



OPEN

Optimisation of Tet-On inducible systems for Sleeping Beauty-based chimeric antigen receptor (CAR) applications

S. M. Ali Hosseini Rad^{1,2}, Aarati Poudel^{1,2}, Grace Min Yi Tan^{1,2} & Alexander D. McLellan¹✉

Regulated expression of genetic elements that either encode polypeptides or various types of functional RNA is a fundamental goal for gene therapy. Inducible expression may be preferred over constitutive promoters to allow clinician-based control of gene expression. Existing Tet-On systems represent one of the tightest rheostats for control of gene expression in mammals. However, basal expression in absence of tetracycline compromises the widespread application of Tet-controlled systems in gene therapy. We demonstrate that the order of P2A-linked genes of interest was critical for maximal response and tightness of a chimeric antigen receptor (CAR)-based construct. The introduction of G72V mutation in the activation region of the TetR component of the rtTA further improved the fold response. Although the G72V mutation resulted in a removal of a cryptic splice site within rtTA, additional removal of this splice site led to only a modest improvement in the fold-response. Selective removal of key promoter elements (namely the BRE, TATA box, DPE and the four predicted Inr) confirmed the suitability of the minimal CMV promoter and its downstream sequences for supporting inducible expression. The results demonstrate marked improvement of the rtTA based Tet-On system in Sleeping Beauty for applications such as CART cell therapy.

Inducible-gene expression is one of the most sought-after elements of synthetic gene regulation systems. Engineering mammalian cells to express proteins or RNA in an inducible fashion offers opportunities for the development of safe cellular-based therapies to treat a wide spectrum of inborn and acquired diseases. Compared to prokaryote genetic systems, the development of tight, inducible gene expression in eukaryote cells has been challenging^{1–3}. Unlike prokaryotes, mammalian genetic control is not usually mediated by single or oligo-component regulators, but rather by multiple transcription factors that bind to both promoters, as well as often distant enhancer regions located on different chromosomes. Moreover, both promoters and enhancers may be regulated by epigenetic control mechanisms, and the site of transgene insertion in the genome influences the response profile of transgenes⁴. Tet-On systems utilise a mutant TetR component that binds to tetracycline response elements (TRE) in the presence of tetracycline, or its stable analogue doxycycline⁵. To activate transcription, fusion of the herpes-simplex VP16 transcriptional activator to the C-terminus of the mutant TetR, recruits generalised transcription factors, as well as RNAP II to initiate gene transcription. Modified tetracycline-inducible systems represent the most widely used inducible system in eukaryotic systems, from yeast to human cells⁵. The potential exists for drug inducible systems to be used in cell-based immunotherapy to control the expression of genes or other sequences of interest (GOI). Although 10³ to 10⁶-fold induction of gene expression with tetracycline-based control systems has been reported, basal expression in the absence of inducer can result in undesirable GOI expression⁵. In vivo use would be compromised by such leakiness, particularly if the GOI was involved in T cell survival, or resistance to apoptosis. Unfortunately, compared to Tet-Off systems, Tet-On systems are less sensitive to tetracycline and generally exhibit a higher level of basal expression in the absence of induction⁶. On the other hand, Tet-Off systems are slow to respond to withdrawal of tetracycline and this may be compounded by sequestration of tetracycline in vivo, especially within bones^{2,7}.

¹Department of Microbiology and Immunology, University of Otago, Dunedin, Otago 9010, New Zealand. ²These authors contributed equally: S. M. Ali Hosseini Rad, Aarati Poudel and Grace Min Yi Tan. ✉email: alex.mclellan@otago.ac.nz

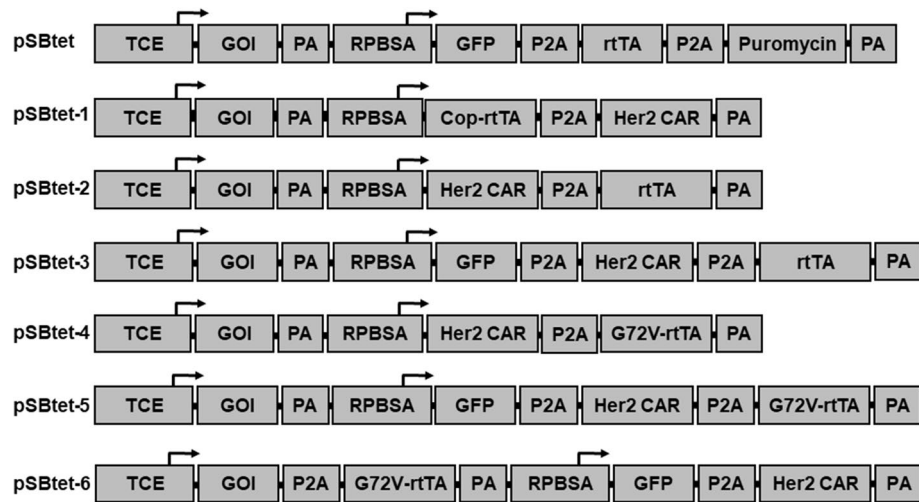


Figure 1. Schematic illustration of SB-based Tet-On systems used in this study. Constructs derived from original pSBtet-GP developed by Kowarz et al. TCE: tet-responsive promoter/ GOI: gene of interest (Mcl-1 or firefly luciferase); PA: polyadenylation site; P2A: 2A self-cleaving; rtTA-M2: reverse tetracycline-controlled transactivator; RPBSA: constitutive promoter comprised of the Rpl13a core promoter and exon 1, plus additional exon and intron elements from Rpl41.

The Sleeping Beauty (SB) transposon system was developed from extinct Salmonid transposons awoken after 10 million years of inactivity through consensus-based correction of accumulated mutations⁸. Compared to retroviral-based insertions into transcriptional units and their regulatory regions, SB vectors insert almost randomly into TA-sites throughout the genome. This property minimises deleterious integrations and helps maintain constitutive or inducible gene expression^{9,10}. SB-based vectors carry a GOI along with optional markers or selection elements flanked by inverted terminal repeats (ITRs)^{9–12}. Although the transposase has been re-engineered to enhance activity^{12,13}, a lower-activity SB-transposase is preferred for human clinical trials to minimise the incidence of multiple genome integrations. Along with piggy bac transposase systems, SB transposase systems have been used in CAR T cell therapy trials for B cell malignancies^{14,15}. To expand the utility of SB-based vectors to express a CAR together with additional GOI under drug-control, we revisited the SB-based Tet-On system, through: (1) alterations in the placement of genes within the P2A-linked CAR cassette, (2) the introduction of a G72V mutation in rtTA-M2 – previously only described for yeast Tet-On control¹⁶ (3) the placement of rtTA under auto-regulatory control, (4) the removal of cryptic splicing sites, and (5) modifications of the proximal promoter. To test the induction of the Tet-On system, we expressed myeloid leukaemia cell differentiation (Mcl-1) as a GOI involved in T cell survival and resistance to apoptosis.

Results

The rtTA location within a multi-gene cassette influences responsiveness of the Tet-On system. We reasoned that placing a codon-optimised rtTA-M2 gene proximal to the RPBSA promoter (pSBtet-1) should result in robust rtTA-M2 expression and therefore tight control of inducible gene expression (Fig. 1).

Surprisingly, this setting led to a decrease in both the fold-expression of luciferase and Mcl-1 mRNA (Fig. 2A,B). It has previously been reported that expression of rtTA-M2 by strong promoters compromises inducible expression^{16,17}. We therefore relocated the original rtTA-M2 sequence distal to the RPBSA and downstream from either one (pSBtet-2) or two (pSBtet-3) additional GOI. However, inducibility of the GOI was still poor (Figs. 1, 2C,D).

Introduction of a G72V mutation in rtTA-M2 enhances the tightness of the Tet-On system.

Roney et al.¹⁶ reported that a GGG to GTG (G72V) missense mutation in rtTA mitigated basal gene expression in the absence of an inducer in *S. cerevisiae* clones. Because of the similarity of transcriptional machinery amongst eukaryote cells, we reasoned that this approach may give similar results in human cells. The G72V mutation was next introduced into pSBtet-2 and pSBtet-3 to create the pSBtet-4 and pSBtet-5 constructs (Fig. 1). The G72V mutation in rtTA-M2 decreased the background expression of TCE promoter in the absence of doxycycline at both the mRNA and protein level ($P < 0.001$, Fig. 3A,B). The G72V mutation also restored the maximal expression of pSBtet-2 and pSBtet-3 constructs following induction with doxycycline (Fig. 3C). As previously reported, G72V-rtTA-M2 appeared less sensitive to doxycycline compared to original rtTA-M2, though this was not statistically significant (Fig. 3C, $P > 0.9$)¹⁶. A similar pattern of results was obtained after two weeks passage of cells to ensure stable integration of the pSBtet-5 (Fig. 3D,E). Note, the GFP expression of the transfected cell lines dropped from ~90 to ~70% after two weeks of culture, most likely due to a shift from transient gene expression to that from integrated cassettes.

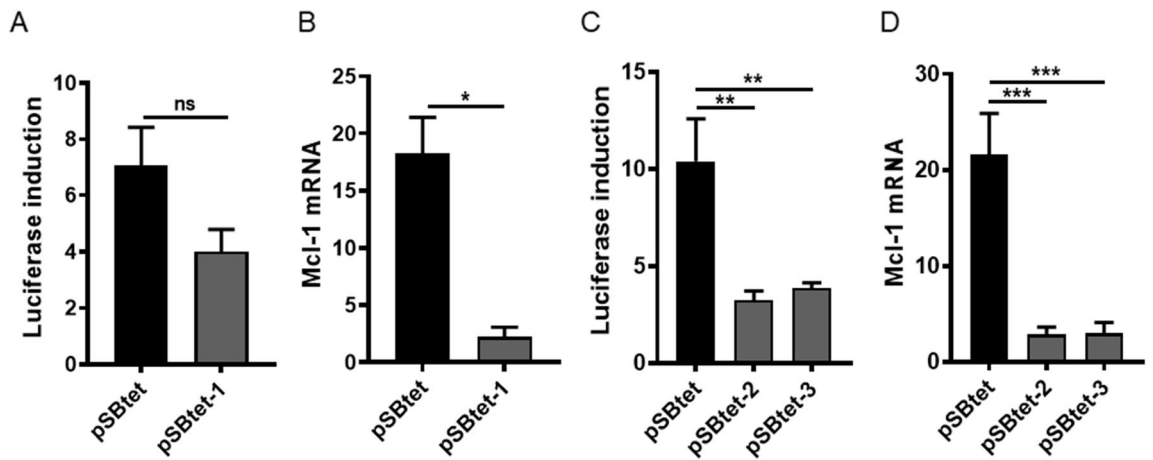


Figure 2. Locating a commercially-codon optimised rtTA-M2 (cop-rtTA; pSBtet-1) proximal to RPBSA increases the leakiness of TCE promoter as measured by (A) luciferase fold induction and (B) qPCR for Mcl-1 mRNA. Relocating the unmodified rtTA-M2 distal to the RPBSA in (pSBtet-2 & -3) did not improve the basal expression measured by (C) luciferase fold induction and (D) qPCR for Mcl-1 mRNA. Experiments were carried out 96 h post-transfection. Statistical analysis: (A, B) two-tailed t-test, (C, D) one-way ANOVA test with Bonferroni post-test correction (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

Investigation of autoregulatory strategy with G72V variant. We next investigated possible improvements in inducibility of multiple GOIs using positive feedback control in an autoregulatory cassette. Autoregulation can improve tetracycline-regulation in a retroviral vector¹⁸ and in a bi-directional¹⁹ or uni-directional lentiviral vector⁷. The bi-directional approach appears tight in transient transfection, however, high background was detected when cells were stably transduced¹⁹. Therefore, we utilised a uni-directional strategy with a P2A sequence in place of an Internal Ribosome Entry Site (IRES) sequence, to allow expression of GOI-P2A-G72VrtTA under TCE promoter (pSBtet-6, Fig. 1). We speculated that the leaky expression of the TCE would still allow sufficient levels of G72V-rtTA inducer to respond to doxycycline stimulation. Although the positive feedback system resulted in tight expression at the protein level (Fig. 4A), as previously reported^{7,18,19}, the system was leaky at the mRNA level for Mcl-1 (Fig. 4B). Basal expression of luciferase in pSBtet-6 was lower than for pSBtet ($P < 0.05$) and showed a higher response upon induction ($P < 0.05$, Fig. 4C). Compared to the constitutive expression of G72V-rtTA, the autoregulatory system showed greater sensitivity to doxycycline induction ($P < 0.05$; Fig. 4C), although the basal expression was higher ($P = 0.0066$).

Removing cryptic alternative splice sites within rtTA reduces the background expression. Since the first description of the eukaryotic Tet-On inducible system, most optimisation studies have focused on rtTA mutations: for example, the removal of cryptic splice sites (flanking amino acids 8–144) in the TetR sequence^{20,21}. Using recently-developed software²², we identified eight additional potential cryptic splice sites within the coding region of rtTA-M2 (Fig. 5A and Table S1). Because the G72V mutation resulted in the loss of a cryptic splice site at position 215 (Fig. 5A and Table S2), we determined if the success of the G72V mutation was due to the removal of the potential cryptic splice site at 215. These cryptic splice sites are located in two regions of rtTA; one in a surface residue (215 nt and seven in the dimerisation domains (320 nt, 326 nt, 367 nt, 392 nt, 408 nt, 456 nt and 541 nt; Fig. 5A, Figure S1). We therefore removed all eight cryptic splice sites by silent or conserved missense-mutations in the pSBtet construct (Table 1).

The removal of six cryptic splice sites modestly enhanced the tightness of Tet-On system 7.7–19.6 fold compared with original rtTA ($P < 0.001$, Fig. 5B). The remaining two mutations at position 320 (~twofold, $P = 0.08$) and 367 (~fivefold, $P = 0.08$) did not significantly affect Tet-On performance. The mutation at position 320 produced E106Q, while 367 (Q122) was a silent mutation (Table 1). It is possible these two splice sites are weak 5' acceptor splice sites which are only used if other competing splice sites are removed²³. Indeed, positions 320 and 367 have low score and confidence which represent strength and the probable occurrence of a splice site, respectively (Table S1). Combining all mutations together, improved the leaky background of the Tet-On system ~40 fold compared to original rtTA ($P < 0.0001$, Fig. 5B). However, superior results were still seen with G72V mutation (Fig. 5C).

Surprisingly, combining the G72V mutation and removing all cryptic splice sites abolished responsiveness and inducibility of Tet-On system (Fig. 5C,D). There are four altered amino acid positions within rtTA-M2 that result in a reverse activator phenotype, as compared to the original TetR: E71K, D95N, L101S and G102D². In TetR, E71 is a surface residue amino acid, D95 connects the DNA reading head to the core domain, while L101 and G102 are crucial for dimerisation and the tetracycline response, respectively². In TetR the E71 and G72 amino acids create the turn between α -helix-4 and 5 (Fig. 6). This region bridges the DNA binding domain to the tetracycline binding domain and the combination of both the E71K and G72V mutations might destroy the structure of this critical turn, causing a loss of rtTA-M2 activity. This may also explain the drop in tetracycline-induction observed with the position 320 mutant (E107 to Q107, Fig. 5D), since this residue is close to a 'high sensitivity

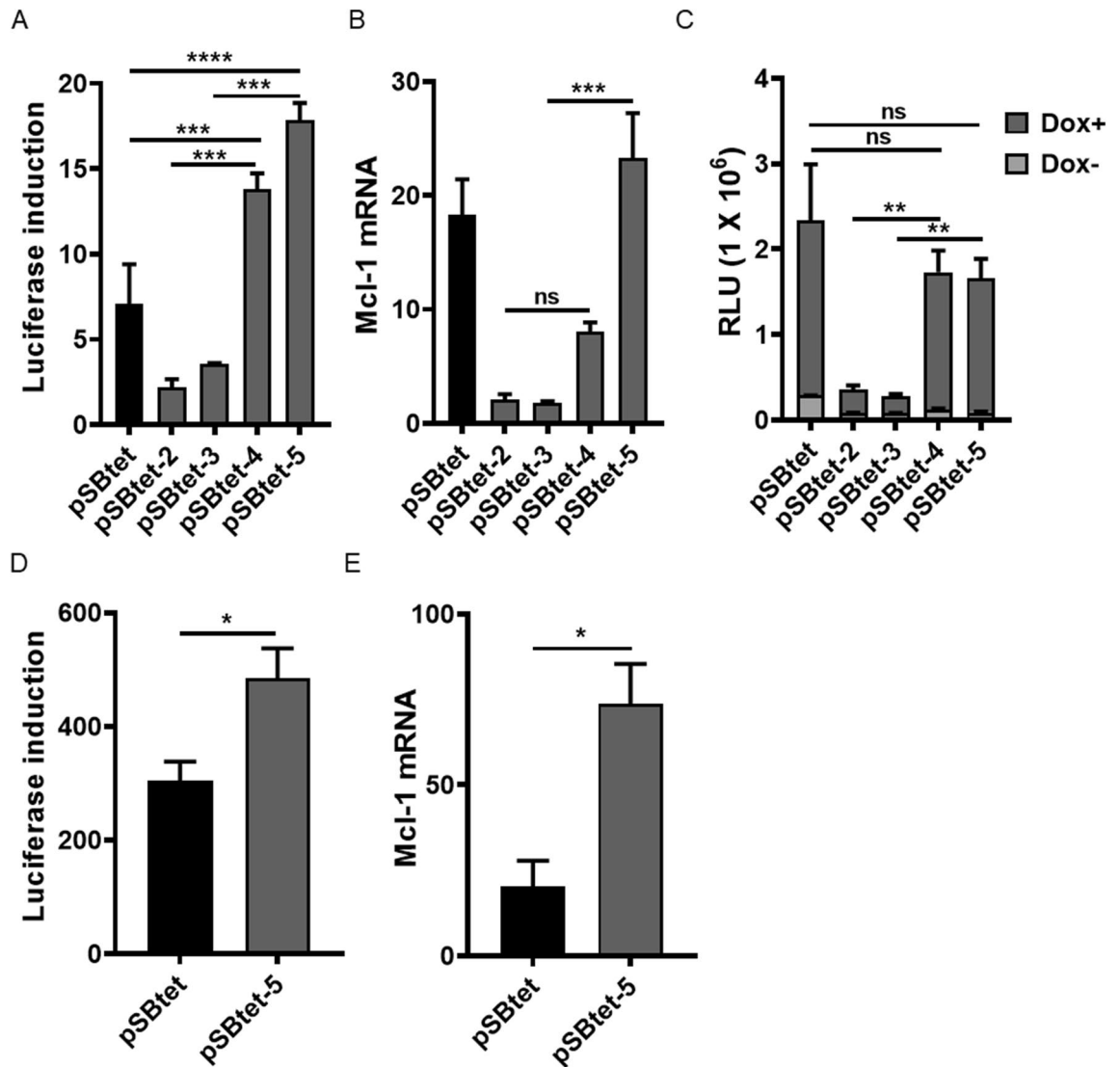


Figure 3. Introducing the G72V mutation into pSBtet-2 & -3 and generation of pSBtet-4 & -5. The efficacy of G72V-rtTA SB based Tet-On system was measured at 96 h by (A) luciferase fold induction and (B) qPCR for Mcl-1 mRNA. (C) The G72V mutation restores the inducibility of pSBtet-2 and pSBtet-3 upon doxycycline induction. (D, E) Confirmation of the improvement of SB-based Tet-On system after two weeks passaging, as quantified by luciferase fold induction and qPCR for Mcl-1 mRNA. Statistical analysis: (A–C) one-way ANOVA test with Bonferroni post-test correction, (D, E) two-tailed t-test (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$).

region' (Figure S2)^{21,24–26}. It is possible that the G72V mutation affected the secondary structure of rtTA¹⁶, rather than simply removing a cryptic splice site. It is interesting to note that further commercial algorithm-mediated codon optimisation of rt-TA-M2 attempted in pSBtet-1 (see Fig. 2) re-introduced 13 cryptic alternative splice sites with high score and confidence (Fig. 5A, Table S3) within the rtTA coding region. This may have contributed to the poor performance of the first pSBtet-1 construct analysed, since cryptic splice sites might be associated with poor performance of Tet-On systems^{20,21} (Fig. 2A,B).

Dissection of the TCE proximal promoter. Modification of the minimal CMV promoter can affect TCE promoter performance²⁷. Removing elements downstream of the TATA box can reduce the maximal expression, whereas deleting the upstream elements can decrease the leakiness²⁷. We therefore revisited the design of pTIGHT to ensure optimal performance in our setting. Core promoter elements were identified using the YAAP program (Fig. 7A). It is possible that the presence of alternative initiator element (Inr) might lead to a loss of control of the TCE-promoter. We therefore removed these elements in single or combinatorial mutation fashion from pSBtet and monitored the tightness and maximal expression of TCE promoter.

Removal of each, or all, downstream elements of the TATA box (Inr-1, -2, -3 and -4 sites, and the DPE element) markedly decreased the tightness of the TCE promoter (Fig. 7B), and also reduced the maximal expression, as previously reported²⁷ (Fig. 7C). Specifically, removing Inr-3 increased the background expression remarkably

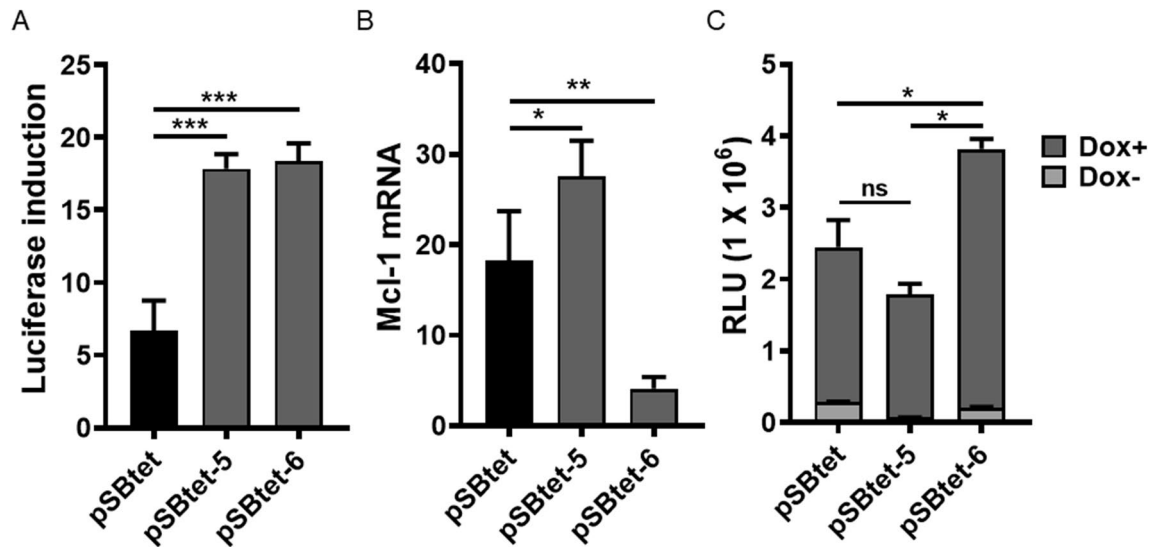


Figure 4. pSBtet-6 autoregulatory system showed superior regulation to pSBtet with (A) luciferase fold induction, but (B) higher background expression at the mRNA level for Mcl-1. (C) Comparison of maximal expression of luciferase in pSBtet-6 versus constitutive expression of G72V (psBtet-5) after induction with doxycycline. pSBtet-6 showed higher sensitivity, but higher background expression in the absence of doxycycline compared to pSBtet-5. Experiments were carried out 96 h post-transfection. Statistical analysis: one-way ANOVA test with Bonferroni post-test correction (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

($P < 0.01$). Inr-3 contains two CTF/NF1 binding sites that bind to DNA as dimers²⁸ (Fig. 7A). CTF/NF1 is an enhancer-blocker element which specifically blocks the interaction of other enhancers with the promoter²⁹. A general explanation would be that deletion of the CTF/NF1 binding site could increase the interaction of neighbouring enhancers to the TCE promoter, resulting in high background.

The only core promoter element found upstream of the TATA box is B recognition element (BREu). As shown in Fig. 7D,E, removing BREu did not improve the tightness and the basal expression of the TCE promoter ($P > 0.19$) and optimal transcription through the TCE promoter was dependent on the TATA box. However, deleting the BREu site increased the response of the TCE promoter to doxycycline (Fig. 7E, $P < 0.01$).

Discussion

Several strategies have been proposed to reduce the leakiness and enhance the inducibility of Tet-On systems, with only some tested in human cells. Such approaches include: (1) increased expression of rtTA using a strong promoter and codon optimisation^{21,30,31}, (2) mutation of rtTA to increase binding to doxycycline or DNA^{16,20,21,26}, (3) autoregulatory systems^{7,18,19,32}, (4) removing a cryptic splice sites in the rtTA coding region²⁰, and (5) alteration of the core promoter elements within the proximal region of the TCE promoter²⁷. We revisited these strategies for use in the SB-based Tet-On system in a human cell line for future investigation in CAR T cell therapy.

The introduction of a single mutation G72V, gave the optimal induction results at both mRNA and protein level, as reported in *S. cerevisiae*¹⁶. Future studies may explore the use of a G72P instead of G72V in our system as a candidate amino acid at position G72, though G72P appeared to result in a small loss in sensitivity, as compared to G72V¹⁶. It is interesting that independent efforts into the rtTA structure have resulted in distinct amino acid changes in different studies, but with similar outcomes. For example, mutations introduced into the rtTA-M2 gene used here are present in distinct positions, as compared to the original four mutations in rtTA^{6,20}. Moreover, introducing sensitivity enhancing (SE) mutations^{24–26} (V9I, F67S, G72P, F86Y, and R171K) could further increase the sensitivity to doxycycline, without increasing the background, as demonstrated in yeast¹⁶.

Autoregulatory systems have recently generated interest, with both the rtTA and GOI transcribed by a single TCE promoter, using either a bi-directional promoter¹⁹ or an IRES sequence^{7,18,32}. However, our constitutive expression of G72V-rtTA gave tighter expression, but was less sensitive to doxycycline compared to the autoregulatory system. The autoregulatory system may be preferred for controllable expression of a toxic rtTA or toxic GOI in mammalian protein production^{32–35}.

Next, our analysis found evidence of cryptic splice sites within an rtTA, a sequence that was previously optimised for mammalian expression by Urlinger et al.³⁰ Removing these splice sites reduced the basal expression and further increased the maximal expression. Unfortunately, using the combination of the G72V mutation with all splice sites removed (using predominantly silent, but with two necessarily non-silent, mutations) created a non-responsive system. It appears likely that the combination of the E71Q and G72V mutations disrupted the turn between two critical α -helices 4 and 5.

It is also noteworthy that different programs identified other possible splice sites that needed to be investigated (Table S4). At least three independent approaches for codon optimisation of rtTA have been reported to enhance Tet-On function^{21,30,31}. For example, Urlinger et al. modified the *S. cerevisiae*-developed rtTA-M2 sequence to remove potential hairpin, splice, and endonuclease sites, as well as codon optimising the sequence for use within mammalian systems²¹.

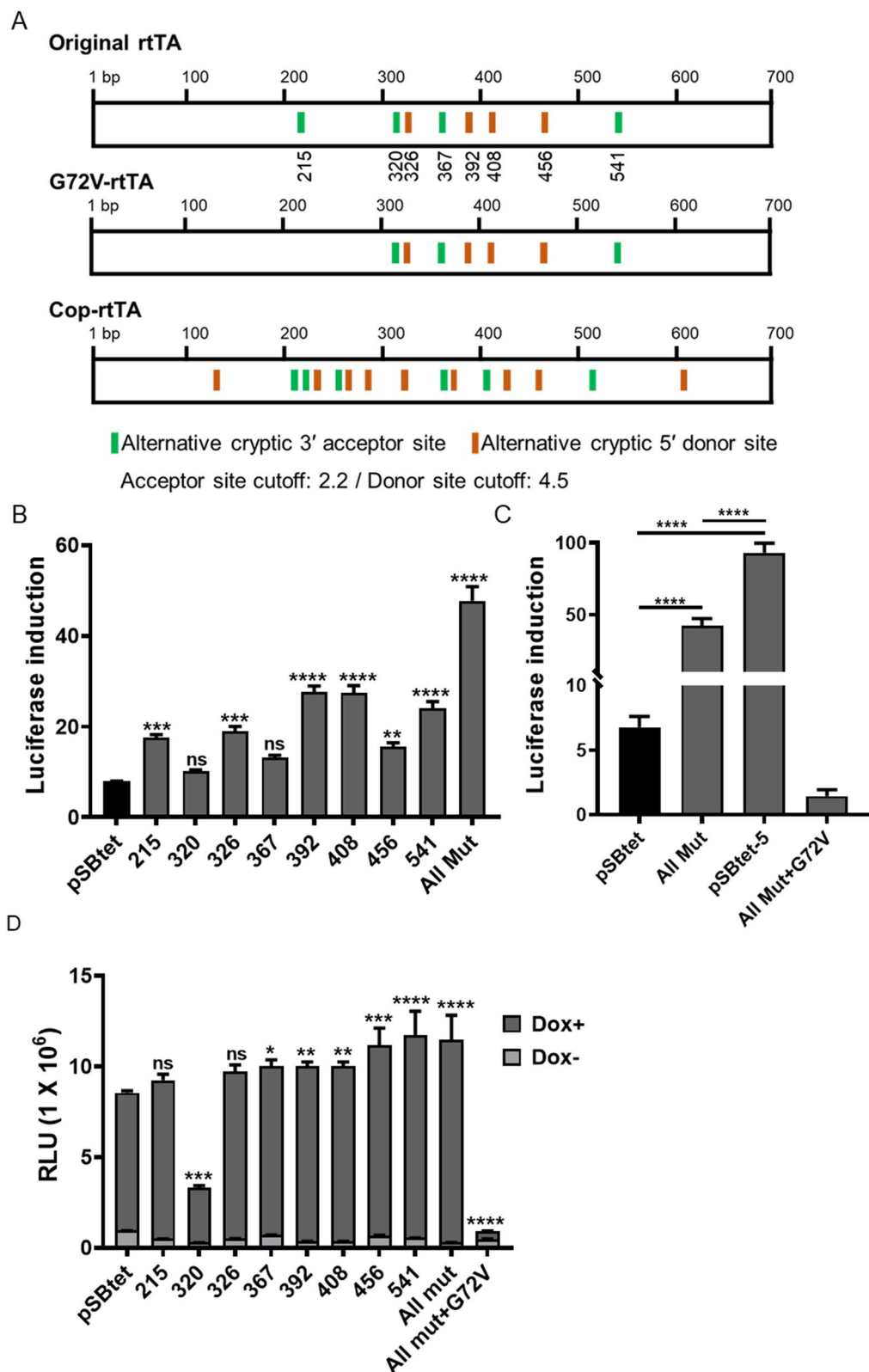


Figure 5. Investigation of potential cryptic splice sites within rtTA. (A) Predicted cryptic alternative splice sites within (top): original rtTA-M2, (middle): G72V-rtTA-M2 and (below): commercially codon optimised (cop)-rtTA-M2 using ASSP program. The default cut-off values of the ASSP program was used. The cut-off 2.2 for acceptor sites and 4.5 for donor sites have shown to correctly predict 75 to 80% of cryptic splice sites (Wang M and et al. 2006). (B) Removing the eight potential cryptic splice sites alone, or combination of all eight, improved the tightness of Tet-On system. (C) Comparison of the fold induction of pSBtet-5 with removing all cryptic splice sites in pSBtet-5. Combining G72V mutation with the eight cryptic splice sites removed, resulted in a non-responsiveness Tet-On system (D) Induction of luciferase expression in mutated rtTA-M2 proteins upon doxycycline induction. The missense E107Q mutation at position 320 bp showed lower induction. Experiments were carried out 96 h post-transfection. Statistical analysis: one-way ANOVA test with Bonferroni post-test correction (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$).

Name	Type	Original sequence	Mutated sequence	Mutation type
215	Acceptor	CCC CTG GAA GGC GAG TCA Pro Leu Glu Gly Glu Ser	CCC CTG GAT <u>GGC</u> GAG TCA Pro Leu Asp Gly Glu Ser	Missense (conservative)
320	Acceptor	CCA ACA GAG AAA CAG TAC Pro Thr Glu Lys Gln Tyr	CCA ACA <u>CAA</u> AAA CAG TAC Pro Thr Gln Lys Gln Tyr	Missense (conservative)
326	Donor	CCA ACA GAG AAA CAG TAC Pro Thr Glu Lys Gln Tyr	CCA ACA GAG AAA <u>CAA</u> TAC Pro Thr Glu Lys Gln Tyr	Silent
367	Acceptor	TGT CAG CAA GGC TTC TCC Cys Gln Gln Gly Phe Ser	TGT <u>CAA</u> CAA GGC TTC TCC Cys Gln Gln Gly Phe Ser	Silent
392	Donor	AAC GCA CTG TAC GCT CTG Asn Ala Leu Tyr Ala Leu	AAC GCA <u>TTA</u> TAC GCT CTG Asn Ala Leu Tyr Ala Leu	Silent
408	Donor	TCC GCC GTG GGC CAC TTT Ser Ala Val Gly His Phe	TCC GCC <u>ATC</u> GGC CAC TTT Ser Ala Ile Gly His Phe	Missense (conservative)
456	Donor	GAG CAT CAA GTA GCA AAA Glu His Gln Val Ala Lys	GAG CAT CAA <u>GTG</u> GCA AAA Glu His Gln Val Ala Lys	Silent
541	Acceptor	GAC CGG CAG GGA GCC GAA Asp Arg Gln Gly Ala Glu	GAC CGG <u>CAA</u> GGA GCC GAA Asp Arg Gln Gly Ala Glu	Silent

Table 1. Putative cryptic acceptor and donor splice sites within rtTA. Dinucleotide splice sites (AG or GT) are highlighted in bold and mutated sequences are underlined.

In the proximal region of the TCE promoter, we confirmed that the TATA box was essential for the function of the TCE promoter²⁷. Removal of the BREu element increased maximal expression, but did not markedly affect the tightness of the TCE. The deletion of the BRE site might enhance the elongation and / or reduce the TFIIB-rtTA sequestration. BRE plays a role in the preinitiation complex (PIC), leading to the dissociation of TFIIB from the promoter which is necessary for RNA polymerase II to initiate the elongation step^{36–38}. Hence, interrupting the BRE-TFIIB interaction may enhance transcription via the enhancement of elongation³⁹. Alternatively, direct sequestration VP16 on TFIIB has been reported^{40,41} that may act to reduce VP16-mediated transcriptional activation.

Collectively, our results demonstrate marked improvements to the rtTA-M2 based Tet-On system in a Sleeping Beauty system through the yeast-optimised G72V mutation. The results especially highlight the necessity to investigate the placement of individual GOI and rtTA within an expression cassette. The use of the clinically relevant CAR cassette within this setting offer the possibility to enhance adoptive cell therapy though drug-inducible expression of cell-survival and memory enhancing genes, or death switches to conditionally ablate CAR T cells following the onset of cytokine release syndrome.

Material and methods

Plasmid construction and cloning. The Tet-On SB (pSBtet-GP) contains the tetracycline-inducible pTIGHT promoter upstream of two asymmetric SfiI sites for cloning genes of interest (GOI), with a downstream RPBSA promoter driving GFP-P2A-rtTA-P2A-puromycin. pSBtet-GP and the SB-transposase vector (pCMV(CAT)T7-SB100) were purchased from Addgene. The pTIGHT promoter is a derivative of (Ptet-14) with shorter spacer (16–17 bp) sequences²⁷ between the TRE and the minimal CMV promoter (see Fig. 7)^{9,27,42}. To generate the modified SB plasmids, a multiple cloning site (MCS) with Bsu36I and BstBI sites was cloned into pSBtet to remove GFP-P2A-rtTA-P2A-puromycin to create pSBtet-MCS. The codon optimised rtTA-M2 gene and FRP5 scFv Her2-CAR⁴³ were synthesised as gene blocks (IDT Singapore) and cloned into pSBtet MCS to create pSBtet-1. Other plasmids were generated by splicing by overlap extension (SOE) PCR to fuse the original rtTA2S-M2 (rtTA)^{9,21}, GFP and Her2CAR in different combinations as illustrated in Fig. 1. Mutations into rtTA was introduced using inverse or SOEing PCR. Codon optimised-mouse Mcl-1 (Cop-Mcl-1) was synthesised as a gene block (IDT) with SfiI overhangs to replace the Firefly luciferase gene in the pSB-tet constructs. Both the Mcl-1 and firefly luciferase genes were used as GOI in this study. To modify the core promoter elements (Fig. 5A), the proximal promoter of TCE was PCR amplified from pSBtet and then subcloned into a pUC19 vector (Addgene) using conventional restriction fragment ligation method with EcoRI and NcoI enzymes. Inverse PCR with primers carrying point mutations were used to change the core promoter elements. Finally, each of the modified fragments were PCR amplified from pUC19 and cloned back to pSBtet using PshAI and NcoI restriction sites. To alter the cryptic splice sites within rtTA (Table 1), rtTA was sub-cloned into PUC19 and mutations introduced using inverse PCR.

Bioinformatics analysis. Analysis of the TCE proximal promoter for core promoter elements, including the initiation repeats (Inr1, 2, 3 and 4), TATA box, B recognition element (BRE) site and downstream promoter element (DPE), was carried out using YAPP Eukaryotic Core promoter predictor. TF binding sites were predicted using AliBaba 2.1⁴⁴ and PROMO^{45,46} programs. The transcriptional start site (TSS) was predicted as reported previously for the minimal CMV promoter⁴⁷. Screening of rtTA for cryptic splice sites was carried out using Alternative Splice Site Predictor (ASSP) software and Human Splicing Finder (HSF)^{22,48}. The protein structure of TetR and the prediction of secondary structure were obtained from Protein Data Bank (PDB).

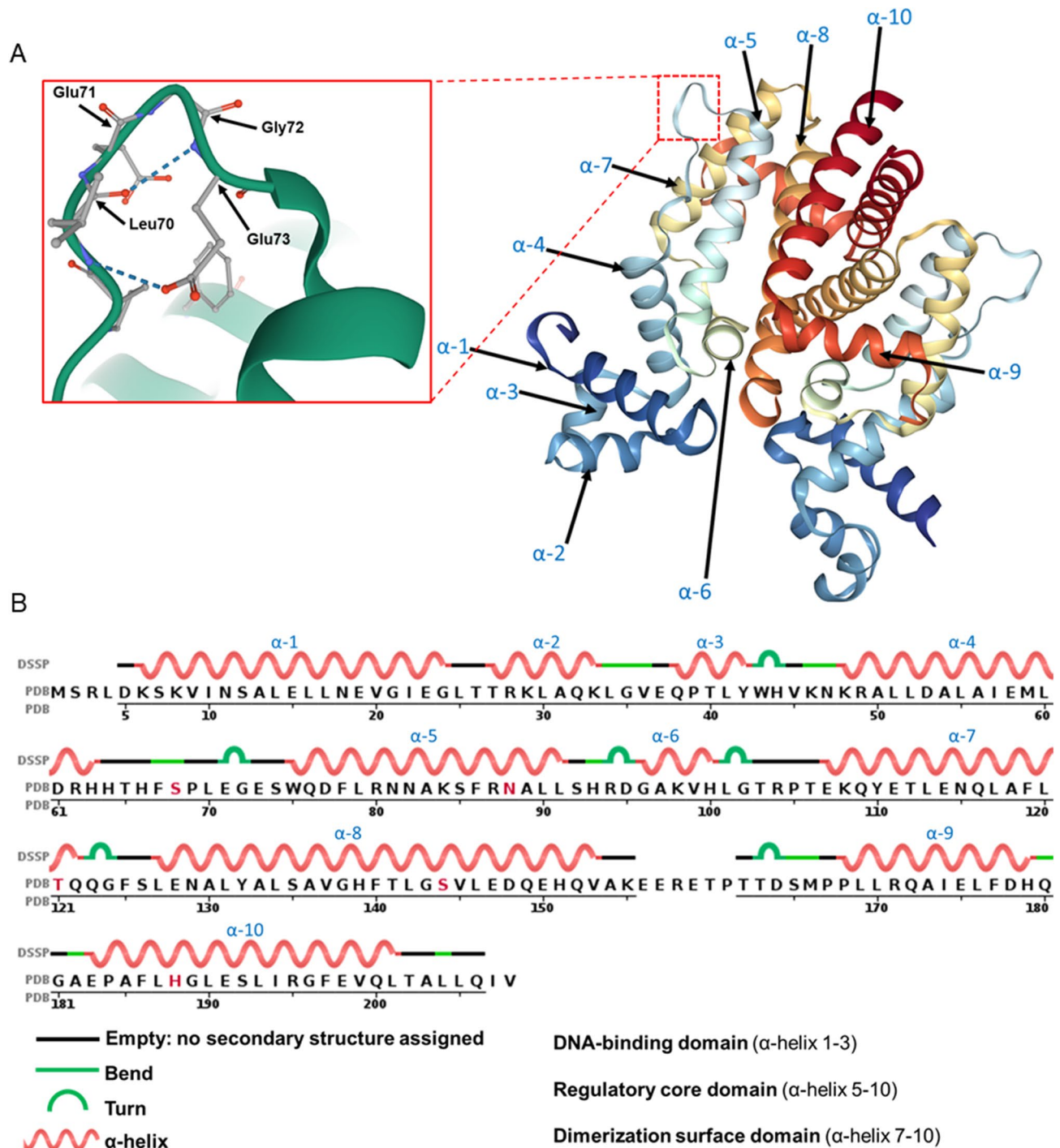


Figure 6. (A) Secondary structure of TetR obtained from protein data bank (PDB) with focus on the E71 and G72 (highlighted by red box) that form a turn between α -helix-4 and 5. Mutating both amino acids may cause a conformational change in rtTA. (B) Annotation of TetR protein sequence and position of the ten α -helices.

Cell culture and transfection. The human embryonic kidney 293 (HEK293; ATCC CRL-1573) cell line was cultured in high glucose Dulbecco's Modified Essential medium (DMEM) supplemented with tetracycline-free 10% foetal bovine serum (FBS; Pan Biotech), Pen-Strep (100 U/mL penicillin and 100 μ g/mL streptomycin) (Gibco) at 37 °C with 5% CO₂. One day prior to transfection, HEK293 cells were cultured in a 24 well plate at 2 \times 10⁵ cells/ mL. A ratio of 5:1 (transfer plasmid: transposase) was used to stably transfect HEK293 cell line using Lipofectamine 3,000 (Thermo Fisher) and the medium was replaced at 24 h post transfection. For induction of the TCE promoter, at 72 h post transfection, cells from each well were detached and divided into four wells in a 96-well plate. Two wells were cultured with DMEM containing 5 μ g / mL of doxycycline (Sigma), while control wells were maintained with only DMEM, for additional 24 h. Doxycycline-induced and control values for each construct are derived from each independent transfection to eliminate the possibility of different

RNA extraction, cDNA synthesis and quantitative PCR (qPCR). Total RNA was extracted using NucleoSpin RNA Plus kit (Macherey–Nagel, Germany) and cDNA prepared using PrimeScript RT Reagent Kit (Takara Bio, USA). QPCR was carried out using Luna Universal qPCR Master Mix (NEB) in a ViiA 6 Real-Time PCR (Applied Biosystems, Foster City, CA). The comparative CT ($2^{-\Delta\Delta ct}$) method was used to analyse the relative expression level of cop-Mcl1, by normalising to β -actin. Primers used for the qPCR reactions were: Mcl1-Fwd: GCA GAA TTG TGA CAC TGA TAA G, Mcl1-Rev: TTT TGT TCT AAC CAA TAC ATC G, β -actin-Fwd: CTT CCT TCC TGG GCA TG, β -actin-Rev: GTC TTT GCG GAT GTC CAC.

Reporter assay. Luciferase assays were carried out using Pierce Firefly Luc one-step glow assay kit (ThermoFisher #16197) with cells at 10^5 cells per 100 μ L in a 96 well plate. Firefly Luc One-Step Glow assay working solution (100 μ L) was added to each well. Cells were incubated at room temperature for one hour before reading with a Varioskan LUX multimode microplate reader (Thermo Fisher, USA). Luciferase data was presented either as relative luminescence units (RLU) or fold change. Fold change was calculated with the following formula:

$$\text{Fold change} = \frac{\text{Luciferase read from cells treated with doxycycline}}{\text{Luciferase read from same cells untreated with doxycycline}}$$

Statistical analysis. All data are presented as mean \pm standard deviation (SD) and pooled from three independent experiments. Statistical analysis was performed by two-tailed t-test or one-way ANOVA test with Bonferroni post-test correction in GraphPad prism (version 8). The P values of ≤ 0.05 were considered statistically significant (* P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001). (* P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001).

Received: 23 April 2020; Accepted: 8 July 2020

Published online: 04 August 2020

References

- Daniel, B., Nagy, G. & Nagy, L. The intriguing complexities of mammalian gene regulation: how to link enhancers to regulated genes. Are we there yet?. *FEBS Lett.* **588**, 2379–2391. <https://doi.org/10.1016/j.febslet.2014.05.041> (2014).
- Gossen, M. *et al.* Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766–1769 (1995).
- Bashor, C. J. *et al.* Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. *Science (New York, N.Y.)* **364**, 593–597 (2019).
- Haruyama, N., Cho, A. & Kulkarni, A. B. Overview: engineering transgenic constructs and mice. *Curr. Protoc. Cell Biol.* **42**, 19.10.11–19.10.19 (2009).
- Das, A. T., Tenenbaum, L. & Berkhout, B. Tet-on systems for doxycycline-inducible gene expression. *Curr. Gene Ther.* **16**, 156–167 (2016).
- Berens, C. & Hillen, W. Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *Eur. J. Biochem.* **270**, 3109–3121. <https://doi.org/10.1046/j.1432-1033.2003.03694.x> (2003).
- Markusic, D., Oude-Elferink, R., Das, A. T., Berkhout, B. & Seppen, J. Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promoter. *Nucl. Acids Res.* **33**, e63–e63 (2005).
- Ivics, Z., Hackett, P. B., Plasterk, R. H. & Izsvak, Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510. [https://doi.org/10.1016/s0092-8674\(00\)80436-5](https://doi.org/10.1016/s0092-8674(00)80436-5) (1997).
- Kowarz, E., Loscher, D. & Marschalek, R. Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol. J.* **10**, 647–653. <https://doi.org/10.1002/biot.201400821> (2015).
- Vink, C. A. *et al.* Sleeping beauty transposition from nonintegrating lentivirus. *Mol. Ther.* **17**, 1197–1204 (2009).
- Moldt, B. *et al.* Comparative genomic integration profiling of Sleeping Beauty transposons mobilized with high efficacy from integrase-defective lentiviral vectors in primary human cells. *Mol. Ther.* **19**, 1499–1510. <https://doi.org/10.1038/mt.2011.47> (2011).
- Mates, L. *et al.* Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat. Genet.* **41**, 753–761. <https://doi.org/10.1038/ng.343> (2009).
- Geurts, A. M. *et al.* Gene transfer into genomes of human cells by the sleeping beauty transposon system. *Mol. Ther.* **8**, 108–117 (2003).
- Magnani, C. F. *et al.* Preclinical efficacy and safety of CD19CAR cytokine-induced killer cells transfected with sleeping beauty transposon for the treatment of acute lymphoblastic leukemia. *Hum. Gene Ther.* **29**, 602–613. <https://doi.org/10.1089/hum.2017.207> (2018).
- Bishop, D. C. *et al.* PiggyBac-engineered T cells expressing CD19-specific CARs that lack IgG1 Fc spacers have potent activity against B-ALL xenografts. *Mol. Ther.* **26**, 1883–1895. <https://doi.org/10.1016/j.ymthe.2018.05.007> (2018).
- Roney, I. J., Rudner, A. D., Couture, J. F. & Kaern, M. Improvement of the reverse tetracycline transactivator by single amino acid substitutions that reduce leaky target gene expression to undetectable levels. *Sci. Rep.* **6**, 27697. <https://doi.org/10.1038/srep27697> (2016).
- Kang, K. *et al.* An improved Tet-on system in microRNA overexpression and CRISPR/Cas9-mediated gene editing. *J. Anim. Sci. Biotechnol.* **10**, 43 (2019).
- Hofmann, A., Nolan, G. P. & Blau, H. M. Rapid retroviral delivery of tetracycline-inducible genes in a single autoregulatory cassette. *Proc. Natl. Acad. Sci.* **93**, 5185–5190 (1996).
- Yang, T., Burrows, C. & Park, J. H. Development of a doxycycline-inducible lentiviral plasmid with an instant regulatory feature. *Plasmid* **72**, 29–35. <https://doi.org/10.1016/j.plasmid.2014.04.001> (2014).
- Stebbins, M. J. *et al.* Tetracycline-inducible systems for Drosophila. *Proc. Natl. Acad. Sci.* **98**, 10775–10780 (2001).
- Urlinger, S. *et al.* Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. USA* **97**, 7963–7968. <https://doi.org/10.1073/pnas.130192197> (2000).
- Wang, M. & Marin, A. Characterization and prediction of alternative splice sites. *Gene* **366**, 219–227 (2006).
- Roca, X., Sachidanandam, R. & Krainer, A. R. Determinants of the inherent strength of human 5' splice sites. *RNA* **11**, 683–698 (2005).

24. Das, A. T. *et al.* Viral evolution as a tool to improve the tetracycline-regulated gene expression system. *J. Biol. Chem.* **279**, 18776–18782 (2004).
25. Hinrichs, W. *et al.* Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science (New York, N.Y.)* **264**, 418–420 (1994).
26. Zhou, X., Vink, M., Klaver, B., Berkhout, B. & Das, A. Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther.* **13**, 1382–1390 (2006).
27. Agha-Mohammadi, S. *et al.* Second-generation tetracycline-regulatable promoter: repositioned Tet operator elements optimize transactivator synergy while shorter minimal promoter offers tight basal leakiness. *J. Gene Med.* **6**, 817–828. <https://doi.org/10.1002/jgm.566> (2004).
28. Gronostajski, R. M. Roles of the NFI/CTF gene family in transcription and development. *Gene* **249**, 31–45 (2000).
29. Gaussin, A. *et al.* CTF/NFI transcription factors act as potent genetic insulators for integrating gene transfer vectors. *Gene Ther.* **19**, 15–24 (2012).
30. Valencik, M. L. & McDonald, J. A. Codon optimization markedly improves doxycycline regulated gene expression in the mouse heart. *Transgenic Res.* **10**, 269–275 (2001).
31. Wells, K. D., Foster, J. A., Moore, K., Pursel, V. G. & Wall, R. J. Codon optimization, genetic insulation, and an rtTA reporter improve performance of the tetracycline switch. *Transgenic Res.* **8**, 371–382 (1999).
32. Anastasiadis, K. *et al.* A predictable ligand regulated expression strategy for stably integrated transgenes in mammalian cells in culture. *Gene* **298**, 159–172 (2002).
33. Balasubramanian, S. *et al.* Rapid recombinant protein production from piggyBac transposon-mediated stable CHO cell pools. *J. Biotechnol.* **200**, 61–69 (2015).
34. Barnes, L. M., Bentley, C. M. & Dickson, A. J. Characterization of the stability of recombinant protein production in the GS-NSO expression system. *Biotechnol. Bioeng.* **73**, 261–270 (2001).
35. Massie, B. *et al.* Inducible overexpression of a toxic protein by an adenovirus vector with a tetracycline-regulatable expression cassette. *J. Virol.* **72**, 2289–2296 (1998).
36. Roberts, S. G., Choy, B., Walker, S. S., Lin, Y.-S. & Green, M. R. A role for activator-mediated TFIIB recruitment in diverse aspects of transcriptional regulation. *Curr. Biol.* **5**, 508–516 (1995).
37. Zewel, L., Kumar, K. P. & Reinberg, D. Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* **9**, 1479–1490 (1995).
38. Buratowski, S. & Zhou, H. Functional domains of transcription factor TFIIB. *Proc. Natl. Acad. Sci.* **90**, 5633–5637 (1993).
39. Evans, R., Fairley, J. A. & Roberts, S. G. Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes Dev.* **15**, 2945–2949. <https://doi.org/10.1101/gad.206901> (2001).
40. Roberts, S. G., Ha, I., Maldonado, E., Reinberg, D. & Green, M. R. Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. *Nature* **363**, 741–744 (1993).
41. Hori, R., Pyo, S. & Carey, M. Protease footprinting reveals a surface on transcription factor TFIIB that serves as an interface for activators and coactivators. *Proc. Natl. Acad. Sci.* **92**, 6047–6051 (1995).
42. Loew, R., Heinz, N., Hampf, M., Bujard, H. & Gossen, M. Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol.* **10**, 81. <https://doi.org/10.1186/1472-6750-10-81> (2010).
43. Schonfeld, K. *et al.* Selective inhibition of tumor growth by clonal NK cells expressing an ErbB2/HER2-specific chimeric antigen receptor. *Mol. Ther.* **23**, 330–338. <https://doi.org/10.1038/mt.2014.219> (2015).
44. Grabe, N. AliBaba2: context specific identification of transcription factor binding sites. *In Silico Biol.* **2**, S1–S15 (2002).
45. Farré, D. *et al.* Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucl. Acids Res.* **31**, 3651–3653 (2003).
46. Messeguer, X. *et al.* PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* **18**, 333–334 (2002).
47. Even, D. Y. *et al.* Engineered promoters for potent transient overexpression. *PLoS ONE* **11**, 1–19 (2016).
48. Desmet, F.-O. *et al.* Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucl. Acids Res.* **37**, e67–e67 (2009).

Acknowledgements

We thank Dr. Sarah Saunderson for critical reading of the manuscript.

Author contributions

A.R. and A.M. wrote the paper and supervised the study. A.R. contributed to experimental data in all figures, prepared the figures, tables and performed the bioinformatics analysis. G.T. and A.P. performed experiments.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-70022-0>.

Correspondence and requests for materials should be addressed to A.D.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020