Small uORFs favor translation re-initiation but do not protect mRNAs from nonsense-mediated decay

Paul J. Russell[1,2,3,4], Jacob A. Slivka[3,5], Elaina P. Boyle[3,4], Arthur H. M. Burghes[3], and Michael G. Kearse[1,2,3,4]

[1]Cellular, Molecular, and Biochemical Sciences Program, [2]The Ohio State Biochemistry Program, [3]Department of Biological Chemistry and Pharmacology, [4]Center for RNA Biology, [5]Department of Computer Science and Engineering, The Ohio State University, Columbus, Ohio 43210 USA

Corresponding author: michael.kearse@osumc.edu

ABSTRACT

It is estimated that nearly 50% of mammalian transcripts contain at least one upstream open reading frame (uORF), which are typically one to two orders of magnitude smaller than the downstream main ORF. Most uORFs are thought to be inhibitory as they sequester the scanning ribosome, but in some cases allow for translation re-initiation. However, termination in the 5' UTR at the end of uORFs resembles pre-mature termination that is normally sensed by the nonsense-mediated mRNA decay (NMD) pathway. Translation re-initiation has been proposed as a method for mRNAs to prevent NMD. Here we test how uORF length influences translation re-initiation and mRNA stability. Using custom 5' UTRs and uORF sequences, we show that re-initiation can occur on heterologous mRNA sequences, favors small uORFs, and is supported when initiation occurs with more initiation factors. After determining reporter mRNA half-lives and mining available mRNA half-life datasets for cumulative uORF length, we conclude that translation re-initiation after uORFs is not a robust method for mRNAs to evade NMD. Together, these data support a model where uORFs have evolved to balance coding capacity, translational control, and mRNA stability.

RUNNING TITLE: Translational control by uORF length

KEYWORDS: eIF; mRNA decay; non-canonical initiation; ribosome; translational control

INTRODUCTION

Canonical eukaryotic translation initiation follows a cap- and scanning-dependent mode [1,2]. First, the 5' m^7G cap recruits the eIF4F complex via the eIF4E cap-binding protein. eIF4F also provides an opportunity for the mRNA to form a close-loop conformation with the poly(A) tail and PABP that is thought to increase translation efficiency. eIF4F then recruits the 43S pre-initiation complex (PIC), which is comprised of the 40S small ribosomal subunit, eIF1, eIF1A, eIF3, eIF5, and the ternary complex (eIF2•GTP•Met-$tRNA_i^{Met}$). The 43S PIC then scans 5' to 3' in search of an AUG start codon. The 48S initiation complex is formed after start codon recognition and eIF2 subsequently hydrolyzes GTP to release Met-$tRNA_i^{Met}$. After initiation factors dissociate, initiation is completed once the 60S large ribosomal subunit joins with aid from eIF5B•GTP to form the complete 80S ribosome.

Due to the 5' to 3' nature of the scanning 43S PIC, the first AUG start codon in the mRNA is often recognized most efficiently and is primarily used for protein synthesis—albeit the surrounding context (i.e., Kozak sequence) and distance from the 5' end influences start codon recognition. A major hurdle for the ribosome from recognizing the AUG start codon of the protein encoding main open reading frame (ORF) is the presence of upstream start codons and upstream open reading frames (uORFs) in the 5' untranslated region (UTR). Nearly 50% of mammalian 5' UTRs harbor at least one uORF [3]. Some uORFs do not present much of an obstacle as recent work has shown that the footprint of the eIF4F complex creates a “blind spot” in the first ~50 nucleotides (nts) for the ribosome and only AUG start codons after this point are efficiently recognized [4]. However, 5' UTRs are often larger than the length of this blind spot and contain multiple uORFs. In fact, stress response mRNAs contain multiple evolutionary conserved uORFs that are thought to be key regulators of translation [5,6]. For example, the *ATF4* mRNA 5' UTR contains at least two uORFs. It has been proposed for *ATF4* mRNA that most scanning PICs initiate at the start codon of uORF1 and that translation of uORF1 stimulates or

favors translation re-initiation at uORF2 under normal physiological conditions. Re-initiation occurs when a ribosome terminates and releases the polypeptide but fails to have the small 40S subunit recycled off and subsequently continues to scan downstream in search of a start codon [5,7]. Since uORF2 overlaps the ATF4 coding sequence and ends downstream from the beginning of the ATF4 coding sequence, ATF4 protein is not efficiently synthesized unless uORF2 is skipped during cell stress [5-7].

It has been proposed that re-initiation could negate the ribosome from triggering nonsense mediated mRNA decay (NMD) during termination at uORFs in the 5' UTR or at premature termination codons within the major ORF since the late steps of termination and/or ribosome recycling are not completed [8-15]. In mammalian cells, NMD is triggered by termination ≥50 nts upstream of an exon-junction complex (EJC) or upstream of long 3' UTRs [16,17]. Ribosomes that re-initiate would also displace EJCs and mRNA-bound Upf1 further preventing NMD during subsequent rounds of translation [15]. In this report, we test how uORF length affects translation re-initiation and mRNA stability. Our data show that uORF length robustly controls re-initiation but higher re-initiation efficiency does not protect mRNAs from NMD. Reporter mRNAs with small or large uORFs had orders of magnitude difference in allowing re-initiation but had remarkably similarly shortened half-lives. Transcriptome-wide analysis of cumulative uORF length and mRNA half-life also suggests whether an uORF is translated, not its length or re-initiation efficiency, is the primary influence if termination in the 5' UTR stimulates mRNA decay.

RESULTS AND DISCUSSION

To specifically measure re-initiation and how it impacts mRNA stability, we designed nanoLuciferase (nLuc) reporters to harbor an uORF that maximally prevents leaky scanning of

the PIC and avoids reporter signal from ribosome readthrough. This was achieved by using a synthetic 5' UTR with three important elements: i) a 72 nt CAA-repeat leader sequence to allow for an unstructured sequence that is in the optimal length window for cap- and scanning-dependent initiation [4], ii) an uORF comprised of three AUG start codons in perfect Kozak context (3XAUG), iii) and a 16 nt unstructured CAA-repeat linker between the uORF and nLuc start codon to separate and frameshift the ORFs (**Fig. 1A**).

We next used ribosome toeprinting to confirm that the 3XAUG start codon uORF prevents leaky scanning and sequesters all detectable scanning PICs from the nLuc ORF (**Fig. 1B, C**). *In vitro* translation extracts were pre-incubated with lactimidomycin (binds to the E site of the 60S subunit) to robustly inhibit the first translocation cycle of the 80S ribosome after initiation [18]. The 20 nt FAM-labeled reverse primer targeted a region in the nLuc coding sequence that was downstream enough to detect toeprints of 80S ribosomes at the AUG start codon of nLuc (175 nt), which would be present if leaky scanning past the uORF occurred, and the uORF (218 nt). Indeed, for the uORF reporter mRNA, toeprint signal mapped primarily to the first and second AUG start codon of the uORF (218 and 210 nt, respectively), with some detectable signal mapping to the third uORF start codon (**Fig. 1C**). Signal from unused primer is seen at 20 nt. Importantly, no signal was mapped to the start codon of nLuc (175 nt) which is in alignment with the design for the uORF to trap all scanning PICs.

As an additional control, we mutated the AUG start codons of the uORF to AAA (which does not support initiation) to allow all scanning PICs to bypass the uORF and initiate at the AUG start codon of nLuc (**Fig. 1A**). As expected, ribosome toeprints of this mutant uORF reporter produced signal that only mapped to the AUG start codon of nLuc (**Fig. 1D**). Similar results were seen with the uORF consisting of 10 consecutive AUG codons, but because the start codons were not in optimal context more initiation within the uORF was observed

(**Supplemental Fig. S1**). Together, these data support that the designed uORF can robustly capture all detectable scanning PICs and allows all luciferase signal to be generated from re-initiation.

Consistent with uORFs generally being translational repressive elements, the small uORF repressed translation of nLuc 5-10-fold *in vitro* and in HeLa cells (**Fig. 2A, B**). Translation of nLuc was rescued when the AUG start codons in the uORF were mutated to AAA codons. Given that the average uORF in mammalian transcripts are much smaller than the main annotated protein coding ORFs (16 codons vs 460 codons) [3], we tested how expanding the uORF length would affect re-initiation. The uORF was expanded from 9 codons to 561 codons by inserting the HaloTag-GFP (HT-GFP) coding sequence immediately downstream from the 3XAUG start codons and upstream of the stop codon. In both *in vitro* and in HeLa cells, the large uORF repressed translation of nLuc orders of magnitude more than the small uORF (**Fig. 2A, B**). To confirm this observation was not due to differences in the sequence that the ribosome occupies during initiation and termination in the large uORF, we made a truncated HT-GFP uORF (total of 24 codons) that preserved the first and last 24 nt of the HT-GFP coding sequence (**Fig. 1A**). The truncated uORF rescued expression compared to the larger uORF *in vitro* and in HeLa cells (**Fig. 2A, B**). However, because the truncated uORF is still ~3X larger than the small uORF, it was more repressive than the small uORF.

It remains possible that the greater repression of the large uORF could be due to the HT-GFP coding sequence forming an unexpected secondary structure that inhibited the ribosome or sequestered the 5' cap. This seems unlikely because we obtained equivalent results with a different but equally large uORF sequence (**Supplemental Fig. S2 and Supplemental Fig. S3**). Additionally, the sequence that could influence an initiating ribosome based off the known ribosome footprint size was preserved between the large and truncated

uORFs. Nevertheless, we further tested this by fusing the 3XAUG start codon to a P2A "ribosome skipping motif" and the nLuc ORF (**Fig. 2C**). The P2A motif allows the ribosome to release the nascent polypeptide but continue elongation [19,20]. Thus, in this test, the same nLuc polypeptide from both reporters is being synthesized and assayed. Consistent with the HT-GFP-P2A reporter sequence being 3.5X larger than the control (776 codons vs 224 codons), the larger HT-GFP-P2A reporter was translated ~4X less (**Fig. 2D**). When the nLuc signal was normalized for ORF length, the difference between the two sized reporters was only ~10% (**Fig. 2D**). This small difference suggests that ribosomes initiate at the small and large uORFs almost equally and does not rationally explain the 100-1000-fold difference seen in the re-initiation reporters (**Fig. 2A, B**). Together, these data support that large uORFs are more repressive than small uORFs because they allow less translation re-initiation.

Recent reports using crosslinking and immunocapture of initiation factors (eIFs) have provided evidence that some eIFs may linger on the ribosome after initiation and could aid re-initiation if present after termination of a small uORF [21,22]. With this in mind, we next tested if re-initiation after small uORFs is as efficient if they are translated by ribosomes requiring less eIFs during initiation. We achieved this by taking advantage of class I-IV cap-independent viral internal ribosome entry sites (IRES) that require subsets of eIFs [23] (**Fig. 3A, B**). In certain cases, IRES can stimulate initiation without requiring any eIFs or the initiator tRNA (**Fig. 3A, B**). It was apparent that IRES-mediated translation is less efficient than canonical translation and we were only able to test the effect of the small uORF and still have luciferase signal above background in the linear range. In alignment with the model that eIFs that are stubbornly bound to the ribosome aid in re-initiation [21,22], we observed that the uORFs repressed translation greater when the IRES utilized less initiation factors (**Fig. 3B,C**). For example, the uORF translated by the Cricket Paralysis Virus Intergenic Region (CrPV IGR) IRES, which does not require any eIFs or the initiator tRNA [24,25], repressed translation the most (**Fig. 3C**).

The presence of uORFs presents a major challenge to mRNAs as termination in the 5' UTR resembles how ribosomes recognize deleterious premature termination codons (PTCs) through nonsense-mediated mRNA decay (NMD). Termination ≥50 nts upstream of an exon junction complex (EJC) typically robustly triggers NMD in mammalian cells [16,17]. However, termination at uORFs almost certainly occurs upstream of an EJC if the uORF is translated during the pioneering round of translation. Others have postulated that re-initiation could be a method that ribosomes use to bypass triggering NMD at PTCs within the major ORF [8-15]. We next asked if large uORFs, which do not favor re-initiation (**Fig. 2**, **Supplemental Fig. S2, Supplemental Fig. S3**), stimulate mRNA decay more than small uORFs in cells. Using a Tet-Off system to selectively and robustly turn off reporter transcription, we measured reporter mRNA levels over an 8 hr time course after the addition of doxycycline. In this experiment, we used the same reporter design as described in **Fig. 1** with the addition of a small functional intron in nLuc ORF. As a positive control, we included a reporter that harbors an intron in the 3' UTR which should stimulate NMD and have a shorter mRNA half-life. As expected, the no uORF + 3' UTR intron reporter had a ~2hr shorter half-life than the control no uORF reporter (3.71 ± 0.52 hr vs. 5.84 ± 1.12 hr, respectively) (**Fig. 4B**). The small and large uORF reporters had even shorter mRNA half-lives of 3.16± 0.34 hr and 2.49 ± 0.33 hr, respectively, (**Fig. 4C, D**). Although we see 100-fold difference in re-initiation between the small and large uORF in cells (**Fig. 2B**), the half-life of the large uORF reporter was only ~40 min shorter than the small uORF reporter (**Fig. 4C, D**). We further investigated the connection between uORF length and mRNA stability on a transcriptome-wide scale by mining published BRIC-seq datasets [26,27] and determined the cumulative uORF length for each mRNA. We found no statical difference between the mean mRNA half-life of transcripts that contained varying cumulative uORF lengths in multiple data sets (**Fig. 4E** and **Supplemental Fig. S4**). However, transcripts without an uORF were on average more stable and had longer half-lives (**Fig. 4E** and **Supplemental Fig.**

S4A). Collectively, these data suggest translation re-initiation after uORFs is not a robust method for mRNAs to evade NMD.

Together, our data support a model where uORF length controls re-initiation but not termination-dependent decay. Whether an uORF is translated, not its length or re-initiation efficiency, is the primary influence if termination in the 5' UTR stimulates mRNA decay. The coding capacity of a single mRNA can be increased with larger uORFs, but a trade-off ensues as they do not favor re-initiation and provide less translational regulation (**Supplemental Fig. S4B**). This control can be regulated by leaky scanning, by complete bypass of the uORF if the scanning 40S ribosome lacks a ternary complex (TC), or if the uORF harbors a start codon in sub-optimal Kozak context. For example, many stress response mRNAs harbor uORFs that are not translated during stress conditions [5,6]. During cell stress, phosphorylation of eIF2α at Ser51 inhibits translation but does not block 40S ribosomes lacking a TC from being loaded and scanning. Current models suggest that TCs can then rejoin the scanning 40S ribosome downstream of the uORFs but upstream of the main ORF. Consistent with these models and uORFs being skipped prevents NMD, stress response mRNAs are often elevated during cell stress [28,29]. Clearly, mammalian evolution has not favored large uORFs as the average mammalian uORF is 16 codons [3]. This may be partially explained not only by the fact that large uORFs provide higher chances of ribosome collisions and non-optimal codons that also stimulate mRNA decay [30-33], but they also prevent translation of the downstream main ORF by decreased re-initiation.

MATERIALS AND METHODS

Plasmids

Complete sequences of reporter plasmid inserts are located in the Supplementary Material. All plasmids were derived from previously described pcDNA3.1(+)/AUG-nLuc-3XFLAG and pcDNA3.1-D/CrPV IGR IRES nLuc-3XFLAG [34]. The CrPV IGR IRES reporter was additionally modified to contain a strong hairpin upstream of IRES element to block scanning pre-initiation complexes. The HT-GFP ORF was taken from pHaloTag-EGFP (a gift from Thomas Leonard and Ivan Yudushkin; Addgene plasmid # 86629) [35]. pGL4.13 (encodes Firefly Luciferase [FFLuc]) was obtained from Promega (# E6681). IRES-containing nLuc reporters were generated using an overlapping PCR method and cloned into pcDNA3.1(+) or pcDNA3-1D. The PV IRES template was pcDNA3 RLUC POLIRES FLUC and was a gift from Nahum Sonenberg (Addgene plasmid # 45642) [36]. The EMCV IRES and HCV IRES templates were kind gifts from Aaron Goldstrohm. 5' UTRs, uORFs, introns, and mutations were introduced using the Q5 Site-Directed Mutagenesis Kit (NEB # E0554S).

To make pTet-Off All-In-One plasmids, pcDNA3.1(+)/no uORF nLuc plasmid was subjected to two rounds of mutagenesis using the NEBuilder HiFi DNA Assembly Master Mix with 25 bp overhangs. First, the complete CMV promoter was replaced with the tetracycline-responsive P_{Tight} promoter from pCW57.1-MAT2A (a gift from David Sabatini; Addgene plasmid # 100521) [37]. Second, the neomycin resistance gene coding sequence was replaced with the tTA-Advanced coding sequence from pCW57.1-MAT2A. The different uORF nLuc inserts were then subcloned into this pTet-Off All-In-One backbone at SacI and XbaI sites. The 133 bp chimeric intron from pCI-neo (Promega # E1841) was inserted into the nLuc ORF by using the Q5 Site-Directed Mutagenesis Kit.

All oligonucleotides were obtained from Integrated DNA Technologies. TOP10 *E. coli* cells were used for all plasmid propagation and cloning. Reporters and any mutated sites were fully Sanger sequenced at The Ohio State Comprehensive Cancer Center Genomics Shared

Resource (OSUCCC GSR). All plasmids are available upon request and will be deposited into Addgene upon publication.

In vitro transcription

Reporter mRNAs were synthesized using linearized plasmids as templates for run off transcription with T7 RNA polymerase as previously described [34] with the single exception that XbaI was used to linearize all plasmids. All mRNAs were transcribed at 30°C for 2 hrs using HiScribe T7 High Yield RNA Synthesis Kit (NEB # E2040S) and were co-transcriptionally capped and post-transcriptionally polyadenylated. Non-IRES mRNAs were capped the 3´-O-Me-m7G(5')ppp(5')G anti-reverse cap analog (NEB # S1411L). IRES mRNAs were capped with the A(5')ppp(5')G cap analog (NEB # S1406L). Post-transcriptional polyadenylation was performed using *E. coli* Poly(A) Polymerase (NEB # M0276L). mRNAs were purified using the Zymo RNA Clean and Concentrator-25 (Zymo Research # R1018), eluted in RNase-free water, aliquoted in single use volumes and stored at -80°C.

In vitro translation and luciferase assay

10 µL *in vitro* translation reactions were performed in the linear range using 3 nM mRNA in the Flexi Rabbit Reticulocyte Lysate (RRL) System (Promega # L4540) with final concentrations of reagents at 30% RRL, 10 mM amino acid mix minus Leucine, 10 mM amino acid mix minus Methionine, 0.5 mM MgOAc, 100 mM KCl, and the addition of 8 U murine RNase inhibitor (NEB # M0314L) [34]. Reactions were incubated for 30 min at 30°C, terminated by incubation on ice, and then diluted with 40 µL Glo Lysis Buffer (Promega # E2661). 25 µL of diluted reaction was mixed with 25 µL of prepared Nano-Glo Luciferase Assay System (Promega # N1120) for 5 min in the dark on an orbital shaker. Luminescence was measured using a Promega GloMax Discover Microplate Reader.

Fluorescent ribosome toeprinting

60 µL *in vitro* translation RRL reactions (same final concentrations of reagents as above) were pre-incubated with 50 µM lactimidomycin (5 mM stock in DMSO) for 10 min at 30°C then placed on ice. 25 nM capped and polyadenylated mRNA was added (increased concentration was important to detect weaker signals), gently mixed, and incubated for an additional 10 min at 30°C to allow inhibited 80S ribosomes to form after start codon recognition. To each reaction, 20 µL 5X AMV RT buffer (final: 50mM Tris-HCl (pH 8.3), 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM DTT), 10 µL 10 µM 5'-FAM labeled-reverse primer (20 nt), 2 µL 25mM complete dNTP set (final of each dNTP at 0.5 mM), 6 µL nuclease-free water, and 2 µL AMV Reverse Transcriptase (stock at 20-25 U/µL) was added and reverse transcriptase (RT) was allowed to progress for 35 min at 30°C. The higher $MgCl_2$ concentration in the RT reaction inhibits new initiation complex formation. Control reactions with 25 mM reporter mRNA in water were treated identically and were used to determine background from the RT reaction. FAM-labeled cDNA was extracted by transferring the 100 µL RT reaction to a new microcentrifuge tube with 150 µL nuclease-free water and adding 250 µL saturated Phenol:Chloroform:Isoamyl Alcohol (25:24:1), pH 8. After vigorous mixing for 1 min, samples were centrifuged at room temperature at 16,000 rcf for 5 min. The top aqueous phase was transferred, and re-extracted with saturated Phenol:Chloroform:Isoamyl Alcohol (25:24:1), pH 8. The final aqueous supernatant was then concentrated using the Zymo DNA Clean and Concentrator-5 using a 7:1 ratio following the manufacture's recommendation. FAM-labeled cDNA was eluted in 7 µL nuclease-free water. 5 µL of each eluate was mixed with 10 µL Hi-Di Formamide (Thermo # 4440753), spiked with a LIZ 500 size standard, and subjected to fragment analysis using Applied Biosystems 3130xl Genetic Analyzer with POP-7 polymer with all fragments being reported. To determine which signals were caused by inhibited 80S ribosomes at start codons, signal from the control samples (RNA in water + RT reaction) were subtracted from the reactions with RRL and inhibitor. Primer sequence is included in the Supplemental Table 1.

Cell culture, transfection, and luciferase assay

HeLa cells were obtained from ATCC and maintained in high glucose DMEM supplemented with 10% heat-inactivated FBS, 1% penicillin-streptomycin, and 1% non-essential amino acids in standard tissue culture-treated plastics. HeLa cells were seeded 24 hr before transfection so that on the day of transfection they were at 50% confluency. ViaFect (Promega # E4982) was used at a 3:1 ratio with 1 µg total plasmid (500 ng nLuc plasmid + 500 ng pGL4.13) in 100 µL in Opti-MEM (Thermo Fisher # 31985062). For 96-well plates, HeLa cells were transfected with a total of 100 ng (10 µL of the transfection mix). 24 hr post transfection, media was aspirated, and cells were lysed in 100 µL Glo Lysis Buffer (Promega # E2661) for 10 min on an orbital shaker. 25 µL of lysate was then mixed with 25 µL of ONE-Glo (Promega # E6120) or 25 µL of prepared Nano-Glo Luciferase Assay System (Promega # N1120) and detected as described above. nLuc signal was then normalized to FFLuc signal of the same sample to normalize for transfection efficiency.

Tet-Off System and mRNA decay measurements

HeLa cells were seeded and maintained in complete media as described above, but supplemented with 10% Tet-approved FBS (Thermo Fisher # A47364-01). 24 hr post seeding in a 10 cm plate, 50% confluent cells were transfected with 6 µg of total plasmid (3 µg pTet-Off All-In-One plasmids + 3 µg pGL4.13) using ViaFect. 24 hr post transfection, cells were trypsinized, diluted, and seeded in five 12-well dishes. 48 hrs later, when cells were ~75% confluent, media was replaced with media containing 2 µg/mL doxycycline (MP Biomedicals # 195044) (stock at 1 mg/mL in water) in the dark. At the indicated time points, total RNA was extracted using TRIzol (Thermo Fisher # 15596018) following the manufacture's recommendations. 500 ng of total RNA was then DNase-treated with amplification grade DNase (Thermo Fisher # 18068015). The entire final 11 µL DNase reaction was then used to synthesize cDNA with oligo-

dT and random hexamers using the iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad # 1708841) following the manufacture's protocol. cDNA was then diluted 1:100 and 1.5 µL was use per 15 µL reaction with iTaq Universal SYBR Green Supermix (Bio-Rad # 1725122) and 250 nM primers (final) on a Bio-Rad CFX Connect Real-Time PCR Detection System using Bio-Rad CFX Maestro software to calculate expression levels. Reporter levels were normalized to RPS17 and half-lives were calculated using first order exponential decay trend lines, calculated by non-linear regression in GraphPad Prism 9.1.2. The 95% confidence intervals were plotted along the mean of three biological replicates. Reverse transcriptase minus reactions were used to confirm less than 2% of reporter signal is from contaminating plasmid DNA. All primer sequences are available in **Supplemental Table S1**.

Bioinformatic analysis of uORF length and mRNA decay

mRNA half-lives measured using BRIC-seq was obtained from published literature [26,27,38]. Custom phyton scripts were written to first calculate the cumulative uORF length in each mRNA and then assigned it to the previously defined half-life. Both scripts have been deposited in GibHub (github.com/michaelkearse/uORF_Half-life). The scripts utilize the RefSeq Transcript and RefSeq Reference Genome Annotation files for human genome build 38 from NCBI for reference. If a gene had multiple transcripts listed, then the longest transcript (usually isoform 1) was used. Only AUG-encoded uORFs were determined. Transcripts were then binned by cumulative uORF length in five codon increments with at least 100 mRNAs in each bin. The total sample number per bin is included in **Supplemental Table S2**, along with the cumulative uORF length and half-life for each mRNA.

ACKNOWLEDGEMENTS

Experiments were conceived and performed by PJR and MGK with input and help from LPB. The bioinformatics analysis was performed by JAS with input from AHMB. The manuscript was written by PJR and MGK with input from JAS, LPB, and AHMB. PJR was supported by NIH grants T32GM086252 and T32GM141955. This work was supported by NIH grant R00GM126064 (to MGK). We thank members of the Kearse lab for critically reading the manuscript. We also thank Christine Daugherty at The Ohio State Comprehensive Cancer Center Genomics Shared Resource (OSUCCC GSR) for her critical help with the fluorescent toeprinting analysis. The OSUCCC GSR is supported by NIH grant P30CA016058.

REFERENCES

1. Shirokikh NE & Preiss T. (2018) Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscip Rev RNA,* 9: e1473.
2. Hinnebusch AG. (2017) Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem Sci,* 42: 589-611.
3. Calvo SE, Pagliarini DJ & Mootha VK. (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci U S A,* 106: 7507-12. PMC2669787
4. Brito Querido J, Sokabe M, Kraatz S, Gordiyenko Y, Skehel JM, Fraser CS *et al.* (2020) Structure of a human 48S translational initiation complex. *Science,* 369: 1220-1227. PMC7116333
5. Young SK & Wek RC. (2016) Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. *J Biol Chem,* 291: 16927-35. PMC5016099
6. Wek RC. (2018) Role of eIF2alpha Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harb Perspect Biol,* 10: PMC6028073
7. Vattem KM & Wek RC. (2004) Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A,* 101: 11269-74. PMC509193
8. Zhang J & Maquat LE. (1997) Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. *EMBO J,* 16: 826-33. PMC1169683
9. Buisson M, Anczukow O, Zetoune AB, Ware MD & Mazoyer S. (2006) The 185delAG mutation (c.68_69delAG) in the BRCA1 gene triggers translation reinitiation at a downstream AUG codon. *Hum Mutat,* 27: 1024-9.
10. Paulsen M, Lund C, Akram Z, Winther JR, Horn N & Moller LB. (2006) Evidence that translation reinitiation leads to a partially functional Menkes protein containing two copper-binding sites. *Am J Hum Genet,* 79: 214-29. PMC1559486

11. Stump MR, Gong Q, Packer JD & Zhou Z. (2012) Early LQT2 nonsense mutation generates N-terminally truncated hERG channels with altered gating properties by the reinitiation of translation. *J Mol Cell Cardiol,* 53: 725-33. PMC3518410
12. Pereira FJ, Teixeira A, Kong J, Barbosa C, Silva AL, Marques-Ramos A *et al.* (2015) Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity. *Nucleic Acids Res,* 43: 6528-44. PMC4513866
13. Hulsebos TJ, Kenter S, Verhagen WI, Baas F, Flucke U & Wesseling P. (2014) Premature termination of SMARCB1 translation may be followed by reinitiation in schwannomatosis-associated schwannomas, but results in absence of SMARCB1 expression in rhabdoid tumors. *Acta Neuropathol,* 128: 439-48.
14. Jagannathan S & Bradley RK. (2016) Translational plasticity facilitates the accumulation of nonsense genetic variants in the human population. *Genome Res,* 26: 1639-1650. PMC5131816
15. Dyle MC, Kolakada D, Cortazar MA & Jagannathan S. (2020) How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. *Wiley Interdiscip Rev RNA,* 11: e1560.
16. Lykke-Andersen S & Jensen TH. (2015) Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol,* 16: 665-77.
17. Yi Z, Sanjeev M & Singh G. (2021) The Branched Nature of the Nonsense-Mediated mRNA Decay Pathway. *Trends Genet,* 37: 143-159. PMC7854845
18. Schneider-Poetsch T, Ju J, Eyler DE, Dang Y, Bhat S, Merrick WC *et al.* (2010) Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol,* 6: 209-217. PMC2831214
19. Szymczak AL & Vignali DA. (2005) Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther,* 5: 627-38.
20. Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L *et al.* (2017) Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep,* 7: 2193. PMC5438344
21. Bohlen J, Fenzl K, Kramer G, Bukau B & Teleman AA. (2020) Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. *Mol Cell,* 79: 561-574 e5.
22. Wagner S, Herrmannova A, Hronova V, Gunisova S, Sen ND, Hannan RD *et al.* (2020) Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes. *Mol Cell,* 79: 546-560 e7. PMC7447980
23. Jackson RJ, Hellen CU & Pestova TV. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol,* 11: 113-27. PMC4461372
24. Pestova TV & Hellen CU. (2003) Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev,* 17: 181-6. PMC195975
25. Petrov A, Grosely R, Chen J, O'Leary SE & Puglisi JD. (2016) Multiple Parallel Pathways of Translation Initiation on the CrPV IRES. *Mol Cell,* 62: 92-103. PMC4826567
26. Maekawa S, Imamachi N, Irie T, Tani H, Matsumoto K, Mizutani R *et al.* (2015) Analysis of RNA decay factor mediated RNA stability contributions on RNA abundance. *BMC Genomics,* 16: 154. PMC4359779
27. Tani H, Mizutani R, Salam KA, Tano K, Ijiri K, Wakamatsu A *et al.* (2012) Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res,* 22: 947-56. PMC3337439
28. Gardner LB. (2008) Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. *Mol Cell Biol,* 28: 3729-41. PMC2423288

29. Karam R, Lou CH, Kroeger H, Huang L, Lin JH & Wilkinson MF. (2015) The unfolded protein response is shaped by the NMD pathway. *EMBO Rep,* 16: 599-609. PMC4428047
30. Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N *et al.* (2015) Codon optimality is a major determinant of mRNA stability. *Cell,* 160: 1111-24. PMC4359748
31. Simms CL, Yan LL & Zaher HS. (2017) Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol Cell,* 68: 361-373 e5. PMC5659757
32. Hanson G & Coller J. (2018) Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol,* 19: 20-30. PMC6594389
33. D'Orazio KN & Green R. (2021) Ribosome states signal RNA quality control. *Mol Cell,* 81: 1372-1383. PMC8041214
34. Kearse MG, Green KM, Krans A, Rodriguez CM, Linsalata AE, Goldstrohm AC *et al.* (2016) CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol Cell,* 62: 314-322. PMC4854189
35. Ebner M, Lucic I, Leonard TA & Yudushkin I. (2017) PI(3,4,5)P3 Engagement Restricts Akt Activity to Cellular Membranes. *Mol Cell,* 65: 416-431 e6.
36. Poulin F, Gingras AC, Olsen H, Chevalier S & Sonenberg N. (1998) 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. *J Biol Chem,* 273: 14002-7.
37. Gu X, Orozco JM, Saxton RA, Condon KJ, Liu GY, Krawczyk PA *et al.* (2017) SAMTOR is an S-adenosylmethionine sensor for the mTORC1 pathway. *Science,* 358: 813-818. PMC5747364
38. Imamachi N, Salam KA, Suzuki Y & Akimitsu N. (2017) A GC-rich sequence feature in the 3' UTR directs UPF1-dependent mRNA decay in mammalian cells. *Genome Res,* 27: 407-418. PMC5340968

Figure 1

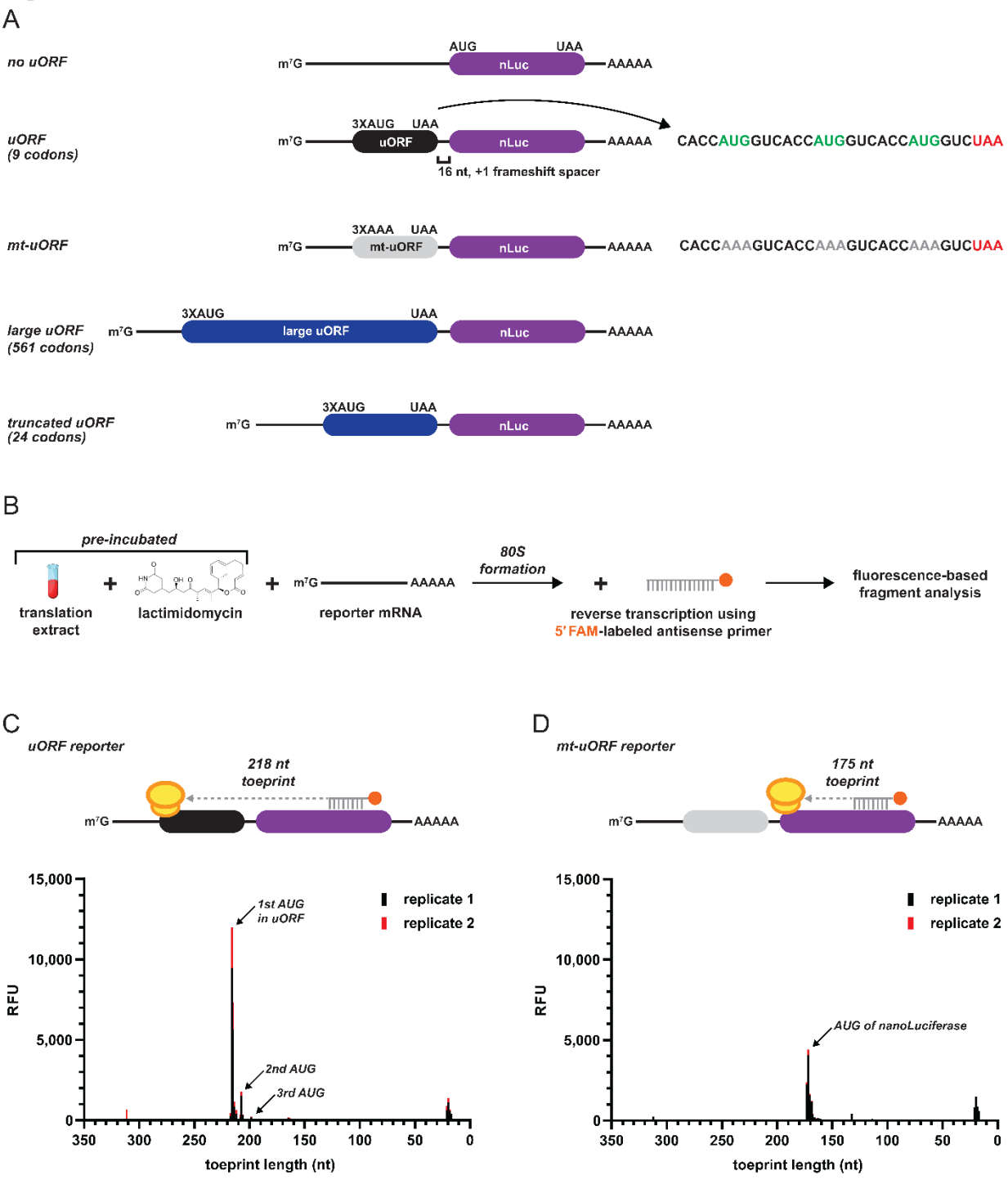


Figure 1. A small uORF with three start codons in perfect context is able to sequester all scanning initiation complexes. A) Design of re-initiation specific nanoLuciferase (nLuc) reporters used in this study. A 16 nt spacer between the variable-sized uORF and nLuc ORF

allows specific detection of re-initiation. B) Schematic of ribosome toeprinting with FAM-labeled primers to detect sites of initiation with lactimidomycin pre-incubation. C-D) Ribosome toeprinting of 80S ribosomes after start codon recognition on small uORF nLuc reporter mRNA (C) and mutated uORF nLuc reporter mRNA (D). Signal from unused primer is seen at 20 nt. Signal from duplicate samples is shown in black and red.

Figure 2

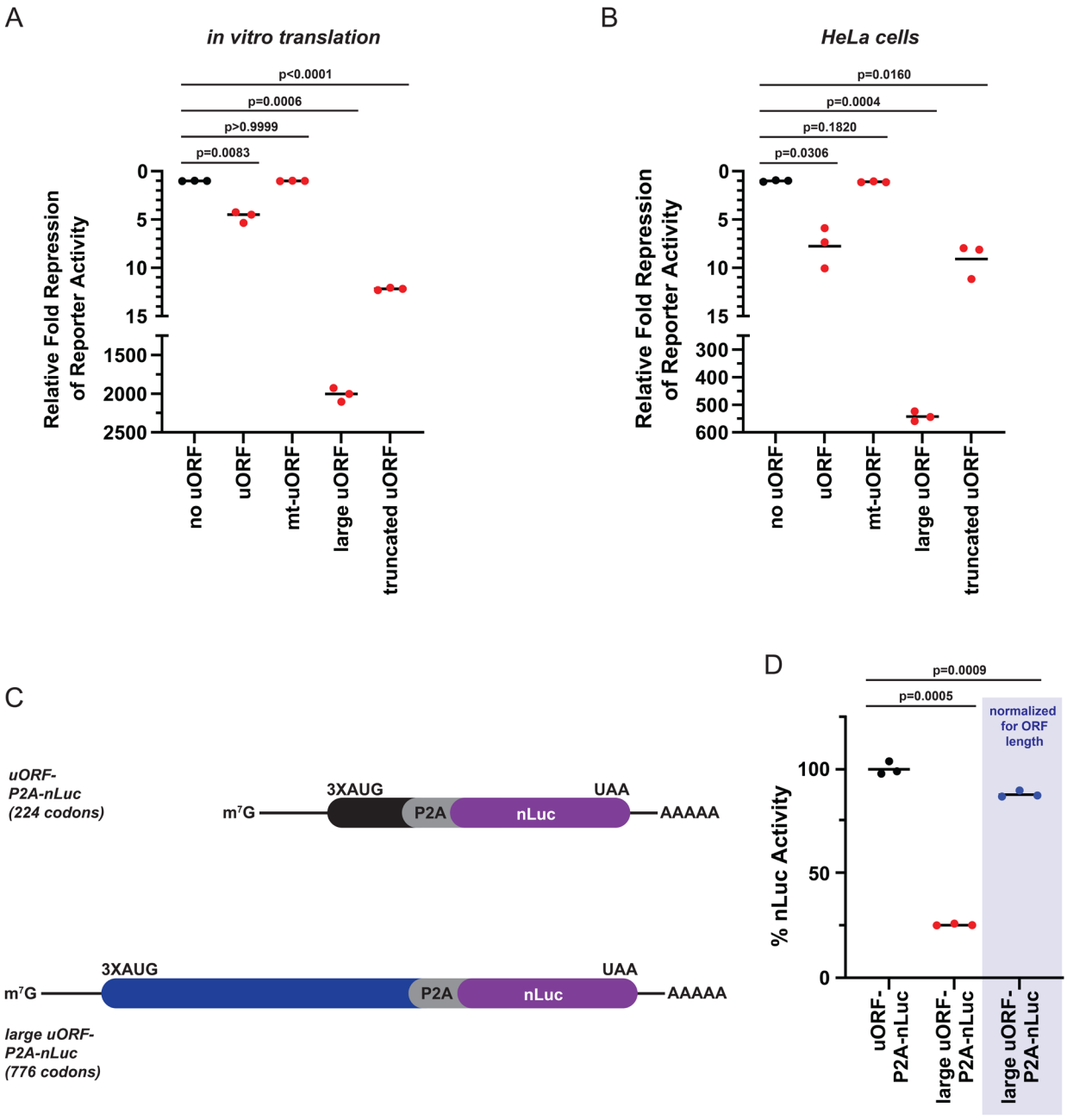


Figure 2. Translation re-initiation is more efficient after small uORFs *in vitro* and in HeLa cells. A-B) Response of nLuc reporters that harbor a small, mutant, large, or truncated uORF from *in vitro* translation (A) and in HeLa cells (B). n=3 biological replicates. Bar represents the mean. C) Insertion of the P2A "ribosome skipping motif" (gray) was used to assess the relative translation efficiency of reporters that harbored a 3XAUG start codon sequence without and with

the large HTGFP (blue) sequence upstream of the nLuc coding sequence (purple). D) Relative nLuc activity of the small and larger ORF reporters from *in vitro* translation. Shaded in blue is the signal of the large uORF-P2A-nLuc reporter normalized for ORF length. n=3 biological replicates. Bar represents the mean. Comparisons were made using a two-tailed unpaired *t*-test with Welch's correction.

Figure 3

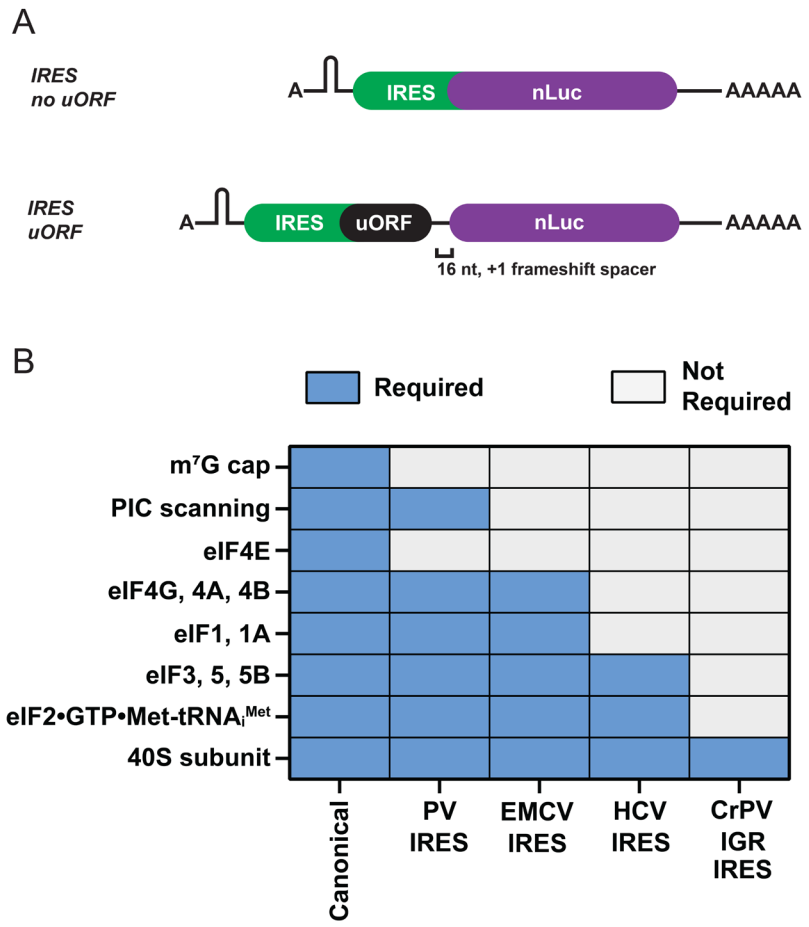


Figure 3. uORFs translated by IRESs that require less initiation factors permit less re-initiation. A) Schematic of A-capped IRES-mediated re-initiation reporters. A stable hairpin was insert upstream of the IRES to block scanning ribosomes B) Requirements of canonical initiation and class I-IV viral IRES-mediated initiation. C) Response of canonical initiation and viral IRES-dependent nLuc reporters without and with small uORFs from *in vitro* translation. n=3 biological replicates. Bar represents the mean. PV=poliovirus; EMCV=encephalomyocarditis virus; HCV=hepatitis C virus; CrPV IGR=cricket paralysis virus intergenic region; PIC=pre-initiation complex; eIF=eukaryotic initiation factor. Comparisons were made using a two-tailed unpaired *t*-test with Welch's correction.

Figure 4

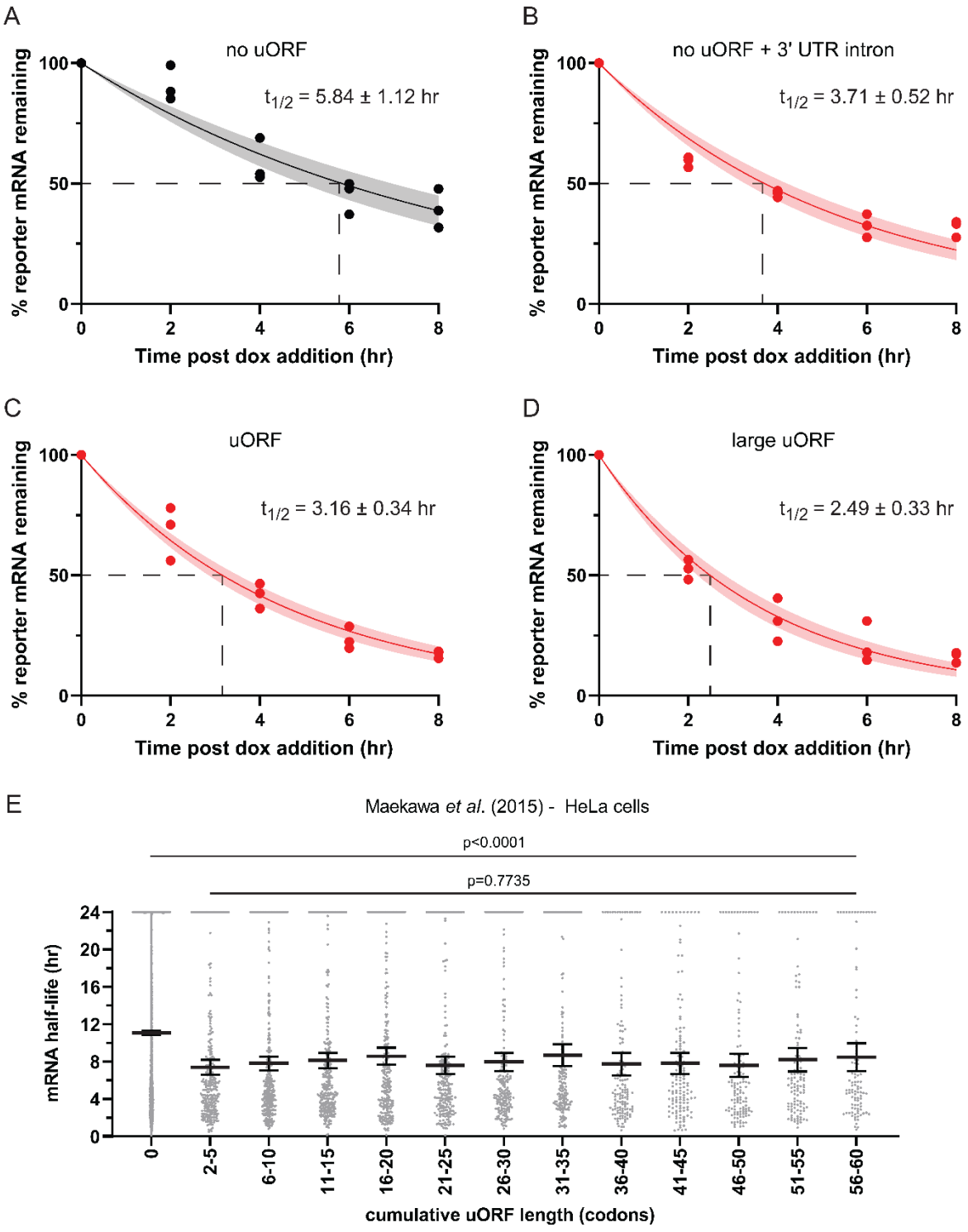
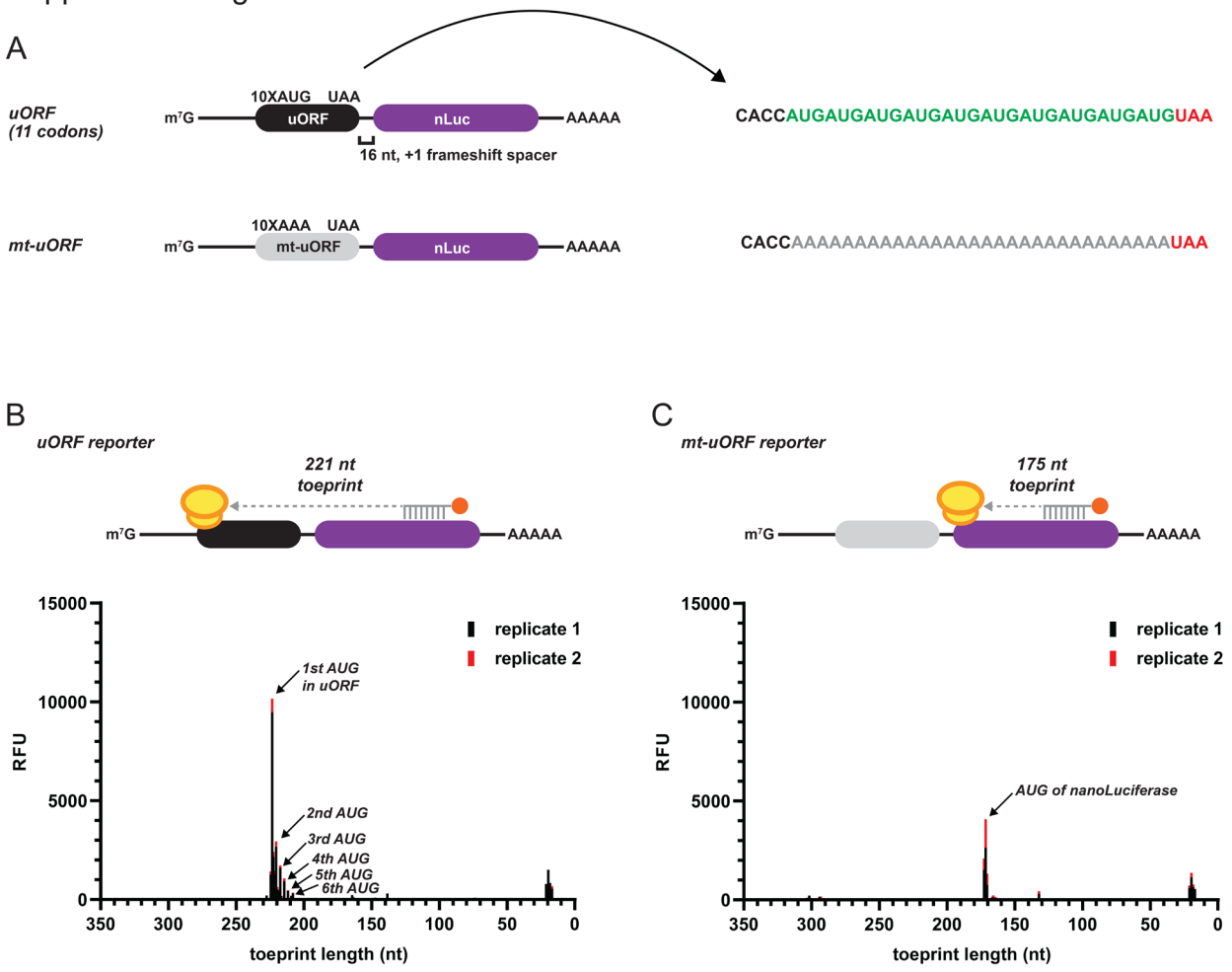


Figure 4. uORF length has minimal effect on mRNA stability in cells. A-D) A Tet-Off system triggered with 2 μg/mL doxycycline (dox) was used to determine reporter mRNA half-lives in

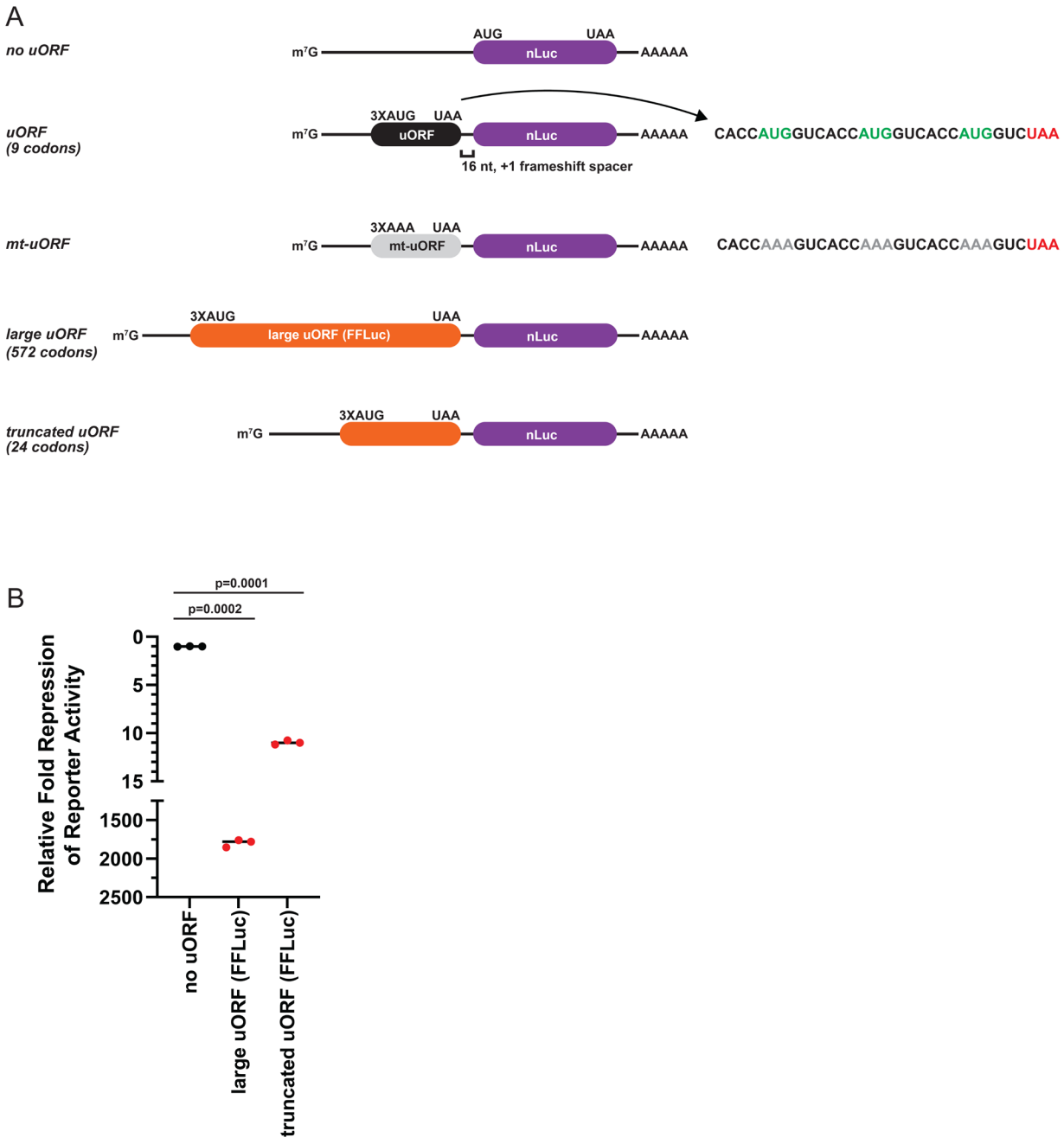
HeLa cells. n=3 biological replicates. A non-linear regression was used to calculate the mRNA half-lives and is shown as the line with the 95% confidence interval included as a watermark. E) Transcriptome-wide comparison of cumulative uORF length and mRNA half-life. Bar represents the mean ± 95% confidence interval. One-way Welch's ANOVA was used to compare between cumulative uORF length bins.

Supplemental Figure S1



Supplemental Figure S1. A small uORF with 10 consecutive start codons is able to sequester all scanning initiation complexes. A) Design of re-initiation specific nanoLuciferase (nLuc) reporters harboring an uORF encoded by 10 consecutive AUG start codons (uORF) or AAA codons (mt-uORF). A 16 nt spacer between the uORF and nLuc ORF allows specific detection of re-initiation. B-C) Ribosome toeprinting of lactimidomycin inhibited 80S ribosomes after start codon recognition on small uORF nLuc reporter mRNA (B) and mutated uORF nLuc reporter mRNA (C). Signal from duplicate samples is shown in black and red.

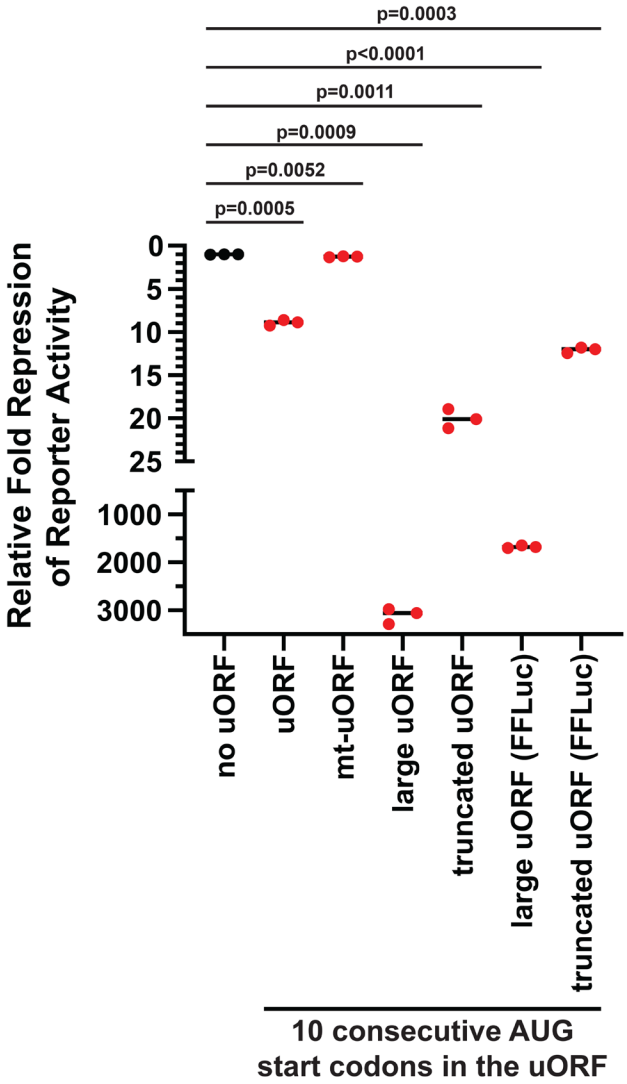
Supplemental Figure S2



Supplemental Figure S2. Less translation re-initiation after large uORFs is not strictly sequence specific. A) The large uORF sequence described in **Fig. 1** was switched from HT-GFP to Firefly Luciferase (FFLuc). Three AUG start codons in perfect context was used to trap all scanning initiation complexes at the uORF. B) Response of nLuc reporters that harbor a large (FFLuc) or truncated uORF (FFLuc) from *in vitro* translation. Similar repression and rescue

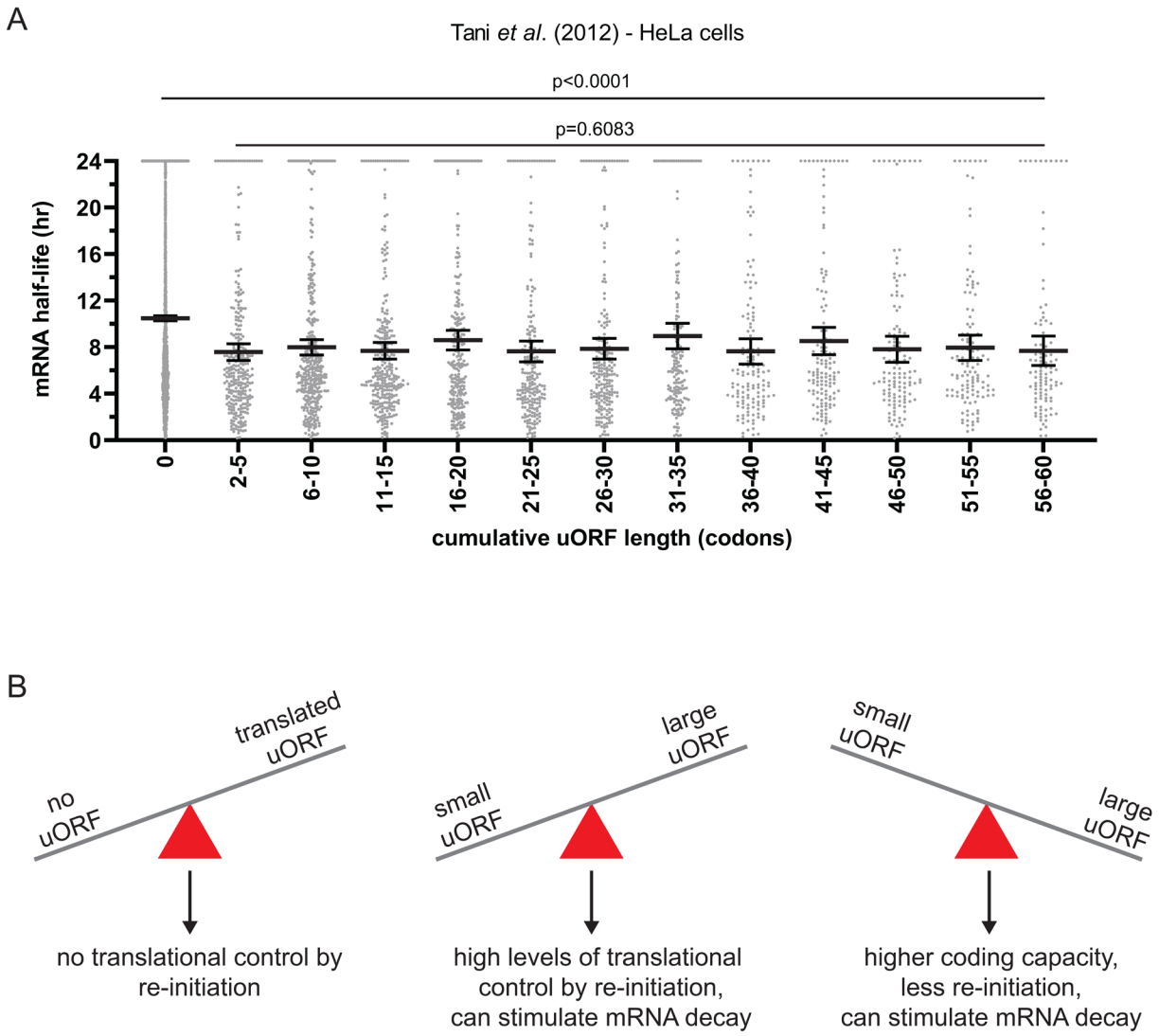
were seen here with FFLuc as the large uORF as observed with HT-GFP sequence in **Fig. 2**. n=3 biological replicates. Bar represents the mean. Comparisons were made using a two-tailed unpaired *t*-test with Welch's correction.

Supplemental Figure S3



Supplemental Figure S3. uORF length control of re-initiation is seen with uORFs harboring 10 consecutive AUG start codons. Response of nLuc reporters that harbor a small, mutant, large, or truncated uORF from *in vitro* translation. All uORFs had 10 consecutive AUG start codons as described in **Supplemental Fig. S1** (instead of the three AUG start codons in perfect Kozak context in **Fig. 1**). Large and truncated uORFs using the FFLuc sequence are clearly labeled from those that used the original HT-GFP sequence. n=3 biological replicates. Bar represents the mean. Comparisons were made using a two-tailed unpaired *t*-test with Welch's correction.

Supplemental Figure S4



Supplemental Figure S4. uORF length has minimal effect on mRNA stability in cells. A) Transcriptome-wide analysis of cumulative uORF length and mRNA half-life from Tani *et al*. (2012). One-way Welch's ANOVA was used to compare between cumulative uORF length bins. B) Model of uORF length controlling re-initiation and mRNA stability.

SUPPLEMENTAL MATERIAL

Supplemental Table S1. Primers used in this study

Primer name	Sequence (5'-3')
F_nLuc qPCR	CAGCGGGCTACAACCTGGAC
R_nLuc qPCR	AGCCCATTTTCACCGCTCAG
F_Hs RPS17 qPCR	GTTCGCACCAAAACCGTGAAG
R_Hs RPS17 qPCR	CGCTTGTTCGTGTGGAAGT
R_FAM nLuc ToePrint	FAM-ACGGGATGATGACATGGATG

Supplemental Table S2. List of cumulative uORF length and mRNA half-life re-analyzed from Tani *et al*. (2012) and Maekawa *et al*. (2015).

Sequences of reporters used in this study

The plasmid backbone, reporter name, and restriction sites are given for each construct. For reference, nanoLuciferase (nLuc) is always highlighted in pink. The multiple start sites for the uORF and single stop codon are highlighted in yellow and red, respectively. The HT-GFP sequence, FFLuc sequence and IRES sequences are highlighted in green, blue, and turquoise, respectively. The unstructured CAA repeats are in lowercase. The P2A sequence is highlighted in gray. The strong hairpins used upstream of the IRES sequences is in **bold**. For the pTet-Off All-In-One (pTet-Off AIO) plasmids, the tight TRE promotor is shown in purple and the intron is underlined.

pcDNA3.1+/no uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCT
GGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGT
AACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTC
ATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGG
TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAAT
CGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGT
GTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGA
CGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGAC
CGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAA
AGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtctaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGT
TGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTG
TGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAG
CGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGG
CGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCAC

TTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCG
ACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAA
CAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCT
CCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTC
TGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGA
CGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/mt-uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccaaagtcaccaaagtcaccaaagtctaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCG
TTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGT
GTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGA
GCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCG
GCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATC
GACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTA
ACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGC
TCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATT
CTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATG
ACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/10AUG uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatgatgatgatgatgatgatgatgatgtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATT
TCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGA
GGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCC
TGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGA
GCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCA
TCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATG
ATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACT
GTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGAC
GGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACG
CATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG
GATGACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/10AAA mt-uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccaaaaaaaaaaaaaaaaaaaaaaaaaaaataaacaacaacaacaacaaATGGTCTTCACACTCGAA
GATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAG
GGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTG
TCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCT
GAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGA
TCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
ATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATC
ACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCC
GACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGA
ACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTAC
AAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/large uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtcGCAGAAATCGGTACTGGCTTTCCATTCGACCCCCATTATG
TGGAAGTCCTGGGCGAGCGCATGCACTACGTCGATGTTGGTCCGCGCGATGGCACCCCT
GTGCTGTTCCTGCACGGTAACCCGACCTCCTCCTACGTGTGGCGCAACATCATCCCGCAT
GTTGCACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAAATCCGACAAA
CCAGACCTGGGTTATTTCTTCGACGACCACGTCCGCTTCATGGATGCCTTCATCGAAGCCC
TGGGTCTGGAAGAGGTCGTCCTGGTCATTCACGACTGGGGCTCCGCTCTGGGTTTCCACT
GGGCCAAGCGCAATCCAGAGCGCGTCAAAGGTATTGCATTTATGGAGTTCATCCGCCCTA
TCCCGACCTGGGACGAATGGCCAGAATTTGCCCGCGAGACCTTCCAGGCCTTCCGCACCA
CCGACGTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTATCGAGGGTACGCTGCCGA
TGGGTGTCGTCCGCCCGCTGACTGAAGTCGAGATGGACCATTACCGCGAGCCGTTCCTGA
ATCCTGTTGACCGCGAGCCACTGTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTGAGC
CAGCGAACATCGTCGCGCTGGTCGAAGAATACATGGACTGGCTGCACCAGTCCCCTGTCC
CGAAGCTGCTGTTCTGGGGCACCCCAGGCGTTCTGATCCCACCGGCCGAAGCCGCTCGC
CTGGCCAAAAGCCTGCCTAACTGCAAGGCTGTGGACATCGGCCCGGGTCTGAATCTGCTG
CAAGAAGACAACCCGGACCTGATCGGCAGCGAGATCGCGCGCTGGCTGTCGACGCTCGA
GATTTCCGGCCACCGGTCGCAACCTTGGGTACCGCGGGCCCGGGATCCACCGGTCGCAA
CCTTGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCAC
CTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCC
CACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAG
AAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGA
CAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCA
CATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTA
CAAGtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAG
ACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAG
AATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTG
AAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAG
ATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCA
CTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCC
GTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAA
CGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGT
AACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGA
CCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGGC
CGCGACTCTAGA

pcDNA3.1+/truncated uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtcGCAGAAATCGGTACTGGCTTTCCAGGCATGGACGAGCTG
TACAAGtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGAC
AGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTC

AGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGC
TGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCC
AGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTG
CACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGG
CCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGG
AACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGA
GTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAA
GACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAG
GCCGCGACTCTAGA

pcDNA3.1+/10AUG large uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatgatgatgatgatgatgatgatgatgGCAGAAATCGGTACTGGCTTTCCATTCGACCCCCAT
TATGTGGAAGTCCTGGGCGAGCGCATGCACTACGTCGATGTTGGTCCGCGCGATGGCACC
CCTGTGCTGTTCCTGCACGGTAACCCGACCTCCTCCTACGTGTGGCGCAACATCATCCCG
CATGTTGCACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAAATCCGAC
AAACCAGACCTGGGTTATTTCTTCGACGACCACGTCCGCTTCATGGATGCCTTCATCGAAG
CCCTGGGTCTGGAAGAGGTCGTCCTGGTCATTCACGACTGGGGCTCCGCTCTGGGTTTCC
ACTGGGCCAAGCGCAATCCAGAGCGCGTCAAAGGTATTGCATTTATGGAGTTCATCCGCC
CTATCCCGACCTGGGACGAATGGCCAGAATTTGCCCGCGAGACCTTCCAGGCCTTCCGCA
CCACCGACGTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTATCGAGGGTACGCTGC
CGATGGGTGTCGTCCGCCCGCTGACTGAAGTCGAGATGGACCATTACCGCGAGCCGTTCC
TGAATCCTGTTGACCGCGAGCCACTGTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTG
AGCCAGCGAACATCGTCGCGCTGGTCGAAGAATACATGGACTGGCTGCACCAGTCCCCTG
TCCCGAAGCTGCTGTTCTGGGGCACCCCAGGCGTTCTGATCCCACCGGCCGAAGCCGCT
CGCCTGGCCAAAAGCCTGCCTAACTGCAAGGCTGTGGACATCGGCCCGGGTCTGAATCTG
CTGCAAGAAGACAACCCGGACCTGATCGGCAGCGAGATCGCGCGCTGGCTGTCGACGCT
CGAGATTTCCGGCCACCGGTCGCAACCTTGGGTACCGCGGGCCCGGGATCCACCGGTCG
CAACCTTGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAG
CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTG
GCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCA
CATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGA
CACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCT
GGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCA
GAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACA
GACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCA
GAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCT
GAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCA
GATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGC
ACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGC
CGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGA
ACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAG
TAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAG

ACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGG
CCGCGACTCTAGA

pcDNA3.1+/10AUG truncated uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatgatgatgatgatgatgatgatgatgGCAGAAATCGGTACTGGCTTTCCAGGCATGGACGAG
CTGTACAAGtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGC
GACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGT
TTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATG
GGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGG
GCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC
CTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGA
CGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTG
TGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTC
CGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTA
CAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAG
TAAGGCCGCGACTCTAGA

pcDNA3.1+/large uORF (FFLuc) nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtcGAAGATGCCAAAAACATTAAGAAGGGCCCAGCGCCATTCT
ACCCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCC
CTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACATTACCTACGCC
GAGTACTTCGAGATGAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACA
AACCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCCCGTGTTGGGT
GCCCTGTTCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGCGCGAGCTG
CTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCGTGAGCAAGAAAGGGCTGCAA
AAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAAGATCATCATCATGGATAGCA
AGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATTTGCCACCCGG
CTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGACCGGGACAAAACCATCGCCCTGAT
CATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGCGTAGCCCTACCGCACCGCACCG
CTTGTGTCCGATTCAGTCATGCCCGCGACCCCATCTTCGGCAACCAGATCATCCCCGACA
CCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCTTCGGCATGTTCACCACGCTGGGCT
ACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTACCGCTTCGAGGAGGAGCTATTCTTGC
GCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTGGTGCCCACACTATTTAGCTTCTTC
GCTAAGAGCACTCTCATCGACAAGTACGACCTAAGCAACTTGCACGAGATCGCCAGCGGC
GGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTGGCCAAACGCTTCCACCTACCAGG
CATCCGCCAGGGCTACGGCCTGACAGAAACAACCAGCGCCATTCTGATCACCCCCGAAGG
GGACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTGG
ACTTGGACACCGGTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGGC
CCCATGATCATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCTCTCATCGACAAG
GACGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCAT
CGTGGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAACT
GGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCCGGGGTCGCCGGCCTGCCCG
ACGACGATGCCGGCGAGCTGCCCGCCGCAGTCGTCGTGCTGGAACACGGTAAAACCATG
ACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAAGAAGCTGCG
CGGTGGTGTTGTGTTCGTGGACGAGGTGCCTAAAGGACTGACCGGCAAGTTGGACGCCC
GCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGTGGGCAAAC
CGATTCCGAACCCGCTGCTGGGCCTGGATAGCACCtaaacaacaacaacaacaaATGGTCTTCAC

ACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCT
TGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAA
AGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATG
AAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGT
GGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACG
CCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAA
AAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATC
AACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCT
GTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATC
GATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/truncated uORF (FFluc) nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtcGAAGATGCCAAAAACATTAAGAAGCTGCTGGGCCTGGATA
GCACCtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACA
GACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCA
GAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCT
GAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCA
GATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGC
ACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGC
CGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGA
ACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAG
TAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAG
ACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGG
CCGCGACTCTAGA

pcDNA3.1+/10AUG large uORF (FFLuc) nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatgatgatgatgatgatgatgatgatgGAAGATGCCAAAAACATTAAGAAGGGCCCAGCGCCA
TTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTA
CGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACATTACCTA
CGCCGAGTACTTCGAGATGAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAA
TACAAACCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCCCGTGTTG
GGTGCCCTGTTCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGCGCGAG
CTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCGTGAGCAAGAAAGGGCTG
CAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAAGATCATCATCATGGATA
GCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATTTGCCACC
CGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGACCGGGACAAAACCATCGCCCT
GATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGCGTAGCCCTACCGCACCGCAC
CGCTTGTGTCCGATTCAGTCATGCCCGCGACCCCATCTTCGGCAACCAGATCATCCCCGA
CACCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCTTCGGCATGTTCACCACGCTGGG
CTACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTACCGCTTCGAGGAGGAGCTATTCTTG
CGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTGGTGCCCACACTATTTAGCTTCTT
CGCTAAGAGCACTCTCATCGACAAGTACGACCTAAGCAACTTGCACGAGATCGCCAGCGG
CGGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTGGCCAAACGCTTCCACCTACCAG
GCATCCGCCAGGGCTACGGCCTGACAGAAACAACCAGCGCCATTCTGATCACCCCCGAAG
GGGACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTG
GACTTGGACACCGGTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGG
CCCCATGATCATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCTCTCATCGACAA

GGACGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCA
TCGTGGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAAC
TGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCCGGGGTCGCCGGCCTGCCC
GACGACGATGCCGGCGAGCTGCCCGCCGCAGTCGTCGTGCTGGAACACGGTAAAACCAT
GACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAAGAAGCTGC
GCGGTGGTGTTGTGTTCGTGGACGAGGTGCCTAAAGGACTGACCGGCAAGTTGGACGCC
CGCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGTGGGCAAA
CCGATTCCGAACCCGCTGCTGGGCCTGGATAGCACCtaaacaacaacaacaacaaATGGTCTTCA
CACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCC
TTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCA
AAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTAT
GAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCT
GTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTA
CGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCA
AAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGA
TCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGG
CTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACA
TCGATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/10AUG truncated uORF (FFLuc) nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatgatgatgatgatgatgatgatgatgGAAGATGCCAAAAACATTAAGAAGCTGCTGGGCCTG
GATAGCACCtaaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGC
GACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGT
TTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATG
GGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGG
GCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC
CTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGA
CGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTG
TGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTC
CGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTA
CAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAG
TAAGGCCGCGACTCTAGA

pcDNA3.1+/uORF-P2A-nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtcGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCT
GGAGACGTGGAGGAGAACCCTGGACCTATGGTCTTCACACTCGAAGATTTCGTTGGGGAC
TGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGT
TTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAA
ATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAA
TGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGT
GATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTC
GGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACC
CTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTG
TTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGA
CTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGAC
AAGTAAGGCCGCGACTCTAGA

pcDNA3.1+/large uORF-P2A-nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGcaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaac
aacaccatggtcaccatggtcaccatggtcGCAGAAATCGGTACTGGCTTTCCATTCGACCCCCATTATG
TGGAAGTCCTGGGCGAGCGCATGCACTACGTCGATGTTGGTCCGCGCGATGGCACCCCT
GTGCTGTTCCTGCACGGTAACCCGACCTCCTCCTACGTGTGGCGCAACATCATCCCGCAT
GTTGCACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAAATCCGACAAA
CCAGACCTGGGTTATTTCTTCGACGACCACGTCCGCTTCATGGATGCCTTCATCGAAGCCC
TGGGTCTGGAAGAGGTCGTCCTGGTCATTCACGACTGGGGCTCCGCTCTGGGTTTCCACT
GGGCCAAGCGCAATCCAGAGCGCGTCAAAGGTATTGCATTTATGGAGTTCATCCGCCCTA
TCCCGACCTGGGACGAATGGCCAGAATTTGCCCGCGAGACCTTCCAGGCCTTCCGCACCA
CCGACGTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTATCGAGGGTACGCTGCCGA
TGGGTGTCGTCCGCCCGCTGACTGAAGTCGAGATGGACCATTACCGCGAGCCGTTCCTGA
ATCCTGTTGACCGCGAGCCACTGTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTGAGC
CAGCGAACATCGTCGCGCTGGTCGAAGAATACATGGACTGGCTGCACCAGTCCCCTGTCC
CGAAGCTGCTGTTCTGGGGCACCCCAGGCGTTCTGATCCCACCGGCCGAAGCCGCTCGC
CTGGCCAAAAGCCTGCCTAACTGCAAGGCTGTGGACATCGGCCCGGGTCTGAATCTGCTG
CAAGAAGACAACCCGGACCTGATCGGCAGCGAGATCGCGCGCTGGCTGTCGACGCTCGA
GATTTCCGGCCACCGGTCGCAACCTTGGGTACCGCGGGCCCGGGATCCACCGGTCGCAA
CCTTGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCAC
CTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCC
CACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAG
AAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGA
CAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCA
CATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTA
CAAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGA
ACCCTGGACCTATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCG
GCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGG
GGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGA
CATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAA
AATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCA
CACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAG
GCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACA
AAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAA
CGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACG
GTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGGCCGCGACTC
TAGA

pcDNA3.1+/PV IRES no uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGCTCGAG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCGG**
aaaaaAAGCTTGGGCTGCAGGTCTTAAAACAGCTCTGGGGTTGTACCCACCCCAGAGGCCC
ACGTGGCGGCTAGTACTCCGGTATTGCGGTACCTTTGTACGCCTGTTTTATACTCCCTTCC
CCCGTAACTTAGAAGCACAATGTCCAAGTTCAATAGGAGGGGGTACAAACCAGTACCACCA
CGAACAAGCACTTCTGTTCCCCCGGTGAGGCTGTATAGGCTGTTTCCACGGCTAAAAGCG

GCTGATCCGTTATCCGCTCATGTACTTCGAGAAGCCTAGTATCACCTTGGAATCTTCGATG
CGTTGCGCTCAACACTCAACCCCAGAGTGTAGCTTAGGTCGATGAGTCTGGACGTTCCTCA
CCGGCGACGGTGGTCCAGGCTGCGTTGGCGGCCTACCTGTGGCCCAAAGCCACAGGACG
CTAGTTGTGAACAAGGTGTGAAGAGCCTATTGAGCTACCTGAGAGTCCTCCGGCCCCTGA
ATGCGGCTAATCCTAACCACGGAGCAGGCAGTGGCAATCCAGCGACCAGCCTGTCGTAAC
GCGCAAGTTCGTGGCGGAACCGACTACTTTGGGTGTCCGTGTTTCCTTTTATTTTTACAAT
GGCTGCTTATGGTGACAATCATTGATTGTTATCATAAAGCAAATTGGATTGGCCATCCGGT
GAGAATTTGATTATTAAATTACTCTCTTGTTGGGATTGCTCCTTTGAAATCTTGTGCACTCAC
ACCTATTGGAATTACCTCATTGTTAAGATAcaacaacaacaacaacaacaacaacaacaacaacaacaacaaca
acaacaacaacaacaacaacaacaacaacaacaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGC
GACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGT
TTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATG
GGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGG
GCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC
CTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGA
CGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTG
TGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTC
CGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTA
CAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAG
TAAGGCCGCGACTCTAGA

pcDNA3.1+/PV IRES uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGCTCGAG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCGG**
aaaaaAAGCTTGGGCTGCAGGTCTTAAAACAGCTCTGGGGTTGTACCCACCCCAGAGGCCC
ACGTGGCGGCTAGTACTCCGGTATTGCGGTACCTTTGTACGCCTGTTTTATACTCCCTTCC
CCCGTAACTTAGAAGCACAATGTCCAAGTTCAATAGGAGGGGGTACAAACCAGTACCACCA
CGAACAAGCACTTCTGTTCCCCCGGTGAGGCTGTATAGGCTGTTTCCACGGCTAAAAGCG
GCTGATCCGTTATCCGCTCATGTACTTCGAGAAGCCTAGTATCACCTTGGAATCTTCGATG
CGTTGCGCTCAACACTCAACCCCAGAGTGTAGCTTAGGTCGATGAGTCTGGACGTTCCTCA
CCGGCGACGGTGGTCCAGGCTGCGTTGGCGGCCTACCTGTGGCCCAAAGCCACAGGACG
CTAGTTGTGAACAAGGTGTGAAGAGCCTATTGAGCTACCTGAGAGTCCTCCGGCCCCTGA
ATGCGGCTAATCCTAACCACGGAGCAGGCAGTGGCAATCCAGCGACCAGCCTGTCGTAAC
GCGCAAGTTCGTGGCGGAACCGACTACTTTGGGTGTCCGTGTTTCCTTTTATTTTTACAAT
GGCTGCTTATGGTGACAATCATTGATTGTTATCATAAAGCAAATTGGATTGGCCATCCGGT
GAGAATTTGATTATTAAATTACTCTCTTGTTGGGATTGCTCCTTTGAAATCTTGTGCACTCAC
ACCTATTGGAATTACCTCATTGTTAAGATAcaacaacaacaacaacaacaacaacaacaacaacaacaacaaca
acaacaacaacaacaacaacaacaacaacaccatggtcaccatggtcaccatggtctaaacaacaacaacaacaaATG
GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGAC
CAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTC
CGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCAT
CCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGT
GTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGAC
GGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTC
GACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAG
CGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGC
TGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGAT
CATGACATCGATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA

pcDNA3.1-D/EMCV IRES no uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGAGACCCAAGCTGG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGC**
GCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCGCCCTCGAGCGGGATCAATTCCG
CCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTA
TATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCT
GTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGT
TGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGC
GACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGC
CACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGA
TAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGC
CCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGT
GTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTG
AAAAACACGatgataatATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAG
CCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCT
CGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGA
AAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATG
GCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGGCGGCCGTATG
AAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCA
ACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCAT
CAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATG
ACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAAAGGGTCA
AGACAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGA

pcDNA3.1-D/EMCV IRES uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGAGACCCAAGCTGG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGC**
GCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCGCCCTCGAGCGGGATCAATTCCG
CCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTA
TATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCT
GTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGT
TGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGC
GACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGC
CACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGA
TAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGC
CCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGT
GTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTG
AAAAACACGatgataatatggtcaccatggtcaccatggtctaaacaacaacaacaacaaATGGTCTTCACACTCG
AAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAAC
AGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGAT
TGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGT
CTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGAT
GATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGA
ACATGATCGACTATTTCGGGCGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGA
TCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACC
CCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGC
GAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATT
ACAAGGATGACGATGACAAGTAAAAGGGTCAAGACAATTCTGCAGATATCCAGCACAGTGG
CGGCCGCTCGAGTCTAGA

pcDNA3.1-D/HCV no uORF nLuc (SacI/XbaI)

GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGAGACCCAAGCTGG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGC**
GCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCCCTGTGAGGAACTACTGTCTTCA
CGCAGAAAGCGCCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAAGA
CTGGGTCCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTGCCCCCGCAAG
ACTGCTAGCCGAGTAGCGTTGGGTTGCGAAAGGCCTTGTGGTACTGCCTGATAGGGCGCT
TGCGAGTGCCCCGGGAGGTCTCGTAGACCGTGCatcatgagcacgaatcctaaacctcaaagaaaaATG
GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGAC
CAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTC
CGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCAT
CCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGT
GTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGAC
GGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTC
GACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAG
CGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGC
TGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGAT
CATGACATCGATTACAAGGATGACGATGACAAGTAAAAGGGTCAAGACAATTCTGCAGATA
TCCAGCACAGTGGCGGCCGCTCGAGTCTAGA

pcDNA3.1-D/HCV IRES uORF nLuc (SacI/XbaI)
GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGAGACCCAAGCTGG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGC**
GCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCCCTGTGAGGAACTACTGTCTTCA
CGCAGAAAGCGCCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAAGA
CTGGGTCCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTGCCCCCGCAAG
ACTGCTAGCCGAGTAGCGTTGGGTTGCGAAAGGCCTTGTGGTACTGCCTGATAGGGCGCT
TGCGAGTGCCCCGGGAGGTCTCGTAGACCGTGCatcatgagcacgaatcctaaacctcaaagaaaataaa
caacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCG
GCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGG
GGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGA
CATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAA
AATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCA
CACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAG
GCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACA
AAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAA
CGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACG
GTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAAAGGGTCAAGA
CAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGA

pcDNA3.1-D/CrPV IGR IRES no uORF nLuc (SacI/XbaI)
GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGAGACCCAAGCTGG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGC**
GCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCAGTACCCTTCACCAAAGCAAAAAT
GTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGCTATTTTTGTA
TTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACAATATCCAG
GAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGAAATTTACCTGCTAC
ATTTCAAGATACCATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCC
GGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCG
GGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCG

ACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAA
AAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGC
ACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAA
GGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAAC
AAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCA
ACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGAC
GGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAAAGGGTCAAG
ACAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGA

pcDNA3.1-D/CrPV IGR IRES uORF nLuc (SacI/XbaI)
GAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCA
CTATAGGGAGACCCAAGCTGG**TTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGC**
GCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCAGTACCCTTCACCAAAGCAAAAT
GTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGCTATTTTTGTA
TTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACAATATCCAG
GAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGAAATTTACCTGCTAC
ATTTCAAGATACCatggtcaccatggtcaccatggtctaaacaacaacaacaacaaATGGTCTTCACACTCGAA
GATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAG
GGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTG
TCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCT
GAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGA
TCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
ATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATC
ACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCC
GACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGA
ACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTAC
AAGGATGACGATGACAAGTAAAAGGGTCAAGACAATTCTGCAGATATCCAGCACAGTGGC
GGCCGCTCGAGTCTAGA

pTet-Off AIO/no uORF nLuc (MluI/XbaI)
ACGCGTTgctcgagcactttggccgcgaatcgatatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccct
atcagtgatagagaacgatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccctatcagtgatagagaacgtat
gtcgagtttactccctatcagtgatagagaacgtatgtcgagtttatccctatcagtgatagagaacgtatgtcgagtttactccctatcagt
gatagagaacgtatgtcgaggtaggcgtgtacggtgggaggcctatataagcaGAGCTCTCTGGCTAACTAGAGAA
CCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGcaacaacaacaacaacaacaac
aacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTT
CGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAG
GTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCT
GAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAG
CGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCAT
CACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGA
TCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGTAAGTATCAAGGTTACAAGACA
GGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGA
TAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAGGAGCGCCTGATC
AACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCT
GTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATC
GATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA

pTet-Off AIO/no uORF nLuc + 3' UTR intron (MluI/PflMI)

ACGCGTTgctcgagcactttggccgcgaatcgatatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccct
atcagtgatagagaacgatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccctatcagtgatagagaacgtat
gtcgagtttactccctatcagtgatagagaacgtatgtcgagtttatccctatcagtgatagagaacgtatgtcgagtttactccctatcagt
gatagagaacgtatgtcgaggtaggcgtgtacggtgggaggcctatataagcaGAGCTCTCTGGCTAACTAGAGAA
CCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGcaacaacaacaacaacaacaac
aacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaaATGGTCTTCACACTCGAAGATTT
CGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAG
GTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCT
GAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAG
CGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCAT
CACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGA
TCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGTAAGTATCAAGGTTACAAGACA
GGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGA
TAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAGGAGCGCCTGATC
AACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCT
GTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATC
GATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGAGGGCCCGTTTAAACCCGC
TGATCAGCCTCGAGCAGAAATCGGTACTGGCTTTCCATTCGACGTAAGTATCAAGGTTACA
AGACAGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTT
TCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAGCCCCATTA
TGTGG

pTet-Off AIO/uORF nLuc (MluI/XbaI)

ACGCGTTgctcgagcactttggccgcgaatcgatatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccct
atcagtgatagagaacgatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccctatcagtgatagagaacgtat
gtcgagtttactccctatcagtgatagagaacgtatgtcgagtttatccctatcagtgatagagaacgtatgtcgagtttactccctatcagt
gatagagaacgtatgtcgaggtaggcgtgtacggtgggaggcctatataagcaGAGCTCTCTGGCTAACTAGAGAA
CCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGcaacaacaacaacaacaacaac
aacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaccatggtcaccatggtcaccatggtctaaa
caacaacaacaacaaATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCG
GCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGG
GGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGA
CATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAA
AATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCA
CACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAG
GCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACA
AAATTATCGACGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGG
CTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATC
CACTTTGCCTTTCTCTCCACAGGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCG
AGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAA
AGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAA
GGCCGCGACTCTAGA

pTet-Off AIO/large uORF nLuc (MluI/XbaI)

ACGCGTTgctcgagcactttggccgcgaatcgatatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccct
atcagtgatagagaacgatgtcgagtttactccctatcagtgatagagaacgtatgtcgagtttactccctatcagtgatagagaacgtat
gtcgagtttactccctatcagtgatagagaacgtatgtcgagtttatccctatcagtgatagagaacgtatgtcgagtttactccctatcagt
gatagagaacgtatgtcgaggtaggcgtgtacggtgggaggcctatataagcaGAGCTCTCTGGCTAACTAGAGAA
CCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGaacaacaacaacaacaacaac
aacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaacaccatggtcaccatggtcaccatggtcGCA

GAAATCGGTACTGGCTTTCCATTCGACCCCCATTATGTGGAAGTCCTGGGCGAGCGCATG
CACTACGTCGATGTTGGTCCGCGCGATGGCACCCCTGTGCTGTTCCTGCACGGTAACCCG
ACCTCCTCCTACGTGTGGCGCAACATCATCCCGCATGTTGCACCGACCCATCGCTGCATT
GCTCCAGACCTGATCGGTATGGGCAAATCCGACAAACCAGACCTGGGTTATTTCTTCGACG
ACCACGTCCGCTTCATGGATGCCTTCATCGAAGCCCTGGGTCTGGAAGAGGTCGTCCTGG
TCATTCACGACTGGGGCTCCGCTCTGGGTTTCCACTGGGCCAAGCGCAATCCAGAGCGCG
TCAAAGGTATTGCATTTATGGAGTTCATCCGCCCTATCCCGACCTGGGACGAATGGCCAGA
ATTTGCCCGCGAGACCTTCCAGGCCTTCCGCACCACCGACGTCGGCCGCAAGCTGATCAT
CGATCAGAACGTTTTTATCGAGGGTACGCTGCCGATGGGTGTCGTCCGCCCGCTGACTGA
AGTCGAGATGGACCATTACCGCGAGCCGTTCCTGAATCCTGTTGACCGCGAGCCACTGTG
GCGCTTCCCAAACGAGCTGCCAATCGCCGGTGAGCCAGCGAACATCGTCGCGCTGGTCG
AAGAATACATGGACTGGCTGCACCAGTCCCCTGTCCCGAAGCTGCTGTTCTGGGGCACCC
CAGGCGTTCTGATCCCACCGGCCGAAGCCGCTCGCCTGGCCAAAAGCCTGCCTAACTGCA
AGGCTGTGGACATCGGCCCGGGTCTGAATCTGCTGCAAGAAGACAACCCGGACCTGATCG
GCAGCGAGATCGCGCGCTGGCTGTCGACGCTCGAGATTTCCGGCCACCGGTCGCAACCT
TGGGTACCGCGGGCCCGGGATCCACCGGTCGCAACCTTGGTGAGCAAGGGCGAGGAGCT
GTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT
TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTC
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTAC
GGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCC
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA
GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAA
CAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA
GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA
CCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCC
GCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGtaaacaacaacaacaacaaATGGTC
TTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGACCAA
GTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTCGGGGTGTCCGTAACTCCGA
TCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTGAAGATCGACATCCATGTCATCATCCC
GTATGAAGGTCTGAGCGGCGACCAAATGGGCCAGATCGAAAAAATTTTTAAGGTGGTGTAC
CCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGG
TTACGCCGAACATGATCGACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACG
GCAAAAAGATCACTGTAACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGTAAGTAT
CAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAG
ACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCA
CAGGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGGAGTG
ACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACTACAAAGACCATGACGGTGATTAT
AAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGGCCGCGACTCTAGA