Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas

Fahim Farzadfard,[†,‡,§] Samuel D. Perli,[†,‡] and Timothy K. Lu*[†,‡,§,‖]

[†]Department of Electrical Engineering & Computer Science and Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

[‡]MIT Synthetic Biology Center, 500 Technology Square, Cambridge, Massachusetts 02139, United States

[§]MIT Microbiology Program, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

[‖]Biophysics Program, Harvard University, Cambridge, Massachusetts 02139, United States

S *Supporting Information*

ABSTRACT: Transcriptional regulation is central to the complex behavior of natural biological systems and synthetic gene circuits. Platforms for the scalable, tunable, and simple modulation of transcription would enable new abilities to study natural systems and implement artificial capabilities in living cells. Previous approaches to synthetic transcriptional regulation have relied on engineering DNA-binding proteins, which necessitate multistep processes for construction and optimization of function. Here, we show that the CRISPR/Cas system of *Streptococcus pyogenes* can be programmed to direct both activation and repression to natural and artificial eukaryotic promoters through the simple engineering of guide RNAs with base-pairing complementarity to target DNA sites. We demonstrate that the activity of CRISPR-based transcription factors (crisprTFs) can be tuned by directing multiple crisprTFs to different positions in natural promoters and by arraying multiple crisprTF-binding sites in the context of synthetic promoters in yeast and human cells. Furthermore, externally controllable regulatory modules can be engineered by layering gRNAs with small molecule-responsive proteins. Additionally, single nucleotide substitutions within promoters are sufficient to render them orthogonal with respect to the same gRNA-guided crisprTF. We envision that CRISPR-based eukaryotic gene regulation will enable the facile construction of scalable synthetic gene circuits and open up new approaches for mapping natural gene networks and their effects on complex cellular phenotypes.

KEYWORDS: *synthetic transcription factors, CRISPR/Cas9, RNA-guided multiplex gene regulation, synthetic gene regulation, endogenous gene regulation, synthetic biology*
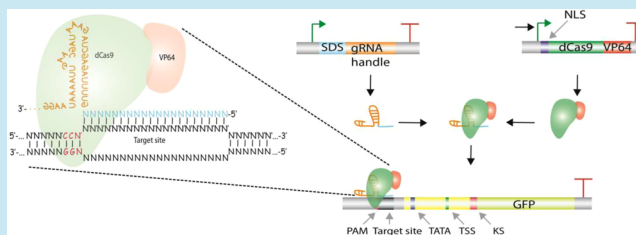
Complex and sophisticated phenotypes in eukaryotic cells manifest from layered regulatory networks and specific expression programs involving the regulated transcription of many genes.[1] As major players in these networks, eukaryotic transcriptional factors (TFs) can integrate multiple signals and perform complex, combinatorial functions on promoters, where regulatory information is encoded in the form of binding sites for TFs and interactions between TFs, to modulate gene expression patterns.[1−3]

Rewiring endogenous transcriptional networks by natural or synthetic TFs is a powerful strategy for interrogating cellular functions and controlling cellular phenotypes.[4−14] Previously, natural DNA-binding domains (DBDs, mainly from bacterial sources, such as TetR, LacI, and LexA) have been used to recruit effector (e.g., activator and repressor) domains to the regulatory regions of eukaryotic genes in order to modulate their transcription.[15−17] This necessitates the placement of DBD-specific operator site(s) in the *cis*-regulatory region of the promoters for specific genes which is a labor- and time-

intensive process, especially if the regulation of multiple genes is desired. Moreover, engineering and modulating complex transcriptional networks requires tunable, extensible, and orthogonal transcription factors. However, only a few orthogonal variants of natural DBDs are well-characterized and changing their specificity has proven to be challenging.[18] As such, the use of natural DBD-based TFs for wiring complex transcriptional networks and synthetic gene circuits has been limited.

To address these limitations, synthetic TFs based on Zinc Fingers (ZFs) and Transcriptional Activator-Like Effectors (TALEs) have been developed.[19−25] The ability to program the specificity of ZFs and TALEs to potentially target any sequence makes these DBDs appealing for designing libraries of orthogonal transcription factors. Synthetic ZF- and TALE-based TFs have been shown to work in a wide range of
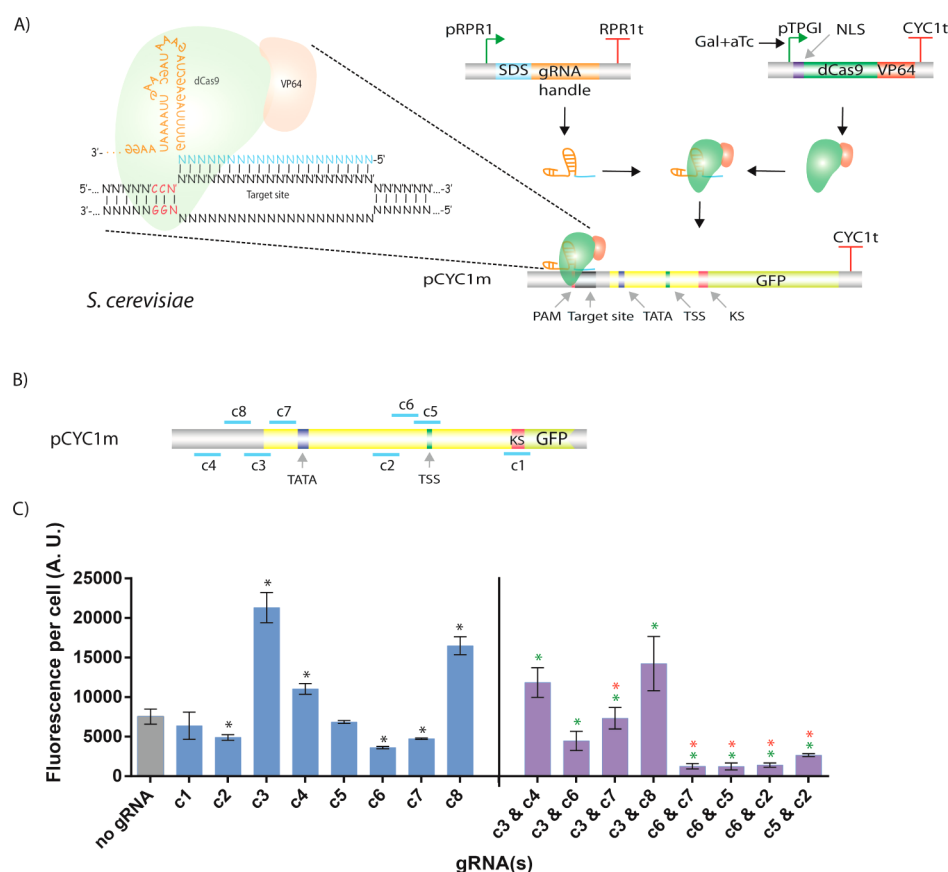
Figure 1. Schematic view of the programmable CRISPR/Cas-based eukaryotic transcriptional regulation system implemented in *S. cerevisiae*. (A) CrisprTF (dCas9_VP64) expression is induced by growing cells in galactose (Gal) + anhydrotetracycline (aTc) media. CrisprTFs are guided to the target sites by guide RNAs (gRNAs), which are constitutively expressed from the pRPR1 promoter and bind to the respective target sites. Specificity of crisprTFs is primarily determined by the 20 bp Specificity Determinant Sequence (SDS) at the 5′-end of the gRNA along with the presence of a Proto-spacer Adjacent Motif (PAM) (NGG) at the target site. (B) Map of pCYC1m illustrating the relative positions of known regulatory elements. TATA: TATA box. TSS: Transcription Start Site. KS: Kozak Sequence. Blue lines indicate target sites for each gRNA (c1-c8). (C) Left panel: Regulation of *gfp* expression from pCYC1m by crisprTFs based on the individual gRNAs shown in (B). Yeast cells expressing crisprTFs and containing the reporter construct were transformed with plasmids expressing gRNAs labeled as shown in the x-axis. Targeting crisprTFs to sequences upstream of the TATA boxes (by c3, c4, and c8 gRNAs) resulted in higher *gfp* expression than the no gRNA control. On the other hand, targeting crisprTFs to sequences spanning the TATA box and the Kozak sequence (by c1, c6, and c7 gRNAs) resulted in reduced *gfp* expression relative to the no gRNA control. Error bars indicate the standard error of the mean for three independent biological replicates. Asterisks (*) on each bar indicate statistically significant changes in *gfp* expression relative to the no gRNA control (based on the two-sided Welch's *t* test, p-value < 0.05). Right panel: Coexpression of multiple gRNAs resulted in synergistic gene regulation. Pairwise combinations of non-neutral gRNAs were expressed from pRPR1 promoters on pRS423 and pRS425 backbones. Green and red asterisks (*) indicate statistically significant changes in *gfp* expression in samples with coexpressed gRNAs relative to the first gRNA only and the second gRNA only, respectively (two-sided Welch's *t* test, p-value < 0.05).

eukaryotes;[22,26,27] however, obtaining a TF for a given target site requires tedious selection processes or multistage DNA assembly protocols.[28,29] Furthermore, the scale of regulation that can be achieved by these TFs is potentially limited by the metabolic burden imposed on the cells and the number of TFs that can be simultaneously encoded in a given cell.[30]

Here, we present a strategy for modulating eukaryotic transcription at natural and synthetic promoters using programmable and tunable synthetic transcription factors based on a bacterial CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas system. Many bacteria use CRISPR-based immune systems to degrade genetic materials of invading phages.[31,32] In these systems, short RNAs expressed from CRISPR loci are used to target an endonuclease protein (Cas9) against invading genetic material. Recently, it has been shown that Cas9 can be used as a programmable tool for genome editing across various organisms.[33−38] In this context, small customizable guide RNAs (gRNAs) can be used to program and target Cas9 endonuclease to specific loci in living cells to induce double (or single)-stranded breaks in DNA. Upon cleavage, error-prone or template-directed repair pathways are triggered, generating variants of the original target loci. Recently, Qi et al.[39] showed that an endonuclease-deficient Cas9 (dCas9, with D10A H841A mutations relative to the wild-type Cas9) can be used as a programmable "CRISPRi" tool for gene silencing in *Escherichia coli*. When targeted to a promoter or ORF of a gene of interest, dCas9 can block progression of RNA polymerase and hence silence expression of the targeted gene. They also provide evidence that CRISPRi is functional in human cells, albeit with much lower efficiency compared with *E. coli*. In addition, Bikard et al. demonstrated that along with programmed transcriptional repression, transcriptional activation can be achieved in *E. coli* by fusing the omega subunit of RNA polymerase to the endonuclease-deficient Cas9.[40]

Here, we achieved versatile, programmable, and multiplexable tools for gene regulation in eukaryotes by functionaliz-
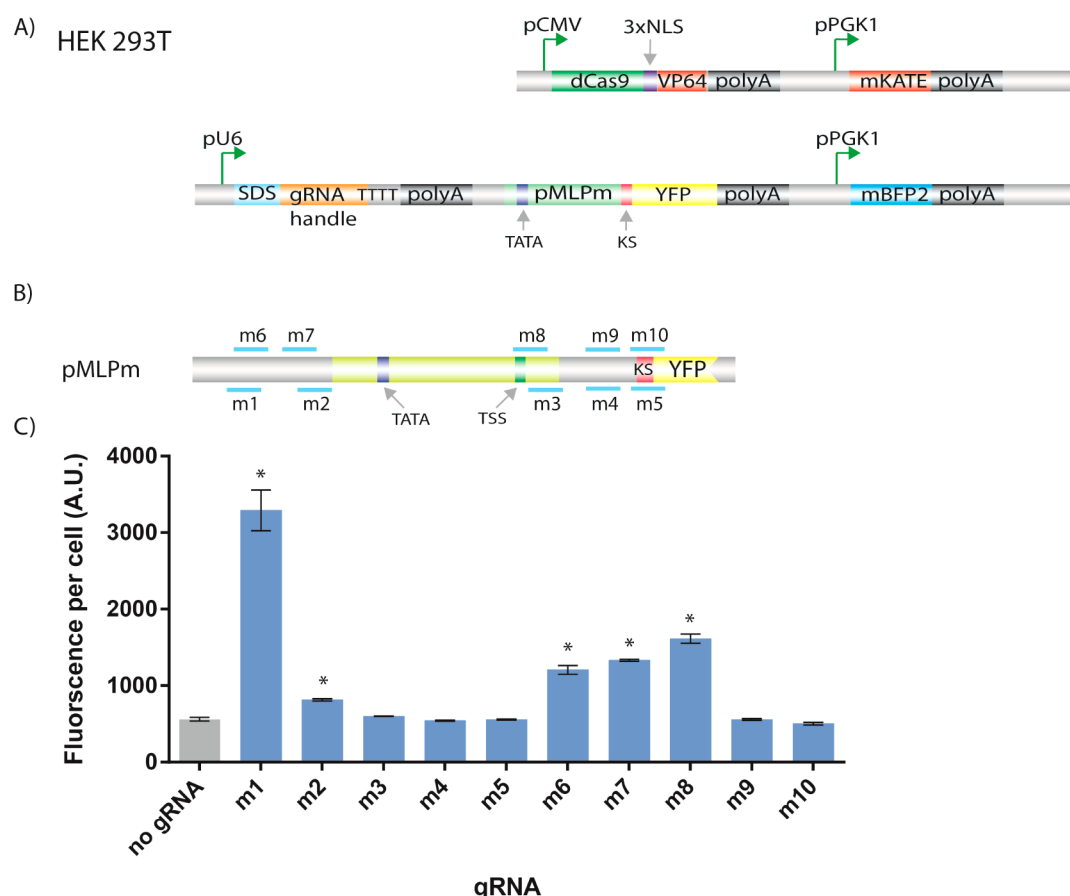
Figure 2. Regulation of *yfp* expression from a minimal MLP promoter (pMLPm) by crisprTFs in HEK293T cells. (A) dCas9_VP64 is expressed in HEK293T cells by the pCMV promoter and directed to target sequences in pMLPm. The mKATE (red) and mBFP2 (blue) fluorophores act as flow-cytometry gating controls for successful plasmid transfections. (B) Map of pMLPm illustrating the relative positions of known regulatory elements. Blue lines indicate target sites for each gRNA. (C) Regulation of *yfp* expression from pMLPm by crisprTFs based on the gRNAs shown in (B). HEK293T cells were cotransfected with the plasmids shown in (A), with specific gRNAs labeled as shown in the x-axis. Targeting crisprTFs to sequences upstream of the TATA box (by m1, m2, m6, and m7 gRNAs) resulted in higher *yfp* expression compared with the no gRNA control. Error bars indicate the standard error of the mean for three independent biological replicates. Asterisks (*) on each bar indicate statistically significant changes in *yfp* expression relative to the no gRNA control (based on the two-sided Welch's *t* test, p-value <0.05).

ing dCas9 with effector domains and targeting both natural and synthetic promoters. As a proof of concept, we made an RNA-guidable transcription factor by fusing dCas9 to an activator domain. Using this CRISPR-based transcription factor (crisprTF), we teased apart the regulatory maps of several natural eukaryotic promoters (in *Saccharomyces cerevisiae* and HEK293T cells) without the need to modify promoter architectures. Unlike previous generations of customizable DBDs (i.e., ZFs and TALEs) that require multistage design and cloning strategies, crisprTFs can be readily customized and retargeted to different loci and regulatory regions *in vivo* using specific gRNAs with homology to target sites (Figure 1). dCas9 thus offers a powerful tool for targeting functions of interest to specific genomic loci in living cells, which can potentially be used to regulate gene expression at will, construct scalable synthetic gene circuits, or rewire endogenous regulatory networks.

To implement crisprTFs in *Saccharomyces cerevisiae*, we fused the SV40 nuclear localization sequence (NLS) and four tandem copies of Herpes Simplex Viral Protein 16 (VP64, a commonly used eukaryotic transcription activator domain) to a codon-optimized *S. pyogenes* dCas9 (Figure 1A).[41] The crisprTF cassette was then cloned under the control of pTPGI, a

synthetic promoter which can be induced by growing cells in galactose + anhydrotetracycline (aTc) media[23] and integrated into the yeast genome. To assess the activity of crisprTF, *gfp* was placed under the control of a minimal CYC1 promoter (pCYC1m) and the whole cassette was integrated into the yeast genome. pCYC1m retains one of the two endogenous TATA boxes of the wild-type CYC1 promoter and lacks binding sites for endogenous regulatory factors in the upstream activating sequence (UAS).[42,43] gRNAs were expressed constitutively from the RNA polymerase III-dependent pRPR1 promoter and the 3′-ends of the gRNAs were defined by the pRPR1 terminator.[44]

The expression of gRNAs targeting different regions in the pCYC1m (as shown in Figure 1B) resulted in various statistically significant levels of reporter fluorescence compared to the no gRNA control (Figure 1C, left panel). Targeting crisprTFs to the sequences upstream of the TATA boxes (by c3, c4, and c8 gRNAs) led to the activation of the reporter. However, targeting crisprTFs to the sequences spanning the TATA box and the Kozak sequence (KS) resulted in the repression of *gfp* expression to various degrees. Stronger repression was achieved when crisprTFs were targeted to the proximity of TATA box (using c7 gRNA) and to the vicinity of

the TATA box and the transcription start site (using c2 and c6 gRNAs), likely due to interference of crisprTFs with the formation of the transcriptional initiation complex.[45] No activation was observed with any of the eight tested gRNAs when dCas9, without a fused activator domain, was targeted to pCYC1m (Figure S1, Supporting Information). All the tested gRNAs in this strain repressed *gfp* expression to some extent (Figure S1, Supporting Information), with highest repression observed with c6 and c7. These results demonstrate that dCas9 is able to repress transcription but requires an activation domain (VP64) to activate transcription of a target locus and further supports the hypothesis that dCas9 (or as a fusion to VP64) can act as a repressor by interfering with the formation of the transcriptional initiation complex. Similar results were achieved with the GAL1 promoter (pGAL1) and its variants, where targeting crisprTFs to sequences upstream and downstream of TATA box led to activation and repression of the GFP reporter, respectively (Figure S2, Supporting Information). These results indicate that a single crisprTF can be programmed to act as both an activator and a repressor by targeting it to different positions across endogenous promoters.

The activity of a promoter is determined by combinatorial interactions between transcriptional regulatory factors bound to that promoter. We thus investigated the effects of binding of multiple crisprTFs targeted to the same promoter. To this end, pairwise combinations of gRNAs with non-neutral effects (those that showed either activation or repression in the left panel of Figure 1C) were coexpressed. As shown in the right panel of Figure 1C, coexpression of repressor gRNAs resulted in synergistic repression of the reporter (up to 7× repression was achieved with coexpression of the c5 and c6 pair as well as the c6 and c7 pair). On the other hand, when a repressor gRNA was coexpressed with an activator gRNA (e.g., the c3 and c6 pair), an intermediate level of GFP expression was achieved, indicating an antagonistic interaction between the two gRNAs. Moreover, the effects of repressor gRNAs were dominant over activator gRNAs, suggesting that interruption of the formation of the transcription initiation complex has a stronger effect than activation.[46] Coexpression of two activator gRNAs (e.g., c3 and c4) did not result in synergistic activation of the reporter, which suggests that the relative positions and interactions of bound activators are important for determining synergistic activation.[46] Consistent with our results, it has been shown that synergistic activation from synthetic promoters with multiple GAL4 operator sites depends on the distance and helical phase of the operator sites.[47] Furthermore, in another study it has been shown that not all of the combinations of TALE-activators targeted to the same promoter result in synergistic activation.[25]

We next sought to investigate the activity of crisprTFs in human cells. To this end, a human-codon-optimized crisprTF cassette was placed on a plasmid under the control of the constitutive cytomegalovirus immediate-early promoter (pCMV). The gRNAs were expressed constitutively from a separate plasmid by the RNA polymerase III-dependent U6 promoter (pU6), as previously described[37,48] (Figure 2A). After transfection of these plasmids into HEK293T cells, we investigated the regulatory architecture of the minimal adenovirus major late promoter (pMLPm)[49] by targeting crisprTFs to different positions across this promoter (Figure 3A). YFP was used as the readout for pMLPm promoter activity.

Consistent with the results obtained in *S. cerevisiae*, crisprTFs activated gene expression when targeted to sequences upstream
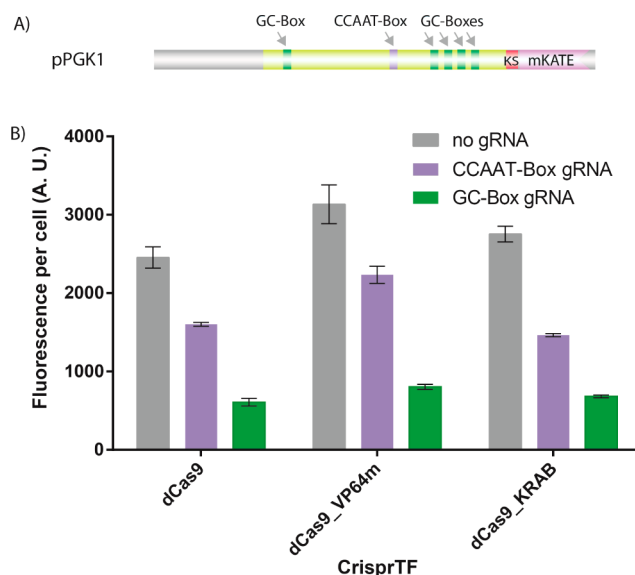

Figure 3. CrisprTF-mediated repression of the constitutive pPGK1 promoter in HEK293T cells. (A) Map of the pPGK1 promoter illustrating the relative positions of known regulatory elements. (B) CrisprTF-based targeted repression of the constitutive pPGK1 promoter. Constructs expressing different dCas9-based proteins (dCas9, dCas9_VP64, and dCas9_KRAB) were cotransfected with plasmids containing pPGK1_mKATE and constructs expressing no gRNAs or gRNAs targeting the CCAAT box or the GC-box gRNA. Significant repression of the pPGK1 promoter relative to the no gRNA control was observed with all of the three different dCas9 constructs (dCas9, dCas9-VP64, and dCas9-KRAB). Error bars indicate the standard error of the mean for three independent biological replicates.

of the pMLPm TATA box (using m1, m2, m6, or m7 gRNA) or downstream of the transcription start site (m8 gRNA) (Figure 2B and C). Since the basal expression level of the pMLPm promoter is low, it was challenging to detect significant repression from this promoter. Thus, to demonstrate that crisprTFs can function as transcriptional repressors in mammalian cells, we placed mKATE under the control of a constitutive mammalian promoter, phosphoglycerate kinase 1 (pPGK1),[50] and targeted crisprTFs to this promoter (Figure 3A). pPGK1 is a strong, constitutive, TATA-less promoter that contains a CCAAT box[51] and five GC-boxes.[52] These sites are the binding sites for the endogenous human transcription factors CBP and SP1, respectively. Targeting dCas9 alone, dCas9 fused to VP64 domain, or dCas9 fused to KRAB domain to the CCAAT box or the GC-boxes resulted in significant repression of the reporter gene (Figure 3B), presumably by preventing endogenous transcription factors from binding to specific DNA recognition elements[51,52] within the pPGK1 promoter.

We further sought to explore the tunability of crisprTFs in the context of synthetic promoters. In order to do so, we engineered multiple artificial binding sites (operators), pCYC1m separated by 20 base-pair sequences, upstream of the pCYC1m promoter in *S. cerevisiae* (Figure 4A). Expression of a gRNA, which targeted these arrayed operator sites resulted in synergistic activation of the *bfp* reporter (Figure 4B). Higher levels of *bfp* expression (up to 70-fold activation with 12× gRNA operator sites) were achieved by increasing the number of gRNA binding sites upstream of the engineered pCYC1m. This level of activation in yeast is comparable to the activation reported for commonly used endogenous yeast promoters (e.g.,
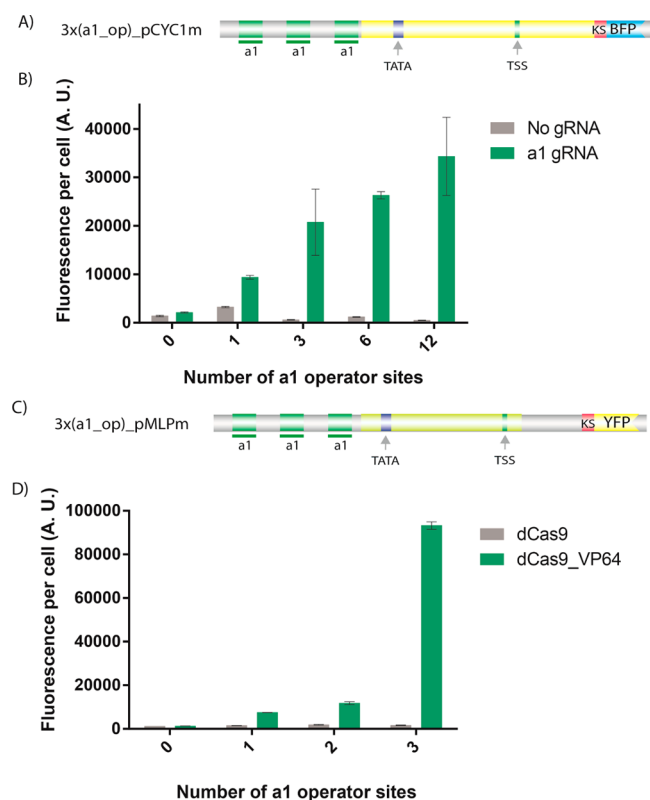
Figure 4. Synergistic and tunable activation of synthetic promoters with arrayed operator sites upstream of pCYC1m in *S. cerevisiae* and pMLPm in HEK293T cells using crisprTFs. (A) A schematic view of the pCYC1m synthetic promoter with three a1_gRNA operator sites (3× DNA sequences recognized by the a1_gRNA) arrayed upstream of pCYC1m, thus named 3x(a1_op)_pCYC1m. (B) Increasing the number of arrayed a1_gRNA operator sites upstream of pCYC1m resulted in higher *bfp* expression in *S. cerevisiae* cells expressing the a1_gRNA compared to the no gRNA controls. Error bars indicate the standard error of the mean for three independent biological replicates. (C) A schematic view of the pMLPm synthetic promoter with three a1_gRNA operator sites arrayed upstream of pMLPm, thus named 3x(a1_op)_pMLPm. (D) Increasing the number of arrayed a1_gRNA operator sites upstream of pMLPm resulted in higher *yfp* expression in HEK293T cells when cotransfected with a1_gRNA and dCas9_VP64 versus when cotransfected with a1_gRNA and dCas9. Error bars indicate the standard error of the mean for three independent biological replicates.

pGAL1[53] and pCUP1[54]) and synthetic promoters that are modulated by engineered bacterial DNA binding domains (e.g., TetON/TetOFF promoters[16]). Our results are consistent with previous observations that arraying multiple binding sites for a transcriptional activator upstream of a promoter results in longer transcriptional bursts from the promoter and thus leads to higher levels of expression from the targeted promoter.[55] We saw similar synergistic activation in HEK293T cells when multiple gRNA operator sites were placed upstream of pMLPm (Figure 4C and D), with up to 56× activation attained with 3× gRNA operator sites. The level of activation that is achieved by crisprTFs in human cells is comparable to the levels of activation reported for ZF- and TALE-activators,[19,24,25] where higher activation levels can be achieved by increasing the number of operator sites or by targeting multiple synthetic transcription factors to the same locus.[25,27] These results demonstrate that crisprTFs can be used to build synthetic

promoters with tunable strengths by the straightforward engineering of gRNA-binding sites.

In many applications, control of the activity of a transcription factor by an inducer (e.g., a small molecule) is desired. With crisprTFs, one viable strategy is to constitutively express the protein component of the system (i.e., dCas9) and then modulate the amount of gRNA available for binding to dCas9 and thus the activity achieved at the target DNA. To test this strategy, we constructed an anhydrotetracycline (aTc)-inducible pRPR1 promoter by placing a TetR operator site (1xTetO) in the pRPR1 promoter, as previously described,[56] and constitutively expressing Tet repressor (TetR) (Figure 5A). Furthermore, to make the expression of dCas9_VP64 independent of aTc, we placed it under the control of pGAL1. We tested this system in cells containing a pCYC1m promoter with six a1_gRNA operator sites, named 6x(a1_op)_pCYC1m, controlling expression of *bfp*. As shown in Figure 5B, *bfp* expression increased about 20-fold when *S. cerevisiae* cells were induced with galactose and aTc, compared with galactose only, thus demonstrating external control of crisprTF activity.

Although the exact parameters that determine the sequence specificity of Cas9 are not yet well-characterized, it has been shown that the specificity of Cas9 for target loci is mainly determined by the PAM motif (NGG) and the 12 base-pairs preceding this motif (seed sequence). Mutations in any of these 15 positions can severely affect the targeting efficiency and binding specificity of Cas9.[35,37,57,58] To test whether this property can be used to create orthogonal synthetic promoters, new PAM motifs or single or multiple point mutations (marked with asterisks, Figure 6A) were introduced into the wild-type pCYC1m sequence to design a modified pCYC1m promoter (pCYC1m(modified), Figure 6A). As shown in the left panel of Figure 6B, the modified promoter did not respond to the wild-type gRNAs (except for c4, which still perfectly matched target sites within this modified promoter). However, this modified promoter responded to a new set of gRNAs (cm1, cm2, and cm5) that were designed to match the modified target sites. The wild-type pCYC1m promoter did not respond to the new set of gRNAs (cm1−cm6) (Figure 6B, left panel). These results demonstrate that as little as a single base-pair mismatch is sufficient to direct the crisprTF to one locus while preventing activity at another locus.

To further demonstrate the potential of crisprTFs toward constructing synthetic promoters and gRNAs that are orthogonal with respect to each other, we tested three randomly designed gRNAs (a1, a2 and a3 gRNAs) for their ability to activate each other's target sequences. As shown in Figure 6C, each of the gRNAs exhibited high activity at their cognate target sequences but low activity at noncognate sequences. These results suggest that one can construct synthetic promoters and gRNAs that are orthogonal with respect to each other and to the host genome, especially within eukaryotes with smaller genomes, such as yeasts.

During the course of the peer-review for this work, similar systems for transcriptional control in eukaryotic cells were described. Gilbert et al.[59] demonstrated that CRISPR-mediated gene repression and activation can be achieved in both yeast and mammalian cells by using fusions of dCas9 with repressor and activator domains respectively. Furthermore, Maeder et al.[60] and Perez-Pinera et al.[61] showed synergistic CRISPR/Cas-based gene activation in human cells with multiple gRNAs. In this paper, we additionally show that both activation and
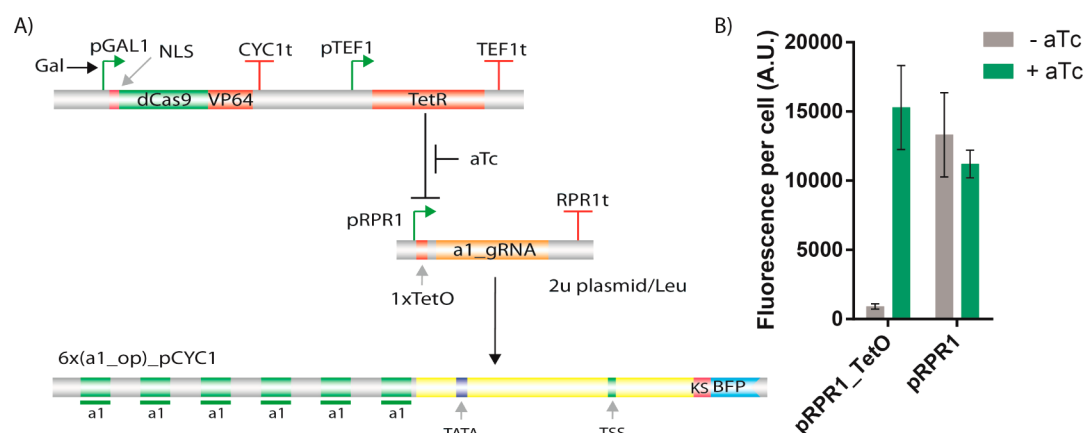
Figure 5. Inducible crisprTF-guided activation of synthetic promoters. (A) Schematic of the aTc-inducible pRPR1_TetO promoter. Expression of dCas9_VP64 is driven by the galactose-inducible pGAL1 promoter. A TetR operator site (1xTetO) was placed in the pRPR1 promoter to make an aTc-responsive pRPR1_TetO promoter. Addition of aTc releases TetR-mediated repression on the pRPR1_TetO promoter and results in a1_gRNA expression. (B) aTc-dependent *bfp* expression from a synthetic 6x(a1_op)_pCYC1m promoter. *S. cerevisiae* cells containing the circuit shown in (A) were grown in galactose media with either 250 ng/mL aTc or no aTc. Error bars indicate the standard error of mean for three biological replicates.

repression functions can be achieved with only one transcription factor, by targeting dCas9_VP64 fusions to different regulatory sequences along a promoter. Our finding that one can activate or repress the expression of a gene of interest by directing a single protein to different positions of a promoter is advantageous for the efficient design of synthetic transcriptional networks or rewiring natural ones. This property obviates the need for using separate orthogonal Cas9 protein fusions as activators and repressors. Furthermore, we demonstrate that more sophisticated regulatory motifs, such as small-molecule responsive modules can be built for crisprTFs, thus enabling external control of crisprTF-based transcriptional circuits. Such synthetic modules can be interfaced with other regulatory elements to achieve more complex regulation for synthetic biology.

Our results show that dCas9 can be used as a customizable RNA-guided DNA-binding platform for the regulation of gene expression at natural and synthetic promoters in eukaryotic cells. The ease of design and expression of customized gRNAs in comparison to ZFs and TALEs make CRISPR-based transcription factors appealing as synthetic TFs for modulating endogenous gene expression as well as for synthetic biology. The ability to customize the target site of dCas9 via the expression of short gRNAs obviates the need to engineer multiple orthogonal DBDs in order to construct complex transcriptional circuits. This could potentially reduce the overall metabolic burden on cells and enable the integration of more complex synthetic computation and logic within living cells.[62,63] More complex regulatory and logic circuits, such as cascades and complex digital logics gates can be built by layering crisprTFs. The possibility of integrating multiple inputs at a single promoter expands the regulatory potential and provides us with increased flexibility that can be leveraged while designing synthetic transcriptional networks or rewiring endogenous pathways.

Furthermore, since both activation and repression functions can be achieved with crisprTFs, the crisprTF platform may be advantageous compared to noncoding RNA-based gene regulatory platforms where only repression can be achieved. In a way, crisprTFs combine the multiplexability of RNA-based regulatory approaches with the flexibility and rich functionality repertoire of protein-based gene regulatory approaches: Cas9

can be functionalized with regulatory domains of interest (e.g., activation, repression, or epigenetic effector) and then be targeted to multiple loci using different gRNAs.

Future work is needed to define the range of effector domains that can be used with dCas9 for a variety of regulatory functions, including transcriptional regulation and epigenetic modifications. In addition, the identification, characterization, and optimization of Cas9 homologues or evolved variants may enable enhanced activity and specificity of this system. Moreover, the ability to synthesize random libraries of gRNAs opens the possibility for high-throughput perturbations of transcriptional networks and screening for desirable phenotypes. Ultimately, we envision that crisprTFs will enable the regulation and perturbation of natural transcriptional networks as well as the construction of complex synthetic circuits at an unprecedented speed and scale.

■ METHODS

Strain and Plasmid Construction. *Saccharomyces cerevisiae.* dCas9 (endonuclease-deficient Cas9, with D10A and H841A mutations relative to the wild-type sequence of *S. pyogenes* Cas9[39]) with an N-terminal SV40 nuclear localization signal (NLS) was codon-optimized for expression in *S. cerevisiae* and cloned into a pRS314 backbone under control of the pTPGI promoter.[23] The RNA-guided transcription factors (crisprTFs) were built by fusing four repeats of the minimal domain of the herpes simplex viral protein 16 (VP16) to the C-terminus of dCas9 (dCas9_VP64). The crisprTF-expressing plasmid was then integrated into the TRP1 locus of *S. cerevisiae* W303.

The reporter plasmids were built by cloning yeast-enhanced *gfp* under the control of the wild-type or modified pCYC1m promoter into pRS406 using one-step Gibson assembly. The reporters for the multiple-gRNA-binding-site experiment (Figure 4A) were built by cloning the corresponding number of binding sites upstream of the pCYC1m promoter driving production of EBFP2. All reporters were integrated into the *bla1* locus of the integrated crisprTF plasmid.

To build gRNA-expressing plasmids, empty gRNA expressing vectors were first made by cloning the pRPR1 promoter (an RNA-polymerase-III-dependent promoter[44]), the gRNA handle (flanked by *Hin*dIII and Xho1 sites), and the RPR terminator
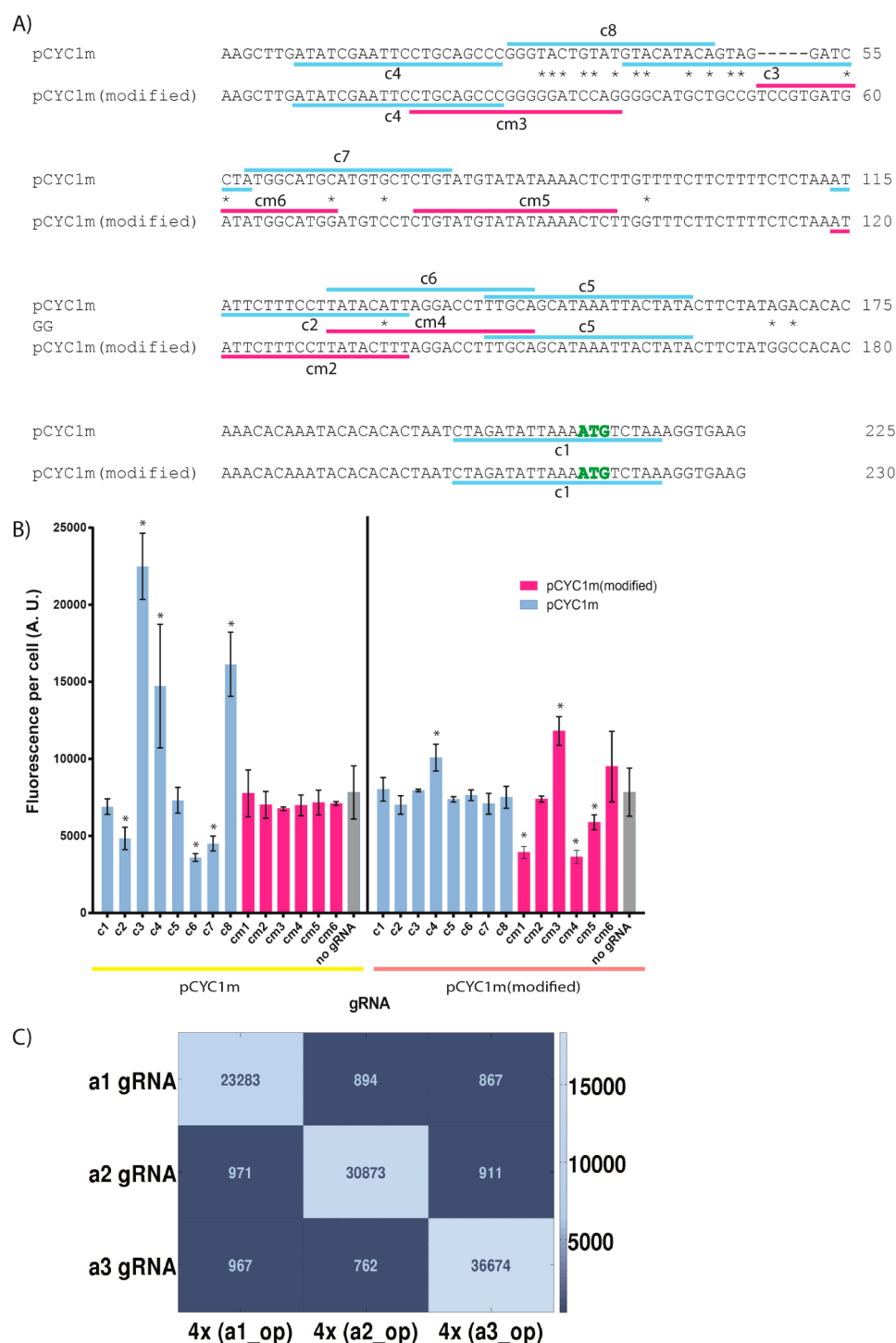
Figure 6. Constructing orthogonal crisprTF-responsive promoters. (A) A schematic view of gRNAs targeting the wild-type and modified pCYC1m promoters in yeast. Only the c1-c8 gRNAs have perfect homology to the target sequences in pCYC1m. On the other hand, c1, c4, c5, and cm1-cm6 gRNAs have perfect homology to the sequences in the pCYC1m(modified) promoter. Mismatches between pCYC1m and pCYC1m(modified) are marked by asterisks (*). (B) pCYC1m only responds to gRNAs that are perfectly matching gRNAs (c1-c8 gRNAs) and not to those that contain mismatches (cm1−cm6 gRNAs). The pCYC1m(modified) promoter responds to the cm1−cm6 gRNAs. Those gRNAs that bind to the sequences upstream of the TATA boxes activate *gfp* expression and those that target sequences downstream of the TATA boxes repress *gfp* expression. The c1 and c5 gRNAs have similarly neutral effects on both the wild-type and modified promoters. Error bars indicate standard error of the mean for three independent biological replicates. Asterisks (*) on each bar indicate statistically significant activation or repression relative to no gRNA controls (based on the two-sided Welch's *t* test, p-value < 0.05). (C) Heat map illustrating the orthogonality of crisprTFs in human cells. Plasmids encoding three orthogonal gRNAs (a1, a2, and a3 gRNAs) were cotransfected into HEK293T cells along with one of the three reporter plasmids (each encoding 4× operator sites for a given gRNA) upstream of pMLPm promoter driving *yfp* expression. Only cognate interactions between gRNAs and target binding sites resulted in significant activation. The standard error of the mean of YFP fluorescence for three independent biological replicates is indicated in each cell of the heat-map plot.

into the SacI and *Kpn*I sites of either the pRS423 or pRS425 plasmid using one-step Gibson assembly. The specificity determinant sequence (SDS) for each gRNA was then cloned into the *Hin*dIII site of these vectors by one-step Gibson assembly. Sequences of the constructs used in this study are listed in Table S1, Supporting Information.

HEK293T Cells. To construct the mammalian dCas9_VP64 expressing plasmid, we first introduced D10A and H841A mutations into hCas9[37] (Addgene, Plasmid #41815). Then, three repeats of SV40 NLS (3xNLS) were fused to the C-terminus of the mutated hCas9 using a PCR-based assembly protocol. Using a multipart Gibson assembly protocol, the immediate-early promoter of cytomegalovirus (pCMV), dCas9_3xNLS, VP64, and SV40 polyA terminator were cloned into the NotI site of the pG5-Luc plasmid (Promega). To monitor successfully transfected cells by flow cytometry, we replaced the original luciferase gene in pG5-Luc with *mKATE* (Evrogen). The resulting pPGK1_*mKATE* cassette served as a constitutive fluorescent protein control that was used to gate for the presence of the crisprTF-expressing plasmid with flow cytometry.

The gRNA expression plasmids were constructed by cloning the 138 bp human U6 promoter (an RNA-polymerase-III-dependent promoter[48]), along with the gRNA handle and terminator into a plasmid containing pPGK1-*eBFP2* flanked by the SV40 polyA terminator (a gift from Lior Nissim). A SacI site was placed at the 3′-end of the U6 promoter to enable the cloning of different specificity determining sequences for each gRNA. The reporters were assembled into the gRNA-expressing plasmid through a one-step Gibson assembly reaction, where the upstream polyadenylation signal and transcriptional pause site from pG5-Luc, along with a 41 bp, minimal adenovirus type 2 major late promoter (pMLPm), mYFP, and HSV polyA signal were cloned into the AatII site of the gRNA-expressing plasmids.

For the synthetic promoter experiments, additional gRNA operator sites were cloned in the NheI site upstream of the pMLPm promoter (see Supporting Information). For the repression experiments, dCas9_KRAB was constructed by cloning a 366 bp KRAB domain to the C-terminus of dCas9. GCCACC was used as the Kozak sequence for the expression of dCas9_VP64, *mYFP*, *eBFP2*, and *mKATE*.

Unless directly targeted by gRNAs for repression assays, the mKATE fluorescent protein on the crisprTF-expression plasmid and the eBFP2 fluorescent protein on the reporter/gRNA plasmid served as our gating controls for flow cytometry analysis.

Fluorescence Assays. To assess expression of the reporter constructs, yeast cells expressing different gRNAs (or no gRNA as control) were grown overnight (900 rpm, 30C) in 96-deep-well plates in yeast minimal media supplemented with glucose with appropriate selection (three independent cultures for each sample). Ten microliters of these cultures were then transferred into fresh media supplemented with galactose +250 ng/mL anhydrotetracycline (aTc) and grown for 20 h (900 rpm, 30C) before analysis by flow cytometry.

For the human cell culture experiments, HEK293T kidney epithelial cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 1% glutamine, and 1% penicillin/streptomycin. Cells were grown under 5% CO_2 at 37 °C. HEK293T cells were transfected with Fugene-HD transfection reagent (Promega)

and assayed for gene expression with flow cytometry at 48 h post transfection.

An LSR Fortessa II flow cytometer equipped with 405 nm, 488 nm, and 561 nm lasers was used for all the experiments. GFP/YFP, BFP, and mKATE levels were detected using 488/FITC, 405/Pacific-Blue, and 561/TX-red laser/filter sets, respectively. All samples were uniformly gated by forward and side scatter. Additional gating for the presence of red and blue fluorophores was applied to the HEK293T samples to ensure only cells successfully transfected with both the crisprTF and the reporter/gRNA plasmids are analyzed. For each gated sample, the mean fluorescence per cell was calculated. Three independent biological samples were used to calculate the mean and standard error of the mean for each data point.

■ ASSOCIATED CONTENT

ⓢ Supporting Information
This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: timlu@mit.edu.

Author Contributions
F.F. built and tested the yeast crisprTF system. S.D.P. built and tested the human crisprTF system. F.F., S.D.P., and T.K.L. designed experiments, analyzed data, and wrote the paper.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Lelli, K. M., Slattery, M., and Mann, R. S. (2012) Disentangling the many layers of eukaryotic transcriptional regulation. *Annu. Rev. Genet. 46*, 43−68.
(2) Ellwood, K., Huang, W., Johnson, R., and Carey, M. (1999) Multiple layers of cooperativity regulate enhanceosome-responsive RNA polymerase II transcription complex assembly. *Mol. Cell. Biol. 19*, 2613−2623.
(3) Chen, L. (1999) Combinatorial gene regulation by eukaryotic transcription factors. *Curr. Opin. Struct. Biol. 9*, 48−55.
(4) Auslander, S., Auslander, D., Muller, M., Wieland, M., and Fussenegger, M. (2012) Programmable single-cell mammalian biocomputers. *Nature 487*, 123−127.
(5) Bacchus, W., Lang, M., El-Baba, M. D., Weber, W., Stelling, J., and Fussenegger, M. (2012) Synthetic two-way communication between mammalian cells. *Nat. Biotechnol. 30*, 991−996.
(6) Bashor, C. J., Horwitz, A. A., Peisajovich, S. G., and Lim, W. A. (2010) Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annu. Rev. Biophys. 39*, 515−537.
(7) Nandagopal, N., and Elowitz, M. B. (2011) Synthetic biology: integrated gene circuits. *Science 333*, 1244−1248.
(8) Guet, C. C., Elowitz, M. B., Hsing, W., and Leibler, S. (2002) Combinatorial synthesis of genetic networks. *Science 296*, 1466−1470.
(9) Bandyopadhyay, S., Mehta, M., Kuo, D., Sung, M. K., Chuang, R., Jaehnig, E. J., Bodenmiller, B., Licon, K., Copeland, W., Shales, M., Fiedler, D., Dutkowski, J., Guenole, A., van Attikum, H., Shokat, K. M., Kolodner, R. D., Huh, W. K., Aebersold, R., Keogh, M. C., Krogan, N.

J., and Ideker, T. (2010) Rewiring of genetic networks in response to DNA damage. *Science 330*, 1385−1389.

(10) Kwon, E. J., Laderoute, A., Chatfield-Reed, K., Vachon, L., Karagiannis, J., and Chua, G. (2012) Deciphering the transcriptional-regulatory network of flocculation in *Schizosaccharomyces pombe*. *PLoS Genet. 8*, e1003104.

(11) Tabor, J. J., Salis, H. M., Simpson, Z. B., Chevalier, A. A., Levskaya, A., Marcotte, E. M., Voigt, C. A., and Ellington, A. D. (2009) A synthetic genetic edge detection program. *Cell 137*, 1272−1281.

(12) Lou, C., Liu, X., Ni, M., Huang, Y., Huang, Q., Huang, L., Jiang, L., Lu, D., Wang, M., Liu, C., Chen, D., Chen, C., Chen, X., Yang, L., Ma, H., Chen, J., and Ouyang, Q. (2010) Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. *Mol. Syst. Biol. 6*, 350.

(13) Leisner, M., Bleris, L., Lohmueller, J., Xie, Z., and Benenson, Y. (2010) Rationally designed logic integration of regulatory signals in mammalian cells. *Nat. Nanotechnol. 5*, 666−670.

(14) Nissim, L., and Bar-Ziv, R. H. (2010) A tunable dual-promoter integrator for targeting of cancer cells. *Mol. Syst. Biol. 6*, 444.

(15) Gossen, M., and Bujard, H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A. 89*, 5547−5551.

(16) Belli, G., Gari, E., Piedrafita, L., Aldea, M., and Herrero, E. (1998) An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucl. Acids Res. 26*, 942−947.

(17) Cronin, C. A., Gluba, W., and Scrable, H. (2001) The lac operator-repressor system is functional in the mouse. *Genes Dev. 15*, 1506−1517.

(18) Urlinger, S., Baron, U., Thellmann, M., Hasan, M. T., Bujard, H., and Hillen, W. (2000) Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. U.S.A. 97*, 7963−7968.

(19) Khalil, A. S., Lu, T. K., Bashor, C. J., Ramirez, C. L., Pyenson, N. C., Joung, J. K., and Collins, J. J. (2012) A synthetic biology framework for programming eukaryotic transcription functions. *Cell 150*, 647−658.

(20) Li, Y., Moore, R., Guinn, M., and Bleris, L. (2012) Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Sci. Rep. 2*, 897.

(21) Bae, K. H., Kwon, Y. D., Shin, H. C., Hwang, M. S., Ryu, E. H., Park, K. S., Yang, H. Y., Lee, D. K., Lee, Y., Park, J., Kwon, H. S., Kim, H. W., Yeh, B. I., Lee, H. W., Sohn, S. H., Yoon, J., Seol, W., and Kim, J. S. (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol. 21*, 275−280.

(22) Morbitzer, R., Romer, P., Boch, J., and Lahaye, T. (2010) Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc. Natl. Acad. Sci. U.S.A. 107*, 21617−21622.

(23) Ellis, T., Wang, X., and Collins, J. J. (2009) Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol. 27*, 465−471.

(24) Lohmueller, J. J., Armel, T. Z., and Silver, P. A. (2012) A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucl. Acids Res. 40*, 5180−5187.

(25) Perez-Pinera, P., Ousterout, D. G., Brunger, J. M., Farin, A. M., Glass, K. A., Guilak, F., Crawford, G. E., Hartemink, A. J., and Gersbach, C. A. (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat. Methods 10*, 239−242.

(26) Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J., and Voytas, D. F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucl. Acids Res. 39*, e82.

(27) Maeder, M. L., Linder, S. J., Reyon, D., Angstman, J. F., Fu, Y., Sander, J. D., and Joung, J. K. (2013) Robust, synergistic regulation of human gene expression using TALE activators. *Nat. Methods 10*, 243−245.

(28) Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., and Zhang, F. (2012) A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc. 7*, 171−192.

(29) Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. A., Sander, J. D., and Joung, J. K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol. 30*, 460−465.

(30) Glick, B. R. (1995) Metabolic load and heterologous gene expression. *Biotechnol. Adv. 13*, 247−261.

(31) Wiedenheft, B., Sternberg, S. H., and Doudna, J. A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature 482*, 331−338.

(32) Bhaya, D., Davison, M., and Barrangou, R. (2011) CRISPR/Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet. 45*, 273−297.

(33) Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J. R., and Joung, J. K. (2013) Efficient genome editing in zebrafish using a CRISPR/Cas system. *Nat. Biotechnol. 31*, 227−229.

(34) Dicarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013) Genome engineering in Saccharomyces cerevisiae using CRISPR/Cas systems. *Nucl. Acids Res. 41*, 4336−4343.

(35) Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science 339*, 819−823.

(36) Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol. 31*, 233−239.

(37) Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. (2013) RNA-guided human genome engineering via Cas9. *Science 339*, 823−826.

(38) Cho, S. W., Kim, S., Kim, J. M., and Kim, J. S. (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol. 31*, 230−232.

(39) Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell 152*, 1173−1183.

(40) Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L. A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.*, DOI: 10.1093/nar/gkt520.

(41) Beerli, R. R., Segal, D. J., Dreier, B., and Barbas, C. F., 3rd. (1998) Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U.S.A. 95*, 14628−14633.

(42) Li, W. Z., and Sherman, F. (1991) Two types of TATA elements for the CYC1 gene of the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol. 11*, 666−676.

(43) Guarente, L., Yocum, R. R., and Gifford, P. (1982) A GAL10-CYC1 hybrid yeast promoter identifies the GAL4 regulatory region as an upstream site. *Proc. Natl. Acad. Sci. U.S.A. 79*, 7410−7414.

(44) Burke, D. H., and Nickens, D. G. (2002) Expressing RNA aptamers inside cells to reveal proteome and ribonome function. *Brief Funct. Genomic Proteomic 1*, 169−188.

(45) Kuras, L., and Struhl, K. (1999) Binding of TBP to promoters *in vivo* is stimulated by activators and requires Pol II holoenzyme. *Nature 399*, 609−613.

(46) Li, X. Y., Virbasius, A., Zhu, X., and Green, M. R. (1999) Enhancement of TBP binding by activators and general transcription factors. *Nature 399*, 605−609.

(47) Huang, Q., Gong, C., Li, J., Zhuo, Z., Chen, Y., Wang, J., and Hua, Z. C. (2012) Distance and helical phase dependence of synergistic transcription activation in cis-regulatory module. *PLoS One 7*, e31198.

(48) Kwak, Y. D., Koike, H., and Sugaya, K. (2003) RNA interference with small hairpin RNAs transcribed from a human U6 promoter-driven DNA vector. *J. Pharmacol. Sci. 93*, 214−217.

(49) Sawadogo, M., and Roeder, R. G. (1985) Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell 43*, 165−175.

(50) Adra, C. N., Boer, P. H., and McBurney, M. W. (1987) Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene 60*, 65−74.

(51) Dorn, A., Bollekens, J., Staub, A., Benoist, C., and Mathis, D. (1987) A multiplicity of CCAAT box-binding proteins. *Cell 50*, 863−872.

(52) Blake, M. C., Jambou, R. C., Swick, A. G., Kahn, J. W., and Azizkhan, J. C. (1990) Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. *Mol. Cell. Biol. 10*, 6632−6641.

(53) Yocum, R. R., Hanley, S., West, R., Jr., and Ptashne, M. (1984) Use of lacZ fusions to delimit regulatory elements of the inducible divergent GAL1-GAL10 promoter in Saccharomyces cerevisiae. *Mol. Cell. Biol. 4*, 1985−1998.

(54) Labbe, S., and Thiele, D. J. (1999) Copper ion inducible and repressible promoter systems in yeast. *Methods Enzymol. 306*, 145−153.

(55) Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., and Tyagi, S. (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS Biol. 4*, e309.

(56) Bak, G., Hwang, S. W., Ko, Y., Lee, J., Kim, Y., Kim, K., Hong, S. K., and Lee, Y. (2010) On-off controllable RNA hybrid expression vector for yeast three-hybrid system. *BMB Rep. 43*, 110−114.

(57) Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science 337*, 816−821.

(58) Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., and Sander, J. D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, DOI: 10.1038/nbt.2623.

(59) Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., and Qi, L. S. (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell. 154*, 442−451.

(60) Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H., and Joung, J. K. (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat Methods*, DOI: 10.1038/nmeth.2598.

(61) Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Leong, K. W., Guilak, F., Crawford, G. E., Reddy, T. E., and Gersbach, C. A. (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods*, DOI: 10.1038/nmeth.2600.

(62) Daniel, R., Rubens, J. R., Sarpeshkar, R., and Lu, T. K. (2013) Synthetic analog computation in living cells. *Nature 497*, 619−623.

(63) Cheng, A. A., and Lu, T. K. (2012) Synthetic biology: an emerging engineering discipline. *Annu. Rev. Biomed. Eng. 14*, 155−178.

"Tunable and Multi-Functional Eukaryotic Transcription Factors Based on CRISPR/Cas"

Figure S1
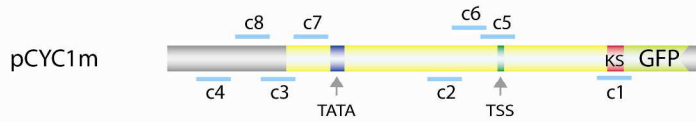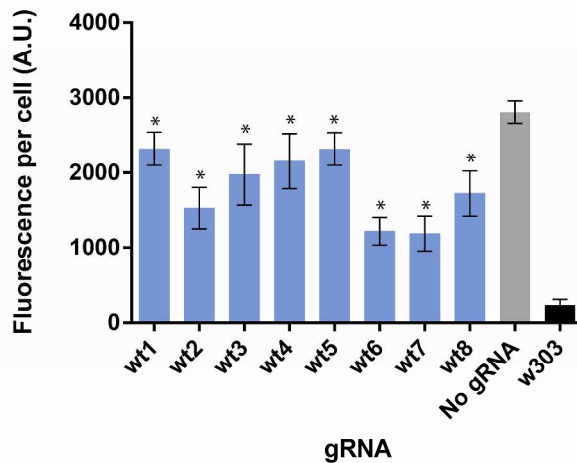
Figure S1. Effect of dCas9 alone (without an attached VP64 domain) on *gfp* expression when targeted to the pCYC1m promoter.

A) A map of gRNA target sites on pCYC1m used in this experiment.

B) Cells expressing dCas9 and containing pCYC1m driving *gfp* expression were transformed with vectors expressing different gRNAs whose target sites are shown in (A). All the tested gRNAs showed various levels of repressive or neutral effects, suggesting that dCas9 by itself cannot activate the pCYC1m promoter. The greatest repression was observed with c6 and c7 gRNAs, as was seen with dCas9_VP64 (Figure 1). Error bars indicate standard error of the mean

for three independent biological replicates. Asterisks (*) on each bar indicate statistically significant change in GFP expression relative to no gRNA control (based on the two-sided Welch's t-test, p-value < 0.05).

Figure S2

Figure S2. Regulation of *gfp* expression from pGAL1, pGAL1 lacking GAL4 operator sites (pGAL1(ΔGAL4_ops)) and minimal pGAL1 (pGAL1m) promoters by crisprTFs.
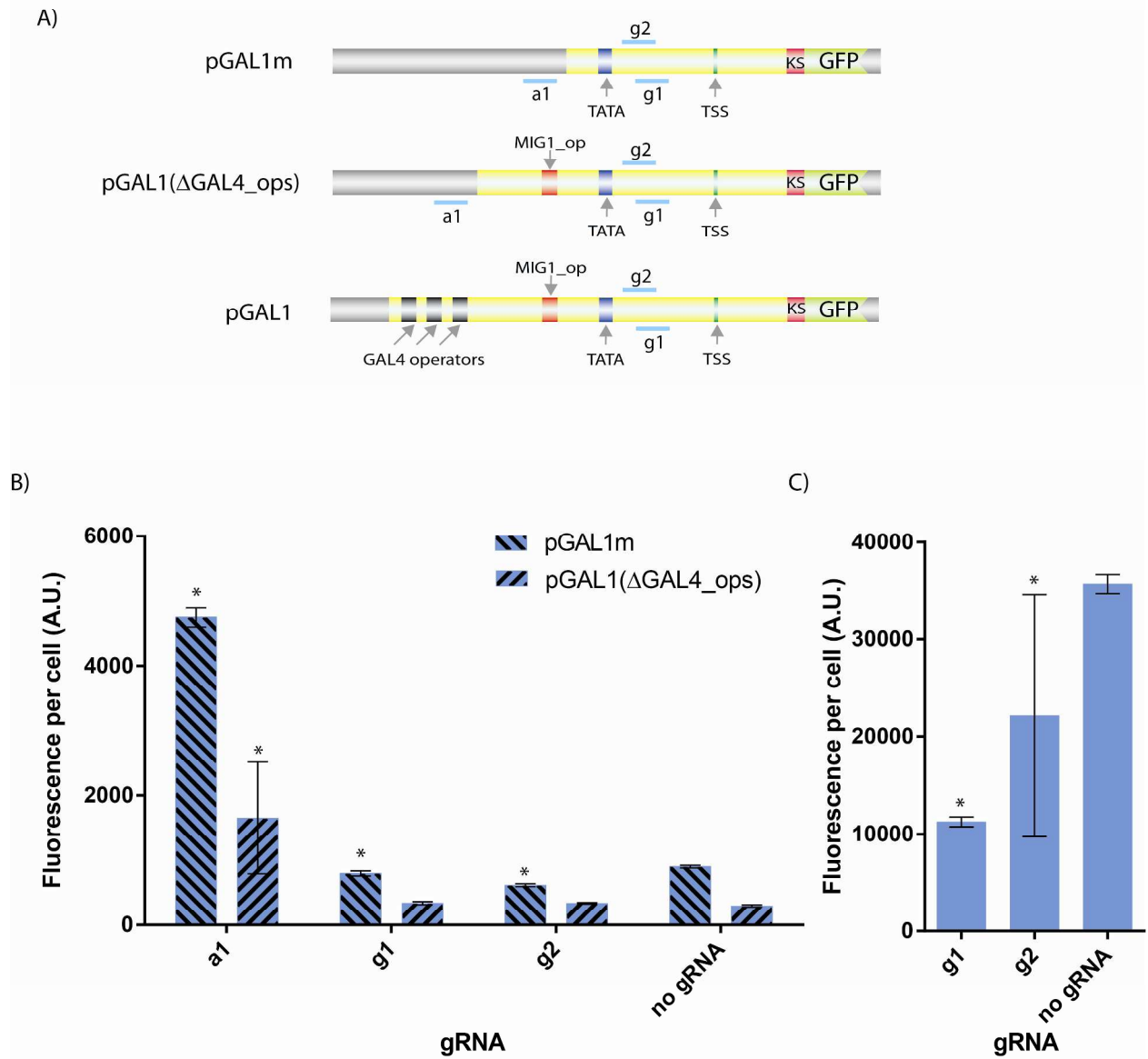
A) Maps of pGAL1m, pGAL1(ΔGAL1_ops) and pGAL1 promoters illustrating the relative positions of binding sites for endogenous pGAL1 regulatory factors (GAL4 and MIG1) along the promoters. Blue lines indicate target sites for tested gRNAs.

B) *S. cerevisiae* cells expressing crisprTFs and containing the pGAL1(ΔGAL4_ops) or pGAL1m driving *gfp* as reporter constructs (as indicated in (A)) were transformed with plasmids expressing gRNAs labeled as shown in the x-axis. Targeting crisprTFs to sequences upstream of the TATA boxes (by a1_gRNA) resulted in higher *gfp* expression than the no gRNA control. On the other hand, targeting crisprTFs to sequences downstream of the TATA boxes (by g1 and g2 gRNAs) resulted in reduced *gfp* expression relative to the no gRNA control. Error bars indicate the standard error of the mean for three independent biological replicates. Asterisks (*) on each bar indicate statistically significant changes in GFP expression relative to the no gRNA control (based on the two-sided Welch's t-test, p-value < 0.05).

C) Targeting crisprTF to sequences downstream of TATA box in pGAL1 (with g1 and g2 gRNAs) result in the repression of *gfp* expression from pGAL1. Error bars indicate the standard error of the mean for three independent biological replicates. Asterisks (*) on each bar indicate statistically significant changes in GFP expression relative to the no gRNA control (based on the two-sided Welch's t-test, p-value < 0.05).

Table S1: Sequences of the constructs used in *S. cerevisiae* experiments

Name	Sequence
NLS_dCas9_VP64	ATGTCTAGAGCCCCAAAGAAGAAGAGAAAAGTTAGACCCGGGGGATAAGAAATAC TCTATTGGTTTGGCTATCGGTACAAACTCTGTTGGTTGGGCTGTTATTACTGATG AATACAAGGTTCCATCCAAGAAGTTCAAGGTTTTGGGTAACACTGATAGACACTC CATCAAAAAGAACTTGATTGGTGCCTTGTTGTTCGATTCTGGTGAAACTGCTGAA GCTACTAGATTGAAAAGAACCGCTAGAAGAAGATACACCAGAAGAAAGAACAGA ATCTGCTACTTGCAAGAAATCTTCTCCAACGAAATGGCCAAGGTTGATGATTCAT TCTTCCACAGATTGGAAGAATCCTTCTTGGTCGAAGAAGATAAGAAGCACGAAAG ACATCCAATCTTCGGTAACATCGTTGATGAAGTTGCTTACCACGAAAAGTACCCA ACTATCTACCATTTGAGAAAGAAGTTGGTTGACTCTACCGATAAGGCTGATTTGA GATTGATCTATTTGGCTTTGGCCCACATGATTAAGTTCAGAGGTCATTTCTTGATC GAAGGTGATTTGAACCCAGATAACTCCGATGTTGATAAGTTGTTCATCCAATTAG TCCAAACCTACAATCAATTATTCGAAGAAACCCAATCAACGCCTCTGGTGTTGA TGCTAAAGCTATTTTGTCTGCCAGATTGTCCAAGTCCAGAAGATTAGAAAATTTGA TCGCCCAATTACCAGGTGAAAAGAAGAATGGTTTGTTCGGTAATTTGATTGCCTT GTCTTTGGGTTTGACTCCAAACTTCAAGTCCAATTTCGATTTGGCTGAAGATGCC AAGTTGCAATTATCTAAGGATACCTACGATGACGATTTGGATAACTTGTTGGCTC AAATCGGTGATCAATACGCTGATTTGTTTTTGGCTGCTAAGAACTTGTCCGATGC CATTTTGTTGTCCGATATTTTGAGAGTCAACACCGAAATTACTAAGGCTCCATTGT CTGCCTCTATGATCAAAAGATACGATGAACACCACCAAGACTTGACTTTGTTGAA GGCTTTGGTCAGACAACAATTACCTGAAAAGTACAAAGAAATTTTCTTCGATCAAT CCAAGAACGGTTACGCCGGTTATATTGATGGTGGTGCTTCTCAAGAAGAATTTTA CAAGTTCATCAAGCCAATCTTGGAAAAGATGGACGGTACTGAAGAATTATTGGTC AAGTTGAACAGAGAAGATTTGTTGAGAAAGCAAAGAACCTTCGACAACGGTTCTA TTCCACATCAAATTCACTTGGGTGAATTGCACGCAATTTTGAGAAGACAAGAAGA TTTTTATCCATTCTTGAAGGACAACAGAGAAAAGATCGAAAAGATTCTGACCTTCA GAATCCCTTACTACGTTGGTCCATTGGCTAGAGGTAATTCAAGATTTGCCTGGAT GACTAGAAAGTCCGAAGAAACTATTACTCCTTGGAACTTCGAAGAAGTTGTAGAT AAGGGTGCTTCTGCCCAATCCTTTATTGAAAGAATGACCAACTTCGACAAGAACT TGCCAAACGAAAAGGTTTTGCCAAAGCACTCTTTGTTGTACGAATACTTCACCGT CTACAACGAATTGACTAAGGTTAAGTACGTCACCGAAGGTATGAGAAAACCAGCT TTTTTATCCGGTGAACAAAAGAAGGCTATCGTCGATTTGTTGTTCAAGACCAACA GAAAGGTTACTGTCAAGCAATTAAAAGAAGATTACTTCAAGAAAATCGAATGCTTC GACTCCGTTGAAATTTCTGGTGTCGAAGATAGATTCAATGCCTCTTTAGGTACTTA CCATGACTTGTTGAAAATCATCAAGGACAAGGATTTCTTGGACAACGAAGAAAAC GAAGATATTTTGGAAGATATTGTCTTGACATTGACCTTGTTTGAAGATAGAGAAAT GATTGAAGAAAGATTGAAAACCTACGCCCACTTGTTCGATGATAAGGTTATGAAG CAATTAAAGAGAAGAAGATACACTGGTTGGGGGTAGATTGTCCAGAAAATTGATTA ACGGTATCAGAGACAAGCAATCCGGTAAGACCATTTTGGACTTTTTGAAGTCTGA TGGTTTCGCTAACAGAAACTTCATGCAATTAATCCACGACGATTCCTTGACTTTCA AAGAAGATATACAAAAGGCCCAAGTCTCTGGTCAAGGTGATTCTTTACATGAACA TATCGCTAACTTGGCTGGTTCTCCAGCTATTAAGAAGGGTATTTTACAAACCGTTA AGGTCGTTGACGAATTGGTCAAAGTTATGGGTAGACATAAGCCAGAAAACATCGT TATCGAAATGGCTAGAGAAAATCAAACCACCCAAAAGGGTCAAAAGAACTCCAGA GAAAGAATGAAGAGAATCGAAGAAGGTATCAAAGAATTGGGTTCCCAAATTTTGA

	AAGAACACCCAGTTGAAAACACCCAATTACAAAACGAAAAGTTGTACTTGTACTA CTTGCAAAACGGTAGAGATATGTACGTTGACCAAGAATTGGACATCAACAGATTG TCTGATTACGATGTTGACGCTATCGTTCCACAATCTTTTTTGAAGGATGACTCCAT TGACAACAAGGTCTTGACTAGATCCGATAAGAATAGAGGTAAGTCCGATAACGTT CCATCTGAAGAAGTCGTTAAGAAAATGAAGAACTATTGGAGACAATTATTGAACG CCAAGTTGATCACCCAAAGAAAGTTTGACAATTTGACCAAGGCTGAAAGAGGTG GTTTGTCTGAATTGGATAAGGCAGGTTTCATCAAAAGACAATTAGTAGAAACCAG ACAAATCACCAAGCACGTTGCTCAAATTTTGGATAGTAGAATGAACACTAAGTAC GACGAAAACGACAAATTGATCAGAGAAGTTAAGGTCATTACCTTGAAGTCCAAGT TGGTTTCCGATTTCAGAAAGGACTTCCAATTCTACAAGGTCAGAGAAATCAACAA CTACCATCATGCACATGATGCTTACTTGAATGCTGTTGTTGGTACTGCCTTGATTA AGAAGTATCCAAAGTTGGAATCCGAATTTGTCTACGGTGATTACAAGGTTTACGA CGTTAGAAAGATGATCGCCAAGTCCGAACAAGAATTGGTAAAGCTACTGCCAAA TACTTCTTCTACTCCAATATTATGAATTTCTTTAAGACCGAAATCACTTTGGCCAA CGGTGAAATTAGAAAAGACCATTGATTGAAACTAATGGTGAAACAGGTGAAATC GTTTGGGATAAGGGTAGAGATTTTGCCACTGTTAGAAAGGTATTGTCCATGCCAC AAGTAAACATCGTCAAAAAGACCGAAGTTCAAACTGGTGGTTTCTCCAAAGAATC CATTTTGCCTAAGAGAAACTCCGATAAGTTGATCGCTAGAAAAAAAGACTGGGAC CCAAAAAAGTACGGTGGTTTTGATTCTCCAACTGTTGCTTACTCTGTTTTGGTTGT TGCTAAGGTCGAAAAGGGTAAGAGTAAGAAGTTGAAGTCCGTCAAAGAATTATTA GGTATCACTATCATGGAAAGATCCTCATTCGAAAAGAATCCTATCGACTTTTTGGA AGCCAAGGGTTACAAAGAAGTCAAGAAGGACTTGATCATTAAGTTGCCAAAGTAC AGTTTGTTCGAATTGGAAAATGGTAGAAAGAGAATGTTGGCTTCTGCCGGTGAAT TACAAAAGGGTAATGAATTGGCTTTGCCATCCAAGTACGTTAATTTCTTATACTTG GCCTCCCACTACGAAAAATTGAAAGGTTCTCCTGAAGATAACGAACAAAGCAAT TATTTGTCGAACAACACAAGCACTACTTGGACGAAATCATTGAACAAATTTCCGAA TTTTCCAAAAGAGTCATTTTGGCTGACGCCAATTTGGACAAAGTTTTGTCAGCTTA CAACAAGCACAGAGATAAGCCAATTAGAGAACAAGCTGAAAACATCATTCACTTG TTCACTTTGACTAACTTGGGTGCTCCAGCTGCTTTTAAGTATTTCGATACCACTAT CGACAGAAAGAGATACACCTCTACCAAAGAAGTTTTGGACGCTACTTTGATCCAC CAATCTATTACTGGTTTGTACGAAACTAGAATCGACTTGTCTCAATTAGGTGGTGA TGGTTCTGGTAGATCTGGAGTCGACGGTGGAGGTTCTGACGCTTTGGACGACTT CGACTTGGATATGCTGGGTTCTGATGCGCTAGATGACTTTGACCTCGACATGCTT GGAAGTGACGCCTTAGATGATTTTGACCTGGATATGCTTGGATCAGACGCTCTG GACGATTTCGACTTAGACATGCTTTCCTAG
pCYC1m	AAGCTTGATATCGAATTCCTGCAGCCCGGGTACTGTATGTACATACAGTAGGATC CTATGGCATGCATGTGCTCTGTATGTATATAAAACTCTTGTTTTCTTCTTTTCTCTA AATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAATTACTATACTTCTATAG ACACACAAACACAAATACACACTAATCTAGATATTAAAATGTCTAAAGGTGAAG
pCYC1m(modifi ed)	AAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCAGGGGCATGCTGCCGTC CGTGATGATATGGCATGGATGTCCTCTGTATGTATATAAAACTCTTGGTTTCTTCT TTTCTCTAAATATTCTTTCCTTATACTTTAGGACCTTTGCAGCATAAATTACTATAC TTCTATGGCCACACAAACACAAATACACACTAATCTAGATATTAAAATGTCTAA AGGTGAAG
0xa1_pCYC1m	AAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCTATGGCATGCATGTGCTC TGTATGTATATAAAACTCTTGTTTTCTTCTTTTCTCTAAATATTCTTTCCTTATACAT

	TAGGACCTTTGCAGCATAAATTACTATACTTCTATAGACACACAAACACAAATACA CACACTAATCTAGATAAAAAATGGTTTCCAAG
1xa1_pCYC1m	AAGCTTGATATCGAATTCCTGCAGCCCGGGTACTGTATGTACATACAGTAGGATC CTATGGCATGCATGTGCTCTGTATGTATATAAAACTCTTGTTTTCTTCTTTTCTCTA AATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAATTACTATACTTCTATAG ACACACAAACACAAATACACACACTAATCTAGATAAAAAATGGTTTCCAAG
3xa1_pCYC1m	AAGCTTGATATCGAATTCCTGCAGCCCGGGTACTGTATGTACATACAGTAGCTCG AAAATATTAATTACTGTATGTACATACAGTAGATCCTCGAAAATATTAATTACTGTA TGTACATACAGTACTATGGCATGCATGTGCTCTGTATGTATATAAAACTCTTGTTT TCTTCTTTTCTCTAAATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAATTA CTATACTTCTATAGACACACAAACACAAATACACACACTAATCTAGATAAAAAATG GTTTCCAAGGGTGAAG
6xa1_pCYC1	AAGCTTGATATCGAATTCCTGCAGCCCGGGTACTGTATGTACATACAGTAGGATC CTCGAAAATATTAATTACTGTATGTACATACAGTAGGATCCTCGAAAATATTAATT ACTGTATGTACATACAGTAGGATCCTACTGTATGTACATACAGTAATTAATATTTT CGAGGATCCTACTGTATGTACATACAGTAATTAATATTTTCGAGGATCCTACTGTA TGTACATACAGTAATTAATATTTTCGAGGATCCTATGGCATGCATGTGCTCTGTAT GTATATAAAACTCTTGTTTTCTTCTTTTCTCTAAATATTCTTTCCTTATACATTAGGA CCTTTGCAGCATAAATTACTATACTTCTATAGACACACAAACACAAATACACACAC TAATCTAGATAAAAAATGGTTTCCAAG
12xa1_pCYC1	AAGCTTGATATCGAATTCCTGCAGCCCGGGTACTGTATGTACATACAGTAGGATC CTCGAAAATATTAATTACTGTATGTACATACAGTAGGATCCTCGAAAATATTAATT ACTGTATGTACATACAGTAGGATCCTCGAAAATATTAATTACTGTATGTACATACA GTAGGATCCTCGAAAATATTAATTACTGTATGTACATACAGTAGGATCCTCGAAAA TATTAATTACTGTATGTACATACAGTAGGATCCTACTGTATGTACATACAGTAATT AATATTTTCGAGGATCCTACTGTATGTACATACAGTAATTAATATTTTCGAGGATC CTACTGTATGTACATACAGTAATTAATATTTTCGAGGATCCTACTGTATGTACATA CAGTAATTAATATTTTCGAGGATCCTACTGTATGTACATACAGTAATTAATATTTTC GAGGATCCTACTGTATGTACATACAGTAATTAATATTTTCGAGGATCCTATGGCAT GCATGTGCTCTGTATGTATATAAAACTCTTGTTTTCTTCTTTTCTCTAAATATTCTT TCCTTATACATTAGGACCTTTGCAGCATAAATTACTATACTTCTATAGACACACAA ACACAAATACACACACTAATCTAGATAAAAAATGGTTTCCAAG
pRPR_TetO	GGGGGATCTGCCAATTGAACATAACATGGTAGTTACATATACTAGTAATATGGTT CGGCACACATTAAAAGTATAAAAACTATCTGAATTACGAATTACATATATTGGTCA TAAAAATCAATCAATCATCGTGTGTTTTATATGTCTCTTATCTAAGTATAAGAATAT CCATAGTTAATATTCACTTACGCTACCTTTTAACCTGTAATCATTGTCAACAGGAT ATGTTAACGACCCACATTGATAAACGCTAGTATTTCTTTTTCCTCTTCTTATTGGC CGGCTGTCTCTATACTCCCCTATAGTCTGTTTCTTTTCGTTTCGATTGTCCCTATC AGTGATAGAGATGGCGCACATGGTACGCTGTGGTGCTCGCGGCTGGGAACGAA ACTCTGGGAGCTGCGATTGGCAG
gRNA expression cassette (SacI_pRPR1	GAGCTCGGGGGATCTGCCAATTGAACATAACATGGTAGTTACATATACTAGTAAT ATGGTTCGGCACACATTAAAAGTATAAAAACTATCTGAATTACGAATTACATATAT TGGTCATAAAAATCAATCAATCATCGTGTGTTTTATATGTCTCTTATCTAAGTATAA GAATATCCATAGTTAATATTCACTTACGCTACCTTTTAACCTGTAATCATTGTCAAC AGGATATGTTAACGACCCACATTGATAAACGCTAGTATTTCTTTTTCCTCTTCTTA

_HindIII_SDS _gRNA backbone_XhoI _RPR1t_KpnI)	TTGGCCGGCTGTCTCTATACTCCCCTATAGTCTGTTTCTTTTCGTTTCGATTGTTT TACGTTTGAGGCCTCGTGGCGCACATGGTACGCTGTGGTGCTCGCGGCTGGGA ACGAAACTCTGGGAGCTGCGATTGGCAGAAGCTTNNNNNNNNNNNNNNNNNNN NGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA AAAGTGGCACCGAGTCGGTGCTTTTTTCTCGAGCCATATCCAACTTCCAATTTAA TCTTTCTTTTTTAATTTTCACTTATTTGCGATACAGAAAGAAAAAAGCGATAGTAAC TATTGAATTTTGTTTGGATTTGGTTAGATTAGATATGGTTTCTCTTTATATTTACAT GCTAAAAATGGGCTACACCAGAGATACATAATTAGATATATATACGCCAGTACAC CTTATCGGCCCAAGCCTTGTCCCAAGGCAGCGTTTTGTTCTTGGAAACGCTGCC CTACACGTTCGCTATGCTTCAAGAACTTTTCTGAGCACTTCATGATGCATGTTTGT TCCTTATTGGTTAGCTTTGATGTTGTGAAGTCATTGACACAGTCTGTGAAACATCT TTCTACCAGATTAGAGTACAAACGCATGAAATCCTTCATTTGCTTTTGTTCCACTA CTTTTTGGAACTCTTGTTGTTCTTTGGTACC
	Where NNNNNNNNNNNNNNNNNNNN is one of the following:
c1 gRNA	CTAGATATTAAAATGTCTAA
c2 gRNA	ATATTCTTTCCTTATACATT
c3 gRNA	GTACATACAGTAGGATCCTA
c4 gRNA	ATATCGAATTCCTGCAGCCC
c5 gRNA	TATAGTAATTTATGCTGCAA
c6 gRNA	TGCAAAGGTCCTAATGTATA
c7 gRNA	ACAGAGCACATGCATGCCAT
c8 gRNA	TGTATGTACATACAGTACCC
cm1 gRNA	ATATTCTTTCCTTATACTTT
cm2 gRNA	CTGTATGTATATAAAACTCT
cm3 gRNA	CTGCAGCCCGGGGGATCCAG
cm4 gRNA	TGCAAAGGTCCTAAAGTATA
cm5 gRNA	AGAGTTTTATATACATACAG
cm6 gRNA	CCATGCCATATCATCACGGA

a1 gRNA	TTACTGTATGTACATACAGT
g1 gRNA	ACTAATACTTTCAACATTTT
g2 gRNA	GTTGAAAGTATTAGTTAAAG

Table S2: Sequences of the constructs used in HEK293T experiments

Name	Sequence
dCas9-3xNLS-VP64	GCCACCATGGACAAGAAGTACTCCATTGGGCTCGCCATCGGCACAAACAGCGTC GGCTGGGCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTT CTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTG TTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCG CAGATATACCCGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAAT GAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAGTCCTTTTTG GTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGGCAATATCGTGGAC GAGGTGGCGTACCATGAAAAGTACCCAACCATATATCATCTGAGGAAGAAGCTT GTAGACAGTACTGATAAGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCAT ATGATCAAATTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAAC AGCGATGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTTTTCG AAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCCTGAGCGCTA GGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGGGAG AAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACCCCC AACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAG ACACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGACCAGTACG CAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTGAGTGATAT TCTGCGAGTGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAA GCGCTATGATGAGCACCACCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACA GCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTAC GCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAG CCCATCTTGGAAAAAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAGA GAAGATCTGTTGCGCAAACAGCGCACTTTCGACAATGGAAGCATCCCCCACCAG ATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCC TTTTTGAAAGATAACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTA CTATGTAGGCCCCCTCGCCCGGGGGAAATTCCAGATTCGCGTGGATGACTCGCAA ATCAGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGTCGTGGATAAGGGGGC CTCTGCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAAC GAAAAGGTGCTTCCTAAACACTCTCTGCTGTACGAGTACTTCACAGTTTATAACG AGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTCCTGT CTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAAG TTACCGTGAAACAGCTCAAAGAAGACTATTTCAAAAAGATTGAATGTTTCGACTCT GTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCAC GATCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAG GACATTCTTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGA TTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTCATGAAACA GCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCA ATGGGATCCGAGACAAGCAGAGTGGAAAGACAATCCTGGATTTTCTTAAGTCCG ATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCATGATGACTCTCTCACCTT TAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCTTCACGA GCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGGGAATACTGCAGAC CGTTAAGGTCGTGGATGAACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAGAA TATCGTTATCGAGATGGCCCGAGAGAACCAAACTACCCAGAAGGGACAGAAGAA CAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCA

	AATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTAC CTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATC AATCGGCTCTCCGACTACGACGTGGATGCCATCGTGCCCCAGTCTTTTCTCAAA GATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAAATAGAGGGAAGA GTGATAACGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCA GCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAGGC TGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCT TGTTGAGACACGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTCACGCAT GAACACCAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAGTTATTACT CTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTGA GAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAG GCACTGCACTTATCAAAAAATATCCCAAGCTTGAATCTGAATTTGTTTACGGAGAC TATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAGCAGGAAATAGGCA AGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAG ATTACACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGA GAAACAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTCCGGAA GGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGACCGG AGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAGCTGATCGC ACGCAAAAAGATTGGGACCCCAAGAAATACGGCGGATTCGATTCTCCTACAGT CGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAAAACT CAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGA AAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGAC CTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCCGGAAAC GAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCC TCTAAATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGT CTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAACACAAACACTACC TTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCGA CGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATC AGGGAGCAGGCAGAAACATTATCCACTTGTTTACTCTGACCAACTTGGGCGCG CCTGCAGCCTTCAAGTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCT ACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTAT GAAACAAGAATCGACCTCTCTCAGCTCGGTGGAGACAGCAGGGCTGACGGGCC CTCACTGGGTTCAGGGTCACCCAAGAAGAAACGCAAAGTCGAGGATCCAAAGAA GAAAAGGAAGGTTGAAGACCCCAAGAAAAGAGGAAGGTGGATGGGATCGGCT CAGGCAGCAACGGCGGTGGAGGTTCAGACGCTTTGGACGATTTCGATCTCGATA TGCTCGGTTCTGACGCCCTGGATGATTTCGATCTGGATATGCTCGGCAGCGACG CTCTCGACGATTTCGACCTCGACATGCTCGGGTCAGATGCCTTGGATGATTTTGA CCTGGATATGCTCTCATGATGA
KRAB	GATGCTAAGTCACTAACTGCCTGGTCCCGGACACTGGTGACCTTCAAGGATGTA TTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTGGACACTGCTCAGCAGATC GTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGTTATC AGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCTGGC TGGTGGAGAGAGAAATTCACCAAGAGACCCATCCTGATTCAGAGACTGCATTTG AAATCAAATCATCAGTTTCCAGCAGGAGCATTTTTAAAGATAAGCAATCCTGTGAC ATTAAAATGGAAGGAATGGCAAGGAATGATCTCTGG
pPGK1	AATTCTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGCTT TAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGCACA

	CATTCCACATCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTC GCGCCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCGCCCCGCAGCTC GCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTCTCACTAGTCTCGTGCAG ATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGGCCA ATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTC CGGGGGCGGGCTCAGGGGCGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTC CTCCGGAGGCCCGGCATTCTGCACGCTTCAAAAGCGCACGTCTGCCGCGCTGT TCTCCTCTTCCTCATCTCCGGGCCTTTCGACCTGCATCCCAAGCTTGCCACCATG
pMLPm	GGGGGGCTATAAAGGGGGTGGGGGCGTTCGTCCTCACTCTAGATCTGCGATC TAAGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCATGGCGCCACCATG
0xa1_pMLPm	ATCGATAGCTAGCGGGGGGGCTATAAAGGGGGTGGGGGCGTTCGTCCTCACTC TAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCAT GGCGCCACCATG
1xa1_pMLPm	ATCGATAGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGTAGGCTAGCG GGGGGCTATAAAGGGGGTGGGGGCGTTCGTCCTCACTCTAGATCTGCGATCTA AGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCATGGCGCCACCATG
2xa1_pMLPm	ATCGATAGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGTAGGCTAGCCT GAAAATATTAATTACTGTATGTACATACAGTAGGCTAGCGGGGGGGCTATAAAGG GGGTGGGGGCGTTCGTCCTCACTCTAGATCTGCGATCTAAGTAAGCTTGGCATT CCGGTACTGTTGGTAAAGCCACCATGGCGCCACCATG
3xa1_pMLPm	ATCGATAGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGTAGGCTAGCCT GAAAATATTAATTACTGTATGTACATACAGTAGGCTAGCCTGAAAATATTAATTAC TGTATGTACATACAGTAGGCTAGCGGGGGGGCTATAAAGGGGGTGGGGGCGTT CGTCCTCACTCTAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGT AAAGCCACCATGGCGCCACCATG
pU6_gRNA	TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCA AGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACG ATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATAT TAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTA AAATTATGTTTTAAAATGGACTATCATTGCTTACCGTAACTTGAAAGTATTTCGATT TCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG<mark>NNNNNNNNNNNNNNNN NNNN</mark>GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACT TGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT
	where <mark>NNNNNNNNNNNNNNNNNNNN</mark> is one of the following:
m1 gRNA	CTGTCCCCAGTGCAAGTGC
m2 gRNA	TTCTCTATCGATAGCTAGC
m3 gRNA	TCTGCGATCTAAGTAAGCT
m4 gRNA	ACTGTTGGTAAAGCCACCA

m5 gRNA	ATGGCGCCACCATGAGCAG
m6 gRNA	GGCACCTGCACTTGCACTG
m7 gRNA	TATCGATAGAGAAATGTTC
m8 gRNA	AGATCGCAGATCTAGAGTG
m9 gRNA	GTGGCTTTACCAACAGTAC
m10 gRNA	GCTCATGGTGGCGCCATGG
CCAAT gRNA	GAAGGAGCAAAGCTGCTAT
GC-box gRNA	GGCGGGCTCAGGGGCGGGG
a1 gRNA	TACTGTATGTACATACAGT
a2 gRNA	AGTCGCGTGTAGCGAAGCAT
a3 gRNA	CAACGCGACGCTAGATAGCA