagtggctcgctccctgtagccggagggttatt ttccaagggttgagtcgcaggacccccggttcgagtctcgggccggccggact gcggcgaacgggggtttgcctccccgtcatgcaagaccccgcttgcaaattcct ccggaaacagggacgagcccctttttgcttttcccagatgcatccggtgctgcg gcagatgcgcccccctcctcagcagcggcaagagcaagagcagcggcaga catgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacatcgatc cagacatgataagatacattgatgagtttggacaaaccacaactagaatgcag tgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataa gctgcaataaacaagtttgtacactctcgggtgattatttacccccacccttgccgt ctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgccactgg cagggacacgttgcgatactggtgtttagtgctccacttaaactcaggcacaacc

		atccgcggcagctcggtgaagttttcactccacaggctgcgcaccatcaccaac gcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgccct gcgcgcgcgagttgcgatacacagggttgcagcactggaacactatcagcgc cgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcgtcca ggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgccttcccaa aaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatcaaaa ggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaagccttg atctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgccgcaa gacttgccggaaaactgattggccggacaggccgcgtcgtgcacgcagcacc ttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttcacgatct tggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcacatcca tttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagctcgcctt cgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtgatg cttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcgccccat catcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgctcc tcgttcagccaggtcttgcatacggccgccagagcttccacttggtcaggcagta gtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagcgcgcgcg cagcctccatgcccttctcccacgcagacacgatcggcacactcagcgggttc atcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcgtccgcata ccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgcttacctccttt gccatgcttgattagcaccggtgggttgctgaaacccaccatttgtagcgccaca tcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgctcgggctt gggagaagggcgcttctttttcttcttgggcgcaatggccaaatccgccgccgag gtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtgatgagtct tcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcgcccgggg aggcggcggcgacggggacggggacgacacgtcctccatggttgggggac gtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcctcttcccg actggccatttccttctcctataggcagaaaaagatccacaaaagcgaagatc agcttcggcgcacgctggaagacgcggaggctctcttcagtaaatactgcgcg ctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaaactacgt catctccagcggccacacccggcgccagcacctgttgtcagcgccattggcgc gcccgcccgccgcgcgcttcgcttttttatagggccgccgccgccgccgcctcgc cataaaaggaaactttcggagcgcgccgctctgattggctgccgccgcacctct ccgcctcgccccgccccgcccctcgccccgccccgccccgcctggcgcgcgc ccccccccccccccccgcccccatcgctgcacaaaataattaaaaaataaata aatacaaaattgggggtggggaggggggggagatggggagagtgaagcag aacgtggggctcacctcgaggccggccgaatatcttcatttaaatgtttaaacat cgatgcGGCCGCgattaggttctacgaagaccaatcccgcccgcctaatg cggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaag ccatcaacaaagcccgccaagagtttctgctacgaaagggacgggggggttta cttggacccccagtccggcgaggagctcaacccaatccccccgccgccgcag ccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaag aagctgcagctgccgccgccacccacggacgaggaggaatactgggacagt caggcagaggaggttttggacgaggaggaggaggacatgatggaagactgg gagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaaca ccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccgg ttccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcg

		ccgacccaaccgtagATGGGACACCACTGGAACCAGGGCC GGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAG CAACAACAGCGCCAAGGCTACCGCTCATGGCGCGGG CACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTG GGGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTA CCATCACGGCGTGGCCTTCCCCCGTAACATCCTGCAT TACTACCGTCATCTCTACAGCCCATACTGCACCGGCG GCAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAA AGGCGACCGGATAGCAAGACTCTGACAAAGCCCAAG AAATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGaataaaagatctttattttcattaga tctgtgtgttggttttttgtgtgGCggccgccgtttgtgttatgtttcaacgtgtttatttttt caattgcagaaaatttcaagtcatttttcattcagtagtatagccccaccaccacat agcttatacagatcaccgtaccttaatcaaactcacagaaccctagtattcaacc tgccacctccctcccaacacacagagtacacagtcctttctccccggctggcctt aaaaagcatcatatcatgggtaacagacatattcttaggtgttatattccacacgg tttcctgtcgagccaaacgctcatcagtgatattaataaactccccgggcagctc acttaagttcatgtcgctgtccagctgctgagccacaggctgctgtccaacttgcg gttgcttaacgggcggcgaaggagaagtccacgcctacatgggggtagagtc ataatcgtgcatcaggatagggcggtggtgctgcagcagcgcgcgaataaact gctgccgccgccgctccgtcctgcaggaatacaacatggcagtggtctcctcag cgatgattcgcaccgcccgcagcataaggcgccttgtcctccgggcacagcag cgcaccctgatctcacttaaatcagcacagtaactgcagcacagcaccacaat attgttcaaaatcccacagtgcaaggcgctgtatccaaagctcatggcgggga ccacagaacccacgtggccatcataccacaagcgcaggtagattaagtggcg acccctcataaacacgctggacataaacattacctcttttggcatgttgtaattcac cacctcccggtaccatataaacctctgattaaacatggcgccatccaccaccat cctaaaccagctggccaaaacctgcccgccggctatacactgcagggaaccg ggactggaacaatgacagtggagagcccaggactcgtaaccatggatcatca tgctcgtcatgatatcaatgttggcacaacacaggcacacgtgcatacacttcct caggattacaagctcctcccgcgttagaaccatatcccagggaacaacccatt cctgaatcagcgtaaatcccacactgcagggaagacctcgcacgtaactcac gttgtgcattgtcaaagtgttacattcgggcagcagcggatgatcctccagtatgg tagcgcgggtttctgtctcaaaaggaggtagacgatccctactgtacggagtgc gccgagacaaccgagatcgtgttggtcgtagtgtcatgccaaatggaacgccg gacgtagtcatatttcctgaagcaaaaccaggtgcgggcgtgacaaacagatc tgcgtctccggtctcgccgcttagatcgctctgtgtagtagttgtagtatatccactct ctcaaagcatccaggcgccccctggcttcgggttctatgtaaactccttcatgcgc cgctgccctgataacatccaccaccgcagaataagccacacccagccaacct acacattcgttctgcgagtcacacacgggaggagcgggaagagctggaaga accatgttttttttttttattccaaaagattatccaaaacctcaaaatgaagatctatta

		agtgaacgcgctcccctccggtggcgtggtcaaactctacagccaaagaaca gataatggcatttgtaagatgttgcacaatggcttccaaaaggcaaacggccct cacgtccaagtggacgtaaaggctaaacccttcagggtgaatctcctctataaa cattccagcaccttcaaccatgcccaaataattctcatctcgccaccttctcaatat atctctaagcaaatcccgaatattaagtccggccattgtaaaaatctgctccaga gcgccctccaccttcagcctcaagcagcgaatcatgattgcaaaaattcaggtt cctcacagacctgtataagattcaaaagcggaacattaacaaaaataccgcg atcccgtaggtcccttcgcagggccagctgaacataatcgtgcaggtctgcacg gaccagcgcggccacttccccgccaggaaccatgacaaaagaacccacact gattatgacacgcatactcggagctatgctaaccagcgtagccccgatgtaagc ttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaaaaatcaggca aagcctcgcgcaaaaaagaaagcacatcgtagtcatgctcatgcagataaag gcaggtaagctccggaaccaccacagaaaaagacaccatttttctctcaaaca tgtctgcgggtttctgcataaacacaaaataaaataacaaaaaaacatttaaac attagaagcctgtcttacaacaggaaaaacaacccttataagcataagacgga ctacggccatgccggcgtgaccgtaaaaaaactggtcaccgtgattaaaaag caccaccgacagctcctcggtcatgtccggagtcataatgtaagactcggtaaa cacatcaggttgattcacatcggtcagtgctaaaaagcgaccgaaatagcccg ggggaatacatacccgcaggcgtagagacaacattacagcccccataggag gtataacaaaattaataggagagaaaaacacataaacacctgaaaaaccctc ctgcctaggcaaaatagcaccctcccgctccagaacaacatacagcgcttcca cagcggcagccataacagtcagccttaccagtaaaaaagaaaacctattaaa aaaacaccactcgacacggcaccagctcaatcagtcacagtgtaaaaaagg gccaagtgcagagcgagtatatataggactaaaaaatgacgtaacggttaaa gtccacaaaaaacacccagaaaaccgcacgcgaacctacgcccagaaac gaaagccaaaaaacccacaacttcctcaaatcgtcacttccgttttcccacgtta cgtcacttcccattttaagaaaactacaattcccaacacatacaagttactccgc ccttaat

126	pEMBR1.4C MVe-33R	TAAggatccGAATTCAGTACTCATATGTTGACAATTAATC ATCGGCATAGTATATCGGCATAGTATAATACGACTCAC TATAGGAGGGCCACCAAGGTGAGGAACTAAACCATGA GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAA TTCCAACATGGATGCTGATTTATATGGGTATAAATGGG CTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTA TCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTT CTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTA CAGATGAGATGGTCAGACTAAACTGGCTGACGGAATT TATGCCACTTCCGACCATCAAGCATTTTATCCGTACTC CTGATGATGCATGGTTACTCACCACTGCGATCCCCGG AAAAACAGCGTTCCAGGTATTAGAAGAATATCCTGATT CAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCT GCGCCGGTTGCACTCGATTCCTGTTTGTAATTGTCCTT TTAACAGCGATCGCGTATTTCGCCTCGCTCAGGCGCA ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATT TTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGT CTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGG ATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAAC CTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGAT CTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTC CTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATT GATAATCCTGATATGAATAAATTGCAGTTTCATTTGAT GCTCGATGAGTTTTTCTAATAAAAACCCGCTTCGGCG GGTTTTTTTATGCATGTGAGCAAAAGGCCAGCAAAAG GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGtcga GGCCGCcacacaaaaaaccaacacacagatctaatgaaaataaagatc ttttattCTAGTCCTTCAGGGTCAGGGCGCAATACTTGCTA

		AACAGAGCCTCGGCATCTTCCAGTGTCCTTCTCAGCT GGTCCTCGCTCTTGTGGTACAGGCAGCTTCTGGTCAG GCTTCTCAGGCTCCTGTTCTTGATCTTCAGCTCCTGCT CCTGGCCCCTGCTCTGCTGAAAGATGGCGTACAGGG TGGGAAAAATCCTGTTTCTCAGCTCCCTGGTGCTCAC GGGCTCGTTGGGGGCCAGTCTCAGGGCTCCGCCCCC TGCGGCAGCGGCGGCGGTGCTTTTCTTGGGCTTGGT CAGGGTCTTGCTATCAGGTCTCCTCTGCTTCTGGGTG GCAGCTGCTGTTGCTGCAGCTGCGGCGGGGGCGGTG TGGGCGGCCCTGGTGCCGGTGGTGTCCCATctcctgttgg gtcttctcacgggcagggcgggggggcctgggggggcggaggtggtggccat gctggagccggtggcgctcttctgggggggcggggctggggaaggccacgga ggggctgggggtctcgtcgctcacctcctccacctcgctggcttcctcgtccagg ctctcccagtcttccatcatgtcctcctcctcctcgtccagcacctcctcggcctgg ctgtcccagtattcctcctcgtcggtgggggggggggggcagctgcagcttcttctt gggggccatcctgggaagcaagggcccgcggctgctgctgatagggctgcg gcggctgtgttctggcggcaaacccgttgcgaaaaagaacgcgttattcgcca gctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatgg ggcggagttgttacgacattttggaaagtcccgttgattttggtgccaaaacaaac tcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgct atccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatga ctaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactggg cataatgccaggcgggccatttaccgtcattgacgtcaatgggggcgtacttg gcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatactccac ccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcatta ttgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttacc gtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaa ttgattactattaataactagtcaataatcaatgtcagttcgacGTTcctatccgta gatgtacctggacatccaggtgatgccggcggcggtggtggaggcgcgcgga aagtcgcggacgcggttccagatgttgcgcagcggcaaaaagtgctccatggt cgggacgctctggccggtgaggcgtgcgcagtcgttgacgctctagaccgtgc aaaaggagagcctgtaagcgggcactcttccgtggtctggtggataaattcgca agggtatcatggcggacgaccggggttcgaaccccggatccggccgtccgcc gtgatccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtcagac aacgggggagcgctccttttggcttccttccaggcgcggcggctgctgcgctag ctttttttggccactggccgcgcgcggcgtaagcggttaggctggaaagcgaaa gcattaagtggctcgctccctgtagccggagggttattttccaagggttgagtcgc aggacccccggttcgagtctcgggccggccggactgcggcgaacgggggttt gcctccccgtcatgcaagaccccgcttgcaaattcctccggaaacagggacga gccccttttttgcttttcccagatgcatccggtgctgcggcagatgcgcccccctcc tcagcagcggcaagagcaagagcagcggcagacatgcagggcaccctccc cttctcctaccgcgtcaggaggggcaacatcgatccagacatgataagatacat tgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgt gaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagtttg tacactctcgggtgattatttacccccaccttgccgtctgcgccgtttaaaaatca aaggggttctgccgcgcatcgctatgcgccactggcagggacacgttgcgata ctggtgtttagtgctccacttaaactcaggcacaaccatccgcggcagctcggtg

		aagttttcactccacaggctgcgcaccatcaccaacgcgtttagcaggtcgggc gccgatatcttgaagtcgcagttggggcctccgccctgcgcgcgcgagttgcga tacacagggttgcagcactggaacactatcagcgccgggtggtgcacgctggc cagcacgctcttgtcggagatcagatccgcgtccaggtcctccgcgttgctcag ggcgaacggagtcaactttggtagctgccttcccaaaaagggcgcgtgcccag gctttgagttgcactcgcaccgtagtggcatcaaaaggtgaccgtgcccggtctg ggcgttaggatacagcgcctgcataaaagccttgatctgcttaaaagccacctg agcctttgcgccttcagagaagaacatgccgcaagacttgccggaaaactgat tggccggacaggccgcgtcgtgcacgcagcaccttgcgtcggtgttggagatct gcaccacatttcggccccaccggttcttcacgatcttggccttgctagactgctcct tcagcgcgcgctgcccgttttcgctcgtcacatccatttcaatcacgtgctccttattt atcataatgcttccgtgtagacacttaagctcgccttcgatctcagcgcagcggtg cagccacaacgcgcagcccgtgggctcgtgatgcttgtaggtcacctctgcaa acgactgcaggtacgcctgcaggaatcgccccatcatcgtcacaaaggtcttgt tgctggtgaaggtcagctgcaacccgcggtgctcctcgttcagccaggtcttgca tacggccgccagagcttccacttggtcaggcagtagtttgaagttcgcctttagat cgttatccacgtggtacttgtccatcagcgcgcgcgcagcctccatgcccttctcc cacgcagacacgatcggcacactcagcgggttcatcaccgtaatttcactttcc gcttcgctgggctcttcctcttcctcttgcgtccgcataccacgcgccactgggtcg tcttcattcagccgccgcactgtgcgcttacctcctttgccatgcttgattagcaccg gtgggttgctgaaacccaccatttgtagcgccacatcttctctttcttcctcgctgtcc acgattacctctggtgatggcgggcgctcgggcttgggagaagggcgcttcttttt cttcttgggcgcaatggccaaatccgccgccgaggtcgatggccgcgggctgg gtgtgcgcggcaccagcgcgtcttgtgatgagtcttcctcgtcctcggactcgata cgccgcctcatccgctttttgggggcgcccggggaggcggcggcgacgggg acggggacgacacgtcctccatggttgggggacgtcgcgccgcaccgcgtcc gcgctcgggggtggtttcgcgctgctcctcttcccgactggccatttccttctcctat aggcagaaaaagatccacaaaagcgaagatcagcttcggcgcacgctgga agacgcggaggctctcttcagtaaatactgcgcgctgactcttaaggactagttt cgcgcccttctcaaatttaagcgcgaaaactacgtcatctccagcggccacac ccggcgccagcacctgttgtcagcgccattggcgcgcccgcccgccgcgcgct tcgctttttatagggccgccgccgccgcctcgccataaaaggaaactttcg gagcgcgccgctctgattggctgccgccgcacctctccgcctcgccccgcccc gcccctcgccccgccccgccccgcctggcgcgcgccccccccccccccg cccccatcgctgcacaaaataattaaaaaataaataaatacaaaattgggggt ggggaggggggggagatggggagagtgaagcagaacgtggggctcacctc gaggccggccgaatatcttcatttaaatgtttaaacatcgatgcGGCCGCga ttaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgcctgcg tcattacccagggccacattcttggccaattgcaagccatcaacaaagcccgcc aagagtttctgctacgaaagggacgggggtttacttggacccccagtccggc gaggagctcaacccaatccccccgccgccgcagccctatcagcagcagccg cgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgccgccg ccacccacggacgaggaggaatactgggacagtcaggcagaggaggttttg gacgaggaggaggaggacatgatggaagactgggagagcctagacgagg aagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgc attcccctcgccggcgccccagaaatcggcaaccggttccagcatggctacaa cctccgctcctcaggcgccgccggcactgcccgttcgccgacccaaccgtag

		ATGGGACACCACTGGAACCAGGGCCGGTAAGTCCAA GCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGCG CCAAGGCTACCGCTCATGGCGCGGGCACAAGAACGC CATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACATC TCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGCG TGGCCTTCCCCCGTAACATCCTGCATTACTACCGTCAT CTCTACAGCCCATACTGCACCGGCGGCAGCGGCAGC AACAGCAGCGGCCACACAGAAGCAAAGGCGACCGGA TAGCAAGACTCTGACAAAGCCCAAGAAATCCACAGCG GCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTGGC GCCCAACGAACCCGTATCGACCCGCGAGCTTAGAAAC AGGATTTTTCCCACTCTGTATGCTATATTTCAACAGAG CAGGGGCCAAGAACAAGAGCTGAAAATAAAAAACAGG TCTCTGCGATCCCTCACCCGCAGCTGCCTGTATCACA AAAGCGAAGATCAGCTTCGGCGCACGCTGGAAGACG CGGAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCT TAAGGACTAGaataaaagatctttattttcattagatctgtgtgttggttttttgt gtgGCggccgccgtttgtgttatgtttcaacgtgtttatttttcaattgcagaaaattt caagtcatttttcattcagtagtatagccccaccaccacatagcttatacagatca ccgtaccttaatcaaactcacagaaccctagtattcaacctgccacctccctccc aacacacagagtacacagtcctttctccccggctggccttaaaaagcatcatat catgggtaacagacatattcttaggtgttatattccacacggtttcctgtcgagcca aacgctcatcagtgatattaataaactccccgggcagctcacttaagttcatgtcg ctgtccagctgctgagccacaggctgctgtccaacttgcggttgcttaacgggcg gcgaaggagaagtccacgcctacatgggggtagagtcataatcgtgcatcag gatagggcggtggtgctgcagcagcgcgcgaataaactgctgccgccgccgc tccgtcctgcaggaatacaacatggcagtggtctcctcagcgatgattcgcacc gcccgcagcataaggcgccttgtcctccgggcacagcagcgcaccctgatctc acttaaatcagcacagtaactgcagcacagcaccacaatattgttcaaaatccc acagtgcaaggcgctgtatccaaagctcatggcggggaccacagaacccac gtggccatcataccacaagcgcaggtagattaagtggcgacccctcataaaca cgctggacataaacattacctcttttggcatgttgtaattcaccacctcccggtacc atataaacctctgattaaacatggcgccatccaccaccatcctaaaccagctgg ccaaaacctgcccgccggctatacactgcagggaaccgggactggaacaat gacagtggagagcccaggactcgtaaccatggatcatcatgctcgtcatgatat caatgttggcacaacacaggcacacgtgcatacacttcctcaggattacaagct cctcccgcgttagaaccatatcccagggaacaacccattcctgaatcagcgta aatcccacactgcagggaagacctcgcacgtaactcacgttgtgcattgtcaaa gtgttacattcgggcagcagcggatgatcctccagtatggtagcgcgggtttctgt ctcaaaaggaggtagacgatccctactgtacggagtgcgccgagacaaccg agatcgtgttggtcgtagtgtcatgccaaatggaacgccggacgtagtcatatttc ctgaagcaaaaccaggtgcgggcgtgacaaacagatctgcgtctccggtctcg ccgcttagatcgctctgtgtagtagttgtagtatatccactctctcaaagcatccag gcgccccctggcttcgggttctatgtaaactccttcatgcgccgctgccctgataa catccaccaccgcagaataagccacacccagccaacctacacattcgttctgc gagtcacacacgggaggagcgggaagagctggaagaaccatgttttttttttttat tccaaaagattatccaaaacctcaaaatgaagatctattaagtgaacgcgctcc

		cctccggtggcgtggtcaaactctacagccaaagaacagataatggcatttgta agatgttgcacaatggcttccaaaaggcaaacggccctcacgtccaagtggac gtaaaggctaaacccttcagggtgaatctcctctataaacattccagcaccttca accatgcccaaataattctcatctcgccaccttctcaatatatctctaagcaaatcc cgaatattaagtccggccattgtaaaaatctgctccagagcgccctccaccttca gcctcaagcagcgaatcatgattgcaaaaattcaggttcctcacagacctgtat aagattcaaaagcggaacattaacaaaaataccgcgatcccgtaggtcccttc gcagggccagctgaacataatcgtgcaggtctgcacggaccagcgcggcca cttccccgccaggaaccatgacaaaagaacccacactgattatgacacgcat actcggagctatgctaaccagcgtagccccgatgtaagcttgttgcatgggcgg cgatataaaatgcaaggtgctgctcaaaaaatcaggcaaagcctcgcgcaaa aaagaaagcacatcgtagtcatgctcatgcagataaaggcaggtaagctccg gaaccaccacagaaaaagacaccatttttctctcaaacatgtctgcgggtttctg cataaacacaaaataaaataacaaaaaaacatttaaacattagaagcctgtct tacaacaggaaaaacaacccttataagcataagacggactacggccatgcc ggcgtgaccgtaaaaaaactggtcaccgtgattaaaaagcaccaccgacag ctcctcggtcatgtccggagtcataatgtaagactcggtaaacacatcaggttga ttcacatcggtcagtgctaaaaagcgaccgaaatagcccgggggaatacata cccgcaggcgtagagacaacattacagcccccataggaggtataacaaaatt aataggagagaaaaacacataaacacctgaaaaaccctcctgcctaggcaa aatagcaccctcccgctccagaacaacatacagcgcttccacagcggcagcc ataacagtcagccttaccagtaaaaaagaaaacctattaaaaaaacaccact cgacacggcaccagctcaatcagtcacagtgtaaaaaagggccaagtgcag agcgagtatatataggactaaaaaatgacgtaacggttaaagtccacaaaaa acacccagaaaaccgcacgcgaacctacgcccagaaacgaaagccaaaa aacccacaacttcctcaaatcgtcacttccgttttcccacgttacgtcacttcccatt ttaagaaaactacaattcccaacacatacaagttactccgcccttaat

127	pEMBR1.4-ef1A-33F	TAAggatccGAATTCAGTACTCATATGTTGACAATTAATC ATCGGCATAGTATATCGGCATAGTATAATACGACTCAC TATAGGAGGGCCACCAAGGTGAGGAACTAAACCATGA GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAA TTCCAACATGGATGCTGATTTATATGGGTATAAATGGG CTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTA TCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTT CTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTA CAGATGAGATGGTCAGACTAAACTGGCTGACGGAATT TATGCCACTTCCGACCATCAAGCATTTTATCCGTACTC CTGATGATGCATGGTTACTCACCACTGCGATCCCCGG AAAAACAGCGTTCCAGGTATTAGAAGAATATCCTGATT CAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCT GCGCCGGTTGCACTCGATTCCTGTTTGTAATTGTCCTT TTAACAGCGATCGCGTATTTCGCCTCGCTCAGGCGCA ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATT TTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGT CTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGG ATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAAC CTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGAT CTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTC CTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATT GATAATCCTGATATGAATAAATTGCAGTTTCATTTGAT GCTCGATGAGTTTTTCTAATAAAAACCCGCTTCGGCG GGTTTTTTTATGCATGTGAGCAAAAGGCCAGCAAAAG GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGtcga ctagatATTTAAATgattaggttctacgaagaccaatcccgcccgcctaatg cggagcttaccggctccggtgcccgtcagtgggcagagcgcacatcgcccac

		agtccccgagaagttgggggagggtcggcaattgaaccggtgcctagaga aggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttc ccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttt tcgcaacgggtttgccgccagaacacagccgccgcagccctatcagcagcag ccgcgggcccttgcttcccaggatggcccccaagaagaagctgcagctgccc cccccccccaccgacgaggaggaatactgggacagccaggccgaggaggt gctggacgaggaggaggaggacatgatggaagactgggagagcctggacg aggaagccagcgaggtggaggaggtgagcgacgagacccccagcccctcc gtggccttccccagccccgccccccagaagagcgccaccggctccagcatgg ccaccacctccgcccccccaggcccccccccgccctgcccgtgagaagaccca acaggagATGGGACACCACCGGCACCAGGGCCGCCCA CACCGCCCCCCGCCGCAGCTGCAGCAACAGCAGCTGC CACCCAGAAGCAGAGGAGACCTGATAGCAAGACCCT GACCAAGCCCAAGAAAAGCACCGCCGCCGCTGCCGC AGGGGGCGGAGCCCTGAGACTGGCCCCCAACGAGC CCGTGAGCACCAGGGAGCTGAGAAACAGGATTTTTCC CACCCTGTACGCCATCTTTCAGCAGAGCAGGGGCCA GGAGCAGGAGCTGAAGATCAAGAACAGGAGCCTGAG AAGCCTGACCAGAAGCTGCCTGTACCACAAGAGCGA GGACCAGCTGAGAAGGACACTGGAAGATGCCGAGGC TCTGTTTAGCAAGTATTGCGCCCTGACCCTGAAGGAC TAGaataaaagatctttattttcattagatctgtgtgttggttttttgtgtgGCGGC CtcgacGTTcctatccgtagatgtacctggacatccaggtgatgccggcggc ggtggtggaggcgcgcggaaagtcgcggacgcggttccagatgttgcgcagc ggcaaaaagtgctccatggtcgggacgctctggccggtgaggcgtgcgcagt cgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgt ggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcgaac cccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaac ccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttccagg cgcggcggctgctgcgctagctttttggccactggccgcgcgcggcgtaagcg gttaggctggaaagcgaaagcattaagtggctcgctccctgtagccggagggtt attttccaagggttgagtcgcaggacccccggttcgagtctcgggccggccgga ctgcggcgaacgggggtttgcctccccgtcatgcaagacccccgcttgcaaattc ctccggaaacagggacgagccccttttttgcttttcccagatgcatccggtgctgc ggcagatgcgcccccctcctcagcagcggcaagagcaagagcagcggcag acatgcagggcaccctcccccttctcctaccgcgtcaggaggggcaacatcgat ccagacatgataagatacattgatgagtttggacaaaccacaactagaatgca gtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata agctgcaataaacaagtttgtacactctcgggtgattatttacccccacccttgcc gtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcgccactg gcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggcacaa ccatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatcacca acgcgtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgc cctgcgcgcgcgagttgcgatacacagggttgcagcactggaacactatcagc gccgggtggtgcacgctggccagcacgctcttgtcggagatcagatccgcgtc caggtcctccgcgttgctcagggcgaacggagtcaactttggtagctgccttccc aaaaagggcgcgtgcccaggctttgagttgcactcgcaccgtagtggcatcaa

		aaggtgaccgtgcccggtctgggcgttaggatacagcgcctgcataaaagcct tgatctgcttaaaagccacctgagcctttgcgccttcagagaagaacatgccgc aagacttgccggaaaactgattggccggacaggccgcgtcgtgcacgcagca ccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttcttcacga tcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgtcacatc catttcaatcacgtgctccttatttatcataatgcttccgtgtagacacttaagctcgc cttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtg atgcttgtaggtcacctctgcaaacgactgcaggtacgcctgcaggaatcgccc catcatcgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgc tcctcgttcagccaggtcttgcatacggccgccagagcttccacttggtcaggca gtagtttgaagttcgcctttagatcgttatccacgtggtacttgtccatcagcgcgc gcgcagcctccatgcccttctcccacgcagacacgatcggcacactcagcgg gttcatcaccgtaatttcactttccgcttcgctgggctcttcctcttcctcttgcgtccg cataccacgcgccactgggtcgtcttcattcagccgccgcactgtgcgcttacct cctttgccatgcttgattagcaccggtgggttgctgaaacccaccatttgtagcgc cacatcttctctttcttcctcgctgtccacgattacctctggtgatggcgggcgctcg ggcttgggagaagggcgcttctttttcttcttgggcgcaatggccaaatccgccgc cgaggtcgatggccgcgggctgggtgtgcgcggcaccagcgcgtcttgtgatg agtcttcctcgtcctcggactcgatacgccgcctcatccgctttttgggggcgccc ggggaggcggcggcgacggggacggggacgacacgtcctccatggttggg ggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctcctctt cccgactggccatttccttctcctataggcagaaaaagatccacaaaagcgaa gatcagcttcggcgcacgctggaagacgcggaggctctcttcagtaaatactgc gcgctgactcttaaggactagtttcgcgccctttctcaaatttaagcgcgaaaact acgtcatctccagcggccacacccggcgccagcacctgttgtcagcgccattg gcgcgcccgcccgccgcgcgcttcgctttttatagggccgccgccgccgccgc ctcgccataaaaggaaactttcggagcgcgccgctctgattggctgccgccgc acctctccgcctcgccccgccccgcccctcgccccgccccgccccgcctggcg cgcgccccccccccccccccgccccatcgctgcacaaaataattaaaaaat aaataaatacaaaattgggggtggggaggggggggagatggggagagtga agcagaacgtggggctcacctcgaggccggccgaatatcttcatttaaatgttta aacatcgatgcGGCCGCgattaggttctacgaagaccaatcccgcccgcc taatgcggagcttaccgcctgcgtcattacccagggccacattcttggccaattg caagccatcaacaaagcccgccaagagtttctgctacgaaagggacggggg gtttacttggacccccagtccggcgaggagctcaacccaatccccccgccgcc gcagccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaa aaagaagctgcagctgccgccgccacccacggacgaggaggaatactggg acagtcaggcagaggaggttttggacgaggaggaggaggacatgatggaag actgggagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacg aaacaccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggc aaccggttccagcatggctacaacctccgctcctcaggcgccgccggcactgc ccgttcgccgacccaaccgtagATGGGACACCACTGGAACCAG GGCCGGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCA AGAGCAACAACAGCGCCAAGGCTACCGCTCATGGCG CGGGCACAAGAACGCCATAGTTGCTTGCTTGCAAGAC TGTGGGGGCAACATCTCCTTCGCCCGCCGCTTTCTTC TCTACCATCACGGCGTGGCCTTCCCCCGTAACATCCT

		GCATTACTACCGTCATCTCTACAGCCCATACTGCACC GGCGGCAGCGGCAGCAACAGCAGCGGCCACACAGAA GCAAAGGCGACCGGATAGCAAGACTCTGACAAAGCC CAAGAAATCCACAGCGGCGGCAGCAGCAGGAGGAGG AGCGCTGCGTCTGGCGCCCAACGAACCCGTATCGAC CCGCGAGCTTAGAAACAGGATTTTTCCCACTCTGTAT GCTATATTTCAACAGAGCAGGGGCCAAGAACAAGAGC TGAAAATAAAAAAACAGGTCTCTGCGATCCCTCACCCG CAGCTGCCTGTATCACAAAAGCGAAGATCAGCTTCGG CGCACGCTGGAAGACGCGGAGGCTCTCTTCAGTAAAT ACTGCGCGCTGACTCTTAAGGACTAGaataaaagatctttattt tcattagatctgtgtgttggttttttgtgtgGCggccgccgtttgtgttatgtttcaacgt gtttatttttcaattgcagaaaatttcaagtcatttttcattcagtagtatagccccacc accacatagcttatacagatcaccgtaccttaatcaaactcacagaaccctagt attcaacctgccacctccctcccaacacacagagtacacagtcctttctccccgg ctggccttaaaaagcatcatatcatgggtaacagacatattcttaggtgttatattc cacacggtttcctgtcgagccaaacgctcatcagtgatattaataaactccccgg gcagctcacttaagttcatgtcgctgtccagctgctgagccacaggctgctgtcc aacttgcggttgcttaacgggcggcgaaggagaagtccacgcctacatgggg gtagagtcataatcgtgcatcaggatagggcggtggtgctgcagcagcgcgcg aataaactgctgccgccgccgctccgtcctgcaggaatacaacatggcagtgg tctcctcagcgatgattcgcaccgcccgcagcataaggcgccttgtcctccggg cacagcagcgcaccctgatctcacttaaatcagcacagtaactgcagcacag caccacaatattgttcaaaatcccacagtgcaaggcgctgtatccaaagctcat ggcggggaccacagaacccacgtggccatcataccacaagcgcaggtagat taagtggcgacccctataaacacgctggacataaacattacctcttttggcatgt tgtaattcaccacctcccggtaccatataaacctctgattaaacatggcgccatc caccaccatcctaaaccagctggccaaaacctgcccgccggctatacactgc agggaaccgggactggaacaatgacagtggagagcccaggactcgtaacc atggatcatcatgctcgtcatgatatcaatgttggcacaacacaggcacacgtg catacacttcctcaggattacaagctcctcccgcgttagaaccatatcccaggg aacaacccattcctgaatcagcgtaaatcccacactgcagggaagacctcgc acgtaactcacgttgtgcattgtcaaagtgttacattcgggcagcagcggatgat cctccagtatggtagcgcgggtttctgtctcaaaaggaggtagacgatccctact gtacggagtgcgccgagacaaccgagatcgtgttggtcgtagtgtcatgccaa atggaacgccggacgtagtcatatttcctgaagcaaaaccaggtgcgggcgtg acaaacagatctgcgtctccggtctcgccgcttagatcgctctgtgtagtagttgta gtatatccactctctcaaagcatccaggcgccccctggcttcgggttctatgtaaa ctccttcatgcgccgctgccctgataacatccaccaccgcagaataagccaca cccagccaacctacacattcgttctgcgagtcacacacgggaggagcgggaa gagctggaagaaccatgttttttttttttattccaaaagattatccaaaacctcaaaat gaagatctattaagtgaacgcgctcccctccggtggcgtggtcaaactctacag ccaaagaacagataatggcatttgtaagatgttgcacaatggcttccaaaaggc aaacggccctcacgtccaagtggacgtaaaggctaaacccttcagggtgaatc tcctctataaacattccagcaccttcaaccatgcccaaataattctcatctcgcca ccttctcaatatatctctaagcaaatcccgaatattaagtccggccattgtaaaaa tctgctccagagcgccctccaccttcagcctcaagcagcgaatcatgattgcaa

		aaattcaggttcctcacagacctgtataagattcaaaagcggaacattaacaaa aataccgcgatcccgtaggtcccttcgcagggccagctgaacataatcgtgca ggtctgcacggaccagcgcggccacttccccgccaggaaccatgacaaaag aacccacactgattatgacacgcatactcggagctatgctaaccagcgtagcc ccgatgtaagcttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaa aaatcaggcaaagcctcgcgcaaaaaagaaagcacatcgtagtcatgctcat gcagataaaggcaggtaagctccggaaccaccacagaaaaagacaccattt ttctctcaaacatgtctgcgggtttctgcataaacacaaaataaaataacaaaaa aacatttaaacattagaagcctgtcttacaacaggaaaaacaacccttataagc ataagacggactacggccatgccggcgtgaccgtaaaaaaactggtcaccgt gattaaaaagcaccaccgacagctcctcggtcatgtccggagtcataatgtaa gactcggtaaacacatcaggttgattcacatcggtcagtgctaaaaagcgacc gaaatagcccgggggaatacatacccgcaggcgtagagacaacattacagc ccccataggaggtataacaaaattaataggagagaaaaacacataaacacct gaaaaaccctcctgcctaggcaaaatagcaccctcccgctccagaacaacat acagcgcttccacagcggcagccataacagtcagccttaccagtaaaaaaga aaacctattaaaaaaacaccactcgacacggcaccagctcaatcagtcacag tgtaaaaaagggccaagtgcagagcgagtatatataggactaaaaaatgacg taacggttaaagtccacaaaaaacacccagaaaaccgcacgcgaacctacg cccagaaacgaaagccaaaaaacccacaacttcctcaaatcgtcacttccgtt ttcccacgttacgtcacttcccattttaagaaaactacaattcccaacacatacaa gttactccgcccttaat

128	pEMBR 1.4-ef1a-33R	TAAggatccGAATTCAGTACTCATATGTTGACAATTAATC ATCGGCATAGTATATCGGCATAGTATAATACGACTCAC TATAGGAGGGCCACCAAGGTGAGGAACTAAACCATGA GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAA TTCCAACATGGATGCTGATTTATATGGGTATAAATGGG CTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTA TCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTT CTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTA CAGATGAGATGGTCAGACTAAACTGGCTGACGGAATT TATGCCACTTCCGACCATCAAGCATTTTATCCGTACTC CTGATGATGCATGGTTACTCACCACTGCGATCCCCGG AAAAACAGCGTTCCAGGTATTAGAAGAATATCCTGATT CAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCT GCGCCGGTTGCACTCGATTCCTGTTTGTAATTGTCCTT TTAACAGCGATCGCGTATTTCGCCTCGCTCAGGCGCA ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATT TTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGT CTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGG ATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAAC CTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGAT CTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTC CTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATT GATAATCCTGATATGAATAAATTGCAGTTTCATTTGAT GCTCGATGAGTTTTTCTAATAAAAACCCGCTTCGGCG GGTTTTTTTATGCATGTGAGCAAAAGGCCAGCAAAAG GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGtcga GGCCGCcacacaaaaaaccaacacacagatctaatgaaaataaagatc ttttattCTAGTCCTTCAGGGTCAGGGCGCAATACTTGCTA

		AACAGAGCCTCGGCATCTTCCAGTGTCCTTCTCAGCT GGTCCTCGCTCTTGTGGTACAGGCAGCTTCTGGTCAG GCTTCTCAGGCTCCTGTTCTTGATCTTCAGCTCCTGCT CCTGGCCCCTGCTCTGCTGAAAGATGGCGTACAGGG TGGGAAAAATCCTGTTTCTCAGCTCCCTGGTGCTCAC GGGCTCGTTGGGGGCCAGTCTCAGGGCTCCGCCCCC TGCGGCAGCGGCGGCGGTGCTTTTCTTGGGCTTGGT CAGGGTCTTGCTATCAGGTCTCCTCTGCTTCTGGGTG GCAGCTGCTGTTGCTGCAGCTGCGGCGGGGGCGGTG TGGGCGGCCCTGGTGCCGGTGGTGTCCCATctcctgttgg gtcttctcacgggcagggcgggggggcctgggggcggaggtggtggccat gctggagccggtggcgctcttctgggggcggggctggggaaggccacgga ggggctgggggtctcgtcgctcacctcctccacctcgctggcttcctcgtccagg ctctcccagtcttccatcatgtcctcctcctcctcgtccagcacctcctcggcctgg ctgtcccagtattcctcctcgtcggtgggggggggggggcagctgcagcttcttctt gggggccatcctgggaagcaagggcccgcggctgctgctgatagggctgcg gcggctgtgttctggcggcaaacccgttgcgaaaaagaacgttcacggcgact actgcacttatatacggttctcccccacctcgggaaaaaggcggagccagta cacgacatcactttcccagtttaccccgcgccaccttctctaggcaccggttcaat tgccgacccctcccccccaacttctcggggactgtgggcgatgtgcgctctgccc actgacgggcaccggagccggtaagctccgcattaggcgggcgggattggtc ttcgtagaacctaatcATTTAAATatctagtcgacGTTcctatccgtagatg tacctggacatccaggtgatgccggcggcggtggtggaggcgcgcggaaagt cgcggacgcggttccagatgttgcgcagcggcaaaaagtgctccatggtcgg gacgctctggccggtgaggcgtgcgcagtcgttgacgctctagaccgtgcaaa aggagagcctgtaagcgggcactcttccgtggtctggtggataaattcgcaagg gtatcatggcggacgaccggggttcgaaccccggatccggccgtccgccgtg atccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtcagacaac gggggagcgctccttttggcttccttccaggcgcggcggctgctgcgctagcttttt tggccactggccgcgcgcggcgtaagcggttaggctggaaagcgaaagcatt aagtggctcgctccctgtagccggagggttattttccaagggttgagtcgcagga cccccggttcgagtctcgggccggccggactgcggcgaacgggggtttgcctc cccgtcatgcaagaccccgcttgcaaattcctccggaaacagggacgagccc ctttttttgcttttcccagatgcatccggtgctgcggcagatgcgcccccctcctcag cagcggcaagagcaagagcagcggcagacatgcagggcaccctcccctct cctaccgcgtcaggaggggcaacatcgatccagacatgataagatacattgat gagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaa atttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagtttgtac actctcgggtgattatttacccccacccttgccgtctgcgccgtttaaaaatcaaa ggggttctgccgcgcatcgctatgcgccactggcagggacacgttgcgatactg gtgtttagtgctccacttaaactcaggcacaaccatccgcggcagctcggtgaa gttttcactccacaggctgcgcaccatcaccaacgcgtttagcaggtcgggcgc cgatatcttgaagtcgcagttggggcctccgccctgcgcgcgcgagttgcgata cacagggttgcagcactggaacactatcagcgccgggtggtgcacgctggcc agcacgctcttgtcggagatcagatccgcgtccaggtcctccgcgttgctcagg gcgaacggagtcaactttggtagctgccttcccaaaaagggcgcgtgcccagg ctttgagttgcactcgcaccgtagtggcatcaaaaggtgaccgtgcccggtctgg

		gcgttaggatacagcgcctgcataaaagccttgatctgcttaaaagccacctga gcctttgcgccttcagagaagaacatgccgcaagacttgccggaaaactgatt ggccggacaggccgcgtcgtgcacgcagcaccttgcgtcggtgttggagatct gcaccacatttcggccccaccggttcttcacgatcttggccttgctagactgctcct tcagcgcgcgctgcccgttttcgctcgtcacatccatttcaatcacgtgctccttattt atcataatgcttccgtgtagacacttaagctcgccttcgatctcagcgcagcggtg cagccacaacgcgcagcccgtgggctcgtgatgcttgtaggtcacctctgcaa acgactgcaggtacgcctgcaggaatcgccccatcatcgtcacaaaggtcttgt tgctggtgaaggtcagctgcaacccgcggtgctcctcgttcagccaggtcttgca tacggccgccagagcttccacttggtcaggcagtagtttgaagttcgcctttagat cgttatccacgtggtacttgtccatcagcgcgcgcgcagcctccatgcccttctcc cacgcagacacgatcggcacactcagcgggttcatcaccgtaatttcactttcc gcttcgctgggctcttcctcttcctcttgcgtccgcataccacgcgccactgggtcg tcttcattcagccgccgcactgtgcgcttacctcctttgccatgcttgattagcaccg gtgggttgctgaaacccaccatttgtagcgccacatcttctctttcttcctcgctgtcc acgattacctctggtgatggcgggcgctcgggcttgggagaagggcgcttcttttt cttcttgggcgcaatggccaaatccgccgccgaggtcgatggccgcgggctgg gtgtgcgcggcaccagcgcgtcttgtgatgagtcttcctcgtcctcggactcgata cgccgcctcatccgcttttttgggggcgcccgggggaggcggcggcgacgggg acggggacgacacgtcctccatggttgggggacgtcgcgccgcaccgcgtcc gcgctcgggggtggtttcgcgctgctcctcttcccgactggccatttccttctcctat aggcagaaaaagatccacaaaagcgaagatcagcttcggcgcacgctgga agacgcggaggctctcttcagtaaatactgcgcgctgactcttaaggactagttt cgcgcccttctcaaatttaagcgcgaaaactacgtcatctccagcggccacac ccggcgccagcacctgttgtcagcgccattggcgcgcccgcccgccgcgcgct tcgcttttatagggccgccgccgccgccgcctcgccataaaaggaaactttcg gagcgcgccgctctgattggctgccgccgcacctctccgcctcgccccgcccc gcccctcgccccgccccgccccgcctggcgcgcgcccccccccccccccg cccccatcgctgcacaaaataattaaaaaataaataaatacaaaattgggggt ggggaggggggggagatggggagagtgaagcagaacgtggggctcacctc gaggccggccgaatatcttcatttaaatgtttaaacatcgatgcGGCCGCga ttaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccgcctgcg tcattacccagggccacattcttggccaattgcaagccatcaacaaagcccgcc aagagtttctgctacgaaagggacgggggtttacttggacccccagtccggc gaggagctcaacccaatccccccgccgccgcagccctatcagcagcagccg cgggcccttgcttcccaggatggcacccaaaaagaagctgcagctgccgccg ccacccacggacgaggaggaatactgggacagtcaggcagaggaggttttg gacgaggaggaggaggacatgatggaagactgggagagcctagacgagg aagcttccgaggtcgaagaggtgtcagacgaaacaccgtcaccctcggtcgc attcccctcgccggcgccccagaaatcggcaaccggttccagcatggctacaa cctccgctcctcaggcgccgccggcactgcccgttcgccgacccaaccgtag ATGGGACACCACTGGAACCAGGGCCGGTAAGTCCAA GCAGCCGCCGCCGTTAGCCCAAGAGCAACAACAGCG CCAAGGCTACCGCTCATGGCGCGGGCACAAGAACGC CATAGTTGCTTGCTTGCAAGACTGTGGGGGCAACATC TCCTTCGCCCGCCGCTTTCTTCTCTACCATCACGGCG TGGCCTTCCCCCGTAACATCCTGCATTACTACCGTCAT

		CTCTACAGCCCATACTGCACCGGCGGCAGCGGCAGC AACAGCAGCGGCCACACAGAAGCAAAGGCGACCGGA TAGCAAGACTCTGACAAAGCCCAAGAAATCCACAGCG GCGGCAGCAGCAGGAGGAGGAGCGCTGCGTCTGGC GCCCAACGAACCCGTATCGACCCGCGAGCTTAGAAAC AGGATTTTTCCCACTCTGTATGCTATATTTCAACAGAG CAGGGGCCAAGAACAAGAGCTGAAAATAAAAAACAGG TCTCTGCGATCCCTCACCCGCAGCTGCCTGTATCACA AAAGCGAAGATCAGCTTCGGCGCACGCTGGAAGACG CGGAGGCTCTCTTCAGTAAATACTGCGCGCTGACTCT TAAGGACTAGaataaaagatctttattttcattagatctgtgtgttggttttttgt gtgGCggccgccgtttgtgttatgtttcaacgtgtttatttttcaattgcagaaaattt caagtcatttttcattcagtagtatagccccaccaccacatagcttatacagatca ccgtaccttaatcaaactcacagaaccctagtattcaacctgccacctccctccc aacacacagagtacacagtcctttctccccggctggccttaaaaagcatcatat catgggtaacagacatattcttaggtgttatattccacacggtttcctgtcgagcca aacgctcatcagtgatattaataaactccccgggcagctcacttaagttcatgtcg ctgtccagctgctgagccacaggctgctgtccaacttgcggttgcttaacgggcg gcgaaggagaagtccacgcctacatgggggtagagtcataatcgtgcatcag gatagggcggtggtgctgcagcagcgcgcgaataaactgctgccgccgccgc tccgtcctgcaggaatacaacatggcagtggtctcctcagcgatgattcgcacc gcccgcagcataaggcgccttgtcctccgggcacagcagcgcaccctgatctc acttaaatcagcacagtaactgcagcacagcaccacaatattgttcaaaatccc acagtgcaaggcgctgtatccaaagctcatggcggggaccacagaacccac gtggccatcataccacaagcgcaggtagattaagtggcgacccctcataaaca cgctggacataaacattacctcttttggcatgttgtaattcaccacctcccggtacc atataaacctctgattaaacatggcgccatccaccaccatcctaaaccagctgg ccaaaacctgcccgccggctatacactgcagggaaccgggactggaacaat gacagtggagagcccaggactcgtaaccatggatcatcatgctcgtcatgatat caatgttggcacaacacaggcacacgtgcatacacttcctcaggattacaagct cctcccgcgttagaaccatatcccagggaacaacccattcctgaatcagcgta aatcccacactgcagggaagacctcgcacgtaactcacgttgtgcattgtcaaa gtgttacattcgggcagcagcggatgatcctccagtatggtagcgcgggtttctgt ctcaaaaggaggtagacgatccctactgtacggagtgcgccgagacaaccg agatcgtgttggtcgtagtgtcatgccaaatggaacgccggacgtagtcatatttc ctgaagcaaaaccaggtgcgggcgtgacaaacagatctgcgtctccggtctcg ccgcttagatcgctctgtgtagtagttgtagtatatccactctctcaaagcatccag gcgccccctggcttcgggttctatgtaaactccttcatgcgccgctgccctgataa catccaccaccgcagaataagccacacccagccaacctacacattcgttctgc gagtcacacacgggaggagcgggaagagctggaagaaccatgttttttttttttat tccaaaagattatccaaaacctcaaaatgaagatctattaagtgaacgcgctcc cctccggtggcgtggtcaaactctacagccaaagaacagataatggcatttgta agatgttgcacaatggcttccaaaaggcaaacggccctcacgtccaagtggac gtaaaggctaaacccttcagggtgaatctcctctataaacattccagcaccttca accatgcccaaataattctcatctcgccaccttctcaatatatctctaagcaaatcc cgaatattaagtccggccattgtaaaaatctgctccagagcgccctccaccttca gcctcaagcagcgaatcatgattgcaaaaattcaggttcctcacagacctgtat

		aagattcaaaagcggaacattaacaaaaataccgcgatcccgtaggtcccttc gcagggccagctgaacataatcgtgcaggtctgcacggaccagcgcggcca cttccccgccaggaaccatgacaaaagaacccacactgattatgacacgcat actcggagctatgctaaccagcgtagccccgatgtaagcttgttgcatgggcgg cgatataaaatgcaaggtgctgctcaaaaaatcaggcaaagcctcgcgcaaa aaagaaagcacatcgtagtcatgctcatgcagataaaggcaggtaagctccg gaaccaccacagaaaaagacaccatttttctctcaaacatgtctgcgggtttctg cataaacacaaaataaaataacaaaaaaacatttaaacattagaagcctgtct tacaacaggaaaaacaacccttataagcataagacggactacggccatgcc ggcgtgaccgtaaaaaaactggtcaccgtgattaaaaagcaccaccgacag ctcctcggtcatgtccggagtcataatgtaagactcggtaaacacatcaggttga ttcacatcggtcagtgctaaaaagcgaccgaaatagcccgggggaatacata cccgcaggcgtagagacaacattacagcccccataggaggtataacaaaatt aataggagagaaaaacacataaacacctgaaaaaccctcctgcctaggcaa aatagcaccctcccgctccagaacaacatacagcgcttccacagcggcagcc ataacagtcagccttaccagtaaaaaagaaaacctattaaaaaaacaccact cgacacggcaccagctcaatcagtcacagtgtaaaaaagggccaagtgcag agcgagtatatataggactaaaaaatgacgtaacggttaaagtccacaaaaa acacccagaaaaccgcacgcgaacctacgcccagaaacgaaagccaaaa aacccacaacttcctcaaatcgtcacttccgttttcccacgttacgtcacttcccatt ttaagaaaactacaattcccaacacatacaagttactccgcccttaat

129	pEMBR-sm12-CMVe-33f	CAGTACTCATATGTTGACAATTAATCATCGGCATAGTA TATCGGCATAGTATAATACGACTCACTATAGGAGGGC CACCAAGGTGAGGAACTAAACCATGAGCCATATTCAA CGGGAAACGTCGAGGCCGCGATTAAATTCCAACATGG ATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT GTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATG GGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGG CAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATG GTCAGACTAAACTGGCTGACGGAATTTATGCCACTTC CGACCATCAAGCATTTTATCCGTACTCCTGATGATGCA TGGTTACTCACCACTGCGATCCCCGGAAAAACAGCGT TCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAT ATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGC ACTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGAT CGCGTATTTCGCCTCGCTCAGGCGCAATCACGAATGA ATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGA GCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAA ATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGT CACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGA CGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGA GTCGGAATCGCAGACCGATACCAGGATCTTGCCATCC TATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAG AAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGA TATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGT TTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTTATG CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT CACGTTAAGGGATTTTGGTCATGtcgaGGCCGCcacacaa aaaaccaacacacagatctaatgaaaataaagatcttttattCTAGTCCTT CAGGGTCAGGGCGCAATACTTGCTAAACAGAGCCTCG

		GCATCTTCCAGTGTCCTTCTCAGCTGGTCCTCGCTCTT GTGGTACAGGCAGCTTCTGGTCAGGCTTCTCAGGCTC CTGTTCTTGATCTTCAGCTCCTGCTCCTGGCCCCTGC TCTGCTGAAAGATGGCGTACAGGGTGGGAAAAATCCT GTTTCTCAGCTCCCTGGTGCTCACGGGCTCGTTGGGG GCCAGTCTCAGGGCTCCGCCCCCTGCGGCAGCGGCG GCGGTGCTTTTCTTGGGCTTGGTCAGGGTCTTGCTAT CAGGTCTCCTCTGCTTCTGGGTGGCAGCTGCTGTTGC TGCAGCTGCGGCGGGGGCGGTGTGGGCGGCCCTGG TGCCGGTGGTGTCCCATctcctgttgggtcttctcacgggcagggcg gggggggcctgggggggcggaggtggtggccatgctggagccggtggcgctct tctgggggggcggggctggggaaggccacggaggggctgggggtctcgtcgct cacctcctccacctcgctggcttcctcgtccaggctctcccagtcttccatcatgtc ctcctcctcctcgtccagcacctcctcggcctggctgtcccagtattcctcctcgtc ggtgggggggggggcagctgcagcttcttcttgggggccatcctgggaagca agggcccgcggctgctgctgatagggctgcggcggctgtgttctggcggcaaa cccgttgcgaaaaagaacgcgttattcgccagctctgcttatatagacctcccac cgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttgg aaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtgg agacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgc caaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtactgcca agtaggaaagtcccataaggtcatgtactgggcataatgccaggcgggccattt accgtcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactg ccaagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtcc ctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtcg ttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgcggaactcc atatatgggctatgaactaatgaccccgtaattgattactattaataactagtcaat aatcaatgtcagttcgacGTTAATTcgagctcggtacctcgcgaatacatct agatatcggatccgatttaATTAAGGGCGGAGTAACTTGTATGT GTTGGGAATTGTAGTTTTCTTAAAATGGGAAGTGACGT AACGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGG CTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTC AATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCT CCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAA TTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTC CGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAG GCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAG TAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCA AATGCCACCATGGTTCTTCCAGCTCTTCCCGCTCCTC CCGTGTGTGACTCGCAGAACGAATGTGTAGGTTGGCT GGGTGTGGCTTATTCTGCGGTGGTGGATGTTATCAGG GCAGCGGCGCATGAAGGAGTTTACATAGAACCCGAA GCCAGGGGGCGCCTGGATGCTTTGAGAGAGTGGATA TACTACAACTACTACACAGAGCGATCTAAGCGGCGAG ACCGGAGACGCAGATCTGTTTGTCACGCCCGCACCTG GTTTTGCTTCAGGAAATATGACTACGTCCGGCGTTCC

		ATTTGGCATGACACTACGACCAACACGATCTCGGTTG TCTCGGCGCACTCCGTACAGTAGGGATCGTCTACCTC CTTTTGAGACAGAAACCCGCGCTACCATACTGGAGGA TCATCCGCTGCTGCCCGAATGTAACACTTTGACAATG CACAACGTGAGTTACGTGCGAGGTCTTCCCTGCAGTG TGGGATTTACGCTGATTCAGGAATGGGTTGTTCCCTG GGATATGGTTCTAACGCGGGAGGAGCTTGTAATCCTG AGGAAGTGTATGCACGTGTGCCTGTGTTGTGCCAACA TTGATATCATGACGAGCATGATGATCCATGGTTACGA GTCCTGGGCTCTCCACTGTCATTGTTCCAGTCCCGGT TCCCTGCAGTGTATAGCCGGCGGGCAGGTTTTGGCC AGCTGGTTTAGGATGGTGGTGGATGGCGCCATGTTTA ATCAGAGGTTTATATGGTACCGGGAGGTGGTGAATTA CAACATGCCAAAAGAGGTAATGTTTATGTCCAGCGTG TTTATGAGGGGTCGCCACTTAATCTACCTGCGCTTGT GGTATGATGGCCACGTGGGTTCTGTGGTCCCCGCCAT GAGCTTTGGATACAGCGCCTTGCACTGTGGGATTTTG AACAATATTGTGGTGCTGTGCTGCAGTTACTGTGCTG ATTTAAGTGAGATCAGGGTGCGCTGCTGTGCCCGGAG GACAAGGCGCCTTATGCTGCGGGCGGTGCGAATCAT CGCTGAGGAGACCACTGCCATGTTGTATTCCTGCAGG ACGGAGCGGCGGCGGCAGCAGTTTATTCGCGCGCTG CTGCAGCACCACCGCCCTATCCTGATGCACGATTATG ACTCTACCCCCATGTAAAtAAatatctttattttCAttacatctGTGT GTTGGTTTTTTGTGTGGCGGCCAACAAGCTTTACcctatc cgtagatgtacctggacatccaggtgatgccggcggcggtggtggaggcgcg cggaaagtcgcggacgcggttccagatgttgcgcagcggcaaaaagtgctcc atggtcgggacgctctggccggtgaggcgtgcgcagtcgttgacgctctagacc gtgcaaaaggagagcctgtaagcgggcactcttccgtggtctggtggataaatt cgcaagggtatcatggcggacgaccggggttcgaaccccggatccggccgtc cgccgtgatccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtc agacaacgggggagcgctccttttggcttccttccaggcgcggcggctgctgcg ctagctttttggccactggccgcgcgcggcgtaagcggttaggctggaaagcg aaagcattaagtggctcgctccctgtagccggagggttattttccaagggttgagt cgcaggacccccggttcgagtctcgggccggccggactgcggcgaacgggg gtttgcctccccgtcatgcaagaccccgcttgcaaattcctccggaaacaggga cgagccccttttttgcttttcccagatgcatccggtgctgcggcagatgcgccccc ctcctcagcagcggcaagagcaagagcagcggcagacatgcagggcaccc tcccccttctcctaccgcgtcaggaggggcaacatcgatccagacatgataagat acattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgcttta tttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaa gtttGTACACTCTCGGGTGATTATTTACCCCCACCCTTG CCGTCTGCGCCGTTTAAAAATCAAAGGGGTTCTGCCG CGCATCGCTATGCGCCACTGGCAGGGACACGTTGCG ATACTGGTGTTTAGTGCTCCACTTAAACTCAGGCACAA CCATCCGCGGCAGCTCGGTGAAGTTTTCACTCCACAG GCTGCGCACCATCACCAACGCGTTTAGCAGGTCGGG

		CGCCGATATCTTGAAGTCGCAGTTGGGGCCTCCGCC CTGCGCGCGCGAGTTGCGATACACAGGGTTGCAGCA CTGGAACACTATCAGCGCCGGGTGGTGCACGCTGGC CAGCACGCTCTTGTCGGAGATCAGATCCGCGTCCAG GTCCTCCGCGTTGCTCAGGGCGAACGGAGTCAACTTT GGTAGCTGCCTTCCCAAAAAGGGCGCGTGCCCAGGC TTTGAGTTGCACTCGCACCGTAGTGGCATCAAAAGGT GACCGTGCCCGGTCTGGGCGTTAGGATACAGCGCCT GCATAAAAGCCTTGATCTGCTTAAAAGCCACCTGAGC CTTTGCGCCTTCAGAGAAGAACATGCCGCAAGACTTG CCGGAAAACTGATTGGCCGGACAGGCCGCGTCGTGC ACGCAGCACCTTGCGTCGGTGTTGGAGATCTGCACCA CATTTCGGCCCCACCGGTTCTTCACGATCTTGGCCTT GCTAGACTGCTCCTTCAGCGCGCGCTGCCCGTTTTCG CTCGTCACATCCATTTCAATCACGTGCTCCTTATTTAT CATAATGCTTCCGTGTAGACACTTAAGCTCGCCTTCGA TCTCAGCGCAGCGGTGCAGCCACAACGCGCAGCCCG TGGGCTCGTGATGCTTGTAGGTCACCTCTGCAAACGA CTGCAGGTACGCCTGCAGGAATCGCCCCATCATCGTC ACAAAGGTCTTGTTGCTGGTGAAGGTCAGCTGCAACC CGCGGTGCTCCTCGTTCAGCCAGGTCTTGCATACGGC CGCCAGAGCTTCCACTTGGTCAGGCAGTAGTTTGAAG TTCGCCTTTAGATCGTTATCCACGTGGTACTTGTCCAT CAGCGCGCGCGCAGCCTCCATGCCCTTCTCCCACGC AGACACGATCGGCACACTCAGCGGGTTCATCACCGTA ATTTCACTTTCCGCTTCGCTGGGCTCTTCCTCTTCCTC TTGCGTCCGCATACCACGCGCCACTGGGTCGTCTTCA TTCAGCCGCCGCACTGTGCGCTTACCTCCTTTGCCAT GCTTGATTAGCACCGGTGGGTTGCTGAAACCCACCAT TTGTAGCGCCACATCTTCTCTTTCTTCCTCGCTGTCCA CGATTACCTCTGGTGATGGCGGGCGCTCGGGCTTGG GAGAAGGGCGCTTCTTTTTCTTCTTGGGCGCAATGGC CAAATCCGCCGCCGAGGTCGATGGCCGCGGGCTGGG TGTGCGCGGCACCAGCGCGTCTTGTGATGAGTCTTCC TCGTCCTCGGACTCGATACGCCGCCTCATCCGCTTTT TTGGGGGCGCCCGGGGAGGCGGCGGCGACGGGGAC GGGGACGACACGTCCTCCATGGTTGGGGGACGTCGC GCCGCACCGCGTCCGCGCTCGGGGGTGGTTTCGCGC TGCTCCTCTTCCCGACTGGCCATttccttctcctataggcagaaa aagatcCACAAAAGCGAAGATCAGCTTCGGCGCACGCT GGAAGACGCGGAGGCTCTCTTCAGTAAATACTGCGCG CTGACTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAAT TTAAGCGCGAAAACTACGTCATCTCCAGCGGCCACAC CCGGCGCCAGCACCTGTTGTCAGCGCCATTGGCGCG CCGGCCGGCCGAAtatcTTCATTTAAATgattaggttctacgaa gaccaatcccgcccgcctaatgcggagcttaccgcctgcgtcattacccaggg ccacattcttggccaattgcaagccatcaacaaagcccgccaagagtttctgct

		acgaaagggacgggggtttacttggacccccagtccggcgaggagctcaa cccaatccccccgccgccgcagccctatcagcagcagccgcgggcccttgctt cccaggatggcacccaaaaagaagctgcagctgccgccgccacccacgga cgaggaggaatactgggacagtcaggcagaggaggttttggacgaggagga ggaggacatgatggaagactgggagagcctagacgaggaagcttccgaggt cgaagaggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgccgg cgccccagaaatcggcaaccggttccagcatggctacaacctccgctcctcag gcgccgccggcactgcccgttcgccgacccaaccgtagATGGGACACC ACTGGAACCAGGGCCGGTAAGTCCAAGCAGCCGCCG CCGTTAGCCCAAGAGCAACAACAGCGCCAAGGCTAC CGCTCATGGCGCGGGCACAAGAACGCCATAGTTGCTT GCTTGCAAGACTGTGGGGGCAACATCTCCTTCGCCCG CCGCTTTCTTCTCTACCATCACGGCGTGGCCTTCCCC CGTAACATCCTGCATTACTACCGTCATCTCTACAGCCC ATACTGCACCGGCGGCAGCGGCAGCAACAGCAGCGG CCACACAGAAGCAAAGGCGACCGGATAGCAAGACTCT GACAAAGCCCAAGAAATCCACAGCGGCGGCAGCAGC AGGAGGAGGAGCGCTGCGTCTGGCGCCCAACGAACC CGTATCGACCCGCGAGCTTAGAAACAGGATTTTTCCC ACTCTGTATGCTATATTTCAACAGAGCAGGGGCCAAG AACAAGAGCTGAAAATAAAAAACAGGTCTCTGCGATC CCTCACCCGCAGCTGCCTGTATCACAAAAGCGAAGAT CAGCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTC TTCAGTAAATACTGCGCGCTGACTCTTAAGGACTAGaa taaaagatctttattttcattagatctgtgtgttggtttttgtgtgGCGGCC

130	pEMBR-sm12-CMVe-33r	CAGTACTCATATGTTGACAATTAATCATCGGCATAGTA TATCGGCATAGTATAATACGACTCACTATAGGAGGGC CACCAAGGTGAGGAACTAAACCATGAGCCATATTCAA CGGGAAACGTCGAGGCCGCGATTAAATTCCAACATGG ATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT GTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATG GGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGG CAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATG GTCAGACTAAACTGGCTGACGGAATTTATGCCACTTC CGACCATCAAGCATTTTATCCGTACTCCTGATGATGCA TGGTTACTCACCACTGCGATCCCCGGAAAAACAGCGT TCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAT ATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGC ACTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGAT CGCGTATTTCGCCTCGCTCAGGCGCAATCACGAATGA ATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGA GCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAA ATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGT CACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGA CGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGA GTCGGAATCGCAGACCGATACCAGGATCTTGCCATCC TATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAG AAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGA TATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGT TTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTTATG CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT CACGTTAAGGGATTTTGGTCATGtcgaactgacattgattattga ctagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagtt ccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacc

		cccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaataggga ctttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagta catcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaat ggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggcag tacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatca atgggcgtggatagcggtttgactcacggggatttccaagtctccaccccattga cgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgta acaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggt ctatataagcagagctggcgaataacgcgttcttttcgcaacgggtttgccgcc agaacacagccgccgcagccctatcagcagcagccgcgggcccttgcttccc aggatggcccccaagaagaagctgcagctgccccccccccccaccgacga ggaggaatactgggacagccaggccgaggaggtgctggacgaggaggag gaggacatgatggaagactgggagagcctggacgaggaagccagcgaggt ggaggaggtgagcgacgagacccccagcccctccgtggccttccccagccc cgccccccagaagagcgccaccggctccagcatggccaccacctccgcccc ccaggcccccccccgccctgcccgtgagaagacccaacaggagATGGGA CACCACCGGCACCAGGGCCGCCCACACCGCCCCCGC CGCAGCTGCAGCAACAGCAGCTGCCACCCAGAAGCA GAGGAGACCTGATAGCAAGACCCTGACCAAGCCCAA GAAAAGCACCGCCGCCGCTGCCGCAGGGGGCGGAG CCCTGAGACTGGCCCCCAACGAGCCCGTGAGCACCA GGGAGCTGAGAAACAGGATTTTTCCCACCCTGTACGC CATCTTTCAGCAGAGCAGGGGCCAGGAGCAGGAGCT GAAGATCAAGAACAGGAGCCTGAGAAGCCTGACCAG AAGCTGCCTGTACCACAAGAGCGAGGACCAGCTGAG AAGGACACTGGAAGATGCCGAGGCTCTGTTTAGCAAG TATTGCGCCCTGACCCTGAAGGACTAGaataaaagatctttat tttcattagatctgtgtgttggtttttgtgtgGCGGCCtcgacGTTAATTcga gctcggtacctcgcgaatacatctagatatcggatccgatttaATTAAGGG CGGAGTAACTTGTATGTGTTGGGAATTGTAGTTTTCTT AAAATGGGAAGTGACGTAACGTGTGTGTCAGTTAGGG TGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGT ATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGT GTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTA TGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGC CCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAAT TTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCC TCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTG GAGGCCTAGGCTTTTGCAAATGCCACCATGGTTCTTC CAGCTCTTCCCGCTCCTCCCGTGTGTGACTCGCAGAA CGAATGTGTAGGTTGGCTGGGTGTGGCTTATTCTGCG GTGGTGGATGTTATCAGGGCAGCGGCGCATGAAGGA GTTTACATAGAACCCGAAGCCAGGGGGCGCCTGGAT GCTTTGAGAGAGTGGATATACTACAACTACTACACAGA GCGATCTAAGCGGCGAGACCGGAGACGCAGATCTGT TTGTCACGCCCGCACCTGGTTTTGCTTCAGGAAATAT

		GACTACGTCCGGCGTTCCATTTGGCATGACACTACGA CCAACACGATCTCGGTTGTCTCGGCGCACTCCGTACA GTAGGGATCGTCTACCTCCTTTTGAGACAGAAACCCG CGCTACCATACTGGAGGATCATCCGCTGCTGCCCGAA TGTAACACTTTGACAATGCACAACGTGAGTTACGTGC GAGGTCTTCCCTGCAGTGTGGGATTTACGCTGATTCA GGAATGGGTTGTTCCCTGGGATATGGTTCTAACGCGG GAGGAGCTTGTAATCCTGAGGAAGTGTATGCACGTGT GCCTGTGTTGTGCCAACATTGATATCATGACGAGCAT GATGATCCATGGTTACGAGTCCTGGGCTCTCCACTGT CATTGTTCCAGTCCCGGTTCCCTGCAGTGTATAGCCG GCGGGCAGGTTTTGGCCAGCTGGTTTAGGATGGTGG TGGATGGCGCCATGTTTAATCAGAGGTTTATATGGTAC CGGGAGGTGGTGAATTACAACATGCCAAAAGAGGTAA TGTTTATGTCCAGCGTGTTTATGAGGGGTCGCCACTT AATCTACCTGCGCTTGTGGTATGATGGCCACGTGGGT TCTGTGGTCCCCGCCATGAGCTTTGGATACAGCGCCT TGCACTGTGGGATTTTGAACAATATTGTGGTGCTGTG CTGCAGTTACTGTGCTGATTTAAGTGAGATCAGGGTG CGCTGCTGTGCCCGGAGGACAAGGCGCCTTATGCTG CGGGCGGTGCGAATCATCGCTGAGGAGACCACTGCC ATGTTGTATTCCTGCAGGACGGAGCGGCGGCGGCAG CAGTTTATTCGCGCGCTGCTGCAGCACCACCGCCCTA TCCTGATGCACGATTATGACTCTACCCCCATGTAAAtA AatatctttattttCAttacatctGTGTGTTGGTTTTTTGTGTGGCG GCCAACAAGCTTTACcctatccgtagatgtacctggacatccaggtga tgccggcggcggtggtggaggcgcgcggaaagtcgcggacgcggttccaga tgttgcgcagcggcaaaaagtgctccatggtcgggacgctctggccggtgagg cgtgcgcagtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgg gcactcttccgtggtctggtggataaattcgcaagggtatcatggcggacgacc ggggttcgaaccccggatccggccgtccgccgtgatccatgcggttaccgccc gcgtgtcgaacccaggtgtgcgacgtcagacaacgggggagcgctcctttttgg cttccttccaggcgcggcggctgctgcgctagctttttttggccactggccgcgcgc ggcgtaagcggttaggctggaaagcgaaagcattaagtggctcgctccctgta gccggagggttattttccaagggttgagtcgcaggacccccggttcgagtctcg ggccggccggactgcggcgaacgggggtttgcctccccgtcatgcaagaccc cgcttgcaaattcctccggaaacagggacgagccccttttttgcttttcccagatg catccggtgctgcggcagatgcgcccccctcctcagcagcggcaagagcaag agcagcggcagacatgcagggcaccctcccctttctcctaccgcgtcaggagg ggcaacatcgatccagacatgataagatacattgatgagtttggacaaaccac aactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttat ttgtaaccattataagctgcaataaacaagtttGTACACTCTCGGGTG ATTATTTACCCCCACCCTTGCCGTCTGCGCCGTTTAAA AATCAAAGGGGTTCTGCCGCGCATCGCTATGCGCCAC TGGCAGGGACACGTTGCGATACTGGTGTTTAGTGCTC CACTTAAACTCAGGCACAACCATCCGCGGCAGCTCGG TGAAGTTTTCACTCCACAGGCTGCGCACCATCACCAA

		CGCGTTTAGCAGGTCGGGCGCCGATATCTTGAAGTCG CAGTTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGA TACACAGGGTTGCAGCACTGGAACACTATCAGCGCCG GGTGGTGCACGCTGGCCAGCACGCTCTTGTCGGAGA TCAGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGG CGAACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAA GGGCGCGTGCCCAGGCTTTGAGTTGCACTCGCACCG TAGTGGCATCAAAAGGTGACCGTGCCCGGTCTGGGC GTTAGGATACAGCGCCTGCATAAAAGCCTTGATCTGC TTAAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGA ACATGCCGCAAGACTTGCCGGAAAACTGATTGGCCGG ACAGGCCGCGTCGTGCACGCAGCACCTTGCGTCGGT GTTGGAGATCTGCACCACATTTCGGCCCCACCGGTTC TTCACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCG CGCGCTGCCCGTTTTCGCTCGTCACATCCATTTCAAT CACGTGCTCCTTATTTATCATAATGCTTCCGTGTAGAC ACTTAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAG CCACAACGCGCAGCCCGTGGGCTCGTGATGCTTGTA GGTCACCTCTGCAAACGACTGCAGGTACGCCTGCAG GAATCGCCCCATCATCGTCACAAAGGTCTTGTTGCTG GTGAAGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCA GCCAGGTCTTGCATACGGCCGCCAGAGCTTCCACTTG GTCAGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTAT CCACGTGGTACTTGTCCATCAGCGCGCGCGCAGCCT CCATGCCCTTCTCCCACGCAGACACGATCGGCACACT CAGCGGGTTCATCACCGTAATTTCACTTTCCGCTTCG CTGGGCTCTTCCTCTTCCTCTTGCGTCCGCATACCAC GCGCCACTGGGTCGTCTTCATTCAGCCGCCGCACTGT GCGCTTACCTCCTTTGCCATGCTTGATTAGCACCGGT GGGTTGCTGAAACCCACCATTTGTAGCGCCACATCTT CTCTTTCTTCCTCGCTGTCCACGATTACCTCTGGTGAT GGCGGGCGCTCGGGCTTGGGAGAAGGGCGCTTCTTT TTCTTCTTGGGCGCAATGGCCAAATCCGCCGCCGAGG TCGATGGCCGCGGGCTGGGTGTGCGCGGCACCAGC GCGTCTTGTGATGAGTCTTCCTCGTCCTCGGACTCGA TACGCCGCCTCATCCGCTTTTTTGGGGGCGCCCGGG GAGGCGGCGGCGACGGGGACGGGGACGACACGTCC TCCATGGTTGGGGGACGTCGCGCCGCACCGCGTCCG CGCTCGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGA CTGGCCATttccttctcctataggcagaaaaagatcCACAAAAGCG AAGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGG CTCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGAC TAGTTTCGCGCCCTTTCTCAAATTTAAGCGCGAAAACT ACGTCATCTCCAGCGGCCACACCCGGCGCCAGCACC TGTTGTCAGCGCCATTGGCGCGCCGGCCGGCCGAAta tcTTCATTTAAATgattaggttctacgaagaccaatcccgcccgcctaatg cggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaag

		ccatcaacaaagcccgccaagagtttctgctacgaaagggacgggggttta cttggacccccagtccggcgaggagctcaacccaatccccccgccgccgcag ccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaag aagctgcagctgccgccgccacccacggacgaggaggaatactgggacagt caggcagaggaggttttggacgaggaggaggaggacatgatggaagactgg gagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaaca ccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccgg ttccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcg ccgacccaaccgtagATGGGACACCACTGGAACCAGGGCC GGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAG CAACAACAGCGCCAAGGCTACCGCTCATGGCGCGGG CACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTG GGGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTA CCATCACGGCGTGGCCTTCCCCCGTAACATCCTGCAT TACTACCGTCATCTCTACAGCCCATACTGCACCGGCG GCAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAA AGGCGACCGGATAGCAAGACTCTGACAAAGCCCAAG AAATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGaataaaagatctttattttcattaga tctgtgtgttggttttttgtgtgGCGGCC

131	pEMBRsm12-ef1a-33F	CAGTACTCATATGTTGACAATTAATCATCGGCATAGTA TATCGGCATAGTATAATACGACTCACTATAGGAGGGC CACCAAGGTGAGGAACTAAACCATGAGCCATATTCAA CGGGAAACGTCGAGGCCGCGATTAAATTCCAACATGG ATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT GTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATG GGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGG CAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATG GTCAGACTAAACTGGCTGACGGAATTTATGCCACTTC CGACCATCAAGCATTTTATCCGTACTCCTGATGATGCA TGGTTACTCACCACTGCGATCCCCGGAAAAACAGCGT TCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAT ATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGC ACTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGAT CGCGTATTTCGCCTCGCTCAGGCGCAATCACGAATGA ATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGA GCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAA ATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGT CACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGA CGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGA GTCGGAATCGCAGACCGATACCAGGATCTTGCCATCC TATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAG AAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGA TATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGT TTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTTATG CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT CACGTTAAGGGATTTTGGTCATGtcgaGGCCGCcacacaa aaaaccaacacacagatctaatgaaaataaagatcttttattCTAGTCCTT CAGGGTCAGGGCGCAATACTTGCTAAACAGAGCCTCG

		GCATCTTCCAGTGTCCTTCTCAGCTGGTCCTCGCTCTT GTGGTACAGGCAGCTTCTGGTCAGGCTTCTCAGGCTC CTGTTCTTGATCTTCAGCTCCTGCTCCTGGCCCCTGC TCTGCTGAAAGATGGCGTACAGGGTGGGAAAAATCCT GTTTCTCAGCTCCCTGGTGCTCACGGGCTCGTTGGGG GCCAGTCTCAGGGCTCCGCCCCCTGCGGCAGCGGCG GCGGTGCTTTTCTTGGGCTTGGTCAGGGTCTTGCTAT CAGGTCTCCTCTGCTTCTGGGTGGCAGCTGCTGTTGC TGCAGCTGCGGCGGGGGCGGTGTGGGCGGCCCTGG TGCCGGTGGTGTCCCATctcctgttgggtcttctcacgggcagggcg gggggggcctgggggggcggaggtggtggccatgctggagccggtggcgctct tctgggggggcggggctggggaaggccacggaggggctgggggtctcgtcgct cacctcctccacctcgctggcttcctcgtccaggctctcccagtcttccatcatgtc ctcctcctcctcgtccagcacctcctcggcctggctgtcccagtattcctcctcgtc ggtgggggggggggcagctgcagcttcttcttgggggccatcctgggaagca agggcccgcggctgctgctgatagggctgcggcggctgtgttctggcggcaaa cccgttgcgaaaaagaacgttcacggcgactactgcacttatatacggttctccc ccaccctcgggaaaaaggcggagccagtacacgacatcactttcccagtttac cccgcgccaccttctctaggcaccggttcaattgccgacccctccccccaacttc tcggggactgtgggcgatgtgcgctctgcccactgacgggcaccggagccggt aagctccgcattaggcgggcgggattggtcttcgtagaacctaatcATTTAA ATatctagtcgacGTTAATTcgagctcggtacctcgcgaatacatctagat atcggatccgatttaATTAAGGGCGGAGTAACTTGTATGTGTT GGGAATTGTAGTTTTCTTAAAATGGGAAGTGACGTAAC GTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTC CCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAAT TAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTA GTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATC CCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCC GAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGT GAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAATG CCACCATGGTTCTTCCAGCTCTTCCCGCTCCTCCCGT GTGTGACTCGCAGAACGAATGTGTAGGTTGGCTGGGT GTGGCTTATTCTGCGGTGGTGGATGTTATCAGGGCAG CGGCGCATGAAGGAGTTTACATAGAACCCGAAGCCAG GGGGCGCCTGGATGCTTTGAGAGAGTGGATATACTAC AACTACTACACAGAGCGATCTAAGCGGCGAGACCGGA GACGCAGATCTGTTTGTCACGCCCGCACCTGGTTTTG CTTCAGGAAATATGACTACGTCCGGCGTTCCATTTGG CATGACACTACGACCAACACGATCTCGGTTGTCTCGG CGCACTCCGTACAGTAGGGATCGTCTACCTCCTTTTG AGACAGAAACCCGCGCTACCATACTGGAGGATCATCC GCTGCTGCCCGAATGTAACACTTTGACAATGCACAAC GTGAGTTACGTGCGAGGTCTTCCCTGCAGTGTGGGAT TTACGCTGATTCAGGAATGGGTTGTTCCCTGGGATAT

		GGTTCTAACGCGGGAGGAGCTTGTAATCCTGAGGAAG TGTATGCACGTGTGCCTGTGTTGTGCCAACATTGATAT CATGACGAGCATGATGATCCATGGTTACGAGTCCTGG GCTCTCCACTGTCATTGTTCCAGTCCCGGTTCCCTGC AGTGTATAGCCGGCGGGCAGGTTTTGGCCAGCTGGT TTAGGATGGTGGTGGATGGCGCCATGTTTAATCAGAG GTTTATATGGTACCGGGAGGTGGTGAATTACAACATG CCAAAAGAGGTAATGTTTATGTCCAGCGTGTTTATGAG GGGTCGCCACTTAATCTACCTGCGCTTGTGGTATGAT GGCCACGTGGGTTCTGTGGTCCCCGCCATGAGCTTT GGATACAGCGCCTTGCACTGTGGGATTTTGAACAATA TTGTGGTGCTGTGCTGCAGTTACTGTGCTGATTTAAGT GAGATCAGGGTGCGCTGCTGTGCCCGGAGGACAAGG CGCCTTATGCTGCGGGCGGTGCGAATCATCGCTGAG GAGACCACTGCCATGTTGTATTCCTGCAGGACGGAGC GGCGGCGGCAGCAGTTTATTCGCGCGCTGCTGCAGC ACCACCGCCCTATCCTGATGCACGATTATGACTCTAC CCCCATGTAAAtAAatatctttattttCAttacatctGTGTGTTGGTT TTTTGTGTGGCGGCCAACAAGCTTTACcctatccgtagatgta cctggacatccaggtgatgccggcggcggtggtggaggcgcgcggaaagtc gcggacgcggttccagatgttgcgcagcggcaaaaagtgctccatggtcggg acgctctggccggtgaggcgtgcgcagtcgttgacgctctagaccgtgcaaaa ggagagcctgtaagcgggcactcttccgtggtctggtggataaattcgcaaggg tatcatggcggacgaccggggttcgaaccccggatccggccgtccgccgtgat ccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtcagacaacg ggggagcgctccttttggcttccttccaggcgcggcggctgctgcgctagcttttt ggccactggccgcgcgcggcgtaagcggttaggctggaaagcgaaagcatt aagtggctcgctccctgtagccggagggttattttccaagggttgagtcgcagga ccccccggttcgagtctcgggccggccggactgcggcgaacgggggtttgcctc cccgtcatgcaagaccccgcttgcaaattcctccggaaacagggacgagccc ctttttttgcttttcccagatgcatccggtgctgcggcagatgcgccccccctcctcag cagcggcaagagcaagagcagcggcagacatgcagggcaccctccccttct cctaccgcgtcaggaggggcaacatcgatccagacatgataagatacattgat gagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaa atttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagtttGTA CACTCTCGGGTGATTATTTACCCCCACCCTTGCCGTC TGCGCCGTTTAAAAATCAAAGGGGTTCTGCCGCGCAT CGCTATGCGCCACTGGCAGGGACACGTTGCGATACT GGTGTTTAGTGCTCCACTTAAACTCAGGCACAACCAT CCGCGGCAGCTCGGTGAAGTTTTCACTCCACAGGCTG CGCACCATCACCAACGCGTTTAGCAGGTCGGGCGCC GATATCTTGAAGTCGCAGTTGGGGCCTCCGCCCTGCG CGCGCGAGTTGCGATACACAGGGTTGCAGCACTGGA ACACTATCAGCGCCGGGTGGTGCACGCTGGCCAGCA CGCTCTTGTCGGAGATCAGATCCGCGTCCAGGTCCTC CGCGTTGCTCAGGGCGAACGGAGTCAACTTTGGTAG CTGCCTTCCCAAAAAGGGCGCGTGCCCAGGCTTTGA

		GTTGCACTCGCACCGTAGTGGCATCAAAAGGTGACCG TGCCCGGTCTGGGCGTTAGGATACAGCGCCTGCATAA AAGCCTTGATCTGCTTAAAAGCCACCTGAGCCTTTGC GCCTTCAGAGAAGAACATGCCGCAAGACTTGCCGGAA AACTGATTGGCCGGACAGGCCGCGTCGTGCACGCAG CACCTTGCGTCGGTGTTGGAGATCTGCACCACATTTC GGCCCCACCGGTTCTTCACGATCTTGGCCTTGCTAGA CTGCTCCTTCAGCGCGCGCTGCCCGTTTTCGCTCGTC ACATCCATTTCAATCACGTGCTCCTTATTTATCATAATG CTTCCGTGTAGACACTTAAGCTCGCCTTCGATCTCAG CGCAGCGGTGCAGCCACAACGCGCAGCCCGTGGGCT CGTGATGCTTGTAGGTCACCTCTGCAAACGACTGCAG GTACGCCTGCAGGAATCGCCCCATCATCGTCACAAAG GTCTTGTTGCTGGTGAAGGTCAGCTGCAACCCGCGGT GCTCCTCGTTCAGCCAGGTCTTGCATACGGCCGCCAG AGCTTCCACTTGGTCAGGCAGTAGTTTGAAGTTCGCC TTTAGATCGTTATCCACGTGGTACTTGTCCATCAGCGC GCGCGCAGCCTCCATGCCCTTCTCCCACGCAGACAC GATCGGCACACTCAGCGGGTTCATCACCGTAATTTCA CTTTCCGCTTCGCTGGGCTCTTCCTCTTCCTCTTGCGT CCGCATACCACGCGCCACTGGGTCGTCTTCATTCAGC CGCCGCACTGTGCGCTTACCTCCTTTGCCATGCTTGA TTAGCACCGGTGGGTTGCTGAAACCCACCATTTGTAG CGCCACATCTTCTCTTTCTTCCTCGCTGTCCACGATTA CCTCTGGTGATGGCGGGCGCTCGGGCTTGGGAGAAG GGCGCTTCTTTTTCTTCTTGGGCGCAATGGCCAAATC CGCCGCCGAGGTCGATGGCCGCGGGCTGGGTGTGC GCGGCACCAGCGCGTCTTGTGATGAGTCTTCCTCGTC CTCGGACTCGATACGCCGCCTCATCCGCTTTTTTGGG GGCGCCCGGGGAGGCGGCGGCGACGGGGACGGGG ACGACACGTCCTCCATGGTTGGGGGACGTCGCGCCG CACCGCGTCCGCGCTCGGGGGTGGTTTCGCGCTGCT CCTCTTCCCGACTGGCCATttccttctcctataggcagaaaaagat cCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAA GACGCGGAGGCTCTCTTCAGTAAATACTGCGCGCTGA CTCTTAAGGACTAGTTTCGCGCCCTTTCTCAAATTTAA GCGCGAAAACTACGTCATCTCCAGCGGCCACACCCG GCGCCAGCACCTGTTGTCAGCGCCATTGGCGCGCCG GCCGGCCGAAtatcTTCATTTAAATgattaggttctacgaagacca atcccgcccgcctaatgcggagcttaccgcctgcgtcattacccagggccacat tcttggccaattgcaagccatcaacaaagcccgccaagagtttctgctacgaaa gggacgggggtttacttggacccccagtccggcgaggagctcaacccaatc cccccgccgccgcagccctatcagcagcagccgcgggcccttgcttcccagg atggcacccaaaaagaagctgcagctgccgccgccacccacggacgagga ggaatactgggacagtcaggcagaggaggttttggacgaggaggaggagga catgatggaagactgggagagcctagacgaggaagcttccgaggtcgaaga ggtgtcagacgaaacaccgtcaccctcggtcgcattcccctcgccggcgcccc

		agaaatcggcaaccggttccagcatggctacaacctccgctcctcaggcgccg ccggcactgcccgttcgccgacccaaccgtagATGGGACACCACTG GAACCAGGGCCGGTAAGTCCAAGCAGCCGCCGCCGT TAGCCCAAGAGCAACAACAGCGCCAAGGCTACCGCT CATGGCGCGGGCACAAGAACGCCATAGTTGCTTGCTT GCAAGACTGTGGGGGCAACATCTCCTTCGCCCGCCG CTTTCTTCTCTACCATCACGGCGTGGCCTTCCCCCGT AACATCCTGCATTACTACCGTCATCTCTACAGCCCATA CTGCACCGGCGGCAGCGGCAGCAACAGCAGCGGCC ACACAGAAGCAAAGGCGACCGGATAGCAAGACTCTGA CAAAGCCCAAGAAATCCACAGCGGCGGCAGCAGCAG GAGGAGGAGCGCTGCGTCTGGCGCCCAACGAACCCG TATCGACCCGCGAGCTTAGAAACAGGATTTTTCCCAC TCTGTATGCTATATTTCAACAGAGCAGGGGCCAAGAA CAAGAGCTGAAAATAAAAAACAGGTCTCTGCGATCCC TCACCCGCAGCTGCCTGTATCACAAAAGCGAAGATCA GCTTCGGCGCACGCTGGAAGACGCGGAGGCTCTCTT CAGTAAATACTGCGCGCTGACTCTTAAGGACTAGaataa aagatctttattttcattagatctgtgtgttggtttttgtgtgGCGGCC

132	pEMBRsm12-ef1a-33R	CAGTACTCATATGTTGACAATTAATCATCGGCATAGTA TATCGGCATAGTATAATACGACTCACTATAGGAGGGC CACCAAGGTGAGGAACTAAACCATGAGCCATATTCAA CGGGAAACGTCGAGGCCGCGATTAAATTCCAACATGG ATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT GTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATG GGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGG CAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATG GTCAGACTAAACTGGCTGACGGAATTTATGCCACTTC CGACCATCAAGCATTTTATCCGTACTCCTGATGATGCA TGGTTACTCACCACTGCGATCCCCGGAAAAACAGCGT TCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAAT ATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGC ACTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGAT CGCGTATTTCGCCTCGCTCAGGCGCAATCACGAATGA ATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGA GCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAA ATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGT CACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGA CGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGA GTCGGAATCGCAGACCGATACCAGGATCTTGCCATCC TATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAG AAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGA TATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGT TTTTCTAATAAAAACCCGCTTCGGCGGGTTTTTTTATG CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT CACGTTAAGGGATTTTGGTCATGtcgactagatATTTAAATg attaggttctacgaagaccaatcccgcccgcctaatgcggagcttaccggctcc ggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgg

		ggggaggggtcggcaattgaaccggtgcctagagaaggtggcgcggggtaa actgggaaagtgatgtcgtgtactggctccgccttttcccgagggtgggggaga accgtatataagtgcagtagtcgccgtgaacgttcttttcgcaacgggtttgccg ccagaacacagccgccgcagccctatcagcagcagccgcgggcccttgcttc ccaggatggcccccaagaagaagctgcagctgccccccccccccaccgacg aggaggaatactgggacagccaggccgaggaggtgctggacgaggagga ggaggacatgatggaagactgggagagcctggacgaggaagccagcgag gtggaggaggtgagcgacgagacccccagcccctccgtggccttccccagcc ccgccccccagaagagcgccaccggctccagcatggccaccacctccgccc cccaggcccccccccgccctgcccgtgagaagacccaacaggagATGGG ACACCACCGGCACCAGGGCCGCCCACACCGCCCCCG CCGCAGCTGCAGCAACAGCAGCTGCCACCCAGAAGC AGAGGAGACCTGATAGCAAGACCCTGACCAAGCCCAA GAAAAGCACCGCCGCCGCTGCCGCAGGGGGCGGAG CCCTGAGACTGGCCCCCAACGAGCCCGTGAGCACCA GGGAGCTGAGAAACAGGATTTTTCCCACCCTGTACGC CATCTTTCAGCAGAGCAGGGGCCAGGAGCAGGAGCT GAAGATCAAGAACAGGAGCCTGAGAAGCCTGACCAG AAGCTGCCTGTACCACAAGAGCGAGGACCAGCTGAG AAGGACACTGGAAGATGCCGAGGCTCTGTTTAGCAAG TATTGCGCCCTGACCCTGAAGGACTAGaataaaagatctttat tttcattagatctgtgtgttggttttttgtgtgGCGGCCtcgacGTTAATTcga gctcggtacctcgcgaatacatctagatatcggatccgatttaATTAAGGG CGGAGTAACTTGTATGTGTTGGGAATTGTAGTTTTCTT AAAATGGGAAGTGACGTAACGTGTGTGTCAGTTAGGG TGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGT ATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGT GTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTA TGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGC CCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAAT TTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCC TCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTG GAGGCCTAGGCTTTTGCAAATGCCACCATGGTTCTTC CAGCTCTTCCCGCTCCTCCCGTGTGTGACTCGCAGAA CGAATGTGTAGGTTGGCTGGGTGTGGCTTATTCTGCG GTGGTGGATGTTATCAGGGCAGCGGCGCATGAAGGA GTTTACATAGAACCCGAAGCCAGGGGGCGCCTGGAT GCTTTGAGAGAGTGGATATACTACAACTACTACACAGA GCGATCTAAGCGGCGAGACCGGAGACGCAGATCTGT TTGTCACGCCCGCACCTGGTTTTGCTTCAGGAAATAT GACTACGTCCGGCGTTCCATTTGGCATGACACTACGA CCAACACGATCTCGGTTGTCTCGGCGCACTCCGTACA GTAGGGATCGTCTACCTCCTTTTGAGACAGAAACCCG CGCTACCATACTGGAGGATCATCCGCTGCTGCCCGAA TGTAACACTTTGACAATGCACAACGTGAGTTACGTGC GAGGTCTTCCCTGCAGTGTGGGATTTACGCTGATTCA

		GGAATGGGTTGTTCCCTGGGATATGGTTCTAACGCGG GAGGAGCTTGTAATCCTGAGGAAGTGTATGCACGTGT GCCTGTGTTGTGCCAACATTGATATCATGACGAGCAT GATGATCCATGGTTACGAGTCCTGGGCTCTCCACTGT CATTGTTCCAGTCCCGGTTCCCTGCAGTGTATAGCCG GCGGGCAGGTTTTGGCCAGCTGGTTTAGGATGGTGG TGGATGGCGCCATGTTTAATCAGAGGTTTATATGGTAC CGGGAGGTGGTGAATTACAACATGCCAAAAGAGGTAA TGTTTATGTCCAGCGTGTTTATGAGGGGTCGCCACTT AATCTACCTGCGCTTGTGGTATGATGGCCACGTGGGT TCTGTGGTCCCCGCCATGAGCTTTGGATACAGCGCCT TGCACTGTGGGATTTTGAACAATATTGTGGTGCTGTG CTGCAGTTACTGTGCTGATTTAAGTGAGATCAGGGTG CGCTGCTGTGCCCGGAGGACAAGGCGCCTTATGCTG CGGGCGGTGCGAATCATCGCTGAGGAGACCACTGCC ATGTTGTATTCCTGCAGGACGGAGCGGCGGCGGCAG CAGTTTATTCGCGCGCTGCTGCAGCACCACCGCCCTA TCCTGATGCACGATTATGACTCTACCCCCATGTAAAtA AatatctttattttCAttacatctGTGTGTTGGTTTTTTGTGTGGCG GCCAACAAGCTTTACcctatccgtagatgtacctggacatccaggtga tgccggcggcggtggtggaggcgcgcggaaagtcgcggacgcggttccaga tgttgcgcagcggcaaaaagtgctccatggtcgggacgctctggccggtgagg cgtgcgcagtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgg gcactcttccgtggtctggtggataaattcgcaagggtatcatggcggacgacc ggggttcgaaccccggatccggccgtccgccgtgatccatgcggttaccgccc gcgtgtcgaacccaggtgtgcgacgtcagacaacgggggagcgctcctttttgg cttccttccaggcgcggcggctgctgcgctagctttttttggccactggccgcgcgc ggcgtaagcggttaggctggaaagcgaaagcattaagtggctcgctccctgta gccggagggttattttccaagggttgagtcgcaggacccccggttcgagtctcg ggccggccggactgcggcgaacgggggtttgcctccccgtcatgcaagaccc cgcttgcaaattcctccggaaacagggacgagccccttttttgcttttcccagatg catccggtgctgcggcagatgcgcccccctcctcagcagcggcaagagcaag agcagcggcagacatgcagggcaccctcccccttctcctaccgcgtcaggagg ggcaacatcgatccagacatgataagatacattgatgagtttggacaaaccac aactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttat ttgtaaccattataagctgcaataaacaagtttGTACACTCTCGGGTG ATTATTTACCCCCACCCTTGCCGTCTGCGCCGTTTAAA AATCAAAGGGGTTCTGCCGCGCATCGCTATGCGCCAC TGGCAGGGACACGTTGCGATACTGGTGTTTAGTGCTC CACTTAAACTCAGGCACAACCATCCGCGGCAGCTCGG TGAAGTTTTCACTCCACAGGCTGCGCACCATCACCAA CGCGTTTAGCAGGTCGGGCGCCGATATCTTGAAGTCG CAGTTGGGGCCTCCGCCCTGCGCGCGCGAGTTGCGA TACACAGGGTTGCAGCACTGGAACACTATCAGCGCCG GGTGGTGCACGCTGGCCAGCACGCTCTTGTCGGAGA TCAGATCCGCGTCCAGGTCCTCCGCGTTGCTCAGGG CGAACGGAGTCAACTTTGGTAGCTGCCTTCCCAAAAA

		GGGCGCGTGCCCAGGCTTTGAGTTGCACTCGCACCG TAGTGGCATCAAAAGGTGACCGTGCCCGGTCTGGGC GTTAGGATACAGCGCCTGCATAAAAGCCTTGATCTGC TTAAAAGCCACCTGAGCCTTTGCGCCTTCAGAGAAGA ACATGCCGCAAGACTTGCCGGAAAACTGATTGGCCGG ACAGGCCGCGTCGTGCACGCAGCACCTTGCGTCGGT GTTGGAGATCTGCACCACATTTCGGCCCCACCGGTTC TTCACGATCTTGGCCTTGCTAGACTGCTCCTTCAGCG CGCGCTGCCCGTTTTCGCTCGTCACATCCATTTCAAT CACGTGCTCCTTATTTATCATAATGCTTCCGTGTAGAC ACTTAAGCTCGCCTTCGATCTCAGCGCAGCGGTGCAG CCACAACGCGCAGCCCGTGGGCTCGTGATGCTTGTA GGTCACCTCTGCAAACGACTGCAGGTACGCCTGCAG GAATCGCCCCATCATCGTCACAAAGGTCTTGTTGCTG GTGAAGGTCAGCTGCAACCCGCGGTGCTCCTCGTTCA GCCAGGTCTTGCATACGGCCGCCAGAGCTTCCACTTG GTCAGGCAGTAGTTTGAAGTTCGCCTTTAGATCGTTAT CCACGTGGTACTTGTCCATCAGCGCGCGCGCAGCCT CCATGCCCTTCTCCCACGCAGACACGATCGGCACACT CAGCGGGTTCATCACCGTAATTTCACTTTCCGCTTCG CTGGGCTCTTCCTCTTCCTCTTGCGTCCGCATACCAC GCGCCACTGGGTCGTCTTCATTCAGCCGCCGCACTGT GCGCTTACCTCCTTTGCCATGCTTGATTAGCACCGGT GGGTTGCTGAAACCCACCATTTGTAGCGCCACATCTT CTCTTTCTTCCTCGCTGTCCACGATTACCTCTGGTGAT GGCGGGCGCTCGGGCTTGGGAGAAGGGCGCTTCTTT TTCTTCTTGGGCGCAATGGCCAAATCCGCCGCCGAGG TCGATGGCCGCGGGCTGGGTGTGCGCGGCACCAGC GCGTCTTGTGATGAGTCTTCCTCGTCCTCGGACTCGA TACGCCGCCTCATCCGCTTTTTTGGGGGCGCCCGGG GAGGCGGCGGCGACGGGGACGGGGACGACACGTCC TCCATGGTTGGGGGACGTCGCGCCGCACCGCGTCCG CGCTCGGGGGTGGTTTCGCGCTGCTCCTCTTCCCGA CTGGCCATttccttctcctataggcagaaaaagatcCACAAAAGCG AAGATCAGCTTCGGCGCACGCTGGAAGACGCGGAGG CTCTCTTCAGTAAATACTGCGCGCTGACTCTTAAGGAC TAGTTTCGCGCCCTTTCTCAAATTTAAGCGCGAAAACT ACGTCATCTCCAGCGGCCACACCCGGCGCCAGCACC TGTTGTCAGCGCCATTGGCGCGCCGGCCGGCCGAAta tcTTCATTTAAATgattaggttctacgaagaccaatcccgcccgcctaatg cggagcttaccgcctgcgtcattacccagggccacattcttggccaattgcaag ccatcaacaaagcccgccaagagtttctgctacgaaagggacgggggttta cttggacccccagtccggcgaggagctcaacccaatccccccgccgccgcag ccctatcagcagcagccgcgggcccttgcttcccaggatggcacccaaaaag aagctgcagctgccgccgccacccacggacgaggaggaatactgggacagt caggcagaggaggttttggacgaggaggaggaggacatgatggaagactgg gagagcctagacgaggaagcttccgaggtcgaagaggtgtcagacgaaaca

		ccgtcaccctcggtcgcattcccctcgccggcgccccagaaatcggcaaccgg ttccagcatggctacaacctccgctcctcaggcgccgccggcactgcccgttcg ccgacccaaccgtagATGGGACACCACTGGAACCAGGGCC GGTAAGTCCAAGCAGCCGCCGCCGTTAGCCCAAGAG CAACAACAGCGCCAAGGCTACCGCTCATGGCGCGGG CACAAGAACGCCATAGTTGCTTGCTTGCAAGACTGTG GGGGCAACATCTCCTTCGCCCGCCGCTTTCTTCTCTA CCATCACGGCGTGGCCTTCCCCCGTAACATCCTGCAT TACTACCGTCATCTCTACAGCCCATACTGCACCGGCG GCAGCGGCAGCAACAGCAGCGGCCACACAGAAGCAA AGGCGACCGGATAGCAAGACTCTGACAAAGCCCAAG AAATCCACAGCGGCGGCAGCAGCAGGAGGAGGAGCG CTGCGTCTGGCGCCCAACGAACCCGTATCGACCCGC GAGCTTAGAAACAGGATTTTTCCCACTCTGTATGCTAT ATTTCAACAGAGCAGGGGCCAAGAACAAGAGCTGAAA ATAAAAAACAGGTCTCTGCGATCCCTCACCCGCAGCT GCCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCAC GCTGGAAGACGCGGAGGCTCTCTTCAGTAAATACTGC GCGCTGACTCTTAAGGACTAGaataaaagatctttattttcattaga tctgtgtgttggttttttgtgtgGCGGCC

133	CBA-UL30-UL42-UL29 with P2A clevage	TTAATTAAAGGCCTtcgaggtgagccccacgttctgcttcactctcccca tctccccccccctccccaccccaattttgtatttatttatttttaattattttgtgcagcg atgggggcgggggggggggggggcgcgcgccaggcggggcggggcgg ggcgaggggcggggcggggcgaggcggagaggtgcggcggcagccaatc agagcggcgcgctccgaaagtttccttttatggcgaggcggcggcggcggcg gccctataaaaagcgaagcgcgcggcgggcgCGCCGGCGcaccgcg ggcccgatccaccggtactgttggtaaagccaccatgttttccggtggcggcgg cccgctgtcccccggaggaaagtcggcggccagggcggcgtccgggtttttg cgcccgccggccctcgcggagccagccggggacccccgccttgtttgaggca aaacttttacaacccctacctcgccccagtcgggacgcaacagaagccgacc gggccaacccagcgccatacgtactatagcgaatgcgatgaatttcgattcatc gccccgcgggtgctggacgaggatgcccccccggagaagcgcgccggggt gcacgacggtcacctcaagcgcgcccccaaggtgtactgcgggggggacga gcgcgacgtcctccgcgtcgggtcgggcggcttctggccgcggcgctcgcgcc tgtggggcggcgtggaccacgccccggcggggttcaaccccaccgtcaccgt ctttcacgtgtacgacatcctggagaacgtggagcacgcgtacggcatgcgcg cggcccagttccacgcgcggtttatggacgccatcacaccgacggggaccgtc atcacgctcctgggcctgactccggaaggccaccgggtggccgttcacgtttac ggcacgcggcagtacttttacatgaacaaggaggaggtcgacaggcacctac aatgccgcgccccacgagatctctgcgagcgcatggccgcggccctgcgcga gtccccgggcgcgtcgttccgcggcatctccgcggaccacttcgaggcggagg tggtggagcgcaccgacgtgtactactacgagacgcgccccgctctgttttacc gcgtctacgtccgaagcgggcgcgtgctgtcgtacctgtgcgacaacttctgcc cggccatcaagaagtacgagggtggggtcgacgccaccacccggttcatcct ggacaaccccgggttcgtcaccttcggctggtaccgtctcaaaccgggccgga acaacacgctagcccagccgcgggccccgatggccttcgggacatccagcg acgtcgagtttaactgtacggcggacaacctggccatcgaggggggcatgag cgacctaccggcatacaagctcatgtgcttcgatatcgaatgcaaggcggggg gggaggacgagctggcctttccggtggccgggcacccggaggacctggtcat ccagatatcctgtctgctctacgacctgtccaccaccgccctggagcacgtcctc ctgttttcgctcggttcctgcgacctccccgaatcccacctgaacgagctggcgg ccaggggcctgcccacgcccgtggttctggaattcgacagcgaattcgagatg ctgttggccttcatgacccttgtgaaacagtacggccccgagttcgtgaccgggt acaacatcatcaacttcgactggcccttcttgctggccaagctgacggacattta caaggtccccctggacgggtacggccgcatgaacggccggggcgtgtttcgc gtgtgggacataggccagagccacttccagaagcgcagcaagataaaggtg aacggcatggtgaacatcgacatgtacgggattataaccgacaagatcaagct ctcgagctacaagctcaacgccgtggccgaagccgtcctgaaggacaagaa gaaggacctgagctatcgcgacatccccgcctactacgccgccgggcccgcg caacgcggggtgatcggcgagtactgcatacaggattccctgctggtgggcca gctgtttttttaagtttttgccccatctggagctctcggccgtcgcgcgcttggcgggt attaacatcacccgcaccatctacgacggccagcagatccgcgtctttacgtgc ctgctgcgcctggccgaccagaagggctttattctgccggacacccaggggcg atttaggggcgccgggggggaggcgcccaagcgtccggccgcagcccggg aggacgaggagcggccagaggaggaggggggaggacgaggacgaacgc gaggagggcgggggcgagcgggagccggagggcgcgcgggagaccgcc ggcaggcacgtggggtaccaggggggccagggtccttgaccccacttccgggtt

		tcacgtgaaccccgtggtggtgttcgactttgccagcctgtaccccagcatcatcc aggcccacaacctgtgcttcagcacgctctccctgagggccgacgcagtggcg cacctggaggcgggcaaggactacctggagatcgaggtgggggggcgacg gctgttcttcgtcaaggctcacgtgcgagagagcctcctcagcatcctcctgcgg gactggctcgccatgcgaaagcagatccgctcgcggattccccagagcagcc ccgaggaggccgtgctcctggacaagcagcaggccgccatcaaggtcgtgtg taactcggtgtacgggttcacgggagtgcagcacggactcctgccgtgcctgca cgttgccgcgacggtgacgaccatcggccgcgagatgctgctcgcgacccgc gagtacgtccacgcgcgctgggcggccttcgaacagctcctggccgatttcccg gaggcggccgacatgcgcgcccccgggccctattccatgcgcatcatctacgg ggacacggactccatctttgtgctgtgccgcggcctcacggccgccgggctgac ggccgtgggcgacaagatggcgagccacatctcgcgcgcgctgtttctgcccc ccatcaaactcgagtgcgaaaagacgttcaccaagctgctgctgatcgccaag aaaaagtacatcggcgtcatctacgggggtaagatgctcatcaagggcgtgga tctggtgcgcaaaaacaactgcgcgtttatcaaccgcacctccagggccctggt cgacctgctgttttacgacgataccgtctccggagcCgccgcGgcgttagccg agcgccccgcggaggagtggctggcgcgacccctgcccgagggactgcag gcgttcggggccgtcctcgtagacgcccatcggcgcatcaccgacccggaga gggacatccaggactttgtcctcaccgccgaactgagcagacacccgcgcgc gtacaccaacaagcgcctggcccacctgacggtgtattacaagctcatggccc gccgcgcgcaggtcccgtccatcaaggaccggatcccgtacgtgatcgtggcc cagacccgcgaggtagaggagacggtcgcgcggctggccgccctccgcga gctagacgccgccgccccaggggacgagcccgcccccccgcggccctgc cctccccggccaagcgcccccgggagacgccgtcgcctgccgacccccgg gaggcgcgtccaagccccgcaagctgctggtgtccgagctggccgaggatcc cgcatacgccattgcccacggcgtcgccctgaacacggactattacttctccca cctgttgggggcggcgtgcgtgacattcaaggccctgtttgggaataacgccaa gatcaccgagagtctgttaaaaaggtttattcccgaagtgtggcacccccggga cgacgtggccgcgcggctccggaccgcagggttcggggcggtgggtgccgg cgctacggcggaggaaactcgtcgaatgttgcatagagcctttgatactctagc agaattcggcagtggagcaacaaacttctctctgctgaaacaagccggagatg tcgaagagaatcctggaccgacggattcccctggcggtgtggcccccgcctcc cccgtggaggacgcgtcggacgcgtccctcgggcagccggaggagggggc gccctgccaggtggtcctgcagggcgccgaacttaatggaatcctacaggcgtt tgccccgctgcgcacgagccttctggactcgcttctggttatgggcgaccgggg catccttatccataacacgatctttggggagcaggtgttcctgcccctggaacact cgcaattcagtcggtatcgctggcgcggacccacggcggcgttcctgtctctcgt ggaccagaagcgctccctcctgagcgtgtttcgcgccaaccagtacccggacc tacgtcgggtggagttggcgatcacgggccaggccccgtttcgcacgctggttc agcgcatatggacgacgacgtccgacggcgaggccgttgagctagccagcg agacgctgatgaagcgcgaactgacgagctttgtggtgctggttccccaggga acccccgacgttcagttgcgcctgacgaggccgcagctcaccaaggtccttaa cgcgaccggggccgatagtgccacgcccaccacgttcgagctcggggttaac ggcaaattttccgtgttcaccacgagtacctgcgtcacctttgctgcccgcgagg agggcgtgtcgtccagcaccagcacccaggtccagatcctgtccaacgcgctc accaaggcgggccaggcCgccgcGaacgccaagacggtgtacggggaa aatacccatcgcaccttctctgtggtcgtcgacgattgcagcatgcgggcggtgc

		tccggcgactgcaggtcggcgggggcaccctcaagttcttcctcacgaccccc gtccccagtctgtgcgtcaccgccaccggtcccaacgcggtatcggcggtatttc tcctgaaaccccagaagatttgcctggactggctgggtcatagccaggggtctc cttcagccgggagctcggcctcccgggcctctgggagcgagccaacagaca gccaggactccgcgtcggacgcggtcagccacggcgatccggaagacctcg atggcgctgcccgggcgggagaggcgggggccttgcatgcctgtccgatgcc gtcgtcgaccacgcgggtcactcccacgaccaagcgggggcgctcgggggg cgaggatgcgcgcgcggacacggccctaaagaaacctaagacggggtcgc ccaccgcacccccgcccgcagatccagtccccctggacacggaggacgact ccgatgcggcggacgggacggcggcccgtcccgccgctccagacgcccgg agcggaagccgttacgcgtgttactttcgcgacctcccgaccggagaagcaag ccccggcgccttctccgccttccgggggggcccccaaaccccgtatggttttgg attccccGCCACGAACTTCTCTCTGTTAAAGCAAGCAGGA GACGTGGAAGAAAACCCCGGTCCTatggagacaaagcccaa gacggcaaccaccatcaaggtccccccccgggcccctgggatacgtgtacgct cgcgcgtgtccgtccgaaggcatcgagcttctggcgttactgtcggcacgcagc ggcgattccgacgtcgccgtggcgcccctggtcgtgggcctgaccgtggagag cggctttgaggccaacgtggccgtggtcgtgggttctcgcacgacggggctcgg gggtaccgcggtgtccctgaaactgacgccctcgcactacagctcgtccgtgta cgtctttcacggcggccggcacctggaccccagcacccaggccccgaacctg acgcgactttgcgagcgggcacgccgccatttttggcttttcggactacacccccc ggcccggcgacctcaaacacgagacgacgggggaggcgctgtgtgagcgc ctcggcctggacccggaccgcgccctcctgtatctggtcgttaccgagggcttca aggaggccgtgtgcatcaacaacacctttctgcacctgggaggctcggacaag gtaaccataggcggggcggaggtgcaccgcatacccgtgtacccgttgcagct gttcatgccggattttagccgtgtcatcgcagagccgttcaacgccaaccaccg atcgatcggggagaatttttacctacccgcttccgttttttaaccgccccctcaaccg cctcctgttcgaggcggtcgtgggacccgccgccgtggcactgcgatgccgaa acgtggacgccgtggcccgcgcCgccgcccacctggcgtttgacgaaaacc acgagggcgccgccctccccgccgacattacgttcacggccttcgaagccag ccagggtaagaccccgcggggcgggcgcgacggcggcggcaagggcccg gcgggcgggttcgaacagcgcctggcctccgtcatggccggagacgccgccc tggccctcgagtctatcgtgtcgatggccgtctttgacgagccgcccaccgacat ctccgcgtggccgctgttcgagggccaggacacggccgcggcccgcgccaa cgccgtcggggcgtacctggcgcgcgccgcgggactcgtgggggccatggta tttagcaccaactcggccctccatctcaccgaggtggacgacgccggcccggc ggacccaaaggaccacagcaaaccctccttttaccgcttcttcctcgtgcccgg gacccacgtggcggccaacccacaggtggaccgcgagggacacgtggtgc ccgggttcgagggtcggcccaccgcgcccctcgtcggcggaacccaggaattt gccggcgagcacctggccatgctgtgtgggttttccccggcgctgctggccaag atgctgttttacctggagcgctgcgacggcggcgtgatcgtcgggcgccagga gatggacgtgtttcgatacgtcgcggactccaaccagaccgacgtgccctgtaa cctatgcaccttcgacacgcgccacgcctgcgtacacacgacgctcatgcgcc tccgggcgcgccatccaaagttcgccagcgccgcccgcggagccatcggcgt cttcgggaccatgaacagcatgtaTagcgactgcgacgtgctgggaaactac gccgccttctcggccctgaagcgcgcggacggatccgagaccgcccggacc atcatgcaggagacgtaccgcgcggcgaccgagcgcgtcatggccgaactc

		gagaccctgcagtacgtggaccaggcggtccccacggccatggggcggctg gagaccatcatcaccaaccgcgaggccctgcatacggtggtgaacaacgtca ggcaggtcgtggaccgcgaggtggagcagctgatgcgcaacctggtggagg ggaggaacttcaagtttcgcgacggtctgggcgaggccaaccacgccatgtcc ctgacgctggacccgtacgcgtgcgggccgtgccccctgcttcagcttctcggg cggcgatccaacctcgccgtgtaccaggacctggccctgagtcagtgccacgg ggtgttcgccgggcagtcggtcgaggggcgcaactttcgcaatcaattccaacc ggtgctgcggcggcgcgtgatggacatgtttaacaacgggtttctgtcggccaa aacgctgacggtcgcgctctcggaggggcggctatctgcgcccccagccta acggcGggccagacggcccccgccgagagcagcttcgagggcgacgttgc ccgcgtgaccctggggtttcccaaggagctgcgcgtcaagagccgcgtgttgtt cgcgggcgcgagcgccaacgcgtccgaggccgccaaggcgcgggtcgcca gcctccagagcgcctaccagaagcccgacaagcgcgtggacatcctcctcgg accgctgggctttctgctCaagcagttccacgcggccatcttccccaacggcaa gcccccggggtccaaccagccgaacccgcagtggttctggacggccctccaa cgcaaccagcttcccgcccggctcctgtcgcgcgaggacatcgagaccatcg cgttcattaaaaagttttccctggactacggcgcgataaactttattaacctggccc ccaacaacgtgagcgagctggcgatgtactacatggcaaaccagattctgcg gtactgcgatcactcgacatacttcatcaacacccttacggccatcatcgcggg gtcccgccgtccccccagcgtgcaggctgcCgccgcgtggtccgcgcagggc ggggcgggcctggaggccggggcccgcgcgctgatggacgccgtggacgc gcatccgggcgcgtggacgtccatgttcgccagctgcaacctgctgcggcccgt catggcggcgcgccccatggtcgtgttggggttgagcatcagcaagtactacg gcatggccggcaacgaccgtgtgtttcaggccgggaactgggccagcctgatg ggcggcaaaaacgcgtgcccgctccttatttttgaccgcacccgcaagttcgtcc tggcctgtccccgggccgggtttgtgtgcgcggcctcaagcctcggcggcgga gcgcacgaaagctcgctgtgcgagcagctccggggcattatctccgagggcg gggcggccgtcgccagtagcgtgttcgtggcgaccgtgaaaagcctggggcc ccgcacccagcagctgcagatcgaggactggctggcgctcctggaggacga gtacctaagcgaggagatgatggagctgaccgcgcgtgccctggagcgcgg caacggcgagtggtcgacggacgcggccctggaggtggcgcacgaggccg aggccctagtcagccaactcggcaacgccggggaggtgtttaactttggggatt ttggctgcgaggacgacaacgcgacgccgttcggcggcccgggggccccgg gaccggcatttgccggccgcaaacgggcgttccacggggatgacccgtttggg gaggggccccccgacaaaaagggagacctgacgttggatatgctgTAGTA AcggcaataaaaagacagaataaaacgcacggtgttgggtcgtttgttcAGG CCTATTAAT

EQUIVALENTS

[0131] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the following claims:

CLAIMS

We claim:

1. An adenoviral helper plasmid comprising a nucleotide sequence encoding:

(a) E2a protein;
(b) an E4 region; and
(c) a VA RNA region;

wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding one or more of:

a Fiber protein or portion thereof;
a L1-52/55K (Packaging Protein 3); and
a Peripentonal Hexon-Associated Protein.

2. The adenoviral plasmid of claim 1, further comprising an engineered L4 region

3. The adenoviral plasmid of claim 2, wherein the engineered L4 region does not express an L4 100K protein.

4. The adenoviral plasmid of claim 2, wherein the engineered L4 region does not comprise a complete L4 100K encoding nucleotide sequence.

5. The adenoviral plasmid of claim 4, wherein the engineered L4 region does not comprise base pairs 24061-25961 relative to GenBank Accession Number M73260.

6. The adenoviral plasmid of claim 2, wherein the engineered L4 region comprises at least one nucleotide sequence encoding L4-33K at least 80% identical to SEQ ID: NO. 5.

7. The adenoviral plasmid of claim 6, wherein the engineered L4 region comprises two nucleotide sequences encoding L4-33K at least 80% identical to SEQ ID: NO. 5.

8. The adenoviral plasmid of claim 2, wherein the engineered L4 region comprises a nucleotide sequence encoding L4-22K at least 80% identical to SEQ ID: NO. 7.

9. The adenoviral plasmid of claim 2, wherein the engineered L4 region comprises a nucleotide sequence encoding L4-33K and a nucleotide sequence encoding L4-22K.

10. The adenoviral plasmid of claim 9, wherein the engineered L4 region comprises 50 nucleotides upstream of the L4-33K/22K region.

11. The adenoviral plasmid of claim 10, wherein the 50 nucleotides upstream of the L4-33K/22K region is identified relative to the L4-33K/22K region of GenBank Accession Number M73260.

12. The adenoviral plasmid of claim 11, further comprising a synthetic polyA signal downstream of the L4-33K/22K region.

13. The adenoviral plasmid of claim 1, wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding L3 23 kDa viral endoprotease.

14. The adenoviral plasmid of claim 1, wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding L4 pVIII.

15. The adenoviral helper plasmid of claim 9, wherein the engineered L4 region comprises a nucleotide sequence comprising an E2a promoter region upstream of the nucleotide sequence encoding L4-33K and the nucleotide sequence encoding L4-22K.

16. The adenoviral helper plasmid of claim 1, wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding DNA Terminal Protein.

17. The adenoviral helper plasmid of claim 1, wherein expression of E2a protein is under the control of an E2a promoter.

18. The adenoviral helper plasmid of claim 1, wherein expression of E2a protein is under the control of an E2a promoter and a chicken β-actin promoter, wherein the chicken β-actin promoter is upstream of the E2a promoter.

19. The adenoviral helper plasmid of claim 1, wherein expression of E2a protein is under the control of the chicken β-actin promoter.

20. The adenoviral helper plasmid of claim 18 or 19, wherein in the chicken β-actin promoter has a nucleotide sequence that is at least 80% identical to SEQ ID. No: 26.

21. The adenoviral helper plasmid of claim 1, wherein the adenoviral helper plasmid comprises an E2a polyadenylation signal downstream of E2a.

22. The adenoviral helper plasmid of claim 1, wherein the adenoviral helper plasmid contains an SV40 polyadenylation signal downstream of E2a.

23. The adenoviral helper plasmid of claim 20, wherein the SV40 polyadenylation signal is downstream of an E2a polyadenylation signal.

24. The adenoviral helper plasmid of claim 21 or 20, wherein the SV40 polyadenylation signal has a sequence that is at least 80% identical to SEQ ID. No: 28.

25. The adenoviral helper plasmid of claim 1, wherein the adenoviral helper plasmid further comprises nucleotide sequences encoding HSV-1 UL30 and HSV-1 UL42,

wherein at least one of the nucleotide sequences is at least 80% identical to SEQ ID NO: 29;

wherein at least one of the nucleotide sequences is at least 80% identical to SEQ ID NO: 31;

wherein UL30 and UL42 are separated by a P2A cleavage site encoded by a nucleic acid sequence that is at least 80% identical to SEQ ID NO: 33.

26. The adenoviral helper plasmid of claim 1, wherein the adenoviral helper plasmid further comprises nucleotide sequences encoding HSV-1 UL29/ICP8.

27. The adenoviral helper plasmid of any one of claims 23 and XX, wherein the HSV-1 UL30; UL42; and UL29 are separated by a P2A cleavage site encoded by a nucleic acid sequence that is at least 80% identical to SEQ ID NO: 139

28. The adenoviral helper plasmid of claim 1, wherein the E4 region does not comprise E4orf1, and wherein the E4 region does not comprise E4orf2.

29. The adenoviral helper plasmid of claim 1, wherein the E4 region is operably linked to the E4 mini promoter, wherein the E4 mini promoter has a nucleotide sequence that is at least 80% identical to SEQ ID NO: 1.

30. The adenoviral helper plasmid of claim 1, wherein the E4 region is operably linked to the SV40 promoter, wherein the SV40 promoter has a nucleotide sequence that is at least 80% identical to SEQ ID NO: 2.

31. The adenoviral helper plasmid of any of the above claims, wherein the adenoviral helper plasmid comprises a resistance gene.

32. The adenoviral helper plasmid of claim 31, wherein the resistance cassette is a kanamycin resistance gene.

33. An adenoviral helper plasmid having 80% sequence identity to any one of SEQ ID NO: 80-105.

34. An adenoviral helper plasmid comprising an engineered E2A region comprising any one of nucleotide sequences SEQ ID NO: 106 or SEQ ID NO: 112.

35. The adenoviral helper plasmid of claim 1 or claim 2, wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding L4 33K.

36. The adenoviral helper plasmid of claim 1 or claim 2, wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding L4 22K.

37. The adenoviral helper plasmid of claim 1 or claim 2, wherein the adenoviral helper plasmid does not comprise a nucleotide sequence encoding L4 22K and does not comprise a nucleotide sequence encoding L4 33K.

38. An L4 *trans* plasmid comprising a nucleotide sequence encoding an engineered L4 region.

39. The L4 *trans* plasmid of claim 38, wherein the engineered L4 region comprises a nucleotide sequence encoding L4 at least 80% identical to SEQ ID NO. 115.

40. The L4 *trans* plasmid of claim 38, wherein the engineered L4 region comprises a nucleotide sequence encoding L4-33K at least 80% identical to SEQ ID NO. 120.

41. The L4 *trans* plasmid of claim 38, wherein the engineered L4 region comprises a nucleotide sequence encoding L4-22K at least 80% identical to SEQ ID NO. 7.

42. A composition comprising:

an adenoviral helper plasmid of any one of claims 35, 36, or 37; and

an L4 *trans* plasmid of any one of claims 38-41.

43. A method of producing a recombinant adenoviral associated viral vector comprising:

transfecting a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, and the adenoviral helper plasmid of any one of claims 1-37.

44. The method of claim 43, wherein the AAV vector plasmid comprises AAV inverted terminal repeats (ITRs) and a transgene of interest.

45. The method of claim 43, wherein the method further comprises transfecting a producer cell with an L4 *trans* plasmid of any one of claims 38-41.

46. The method of any one of claim 43-45 wherein both the L4 *trans* plasmid and the adenoviral helper plasmid comprise an engineered L4 region.

47. The method of claim 46, wherein the engineered L4 region of the L4 *trans* plasmid comprises a nucleotide sequence encoding L33K.

48. A method of producing a recombinant adenoviral associated viral vector comprising:

transfecting a producer cell with an AAV vector plasmid and the adenoviral helper plasmid of any one of claims 1-37, wherein the producer cell stably expresses Rep-Cap and comprises a nucleotide sequence encoding an L4 region or portion thereof.

49. The method of claim 48, wherein the AAV vector plasmid comprises AAV inverted terminal repeats (ITRs) and a transgene of interest.

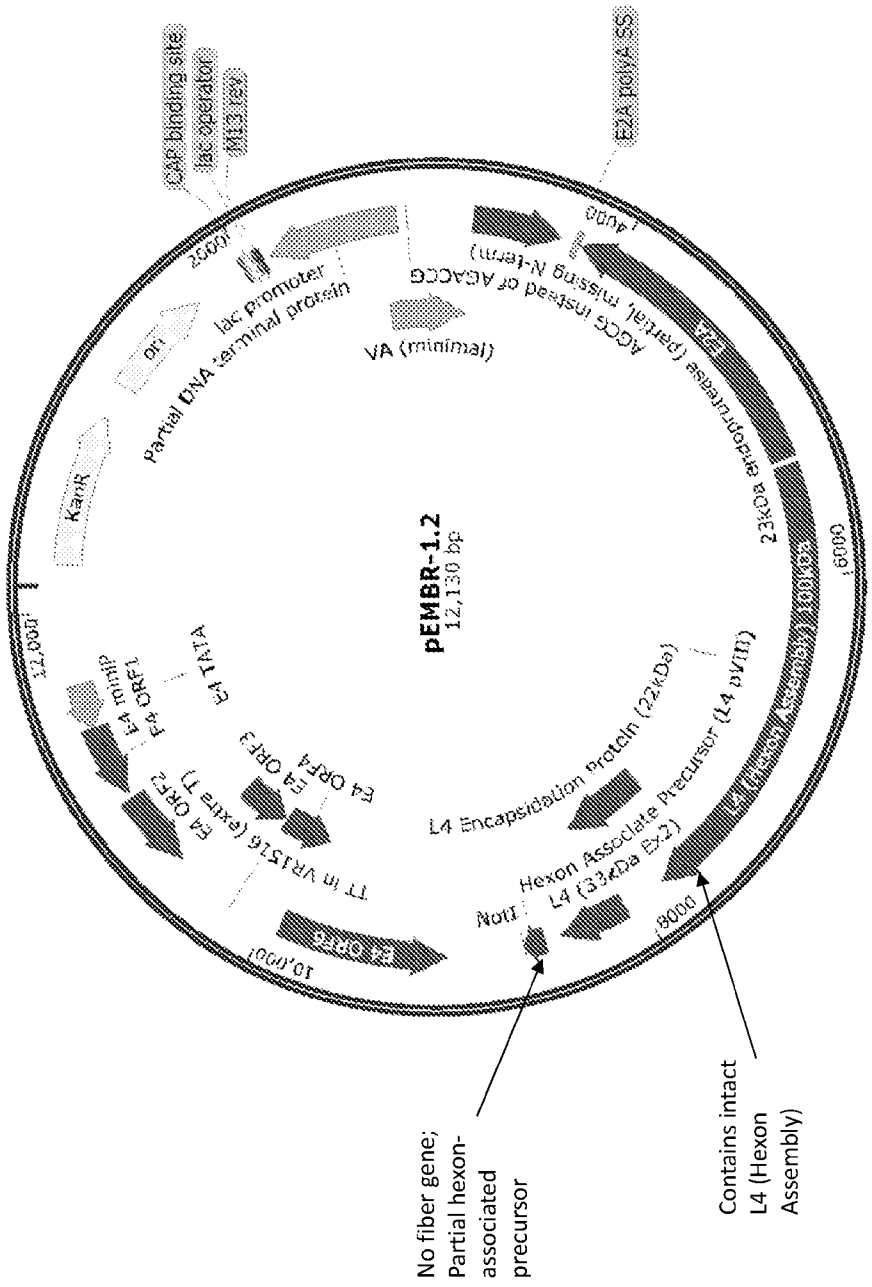
50. The method of claim 49, wherein the method further comprises transfecting a producer cell with an L4 *trans* plasmid of any one of claims 38-41.

51. A method of producing a recombinant adenoviral associated viral vector comprising:

transfecting a producer cell with an AAV vector plasmid, an AAV Rep-Cap expressing plasmid, the adenoviral helper plasmid of any one of claims 1-37, and an L4 *trans* plasmid.

52. The method of 51 wherein, the L4 *trans* plasmid comprises a nucleotide sequence encoding L33K.

53. The method of 51 wherein, the producer cell a nucleotide sequence encoding an L4 region or portion thereof.



No fiber gene; Partial hexon-associated precursor

Contains intact L4 (Hexon Assembly)

Figure 1

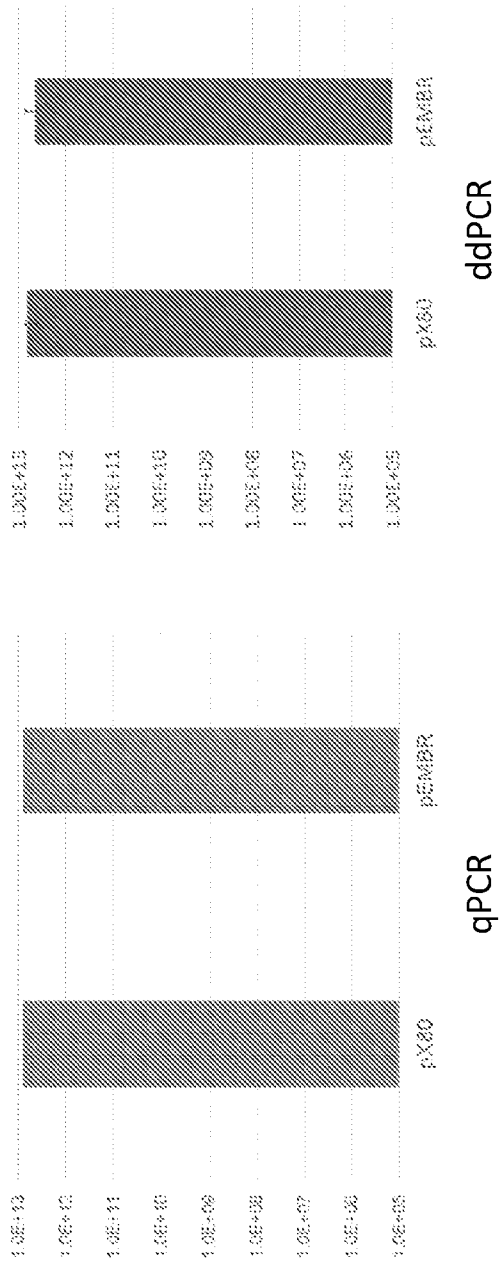

Purified AAVrh.10 Vector Yields
(per liter of production)
qPCR
ddPCR
pXR0
pEMBR
pXR0
pEMBR

Figure 2

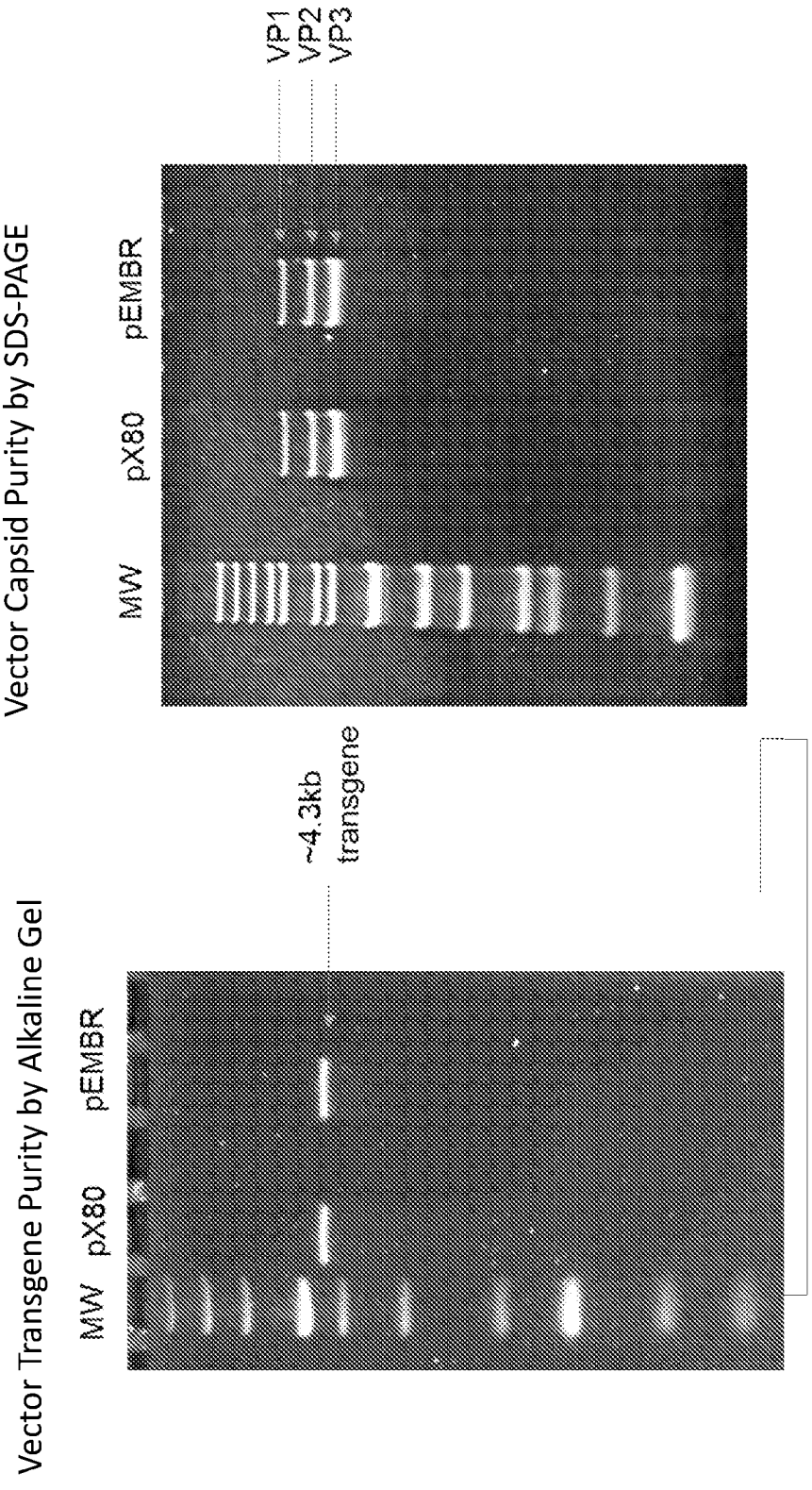

Vector Capsid Purity by SDS-PAGE
MW
pX80
pEMBR
VP1
VP2
VP3
Vector Transgene Purity by Alkaline Gel
MW
pX80
pEMBR
~4.3kb
transgene

Figure 3

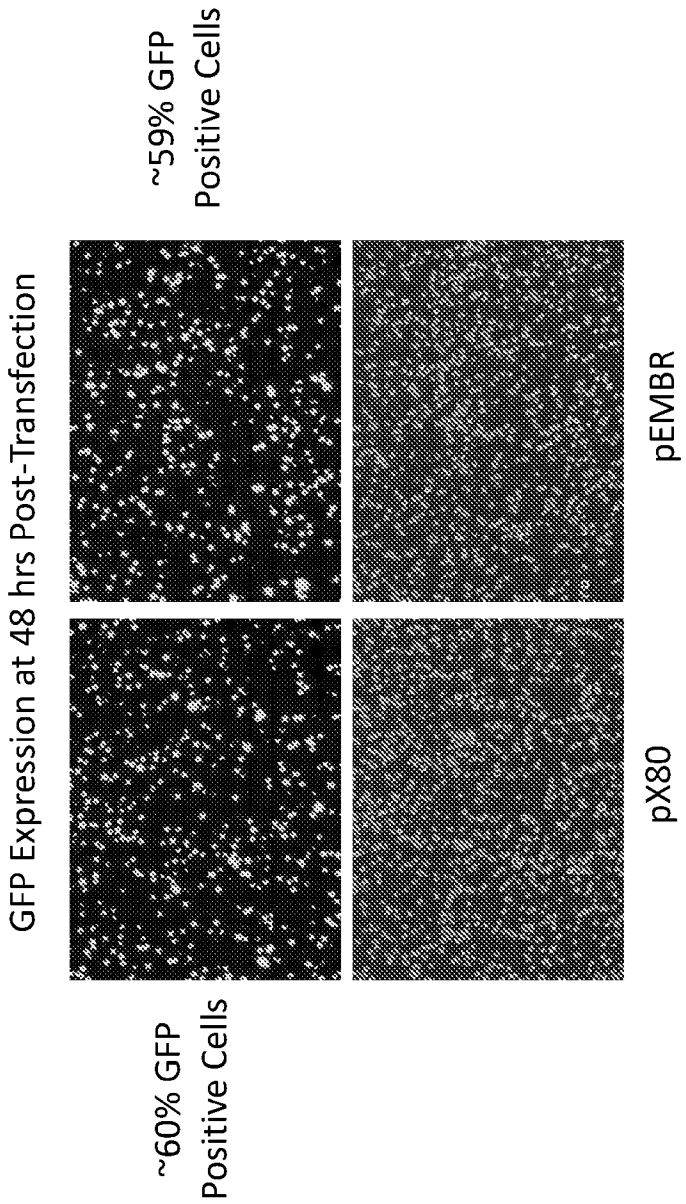

GFP Expression at 48 hrs Post-Transfection
~59% GFP Positive Cells
~60% GFP Positive Cells
pEMBR
pX80

Figure 4

Figure 5

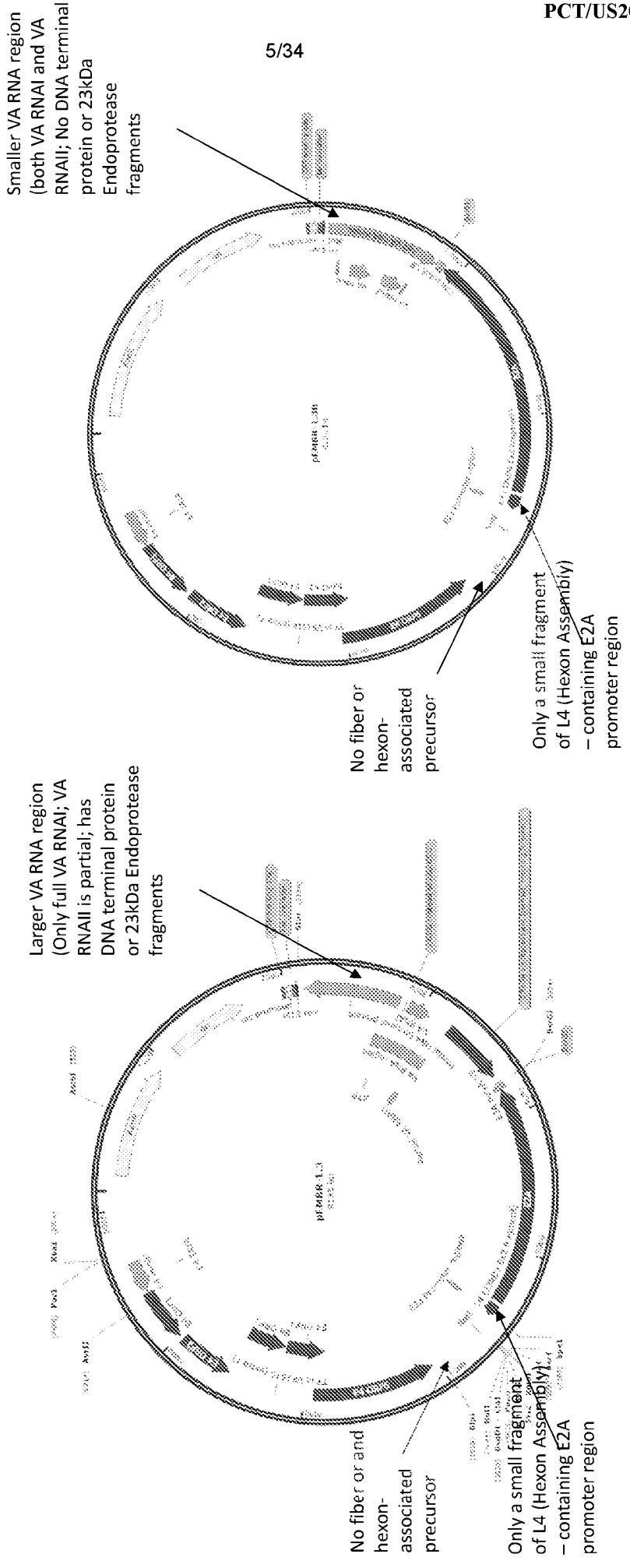
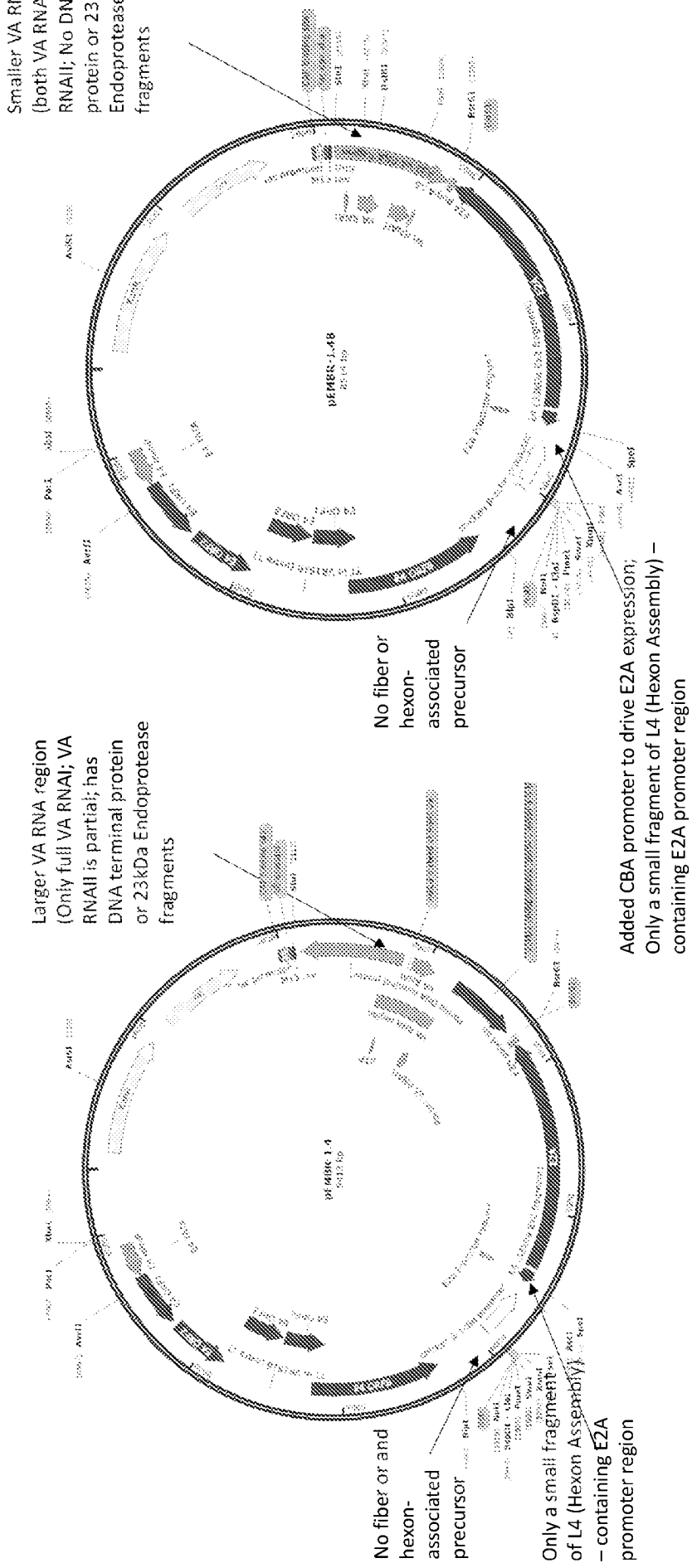

Smaller VA RNA region (both VA RNAI and VA RNAII; No DNA terminal protein or 23kDa Endoprotease fragments
No fiber or hexon-associated precursor
Only a small fragment of L4 (Hexon Assembly) – containing E2A promoter region
Larger VA RNA region (Only full VA RNAI; VA RNAII is partial; has DNA terminal protein or 23kDa Endoprotease fragments
No fiber or and hexon-associated precursor
Only a small fragment of L4 (Hexon Assembly) – containing E2A promoter region

Figure 6


Smaller VA RNA regio
(both VA RNAI and VA
RNAII; No DNA terminal
protein or 23kDa
Endoprotease
fragments
No fiber or
hexon-
associated
precursor
Added CBA promoter to drive E2A expression;
Only a small fragment of L4 (Hexon Assembly) –
containing E2A promoter region
Larger VA RNA region
(Only full VA RNAI; VA
RNAII is partial; has
DNA terminal protein
or 23kDa Endoprotease
fragments
No fiber or and
hexon-
associated
precursor
Only a small fragment
of L4 (Hexon Assembly)
– containing E2A
promoter region

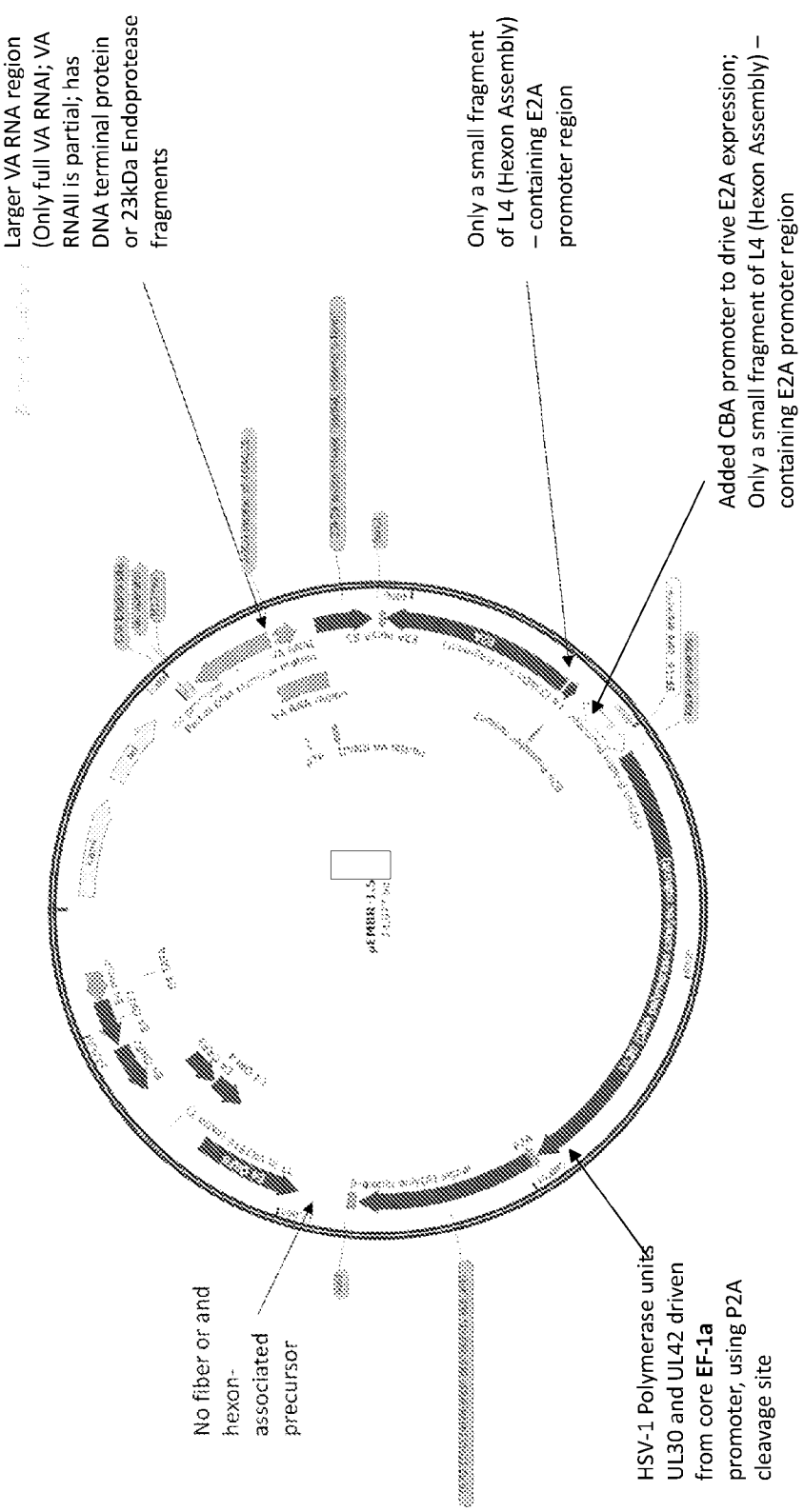

Larger VA RNA region (Only full VA RNAI; VA RNAII is partial; has DNA terminal protein or 23kDa Endoprotease fragments
Only a small fragment of L4 (Hexon Assembly) – containing E2A promoter region
Added CBA promoter to drive E2A expression; Only a small fragment of L4 (Hexon Assembly) – containing E2A promoter region
No fiber or and hexon-associated precursor
HSV-1 Polymerase units UL30 and UL42 driven from core EF-1a promoter, using P2A cleavage site

Figure 7

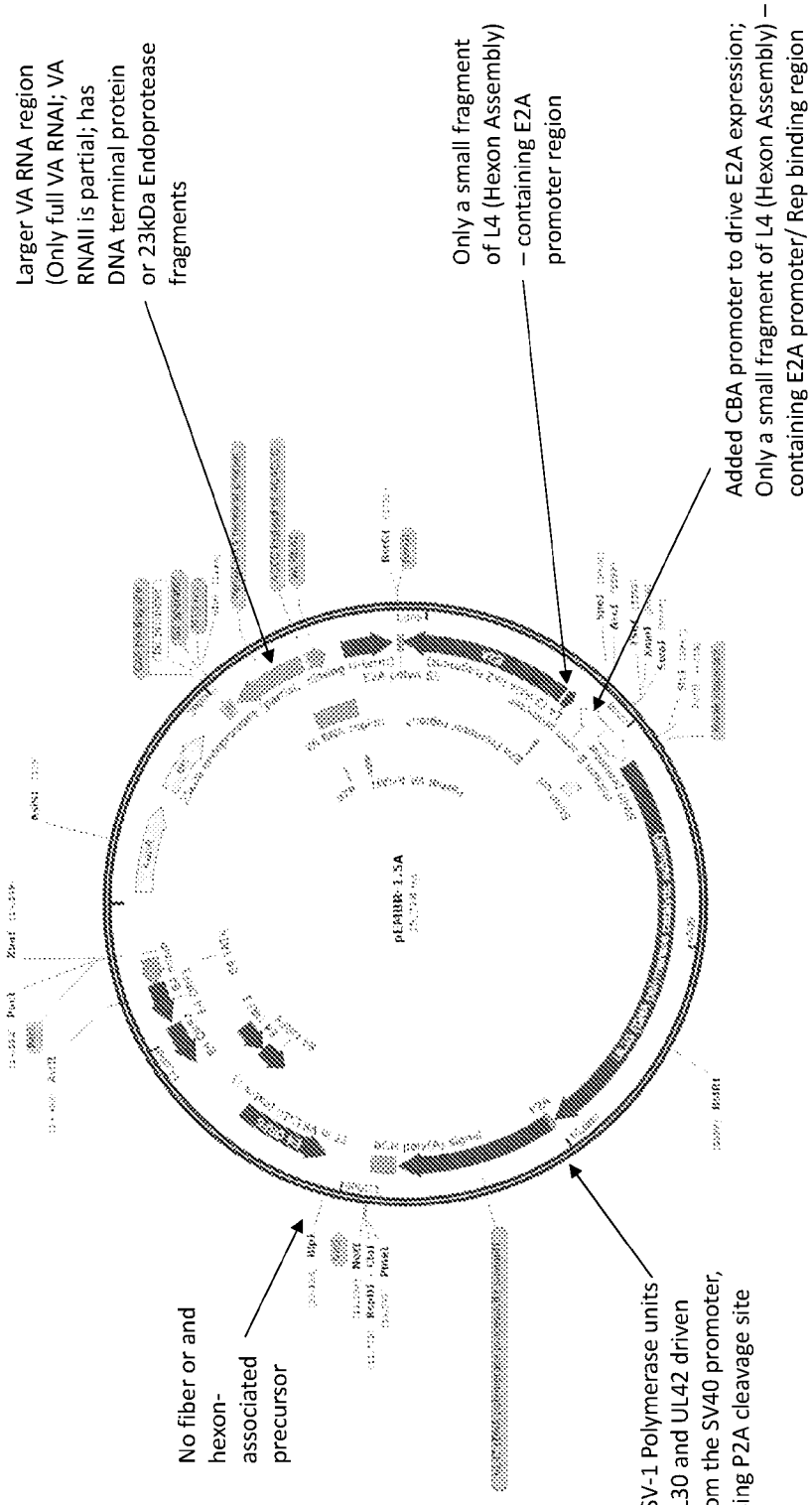

Larger VA RNA region
(Only full VA RNAI; VA
RNAII is partial; has
DNA terminal protein
or 23kDa Endoprotease
fragments
Only a small fragment
of L4 (Hexon Assembly)
– containing E2A
promoter region
Added CBA promoter to drive E2A expression;
Only a small fragment of L4 (Hexon Assembly) –
containing E2A promoter/ Rep binding region
No fiber or and
hexon-
associated
precursor
HSV-1 Polymerase units
UL30 and UL42 driven
from the SV40 promoter,
using P2A cleavage site

Figure 8

Figure 9

A

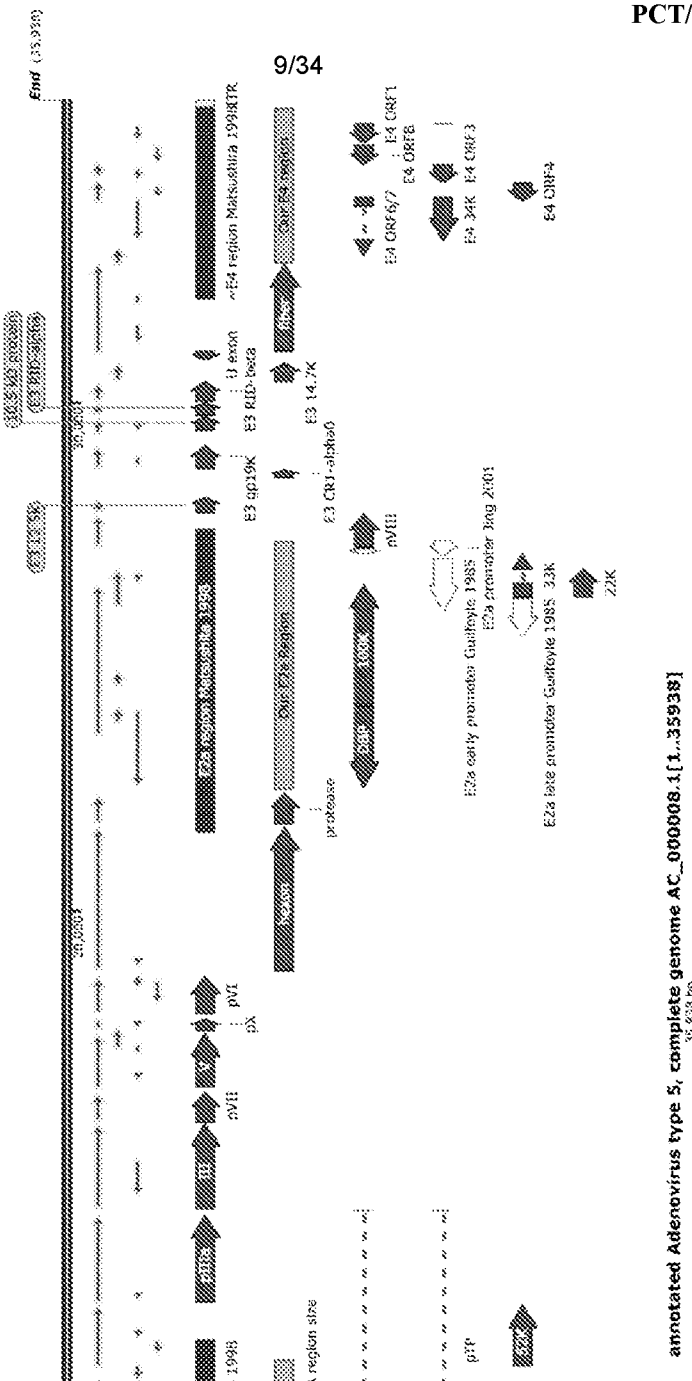

Start (1)
End (35,938)
ITR
E1A
E1B 19K
E1B 55K
IX
IVa2
VA RNA region Matsushita 1998
Our VA RNA region size
pol
pTP
52K
pVIII
pVI
pX
protease
E3 gp19K
E3 RID-beta
E3 14.7K
E3 CR1-alpha6
E4 region Matsushita 1998
E4 ORF6/7
E4 ORF1
E4 ORF6
E4 ORF3
E4 34K
E4 ORF4
E2a early promoter Gaidoyte 1985
E2a promoter Jing Zi01
E2a late promoter Gaidoyte 1985
33K
22K
annotated Adenovirus type 5, complete genome AC_000008.1[1..35938]
35,938 bp

Figure 9 (Continued)

B

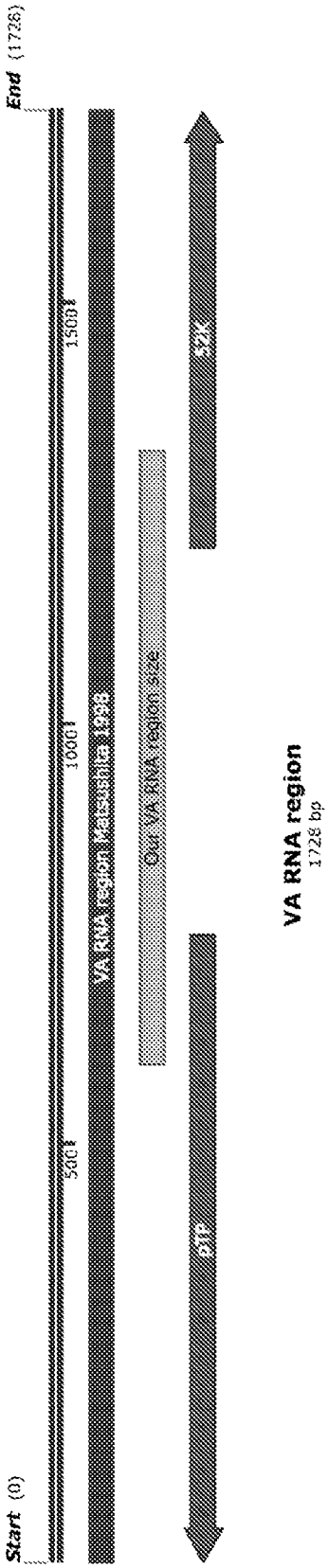


Figure 9 (Continued)

C

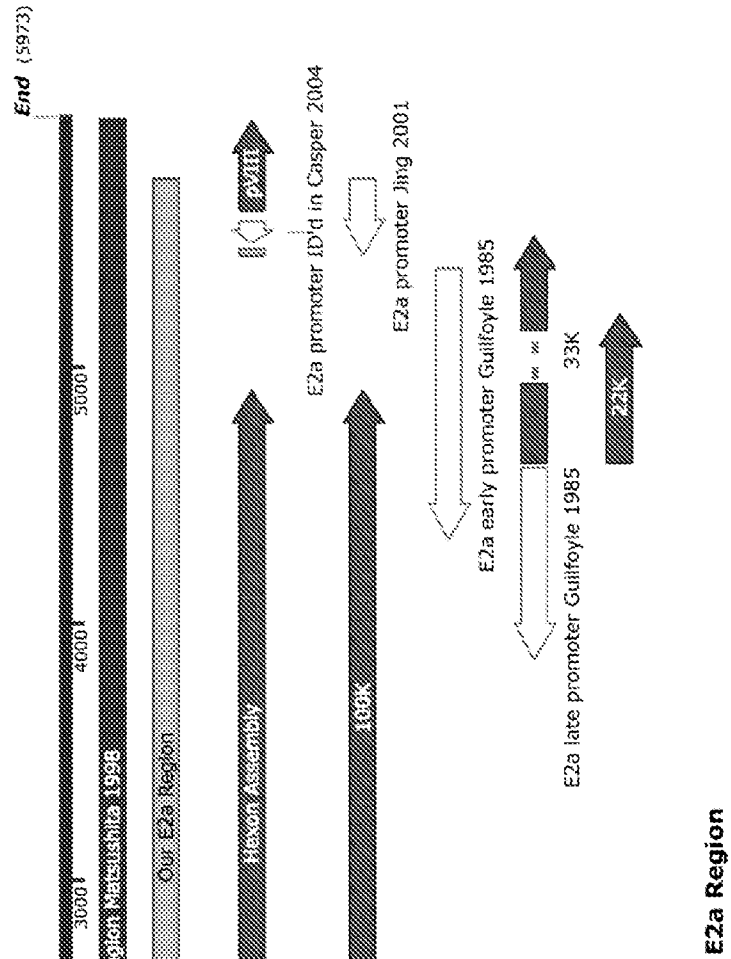
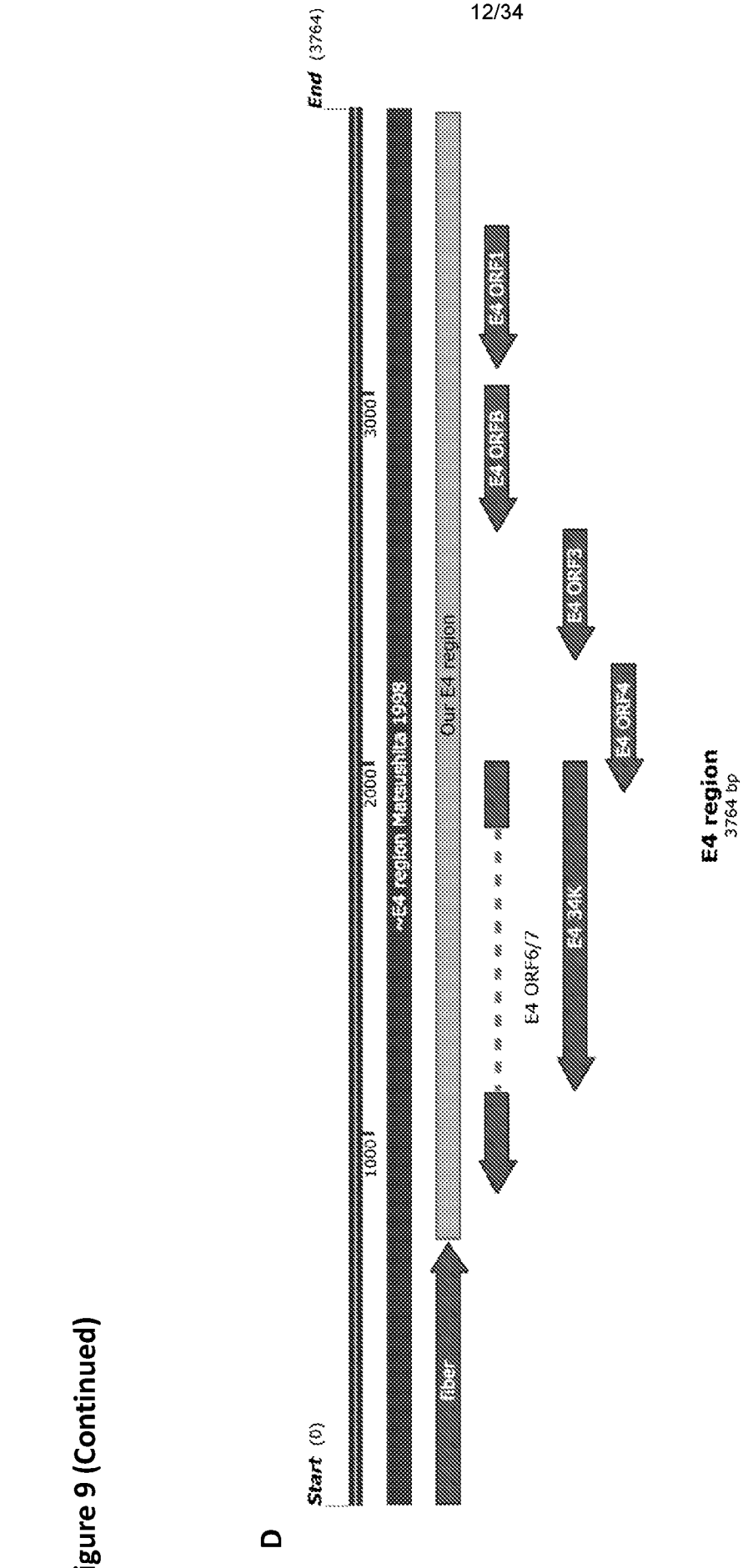


Figure 9 (Continued)



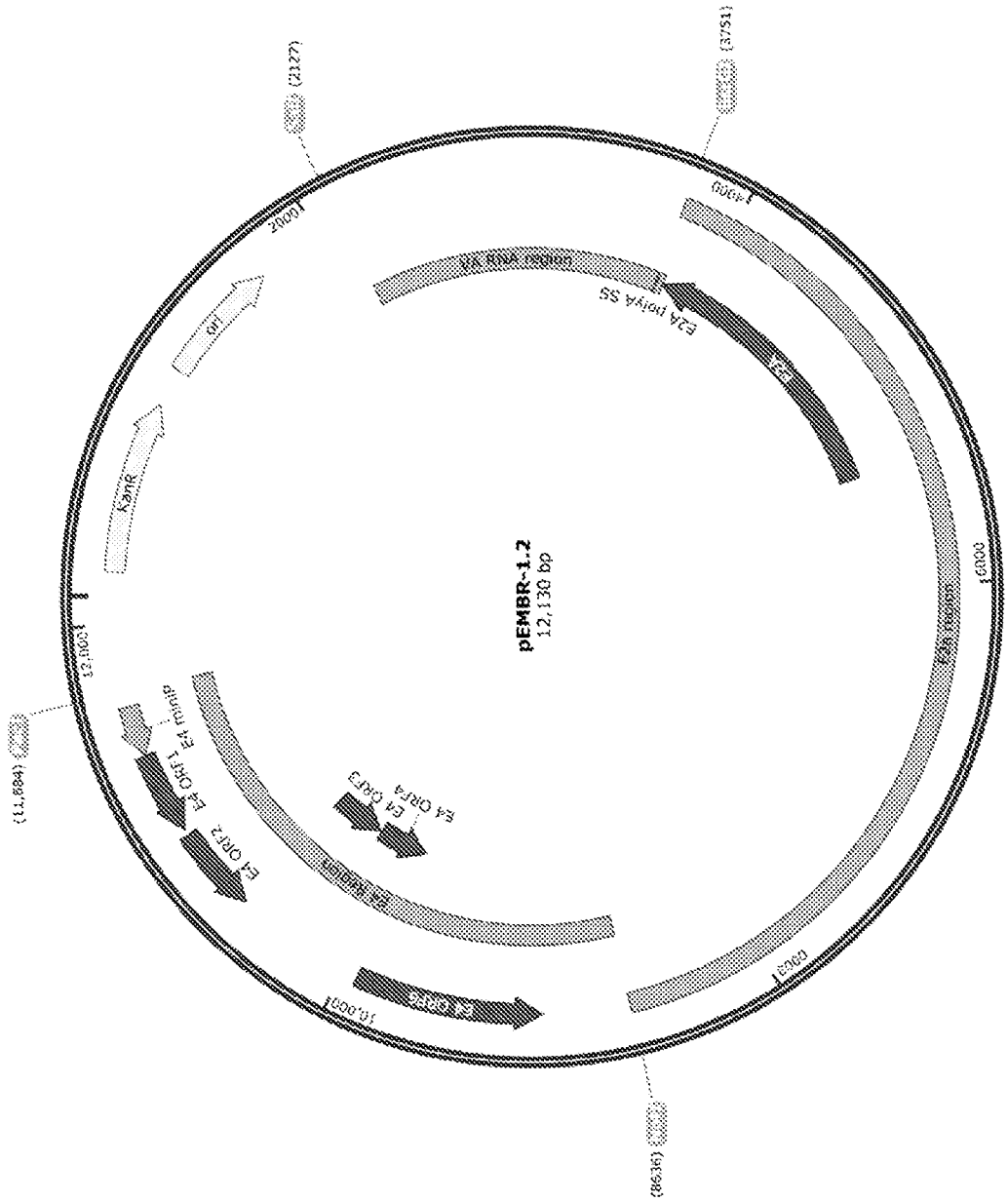

pEMBR-1.2
12,130 bp
(2127)
(3751)
(11,684)
(8636)
2000
4000
6000
8000
10,000
12,000
E2A polyA SS
E2
E4 ORF4
E4 ORF3
E4 ORF2
E4 ORF1
E4 mRNA
E2 mRNA
E3
E1
KanR

Figure 10

Figure 11

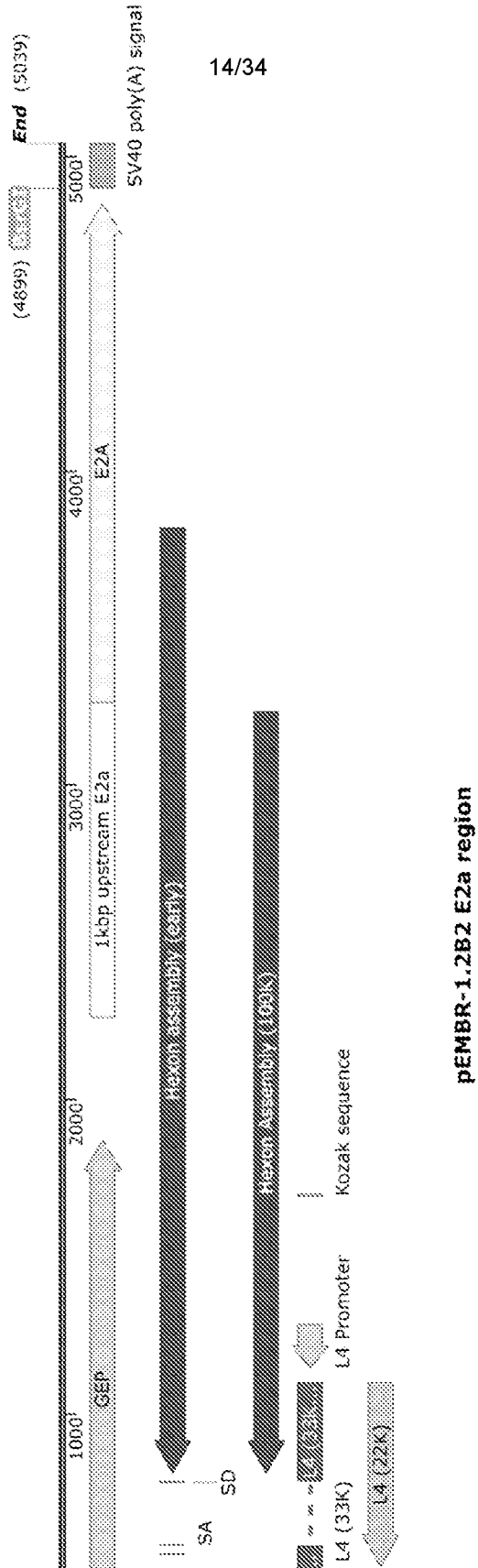

(0) Start
(14)
End (5039)
(4899)
SV40 poly(A) signal
5000
4000
E2A
3000
1kbp upstream E2a
2000
1000
GEP
JEP
Hexon assembly (GEP)
Hexon Assembly (JEP)
NotI
Hexon Associate Precursor
SA
SD
L4 (33K)
L4 Promoter
Kozak sequence
L4 (22K)
pEMBR-1.2B2 E2a region
5039 bp

Figure 12

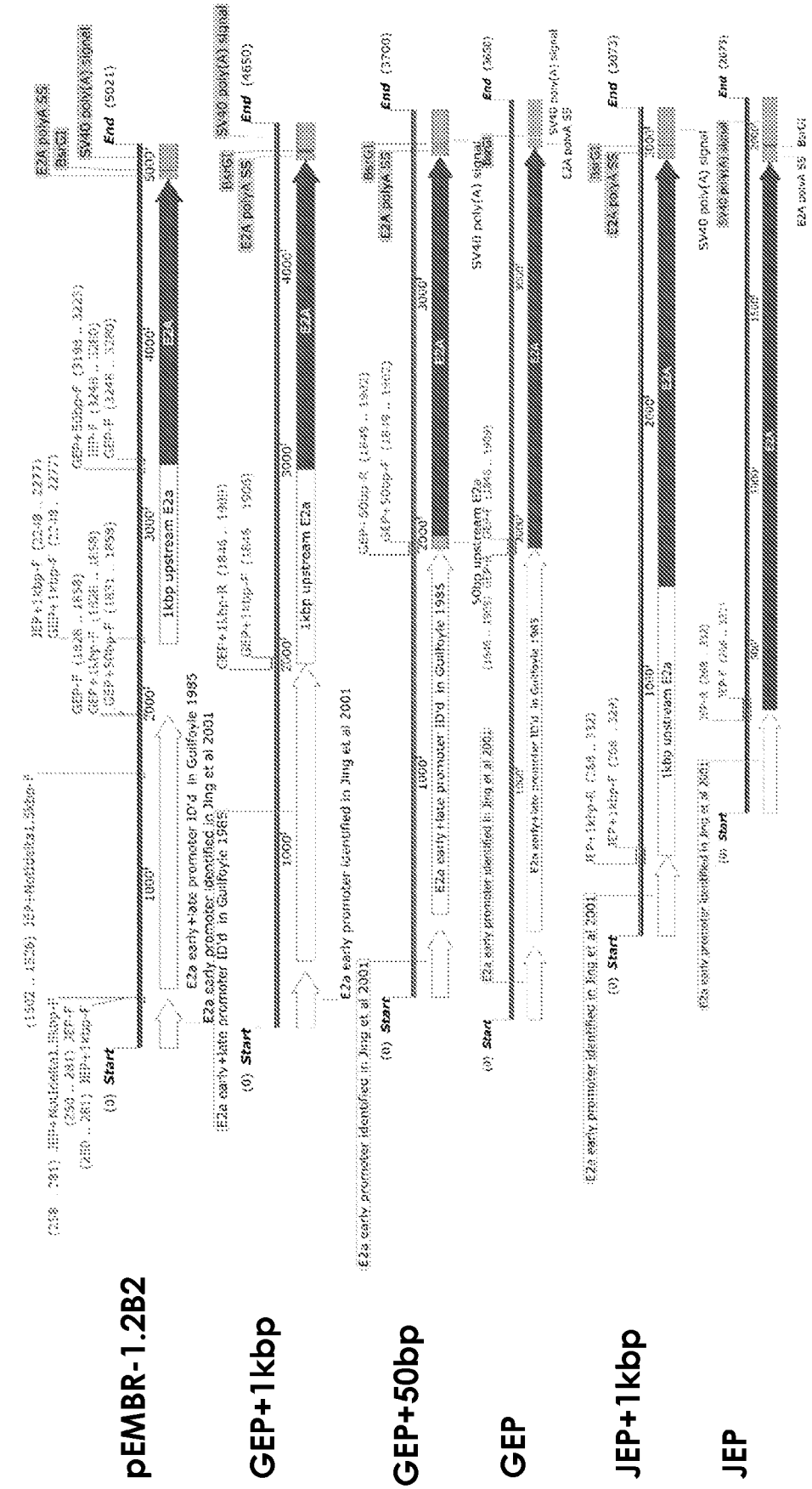

pEMBR-1.2B2
GEP+1kbp
GEP+50bp
GEP
JEP+1kbp
JEP
1kbp upstream E2a
50bp upstream E2a
E2a early promoter identified in Jing et al 2001
E2a early+late promoter ID'd in Guilfoyle 1985
E2a early late promoter ID'd in Guilfoyle 1985
SV40 poly(A) signal
E2A polyA SS
BamHI
(0) Start
End (5021)
End (4650)
End (3700)
End (3658)
End (3673)
End (3673)

Figure 13

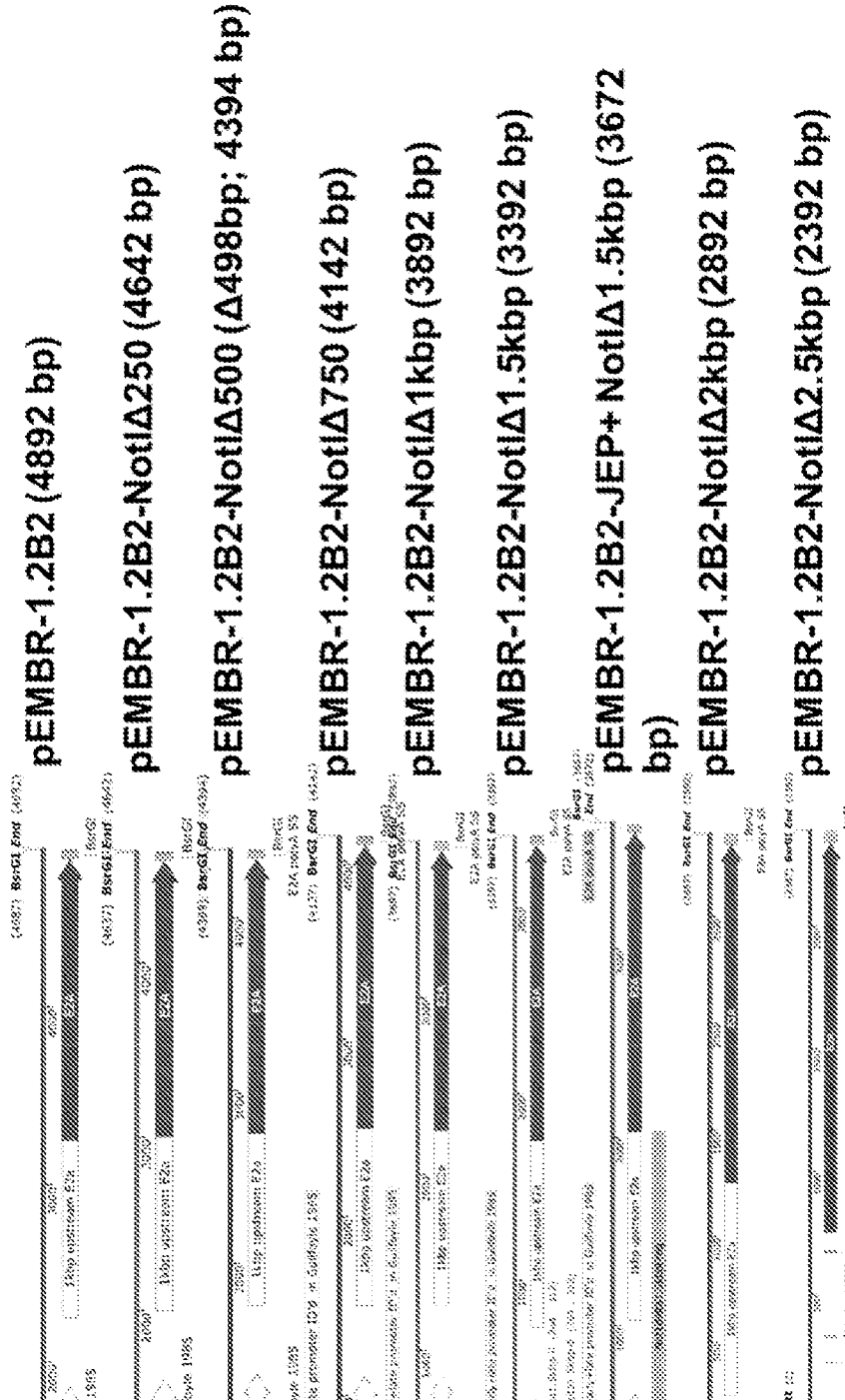
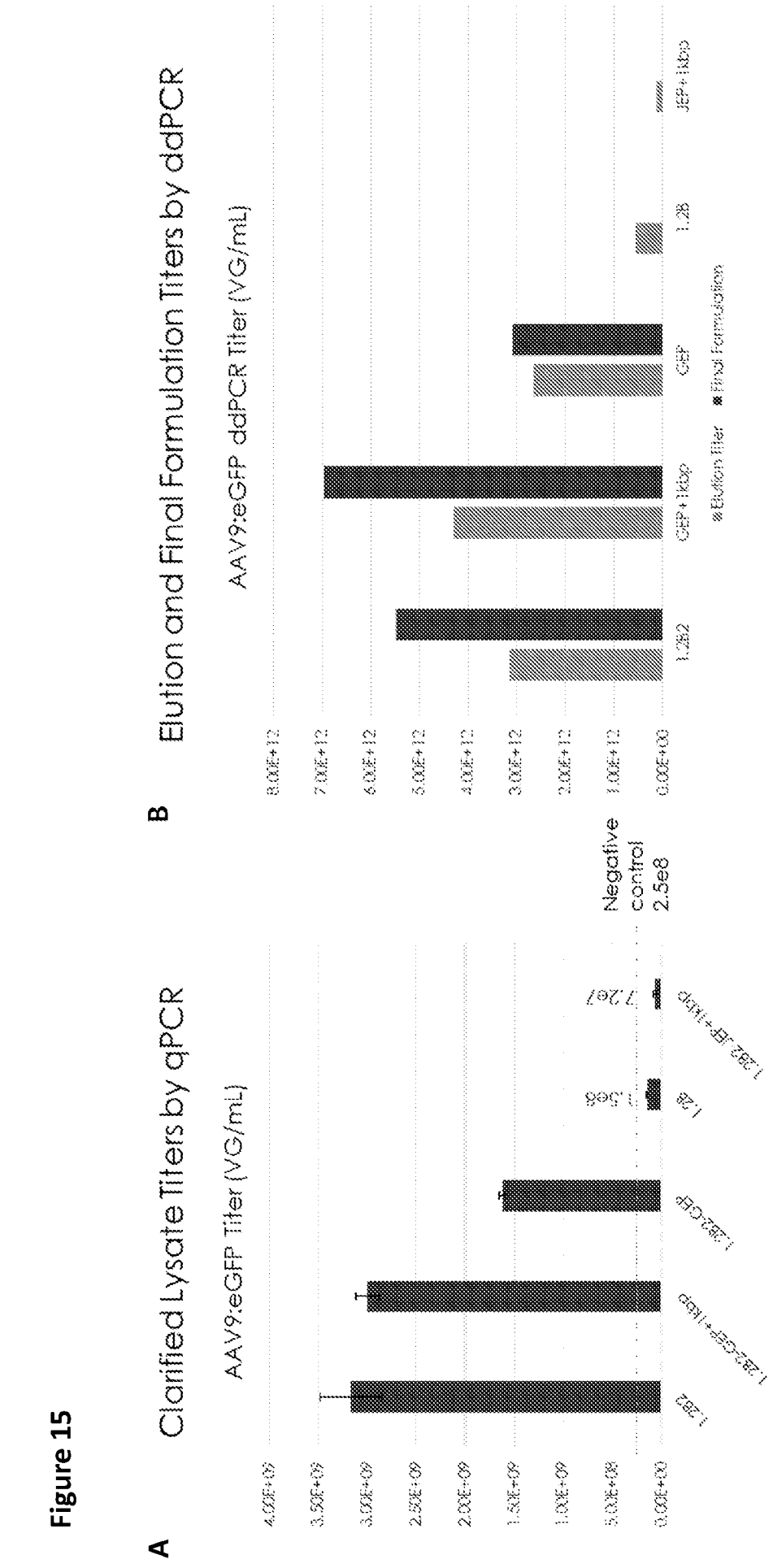

pEMBR-1.2B2 (4892 bp)
pEMBR-1.2B2-NotIΔ250 (4642 bp)
pEMBR-1.2B2-NotIΔ500 (Δ498bp; 4394 bp)
pEMBR-1.2B2-NotIΔ750 (4142 bp)
pEMBR-1.2B2-NotIΔ1kbp (3892 bp)
pEMBR-1.2B2-NotIΔ1.5kbp (3392 bp)
pEMBR-1.2B2-JEP+ NotIΔ1.5kbp (3672 bp)
pEMBR-1.2B2-NotIΔ2kbp (2892 bp)
pEMBR-1.2B2-NotIΔ2.5kbp (2392 bp)

Figure 14

Plasmid	Plasmid Size (bp)	VA RNA region (SalI-BsrGI)	E2a Region (NotI-BsrGI)	E4 Region (PacI-NotI)
pEMBR-1.2	12130	Contains VA RNA I, a portion of VA RNA II, partia lDNA terminal protein	Contains both early and late E2a promoters; includes partial hexon associate precursor, hexon assembly, L4 encapsidation protein	Contains the E4 promoter and all E4 open reading frames
pEMBR-1.2B2	11366	Contains VA RNAI and VA RNAII, no partial DNA terminal protein AND SV40 polyA tail to increase expression of E2a	No change (pEMBR-1.2)	No change (pEMBR-1.2)
pEMBR-GEP+1kbp	10977	Like pEMBR-1.2B2	GEP+1kbp upstream of E2a promoter region	No change (pEMBR-1.2)
pEMBR-GEP+50bp	10027	Like pEMBR-1.2B2	GEP+50bp upstream of E2a promoter region	No change (pEMBR-1.2)
pEMBR-GEP+50bp+Kozak	10033	Like pEMBR-1.2B2	GEP+50bp upstream of E2a promoter region	No change (pEMBR-1.2)
pEMBR-GEP	9977	Like pEMBR-1.2B2	GEP E2a promoter region	No change (pEMBR-1.2)
pEMBR-JEP+1kbp	9400	Like pEMBR-1.2B2	JEP+1kbp upstream of E2a promoter region	No change (pEMBR-1.2)
pEMBR-JEP	8400	Like pEMBR-1.2B2	JEP E2a promoter region	No change (pEMBR-1.2)
pEMBR-JEP+NotIΔ1.5kbp	10146	Like pEMBR-1.2B2	JEP E2a promoter region added to the pEMBR-NotIΔ1.5kbp construct	No change (pEMBR-1.2)

Figure 15


A
Clarified Lysate Titers by qPCR
AAV9:eGFP Titer (VG/mL)
4.00E+09
3.50E+09
3.00E+09
2.50E+09
2.00E+09
1.50E+09
1.00E+09
5.00E+08
0.00E+00
1282
1282-GFP+1Kbp
1282-GFP
1282 GFP+1Kbp
Negative control
2.5e8
7.2e7
B
Elution and Final Formulation Titers by ddPCR
AAV9:eGFP ddPCR Titer (VG/mL)
8.00E+12
7.00E+12
6.00E+12
5.00E+12
4.00E+12
3.00E+12
2.00E+12
1.00E+12
0.00E+00
1.2E2
GFP+1Kbp
GFP
1.2B
JEP+1Kbp
Elution Titer
Final Formulation

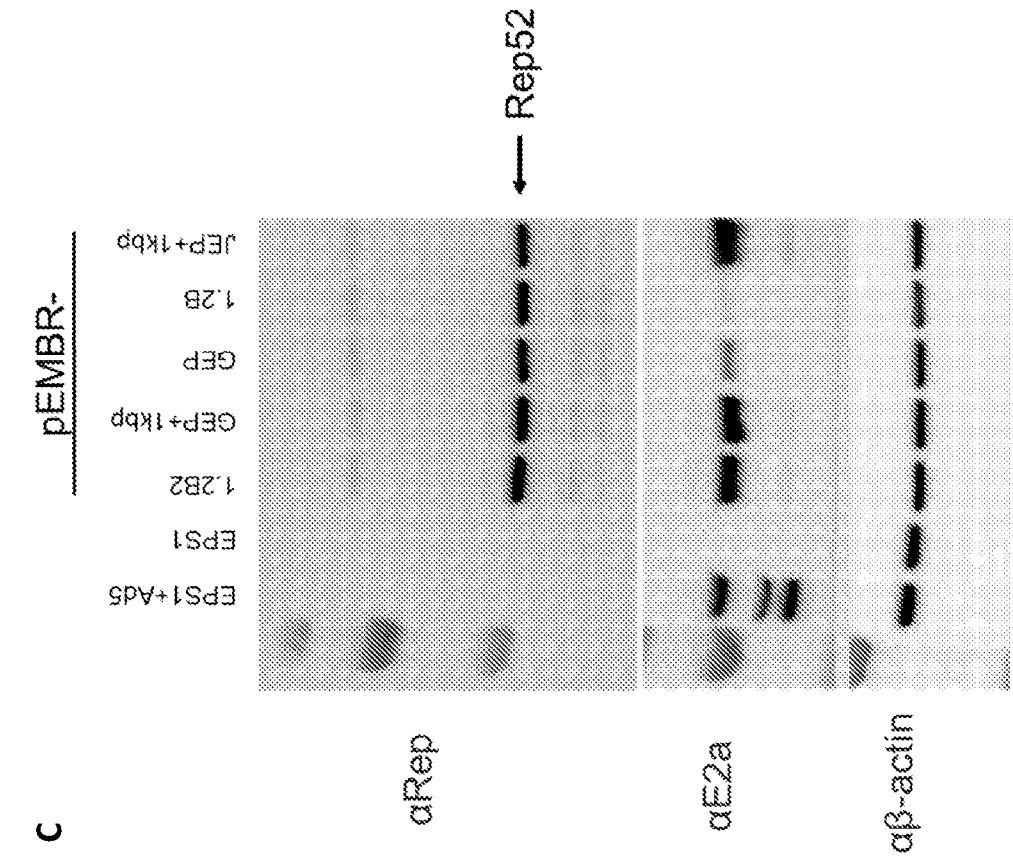


Figure 15 (Continued)

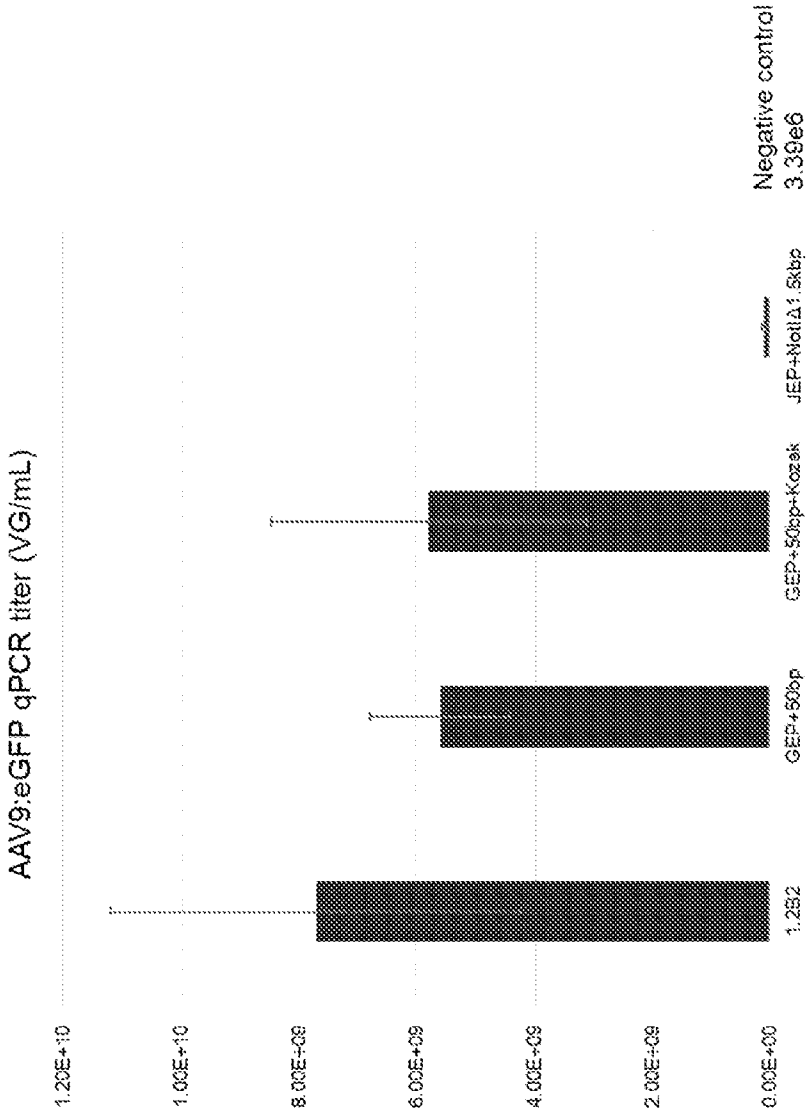

AAV9:eGFP qPCR titer (VG/mL)
1.20E+10
1.00E+10
8.00E+09
6.00E+09
4.00E+09
2.00E+09
0.00E+00
1.2E2
GEP+50bp
GEP+50bp+Kozak
JEP+NotIΔ1.5kbp
Negative control
3.39e6

Figure 16

Figure 17

Plasmid	Contain JEP	Contain GEP	Contain native I4 (33K & 22K)	E2a production	AAV production
pEMBR-1.2	Yes	Yes	Yes	++	+++
pEMBR-1.2B	Yes	Yes	Yes	+	Below LOD
pEMBR-1.2B2	Yes	Yes	Yes	++	+++
pEMBR-1.3, pEMBR-1.4 pEMBR-1.4B2, pEMBR-1.9A	Partial	No	**No**	++	+
pEMBR-1.3B, pEMBR-1.4B	Partial	No	**No**	+	+
pEMBR-1.7, pEMBR-1.8	No	No	**No**	+	+
pEMBR-1.8B, pEMBR-1.9B	No	No	**No**	+	**Below LOD**
pEMBR-1.9	No	No	**No**	++	+
pEMBR-1.8B2, pEMBR-1.9B2	No	No	**No**	++	**Below LOD**
pEMBR-NotI truncations (all)	No	Yes	**No**	+	+
pEMBR-GEP+1kbp	Yes	Yes	Yes	++	+++
pEMBR-GEP	Yes	Yes	Yes	+	+++
pEMBR-JEP+1kbp	Yes	No	**No**	+++	+
pEMBR-JEP	Yes	No	**No**	+++	+
pEMBR-HexΔ	Yes	Yes	Yes	++	+++

Figure 18

Plasmid	Plasmid Size (bp)	VA RNA region (SalI-BsrGI)	E2a Region (NotI-BsrGI)	E4 Region (PacI-NotI)
pEMBR-1.3	9135	No change (pEMBR-1.2)	Contains only early E2a promoter region identified by Casper et al. 2005 as region bound by Rep78 to repress E2a expression	No change (pEMBR-1.2)
pEMBR-1.3 + L4	10311	No change (pEMBR-1.2)	Like pEMBR-1.3 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)
pEMBR-1.4	9413	No change (pEMBR-1.2)	Contains CBA promoter to drive expression of E2a in addition to region identified by Casper et al. 2005	No change (pEMBR-1.2)
pEMBR-1.4 + L4	10589	No change (pEMBR-1.2)	Like pEMBR-1.4 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)
pEMBR-1.4B2	8649	Like pEMBR-1.2B2	Like pEMBR-1.4	No change (pEMBR-1.2)
pEMBR-1.4B2 + L4	9825	Like pEMBR-1.2B2	L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)
pEMBR-1.4B2 + native L4 (BAC)	9978	Like pEMBR-1.2B2	pEMBR-1.4B2 with the native L4 region (1.2B2) added back in.	No change (pEMBR-1.2)
pEMBR-1.5A	15278	No change (pEMBR-1.2)	Like pEMBR-1.4 + UL30 and UL42 expression cassette driven by SV40 promoter	No change (pEMBR-1.2)
pEMBR-1.5A + L4	16454	No change (pEMBR-1.2)	Like pEMBR-1.5A + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)
pEMBR-1.8	9626	No change (pEMBR-1.2)	CBA promoter and CMV enhancer drive expression of E2a	No change (pEMBR-1.2)
pEMBR-1.8 + L4	10802	No change (pEMBR-1.2)	Like pEMBR-1.8 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)
pEMBR-1.9	9759	No change (pEMBR-1.2)	CBA promoter, CMV enhancer, and a chimeric intron drive expression of E2a	No change (pEMBR-1.2)
pEMBR-1.9 + L4	10935	No change (pEMBR-1.2)	Like pEMBR-1.9 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)

Figure 19

Plasmid	Plasmid Size (bp)	VA RNA region (SalI-SbrGI)	E2a Region (NotI-SbrGI)	E4 Region (PacI-NotI)	AAV production vs. 1.2B2
pEMBR-1.2	12130	Contains VA RNA I, a portion of VA RNA II, partial DNA terminal protein	Contains both early and late E2a promoters, excludes partial hexon associate precursor, hexon assembly, L4 encapsidation protein	Contains the E4 promoter and all E4 open reading frames	1
pEMBR-1.2B2*	11366	Contains VA RNAI and VA RNAII, no partial DNA terminal protein AND SV40 polyA tail to increase expression of E2a	No change (pEMBR-1.2)	No change (pEMBR-1.2)	1
pEMBR-GEP+1kbp	10977	Like pEMBR-1.2B2	JEP+GEP+1kbp upstream of E2a promoter region	No change (pEMBR-1.2)	1
pEMBR-GEP+50bp	10027	Like pEMBR-1.2B2	JEP+GEP+50bp upstream of E2a promoter region	No change (pEMBR-1.2)	1
pEMBR-GEP+50bp+Kozak	10033	Like pEMBR-1.2B2	JEP+GEP+50bp upstream of E2a promoter region	No change (pEMBR-1.2)	1
pEMBR-GEP	9977	Like pEMBR-1.2B2	JEP+GEP E2a promoter region	No change (pEMBR-1.2)	0.5
pEMBR-1.4B2 + native L4 (BAC)	9978	Like pEMBR-1.2B2	pEMBR-1.4B2 with the native L4 region (1.2B2) added back in	No change (pEMBR-1.2)	0.5
pEMBR-2.0	9487	Like pEMBR-1.2B2	1879bpΔ from 50bp upstream E2a to 50bp upstream L4 promoter	No change (pEMBR-1.2)	1
pEMBR-1.7B2	10032	Like pEMBR-1.2B2	No change (pEMBR-1.2)	SV40 promoter driven E4 ORFs 4&6 with SV40 polyA	1
pEMBR-1.3 + L4	10311	No change (pEMBR-1.2)	Like pEMBR-1.3 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)	1
pEMBR-1.4 + L4	10589	No change (pEMBR-1.2)	Like pEMBR-1.4 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)	0.5
pEMBR-1.4B2 + L4	9825	Like pEMBR-1.2B2	L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)	1
pEMBR-1.5A + L4	16454	No change (pEMBR-1.2)	Like pEMBR-1.5A + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)	1
pEMBR-1.8 + L4	10802	No change (pEMBR-1.2)	Like pEMBR-1.8 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)	1
pEMBR-1.9 + L4	10935	No change (pEMBR-1.2)	Like pEMBR-1.9 + L4 expression cassette added- native promoter driven with synthetic 49 bp polyA	No change (pEMBR-1.2)	1

Figure 20

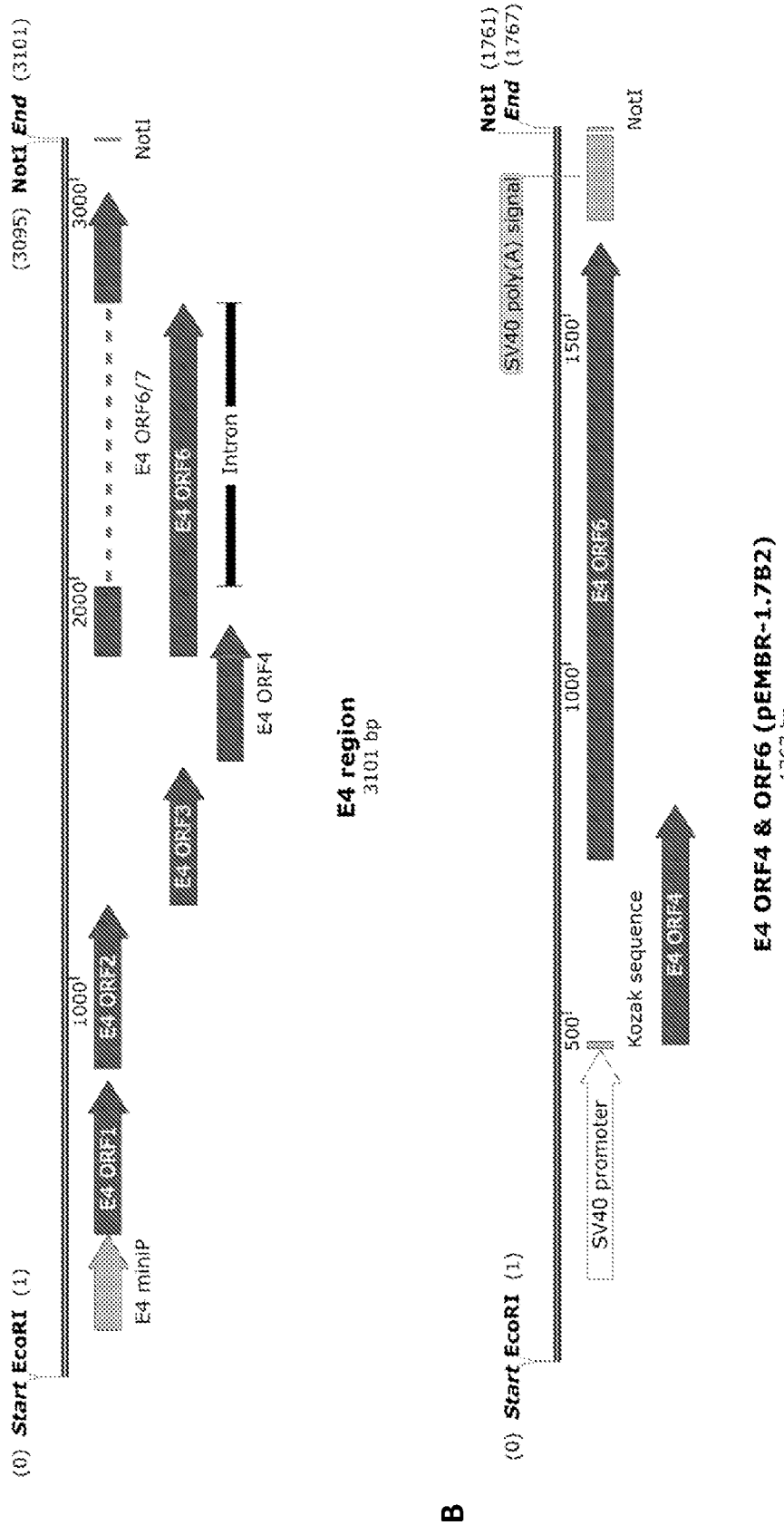
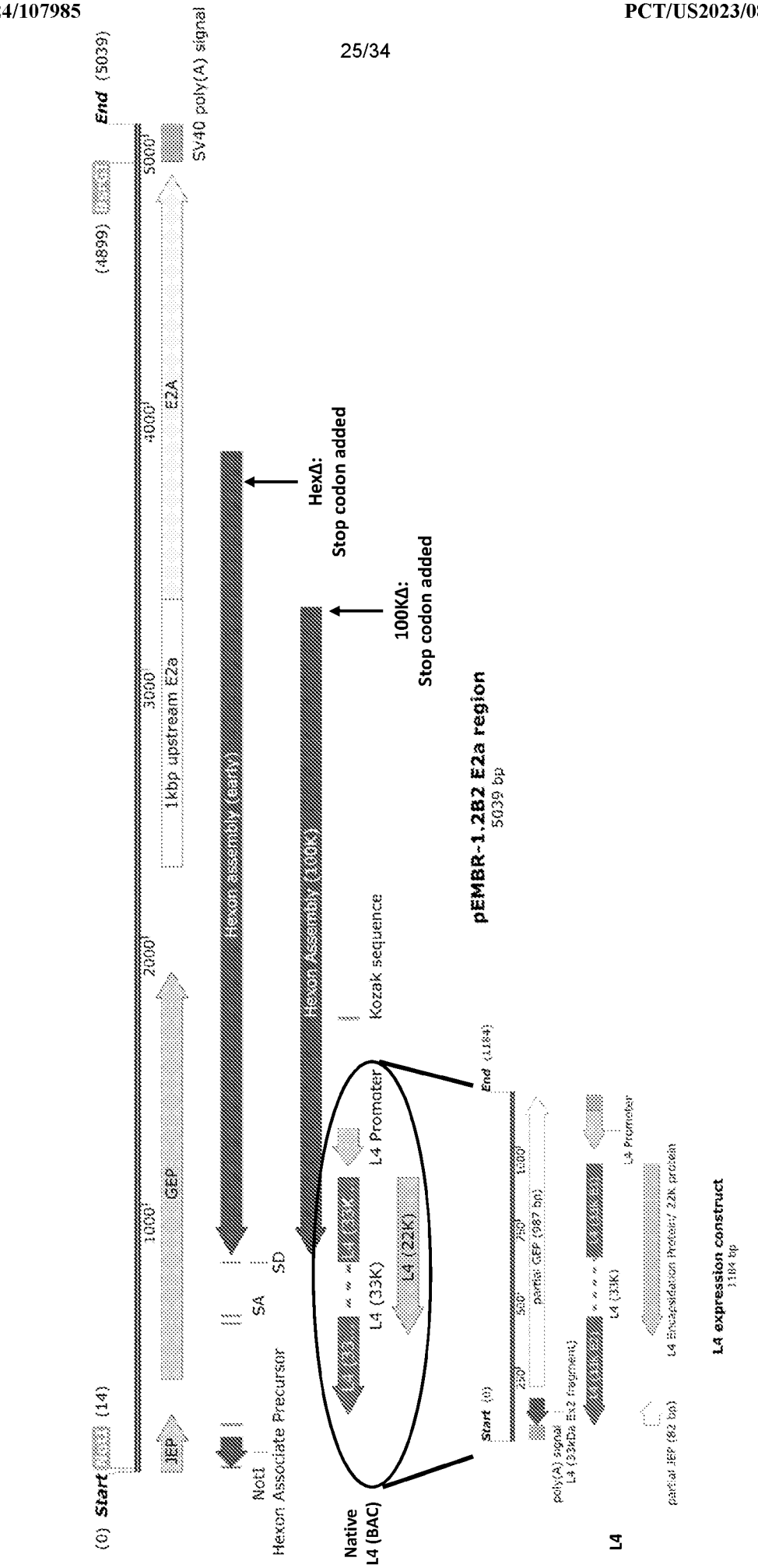

A
(0) Start EcoRI (1)
(3095) NotI End (3101)
1000
2000
3000
E4 miniP
E4 ORF1
E4 ORF2
E4 ORF3
E4 ORF4
E4 ORF6/7
E4 ORF6
Intron
NotI
E4 region
3101 bp
B
(0) Start EcoRI (1)
NotI (1761)
End (1767)
500
1000
1500
SV40 promoter
Kozak sequence
E4 ORF4
E4 ORF6
SV40 poly(A) signal
NotI
E4 ORF4 & ORF6 (pEMBR-1.7B2)
1767 bp

Figure 21


(0) Start
(14)
End (5039)
(4899)
5000
SV40 poly(A) signal
4000
E2A
3000
1kbp upstream E2a
2000
Hexon assembly (GEP)
1000
GEP
JEP
NotI
Hexon Associate Precursor
SA
SD
HexΔ:
Stop codon added
100KΔ:
Stop codon added
Hexon Assembly (100K)
Kozak sequence
pEMBR-1.2B2 E2a region
5039 bp
L4 Promoter
L4 (33K)
L4 (22K)
Native
L4 (BAC)
End (1184)
Start (0)
250
500
750
1000
partial GEP (987 bp)
poly(A) signal
L4 (33K)
L4 (33K)
L4 Promoter
L4 Encapsidation Protein/ 22K protein
partial JEP (82 bp)
L4
L4 expression construct
1184 bp

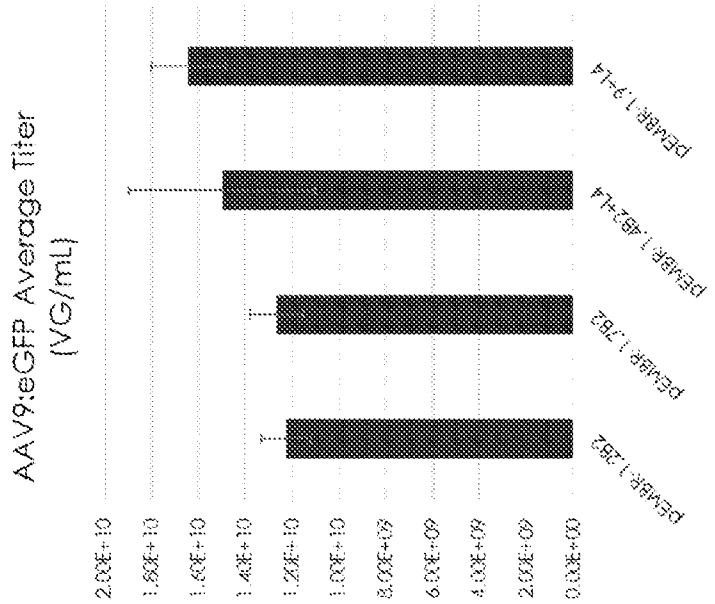
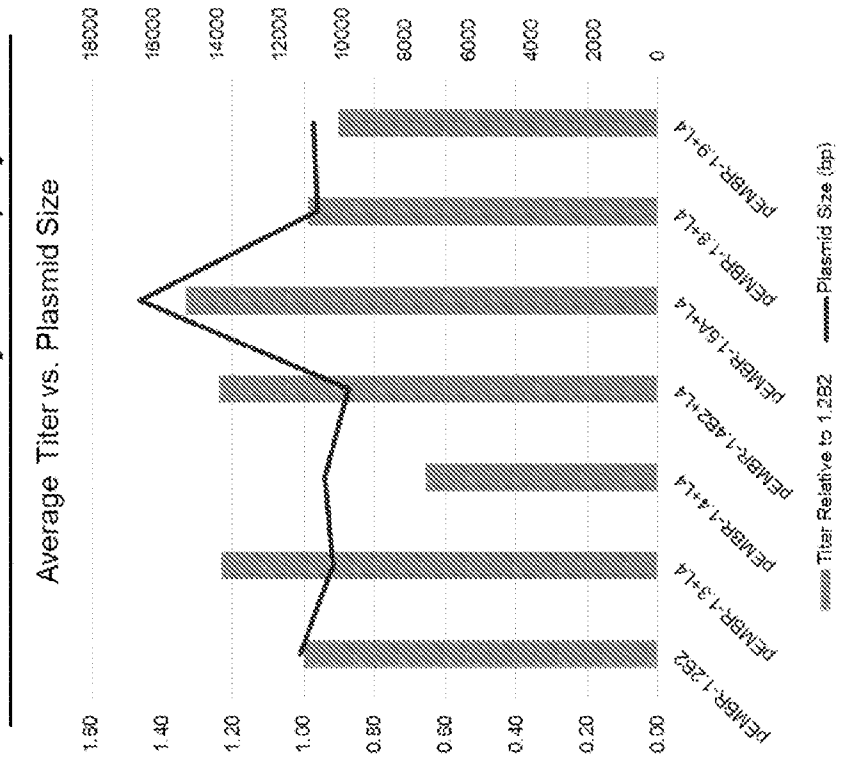

AAV9:eGFP Average Titer
(VG/mL)
2.00E+10
1.80E+10
1.60E+10
1.40E+10
1.20E+10
1.00E+10
8.00E+09
6.00E+09
4.00E+09
2.00E+09
0.00E+00
DEAP6 1.262
DEAP6 1.782
DEAP6 1.4824
DEAP6 1.914

Figure 22


+ L4 construct with synthetic polyA
Average Titer vs. Plasmid Size
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
pEMBR-1.2B2
pEMBR-1.3+L4
pEMBR-1.4+L4
pEMBR-1.4B2+L4
pEMBR-1.5A+L4
pEMBR-1.3+L4
pEMBR-1.9+L4
Titer Relative to 1.2B2
Plasmid Size (bp)

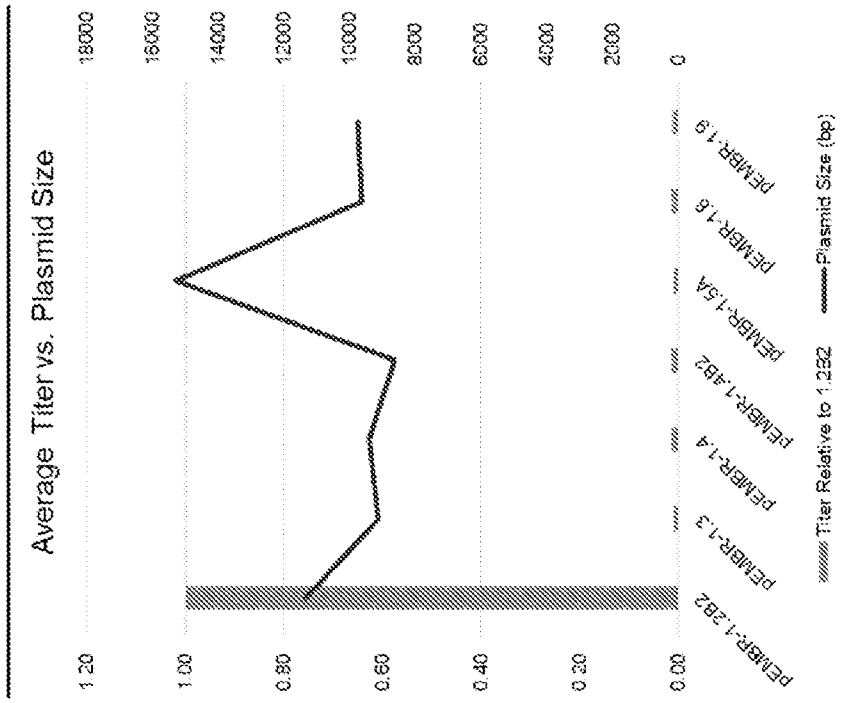

Average Titer vs. Plasmid Size
1.20
1.00
0.80
0.60
0.40
0.20
0.00
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
pEMBR-1.2B2
pEMBR-1.3
pEMBR-1.4
pEMBR-1.4B2
pEMBR-1.5A
pEMBR-1.6
pEMBR-1.9
Titer Relative to 1.2B2
Plasmid Size (bp)

Figure 23

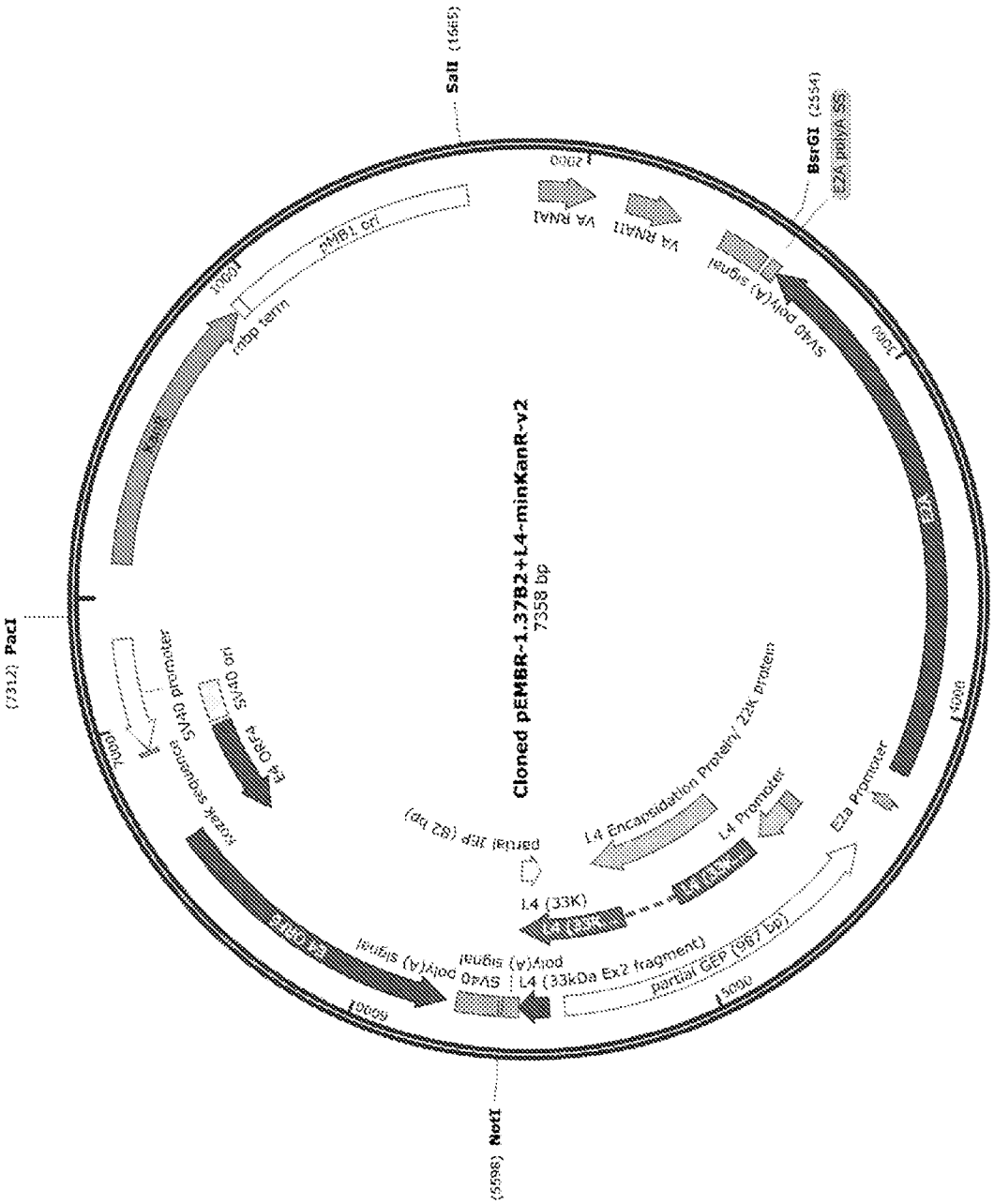

Cloned pEM8R-1.37B2+L4-minKanR-v2
7358 bp
SalI (1685)
BsrGI (2554)
E2A polyA SS
1000
2000
VA RNA1
VA RNA11
SV40 poly(A) signal
E2a
3000
pMB1 ori
ampR
(dep term
4000
E1a
5000
E2a Promoter
L4 Promoter
L4 Encapsidation Protein/ 22K protein
partial IEP (82 bp)
L4 (33K)
SV40 poly(A) signal
L4 (33kDa Ex2 fragment)
partial GFP (967 bp)
6000
6999
kozak sequence
SV40 promoter
E4 ORF4
SV40 ori
7000
PacI (7312)
NotI (5598)

Figure 24

Figure 25

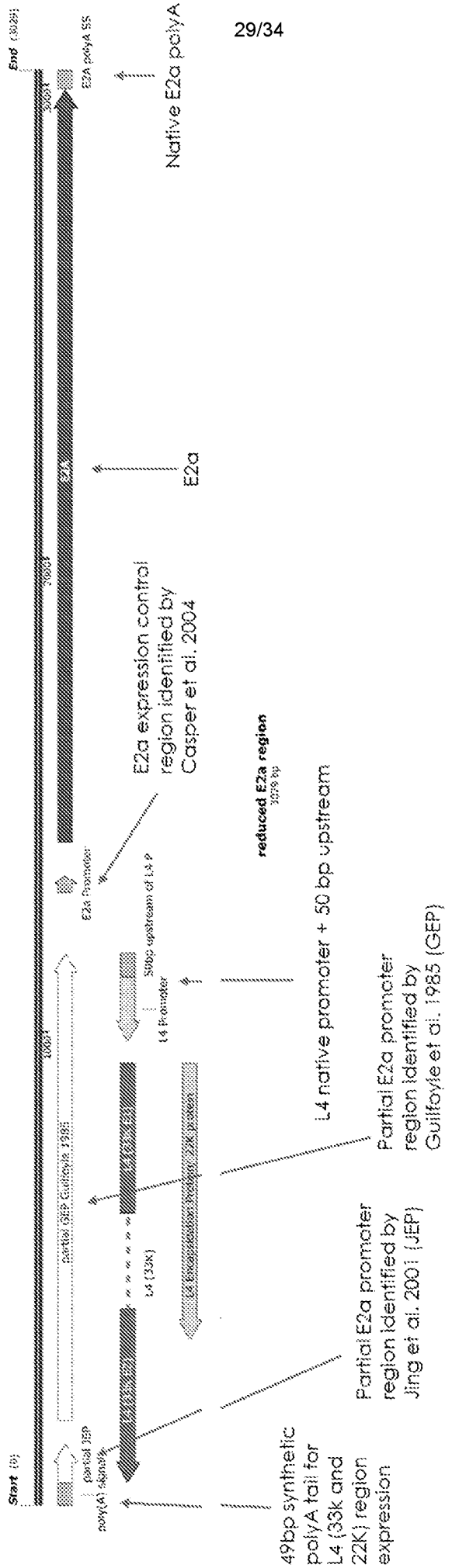
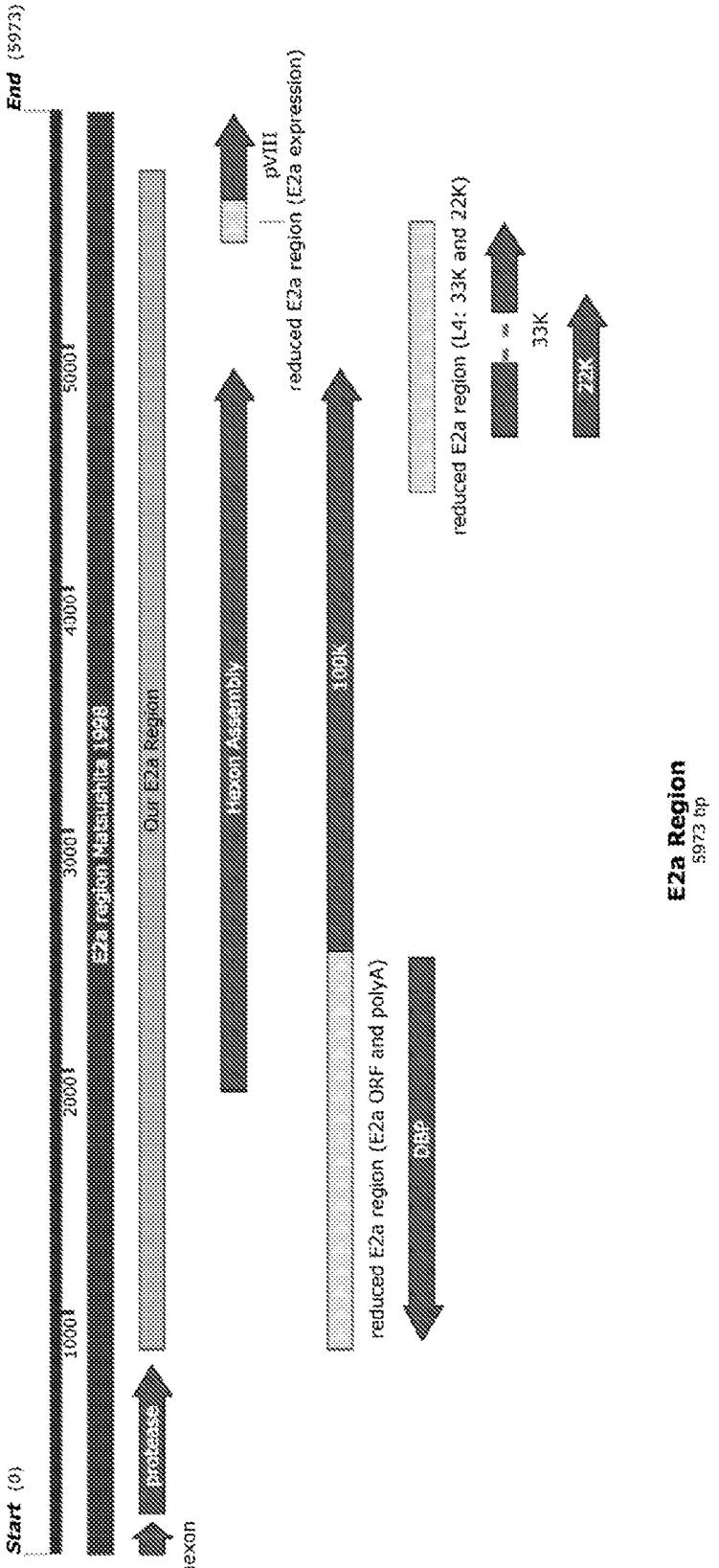

Native E2a polyA
E2a
E2a expression control region identified by Casper et al. 2004
reduced E2a region
L4 native promoter + 50 bp upstream
Partial E2a promoter region identified by Guilfoyle et al. 1985 (GEP)
Partial E2a promoter region identified by Jing et al. 2001 (JEP)
49bp synthetic polyA tail for L4 (33k and 22K) region expression

Figure 26


Start (0)
End (5973)
1000
2000
3000
4000
5000
E2a region Cassette 1998
Our E2a Region
Report Assembly
pVIII
reduced E2a region (E2a expression)
100K
reduced E2a region (E2a ORF and polyA)
pTP
reduced E2a region (L4: 33K and 22K)
33K
22K
hexon
protease
E2a Region
5973 bp

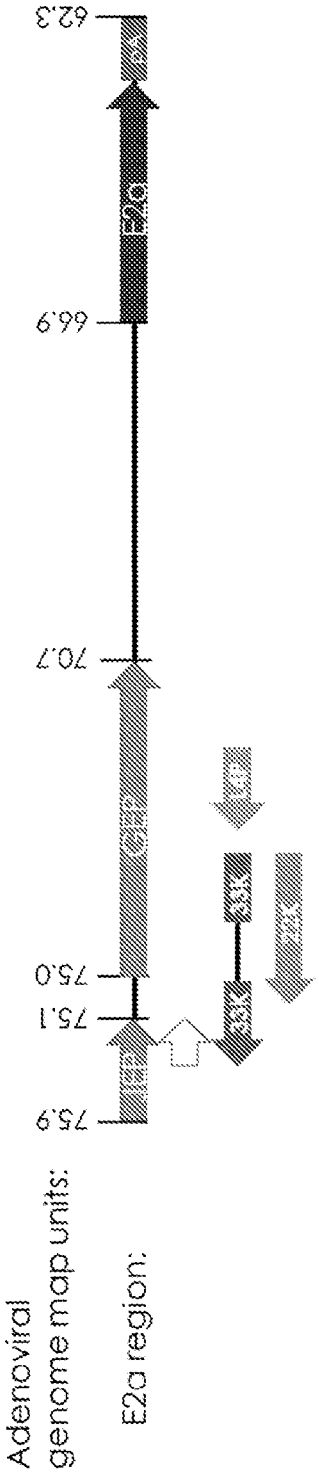
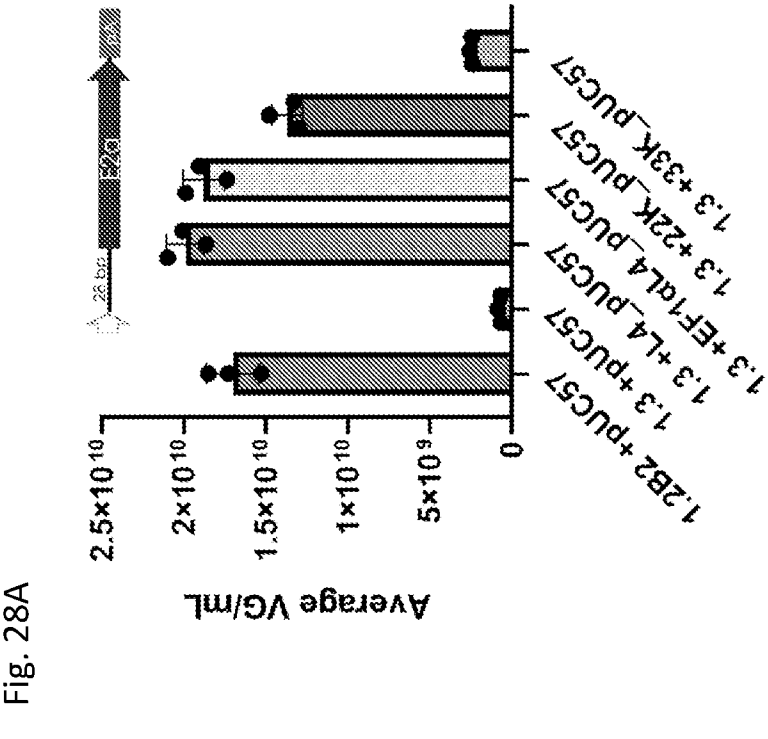
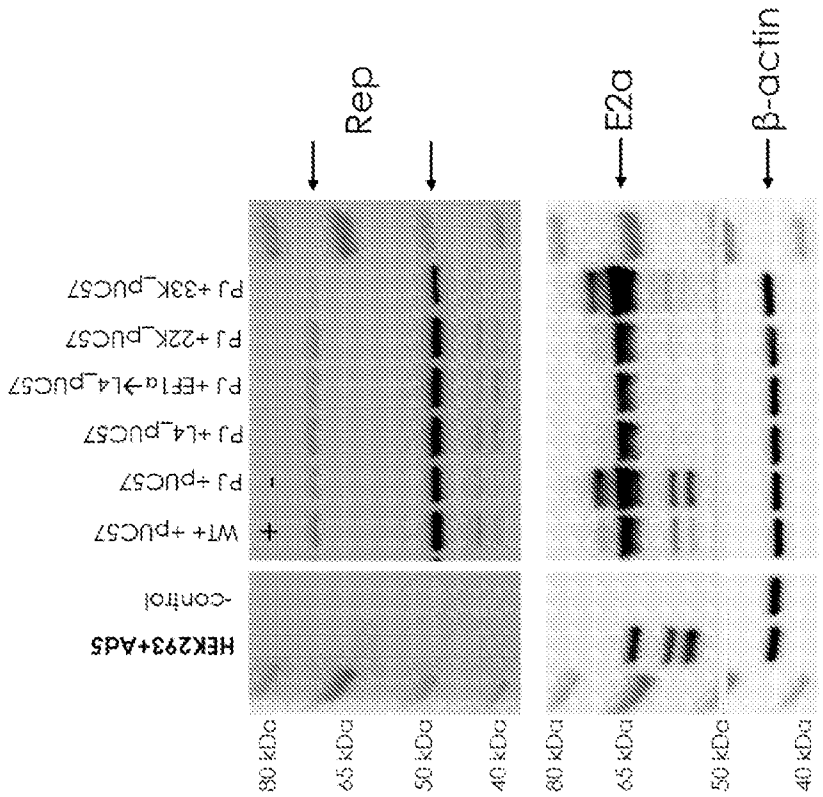

Adenoviral genome map units:
E2a region:
75.9
75.1
75.0
70.7
66.9
62.3
E2a
E2b
33K
22K
pTP
pIIIa

Figure 27

Figure 28


Fig. 28A
Average VG/mL
2.5×10^10
2×10^10
1.5×10^10
1×10^10
5×10^9
0
28 bp
E2a
1.2B2 +pUC57
1.3 +pUC57
1.3 +L4_pUC57
1.3 +EF1αL4_pUC57
1.3 +22K_pUC57
1.3 +33K_pUC57
Quadruple Transfection "Helper" Plasmids pEMBR-X + X_pUC57


Fig. 28B
HEK293+Ad5
-control
WT+ +pUC57
P.J +pUC57
P.J +L4_pUC57
P.J +EF1αL4_pUC57
P.J +22K_pUC57
P.J +33K_pUC57
Rep
E2a
β-actin
80 kDa
65 kDa
50 kDa
40 kDa

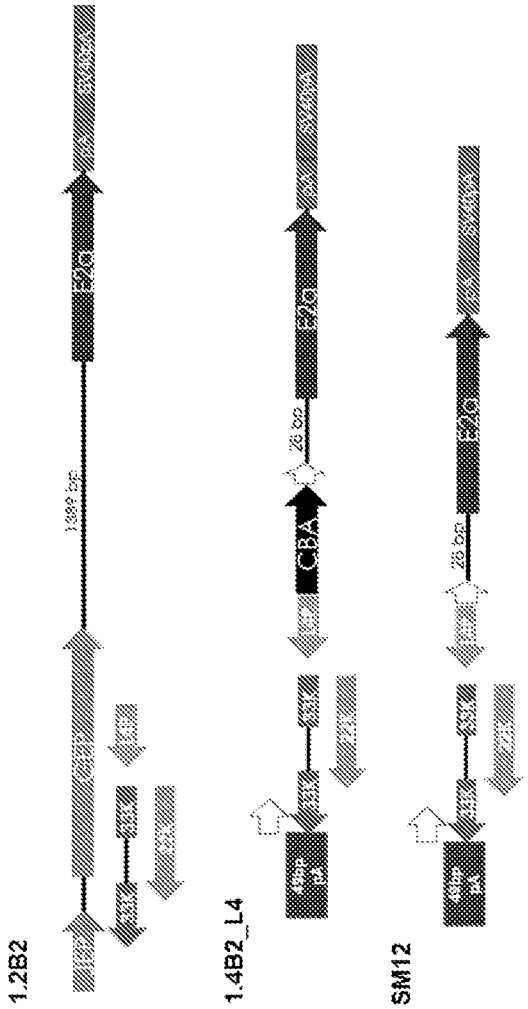


Figure 29

Fig. 29A

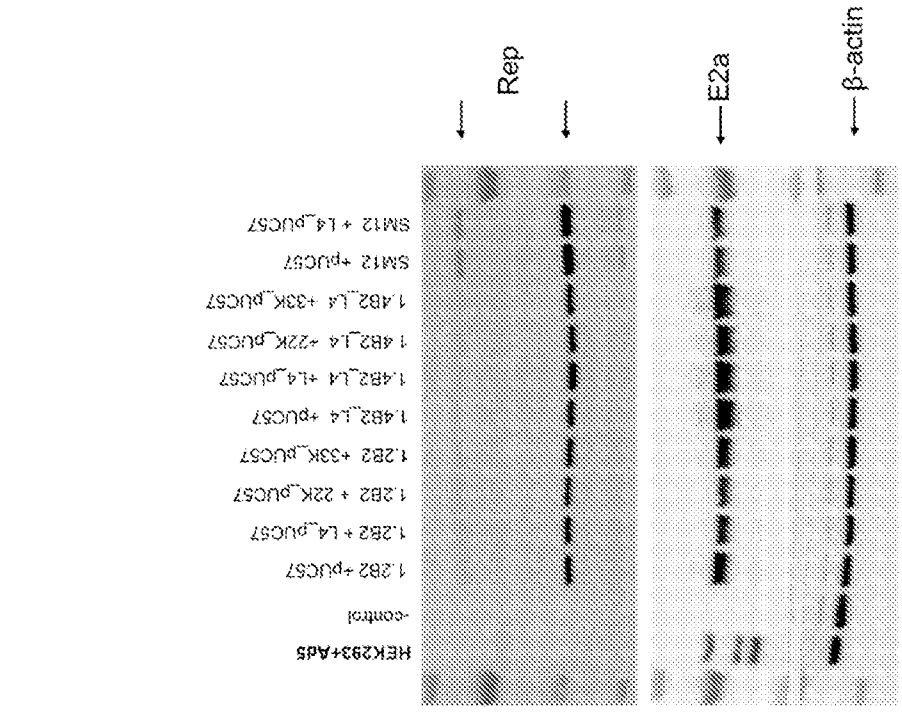


Fig. 29C

Figure 29 (cont.)

Fig. 29B

