Optical control of mammalian endogenous transcription and epigenetic states

Optical control of mammalian endogenous transcription and epigenetic states

A dissertation presented

by

Mark Daniel Brigham

to

The School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University

Cambridge, Massachusetts

March 2014

Dissertation Advisor: Professor Feng Zhang Mark Daniel Brigham

Optical control of mammalian endogenous transcription and epigenetic states

Abstract

The dynamic nature of gene expression enables cellular programming, homeostasis and environmental adaptation in living systems. Dissection of causal gene functions in cellular and organismal processes therefore necessitates approaches that enable spatially and temporally precise modulation of gene expression. Recently, a variety of microbial and plant-derived light-sensitive proteins have been engineered as optogenetic actuators, enabling high-precision spatiotemporal control of many cellular functions[1-11]. However, versatile and robust technologies that enable optical modulation of transcription in the mammalian endogenous genome remain elusive. Here we describe the development of light-inducible transcriptional effectors (LITEs), an optogenetic two-hybrid system integrating the customizable TALE DNA-binding domain[12-14] with the light-sensitive cryptochrome 2 protein and its interacting partner CIB1 from *Arabidopsis thaliana*. LITEs do not require additional exogenous chemical cofactors, are easily customized to target many endogenous genomic loci, and can be activated within minutes with reversibility[6, 15]. LITEs can be packaged into viral vectors and genetically targeted to probe specific cell populations. We have applied this system in primary mouse neurons, as well as in the brain of freely behaving mice *in vivo* to mediate reversible modulation of mammalian endogenous gene expression as well as targeted epigenetic chromatin modifications. We explore the modularity of the LITE approach through the development

of CRISPR/Cas9 transcriptional effectors in either constitutively active or light-inducible contexts. The LITE system establishes a novel mode of optogenetic control of endogenous cellular processes and enables direct testing of the causal roles of genetic and epigenetic regulation in normal biological processes and disease states.

Table of Contents

Acknowledgements

As a graduate student, I have had the great opportunity to work with many outstanding researchers. I have had the fortune to be immersed in an extraordinarily exciting field during a watershed time in its development. Navigating the rapidly evolving landscape of transcriptional engineering with these individuals of unmatched quality has been a true pleasure and unforgettable experience. They have my unending gratitude.

I wish to express my most sincere thanks to my advisor Feng Zhang for his invaluable support. Feng has worked tirelessly to provide an environment where each lab member can exercise their curiosity and pursue their scientific goals without bounds. He aided this work and my own scientific growth by continually contributing intellectual guidance, exciting new ideas, and outstanding scientific resources. Beyond the lab, Feng has never wavered from his complete commitment to the happiness and wellbeing of his students. I will always remember his faith in me as an engineer and as a person.

This work could not have been completed without the superhuman efforts of my teammate Silvana Konermann. I wish to express my gratitude to her for her perseverance through the long nights and troubleshooting of our projects. Her excellence as a collaborator is surpassed only by her qualities as a friend. I cannot thank her enough for her camaraderie, support, and understanding.

The Broad Institute, the Zhang Lab in particular, is a highly collaborative environment, of which I have been a consistent beneficiary. I therefore wish to thank all of the members of the Broad Institute, McGovern Institute for Brain Research, and the Zhang Lab from whom I have received assistance and friendship. Several individuals stand out in my mind for their notable contributions to this work. Alex Trevino aided these projects in innumerable ways, as a key member of team SAM, both at the bench and during the publication process. Le Cong acted as a profound resource in my molecular biology education and developed both the SID4X domain and chemically inducible TALE-

activator system described in this work. Matthias Heidenreich assisted and performed western blotting experiments throughout the development of the LITE system. Patrick Hsu was always willing to lend an extra hand on a variety of molecular experiments and generously provided editing of the written work. Randy Platt synthesized dCas9-effector constructs and performed the first generation of CRISPR/Cas9 transcription experiments. Fei Ann Ran assisted with illustrations on a moment's notice. Michael Yim contributed to in vivo studies and performed many critical laboratory functions. Erin Blackwell provided administrative support to me, as well as the entire lab.

I would also like to thank the members of my thesis committee for their constructive feedback and their flexibility in accommodating my graduate career progress: Professors David Mooney, David Cox, and Dan Needleman, who graciously acted as my SEAS co-advisor. Further acknowledgement goes to the wonderful members of the SEAS Student Affairs Office, including Julie Holbrook and Marie Dooley, with special thanks to William Walker for his thoughtful guidance.

I am grateful to the Friends of the McGovern Institute and the Harvard Stem Cell Institute for financial support, an unquestionably critical component of my research education.

Finally, I must express my deepest gratitude for the love and support of my family. My mom Jean has enabled, encouraged, and believed in all of my pursuits, academic or otherwise. My brother Matthew has brightened my life with humor and empathy for as long as I can remember. My dear friend Austin has uplifted me in challenging times and shared his happiness at moments of celebration. Lastly, my wife Joy has provided companionship, love, understanding, and optimism every day that I have known her.

I dedicate this work to my greatest teacher, my Dad.

List of Figures

Section I: Introduction

The dynamic nature of gene expression enables cellular programming, homeostasis, and environmental adaptation in living systems. Internal and external cellular inputs are functionally integrated and processed by spatiotemporally precise regulation within transcriptional networks. These networks in turn orchestrate the interaction and organization of cells within tissues and whole organisms. Dissecting the contributions of genes to cellular and organismal function therefore requires an approach to enable spatially and temporally controlled modulation of gene expression.

Light provides a fast, reversible, and spatially targeted modality for manipulating a variety of biological processes. Microbial- and plant-derived light-sensitive proteins have been engineered as optogenetic actuators, allowing precise optical control of cellular functions including membrane potential[1, 2, 16], intracellular biochemical signaling[17], protein interactions[3-6], and heterologous gene expression[4, 6-9, 18]. However, the ability to directly modulate endogenous gene expression using light has remained elusive.

An ideal optogenetic approach for controlling endogenous gene transcription would be easily generalizable to target any gene locus, would not require manipulation of the endogenous genomic sequence or the addition of exogenous co-factors, and would exhibit fast and reversible kinetics. The DNA-binding domain of transcription activator-like effectors (TALEs)[12, 13] from *Xanthamonas sp*. can be easily customized to bind a variety of DNA sequences in mammalian cells[14, 19, 20]. TALE DNA-binding domains are

modular and can be fused with a variety of effector domains, including nucleases, transcriptional activators, and transcriptional repressors to edit or modulate endogenous mammalian genomic loci[14, 19-21]. We sought to combine TALEs with light-sensitive plant proteins to create a suite of tools for enabling spatiotemporally precise control of endogenous gene transcription.

This work describes the development of Light-Inducible Transcriptional Effectors (LITEs), a two-component system integrating the customizable TALE DNA-binding domain with the light-sensitive cryptochrome 2 protein and its interacting partner CIB1 from *Arabidopsis thaliana*[6, 15]. LITEs can be engineered to mediate positive and negative regulation of endogenous mammalian gene expression in a reversible manner, and changes in transcript levels occur within minutes after stimulation. Like other optogenetic tools, LITEs can be packaged into viral vectors and genetically targeted to specific cell types to probe gene function within specific cell populations. We demonstrate the application of this system in primary neurons as well as in the mouse brain *in vivo*. We explore the modularity of LITEs by modifying the system to enable RNA targeted transcriptional modulation via the RNA-guided CRISPR/Cas9 system[22, 23]. Finally, we engineer new approaches to Cas9-based activation, in an effort to improve the efficiency and ease of endogenous transcriptional modulation.

Section II: Engineering light-inducible transcriptional effectors

The LITE system employs a modular design consisting of two independent components (**Fig. 1a**): The first component is the genomic anchor and includes a customizable DNA-binding domain, based on transcription activator-like effectors (TALEs)[12, 13] from *Xanthamonas sp.*, fused to the light-sensitive cryptochrome 2 (CRY2) protein from *Arabidopsis thaliana*[6, 15] (TALE-CRY2). The second component includes the CRY2 interacting partner CIB1[6, 15] fused to a desired effector domain (CIB1-effector). In the absence of light (inactive state), TALE-CRY2 binds the promoter region of the target gene while CIB1-effector remains free within the nuclear compartment. Illumination with blue light (peak ~450 nm) triggers a conformational change in CRY2 and subsequently recruits CIB1-effector (VP64 shown in **Fig. 1a**) to the target locus to mediate transcriptional modulation. This modular design allows each LITE component to be independently engineered, allowing the same genomic anchor to be combined with activating or repressing effectors[21, 24] to exert positive and negative transcriptional control over the same endogenous genomic locus. In principle, the genomic anchor may also be replaced with other DNA binding domains such as zinc finger proteins[24] or RNA-guided DNA binding domains based on nucleolytically inactive mutants of Cas9[22, 25-28] (**Section VI**).

In order to identify the most effective architecture, we assessed the efficacy of different LITE designs by measuring blue light illumination induced transcriptional changes of the neural lineage-specifying transcription factor neurogenin 2 (*Neurog2*) (**Fig. 1b**). 3 out of

4 initial LITE pairings produced significant light-induced *Neurog2* mRNA up-regulation in Neuro 2a cells ($p < 0.001$, **Fig. 1b**).

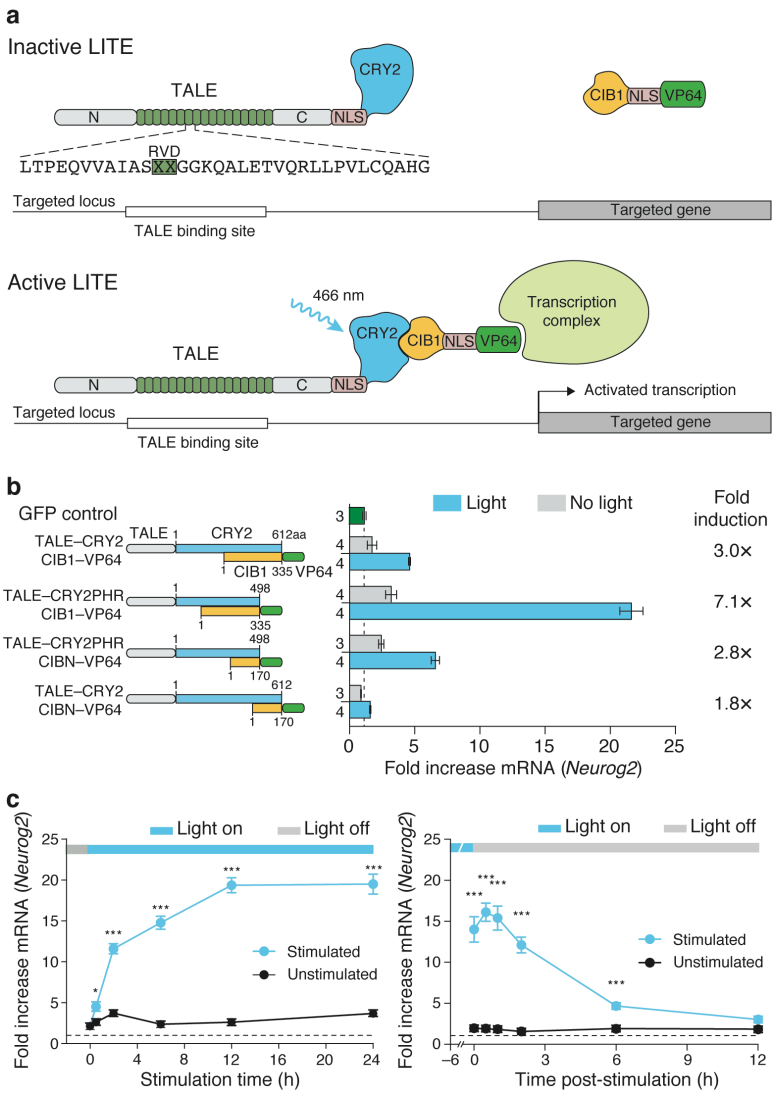

Figure 1 | Design and optimization of the LITE system. a, Schematic of the LITE system. Light stimulation induces dimerization of CRY2 and CIB1, recruiting the effector to the target promoter. **b**, LITE architecture was optimized by fusing TALE and the transcriptional activator VP64[14, 24] to different truncations of CRY2 and CIB1[6] (n next to each bar). **c**, Time-course of light-dependent *Neurog2* upregulation and decay post-illumination (n = 4 biological replicates; *$p < 0.05$; ***$p < 0.001$). Cells were stimulated with 5 mW/cm^2 light (466 nm, 1 s pulses at 0.067 Hz). Mean ± s.e.m. in all panels.

Of the combinations tested, TALE(*Neurog2*)-CRY2PHR::CIB1-VP64 yielded the strongest light-mediated transcription activation as well as the highest induction ratio (light/no light mRNA levels). Therefore TALE-CRY2PHR::CIB1-VP64 was used in subsequent experiments. To ensure optimal function, we also systematically tuned light stimulation parameters and effector domains (wavelength, **Fig. 2a**; duty cycle, **Fig. 2b**; light intensity[29], **Fig. 2c** and **d**; and choice of activation domain, **Fig. 2e**).

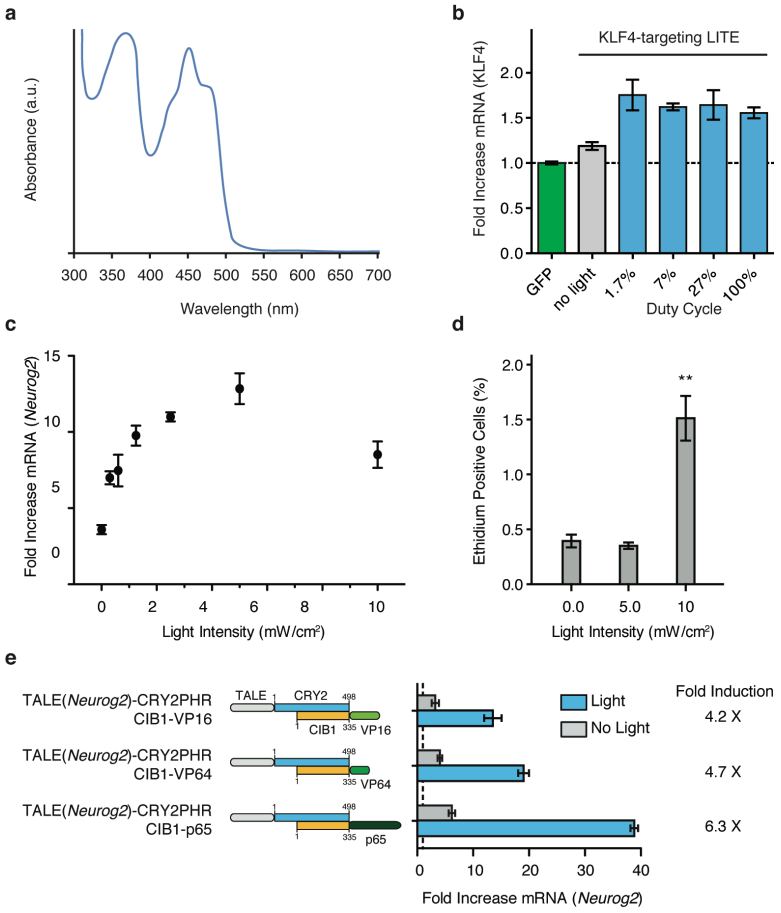

Figure 2 | Engineering of light stimulation parameters and activation domains of LITEs. a, Illustration of the absorption spectrum of CRY2 in vitro. Cryptochrome 2 was optimally activated by 350–475 nm light[29]. A sharp drop in absorption and activation was seen for wavelengths greater than 480 nm. Spectrum was adapted from Banerjee *et al*[29]. **b**, Impact of illumination duty cycle on LITE-mediated gene expression. Varying duty cycles (illumination as percentage of total time) were used to stimulate 293FT cells expressing LITEs targeting the KLF4 gene. KLF4 expression levels were compared to cells expressing GFP only. Stimulation parameters were:

Figure 2 (Continued)

466 nm, 5 mW/cm^2 for 24 h. Pulses were performed at 0.067 Hz with the following durations: 1.7% = 0.25 s pulse, 7% = 1 s pulse, 27% = 4 s pulse, 100% = constant illumination. (mean ± s.e.m.; n = 3–4 biological replicates.) **c**, The transcriptional activity of CRY2PHR/ CIB1 LITE was found to vary according to the intensity of 466 nm blue light. Neuro 2a cells were stimulated for 12 h at a 7% duty cycle (1 s pulses at 0.067 Hz). All *Neurog2* mRNA levels were measured relative to cells expressing GFP only (mean ± s.e.m.; n = 3–4 biological replicates). **d**, Light-induced toxicity measured as the percentage of cells positive for red fluorescent ethidium homodimer-1 versus calcein-positive cells (mean ± s.e.m.; n = 3 biological replicates; **$p < 0.01$). **e**, We compared the activation domains VP16 and p65 in addition to VP64 to test the modularity of the LITE CIB1–effector component. Neuro 2a cells transfected with LITE were stimulated for 24 h with 466 nm light at an intensity of 5mW/cm^2 and a duty cycle of 7% (1 s pulses at 0.067 Hz). All three domains produced a significant light-dependent *Neurog2* mRNA upregulation ($p < 0.001$). We selected VP64 for subsequent experiments due to its lower basal activity in the absence of light-stimulation (mean ± s.e.m.; n = 3–4 biological replicates).

Although the interaction between CRY2 and CIB1 occurs on a sub-second timescale[6], LITE-mediated transcriptional activation is likely dependent on many factors, including rate of transcription, mRNA processing, and transcript stability[30, 31]. We found that LITE-mediated *Neurog2* expression increased considerably as early as 30 min after initial stimulation and rose steadily until saturating at 12 h with approximately 20-fold up-regulation compared to GFP-transfected negative controls (**Fig. 1c**). Interestingly, *Neurog2* transcript levels continued to increase for up to 30 min post-illumination, an effect that may have resulted from residual CRY2PHR-CIB1 dimerization or from previously recruited RNA polymerases. Thereafter, *Neurog2* mRNA returned to baseline levels with a half-life of ~3 h. In contrast, a small-molecule inducible TALE system based on the plant hormone abcisic acid receptor[32] exhibited slower activation kinetics (**Fig. 3**).

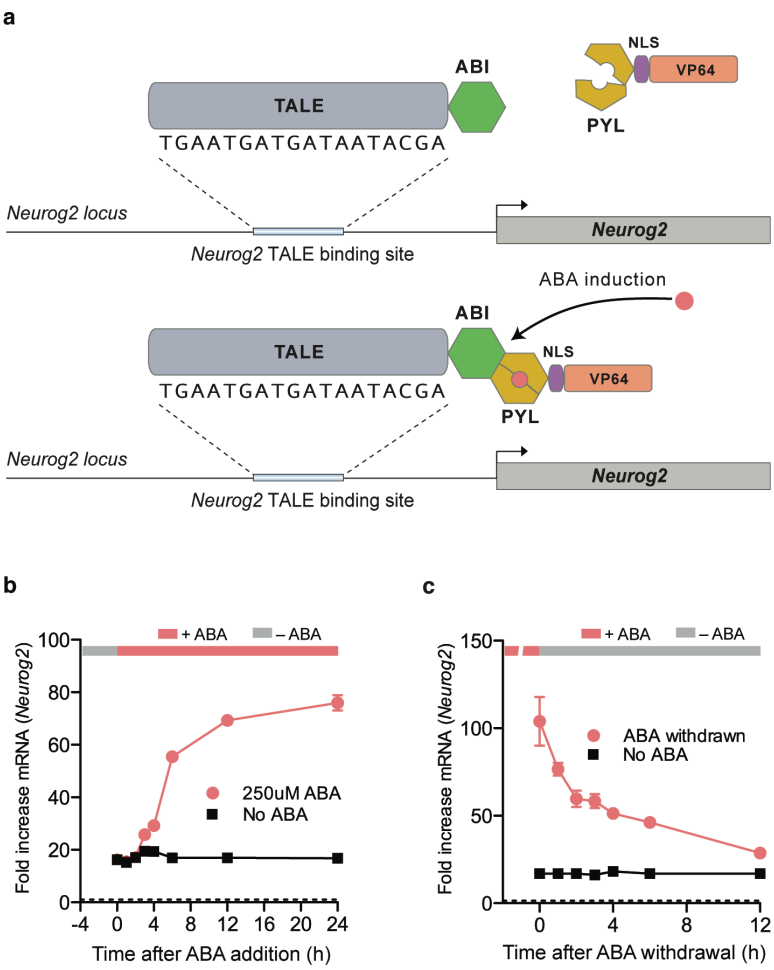

Figure 3 | Chemical induction of endogenous gene transcription. a, Schematic showing the design of a chemically inducible two-hybrid TALE system based on the abscisic acid (ABA) receptor system. ABI and PYL dimerize upon the addition of ABA and dissociate when ABA is withdrawn. **b**, Time-course of ABA-dependent *Neurog2* upregulation. 250 µM of ABA was added to Neuro 2a cells expressing TALE(*Neurog2*)–ABI and PYL–VP64. Fold mRNA increase was measured at the indicated time points after the addition of ABA. **c**, Decrease of *Neurog2* mRNA levels after 24 h of ABA stimulation. All *Neurog2* mRNA levels were measured relative to GFP-expressing control cells (mean ± s.e.m.; n = 3–4 biological replicates).

Section III: TALEs and LITEs in the mammalian brain

To apply LITE for neuronal applications, we developed a simplified process for adeno-associated virus (AAV) production (**Fig. 4**) and an AAV-based vector (**Fig. 5a** and **b**) for the delivery of TALE genes. The ssDNA-based genome of AAV is less susceptible to recombination, providing an advantage over lentiviral vectors[33]. We evaluated a panel of 28 TALE activators targeting the mouse genome in primary neurons and found that most were able to up-regulate transcription in primary neurons (**Fig. 5c**). Moreover, *in vivo* expression of TALE(*Grm2*)-VP64 in the prefrontal cortex (PFC) (**Fig. 5d** and **e**) induced a 2.5-fold increase in *Grm2* mRNA levels compared to GFP-only controls (**Fig. 5f**).

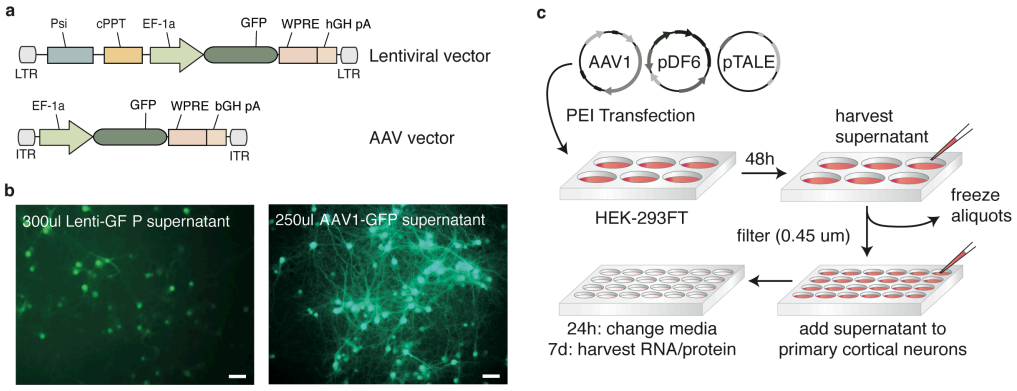

Figure 4 | Efficient AAV production using cell supernatant. a, Lentiviral and AAV vectors carrying GFP were used to test transduction efficiency. **b**, Primary cortical neurons were transduced with supernatant derived from the same number of lentivirus- or AAV-transduced 293FT cells. Representative images of GFP expression were collected at 7 days post infection. Scale bars, 50 µm. **c**, The depicted process was developed for the production of AAV supernatant and subsequent transduction of primary neurons. 293FT cells were transfected with an AAV vector carrying the gene of interest, the AAV1 serotype packaging vector (pAAV1), and helper plasmid (pDF6) using PEI. 48 h later, the supernatant was collected and filtered through a 0.45 µm PVDF membrane. Primary neurons were then transduced with supernatant and remaining aliquots were stored at −80 °C. Stable levels of AAV construct expression were

Figure 4 (Continued)

reached after 5–6 days. AAV supernatant production following this process can be used for production of up to 96 different viral constructs in 96-well format.

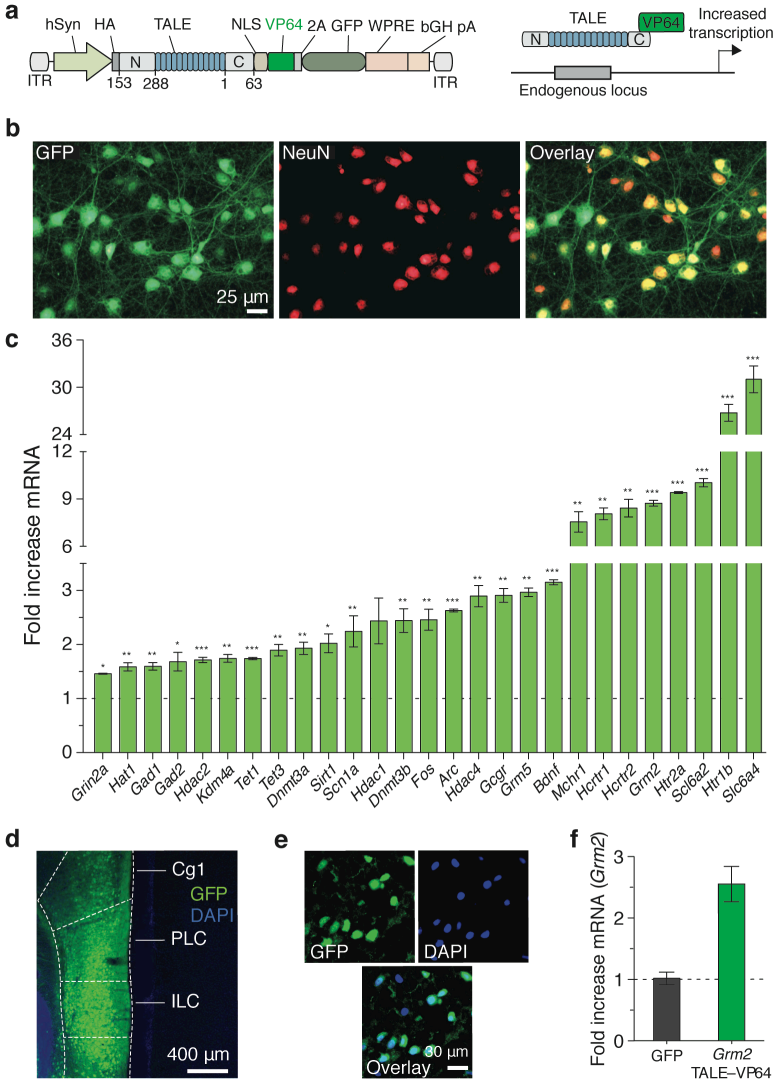

Figure 5 | *In vitro* and *in vivo* AAV-mediated TALE delivery targeting endogenous loci in neurons. a, Schematic of AAV vectors for TALE delivery. **b**, Representative images of primary cortical neurons expressing TALE–VP64. **c**, TALE–VP64 constructs targeting a variety of endogenous genes were screened for transcriptional activation in primary cortical neurons (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; $n = 3$ biological replicates). **d**, TALE–VP64 expression in PFC. DAPI, 4′,6-diamidino-2-phenylindole; Cg1, cingulate cortex area 1; PLC, prelimbic cortex; ILC, infralimbic cortex. **e**, Higher magnification image of TALE–VP64 expressing neurons in

Figure 5 (Continued)

PFC. **f**, *Grm2* mRNA upregulation by TALE–VP64 *in vivo* in PFC ($n = 4$ animals). Mean ± s.e.m. in all panels.

Similarly, we introduced LITEs into primary cortical neurons via co-delivery of two AAVs (**Fig. 6a** and **b**). We tested a *Grm2*-targeted LITE at 2 light pulsing frequencies with a reduced duty cycle of 0.8% to ensure neuron health (**Fig. 7**). Both stimulation conditions achieved a ~7-fold light-dependent increase in *Grm2* mRNA levels (**Fig. 6c**). Further study verified that substantial target gene expression increases could be attained quickly (4-fold up-regulation of mRNA within 4 h; **Fig. 6d**). In addition, we observed significant up-regulation of mGluR2 protein after stimulation, confirming that LITE-mediated transcriptional changes are translated to the protein level ($p < 0.01$ vs GFP control, $p < 0.05$ vs no-light condition; **Fig. 6e**).

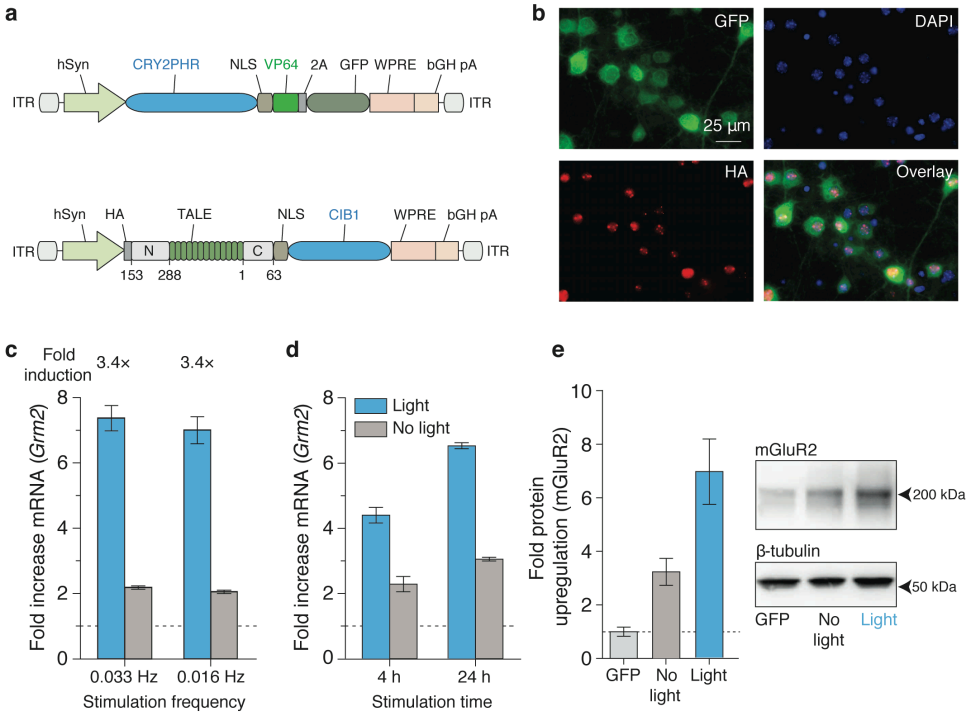

Figure 6 | LITE-mediated optogenetic modulation of endogenous transcription in primary neurons. a, Schematic of AAV LITE constructs. **b**, Images of primary neurons expressing LITE

Figure 6 (Continued)

constructs. HA, haemagglutinin tag. **c**, Light-induced activation of *Grm2* in primary neurons after 24 h of stimulation (250 ms pulses at 0.033 Hz or 500 ms pulses at 0.016 Hz; 5 mW/cm^2 ; n = 4 biological replicates). **d**, Upregulation of *Grm2* in primary cortical neurons after 4 h or 24 h of stimulation. Expression levels are shown relative to neurons transduced with GFP only (n=3-4 biological replicates). **e**, Light-mediated changes in mGluR2 protein levels (n = 7 biological replicates).

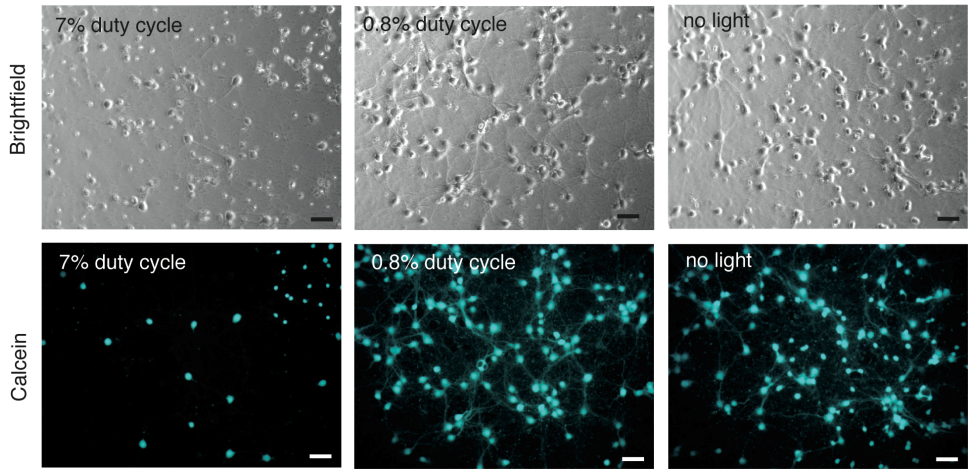

Figure 7 | Impact of light duty cycle on primary neuron health. The effect of light stimulation on primary cortical neuron health was compared for duty cycles of 7%, 0.8%, and no light conditions. Calcein was used to evaluate neuron viability. Brightfield images show cell morphology and integrity. Primary cortical neurons were stimulated with the indicated duty cycle for 24 h with 5 mW/cm^2 of 466 nm light. Representative images, scale bar, 50 μm. Pulses were performed in the following manner: 7% duty cycle = 1 s pulse at 0.067 Hz, 0.8% duty cycle = 0.5 s pulse at 0.0167 Hz.

To test the *in vivo* functionality of the LITE system, we stereotactically delivered a 1:1 mixture of high titer AAV vectors carrying TALE(*Grm2*)-CIB1 and CRY2PHR-VP64 into the PFC. We used a previously established fiber optic cannula system[34] to deliver light to LITE-expressing neurons *in vivo* (**Fig. 8**). After 12 h of stimulation, we observed a significant increase in *Grm2* mRNA compared with unstimulated PFC (**Fig. 8d**, $p \leq$

0.01). Taken together, these results confirm that LITEs enable optical control of endogenous gene expression in cultured neurons and *in vivo*.

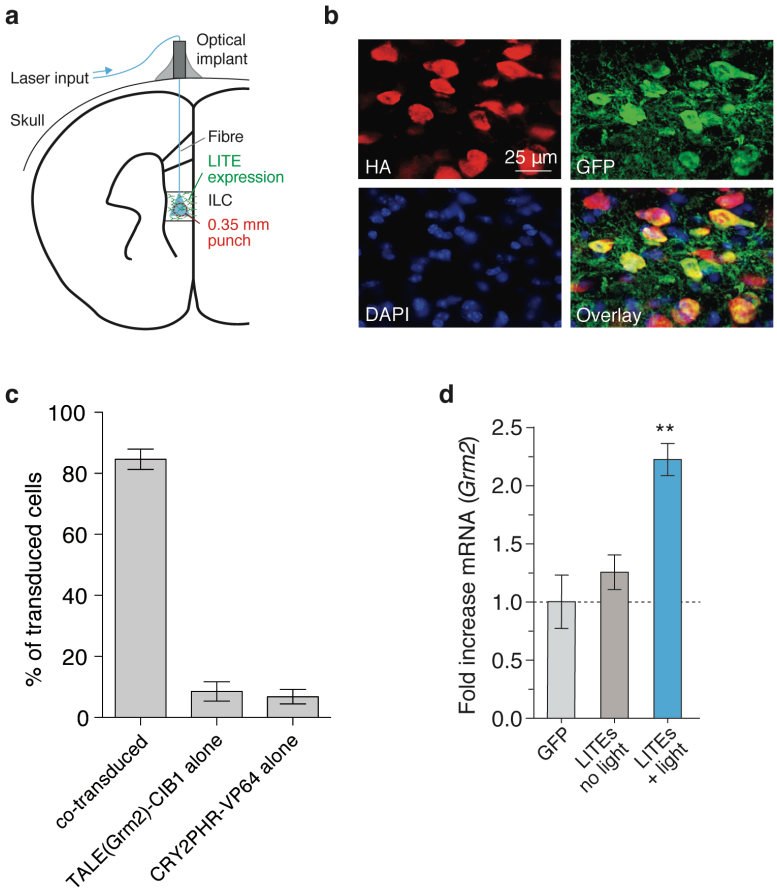

Figure 8 | LITE-mediated optogenetic modulation of endogenous transcription in vivo. a, Schematic of *in vivo* optogenetic stimulation setup. **b**, Representative images of PFC neurons expressing both LITE components. **c**, Co-transduction efficiency of LITE components by AAV1/2 *in vivo* in mouse infralimbic cortex. Cells transduced by TALE(*Grm2*)–CIB1 alone, CRY2PHR–VP64 alone, or co-transduced were calculated as a percentage of all transduced cells (mean ± s.e.m.; $n = 9$ fields from 3 animals). **d**, Light-induced activation of endogenous *Grm2* expression using LITEs transduced into ILC. (**$p < 0.01$; n = 4–6 animals.)

Section IV: Inducible nuclear import and other approaches to LITE optimizations

Having observed persistent baseline up-regulation *in vivo,* we undertook further rounds of optimization to reduce background activity and improve the gene induction ratio of LITEs. We observed that TALE(*Grm2*)-CIB1 alone produced similar levels of up-regulation as background activation, yet CRY2PHR-VP64 alone did not significantly affect transcription (**Fig. 9**). Therefore we rationalized that LITE-dependent background transcriptional activation arises mainly from TALE-CIB1.

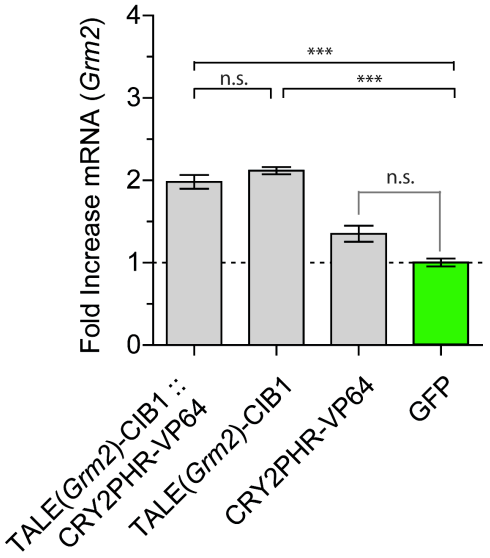

Figure 9 | Basal activation by LITE components in primary neurons. *Grm2* mRNA levels were determined in primary neurons transfected with individual LITE components. Primary neurons expressing TALE(*Grm2*)–CIB1 alone led to a similar increase in *Grm2* mRNA levels as unstimulated cells expressing the complete LITE system (mean ± s.e.m.; n = 3–4 biological replicates).

The subsequent comprehensive screen to reduce baseline TALE-CIB1-mediated up-regulation focuses on two strategies: First, although CIB1 is a plant transcription factor, it may have intrinsic activity in mammalian cells[35]. To address this, we deleted three CIB1

regions conserved amongst basic helix-loop-helix transcription factors of higher plants (**Fig. 10**). Second, to prevent TALE-CIB1 from binding the target locus in absence of light, we engineered TALE-CIB1 to localize in the cytoplasm pending light-induced dimerization with the NLS-containing CRY2PHR-VP64 (**Fig. 11**). To test both strategies independently or in combination, we evaluated 73 distinct LITE architectures and identified 12 effector/targeting-domain pairs (denoted by the "+" column in **Fig. 10**) with both improved light-induction efficiency and reduced background (fold mRNA increase in the no-light condition compared with the original LITE; $p < 0.05$).

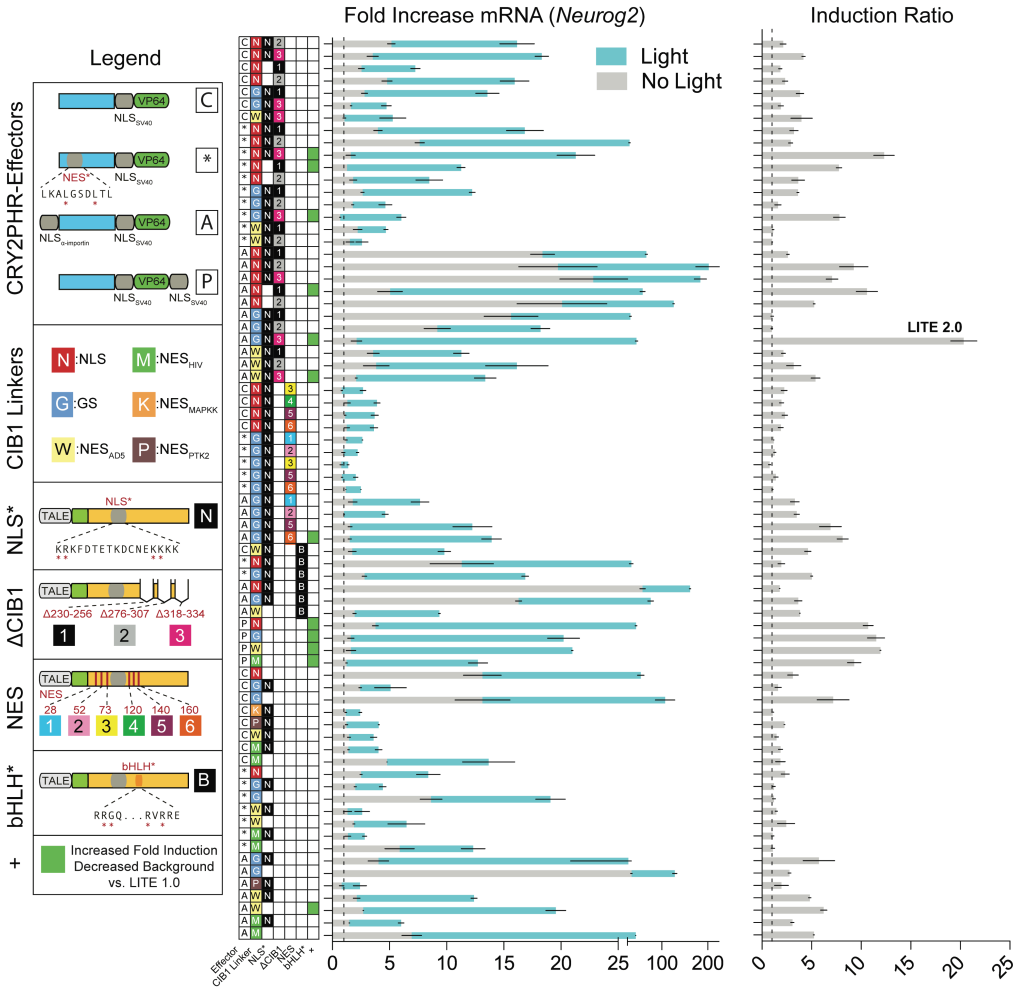

Figure 10 | Effects of LITE component engineering on activation, background signal and fold induction. Protein modifications were used to find LITE components resulting in reduced background transcriptional activation while improving induction ratio by light. In brief, nuclear

Figure 10 (Continued)

localization signals and mutations in an endogenous nuclear export signal were used to improve nuclear import of the CRY2PHR–VP64 component. Several variations of CIB1 intended to either reduce nuclear localization or CIB1 transcriptional activation were pursued to reduce the contribution of the TALE–CIB1 component to background activity. The results of all tested combinations of CRY2PHR–VP64 and TALE–CIB1 are shown. The table to the left of the bar graphs indicates the particular combination of domains/mutations used for each condition. Each row of the table and bar graphs contains the component details, light/no light activity, and induction ratio by light for the particular CRY2PHR/CIB1 combination. Combinations that resulted in both decreased background and increased fold induction compared to LITE1.0 are highlighted in green in the table column marked '+' (*t*-test $p < 0.05$). (mean ± s.e.m.; $n = 2–3$ biological replicates).

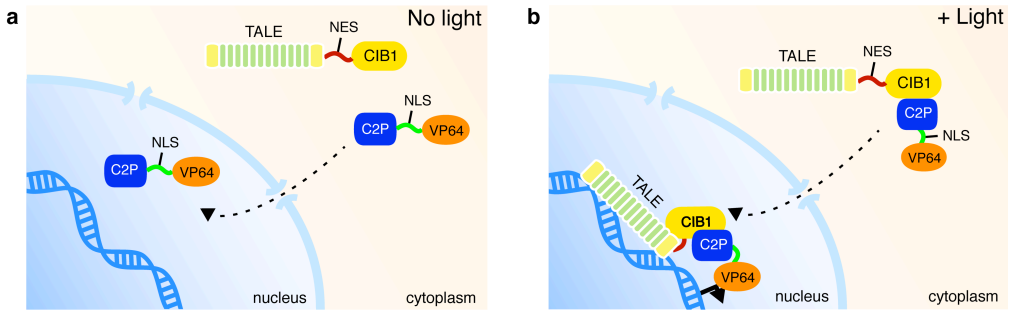

Figure 11 | Inducible nuclear import. **a**, In the absence of light, the TALE–CIB1 LITE component resides in the cytoplasm due to the absence of a nuclear localization signal, NLS (or the addition of a nuclear export signal, NES). The CRY2PHR–VP64 component containing an NLS on the other hand is actively imported into the nucleus on its own. **b**, In the presence of blue light, TALE–CIB1 binds to CRY2PHR. The NLS present in CRY2PHR–VP64 then mediates nuclear import of the complex comprising both LITE components, enabling them to activate transcription at the targeted locus.

One architecture successfully incorporating both strategies, designated LITE2.0, demonstrated the strongest light induction (light/no-light = 20.4) and resulted in greater than 6-fold reduction of background activation compared with the original design (**Fig.**

12a). Another architecture, LITE1.9.1, produced minimal background activation (1.06) while maintaining four-fold light induction (**Fig. 12b**).

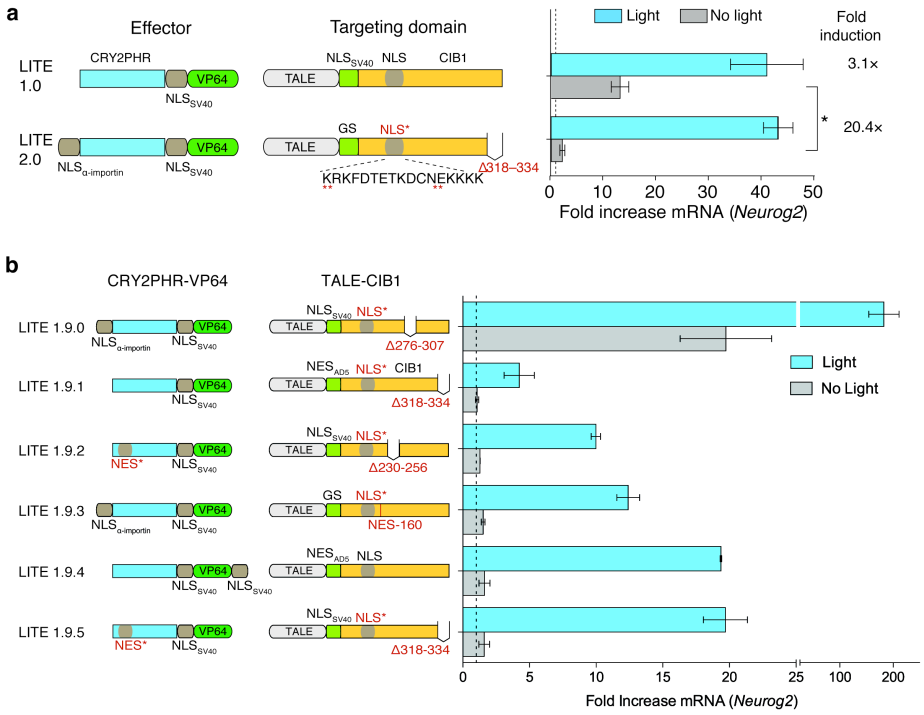

Figure 12 | LITE2.0 and selected LITE1.9 optimizations. a, LITE2.0 significantly reduces the level of background activation in Neuro 2a cells (n = 3 biological replicates). Mean ± s.e.m. in all panels. **b**, Several CRY2PHR–VP64/TALE–CIB1 combinations from the engineered LITE component screen were of particular note. LITE1.9.0, which combined the α-importin NLS effector construct with a mutated endogenous NLS and Δ276–307 TALE–CIB1 construct, exhibited an induction ratio greater than 9 and an absolute light activation of more than 180. LITE1.9.1, which combined the unmodified CRY2PHR–VP64 with a mutated NLS, Δ318–334, AD5 NES TALE–CIB1 construct, achieved an induction ratio of 4 with a background activation of 1.06. A selection of other LITE1.9 combinations with background activations lower than 2 and induction ratios ranging from 7 to 12 were also highlighted (mean ± s.e.m.; n = 2–3 biological replicates).

Modification strategies for background elimination

CRY2PHR-VP64 Constructs: Three new constructs were designed with the goal of improving CRY2PHR-VP64 nuclear import. First, the mutations L70A and L74A within

a predicted endogenous nuclear export sequence of CRY2PHR were created to limit nuclear export of the protein (referred to as '*' in the Effector column of **Fig. 10**). Second, the α-importin nuclear localization sequence was fused to the N-terminus of CRY2PHR-VP64 (referred to as 'A' in the Effector column of **Fig. 10**). Third, the SV40 nuclear localization sequence was fused to the C-terminus of CRY2PHR-VP64 (referred to as 'P' in the Effector column of **Fig. 10**).

TALE-CIB1 Linkers: The SV40 NLS linker between TALE and CIB1 used in LITE 1.0 was replaced with one of several linkers designed to increase nuclear export of the TALE-CIB1 protein (The symbols used in the CIB1 Linker column of **Fig. 10** are shown in parentheses): a flexible glycine-serine linker (G), an adenovirus type 5 E1B nuclear export sequence (W), an HIV nuclear export sequence (M), a MAPKK nuclear export sequence (K), and a PTK2 nuclear export sequence (P).

NLS* Endogenous CIB1 Nuclear Localization Sequence Mutation: A nuclear localization signal exists within the wild type CIB1 sequence. This signal was mutated in NLS* constructs at K92A, R93A, K105A, and K106A in order to diminish TALE-CIB1 nuclear localization (referred to as 'N' in the NLS* column of **Fig. 10**).

ΔCIB1 Transcription Factor Homology Deletions: In an effort to eliminate possible basal CIB1 transcriptional activation, deletion constructs were designed in which regions of high homology to basic helix-loop-helix transcription factors in higher plants were removed. These deleted regions consisted of Δaa230-256, Δaa276-307, Δaa308-334

(referred to as ‘1’ ‘2’ and ‘3’ in the ΔCIB1 column of **Fig. 10**). In each case, the deleted region was replaced with a 3 residue GGS link.

NES Insertions into CIB1: One strategy to facilitate light-dependent nuclear import of TALE-CIB1 was to insert an NES in CIB1 at its dimerization interface with CRY2PHR such that the signal would be concealed upon binding with CRY2PHR. To this end, an NES was inserted at different positions within the known CRY2 interaction domain CIBN (aa 1-170). The positions are as follows (The symbols used in the NES column of **Fig. 10** are shown in parentheses): aa28 (1), aa52 (2), aa73 (3), aa120 (4), aa140 (5), aa160 (6).

***bHLH basic Helix-Loop-Helix Mutation:** To reduce direct CIB1-DNA interactions, several basic residues of the basic helix-loop-helix region in CIB1 were mutated. The following mutations are present in all *bHLH constructs (referred to as ‘B’ in the *bHLH column of **Fig. 10**): R175A, G176A, R187A, and R189A

Section V: Epigenetic modification and gene repression using TALEs and LITEs

We sought to further expand the range of transcriptional processes addressable by TALE and LITE. We developed a TALE-repressor by fusing TALEs to a 4X concatenation of the transcriptional repressor mSin3 interaction domain (SID4X). We observed that, in a manner analogous to TALE-activator constructs, TALE-SID4X could repress endogenous transcription when targeted to the promoter of a gene of interest (**Fig. 13**).

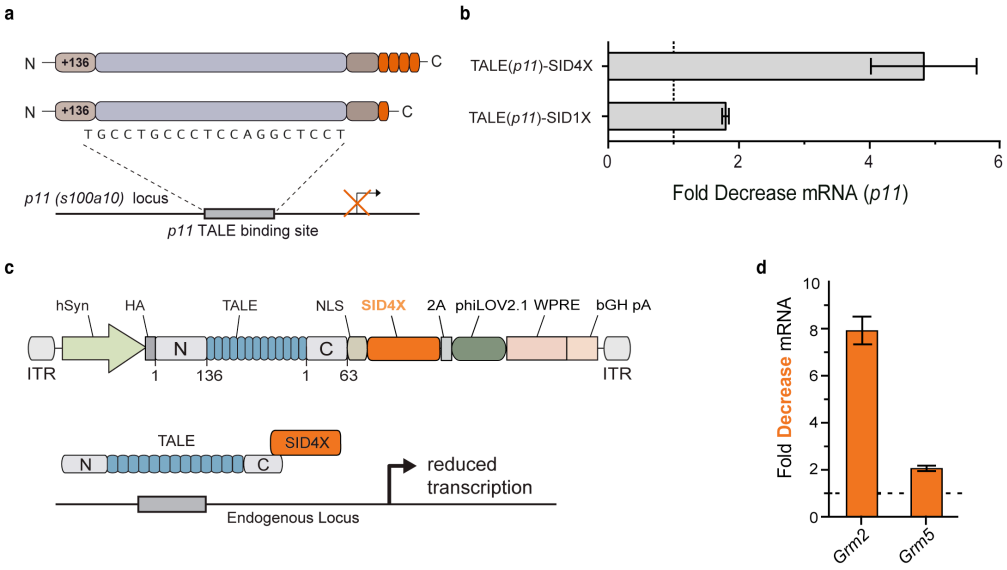

Figure 13 | TALE SID4X repressor characterization and application in neurons. a, A synthetic repressor was constructed by concatenating 4 SID domains (SID4X). To identify the optimal TALE-repressor architecture, SID or SID4X was fused to a TALE designed to target the mouse *p11* (also known as *S100a10*) gene. **b**, Fold decrease in *p11* mRNA was assayed using qRT–PCR (mean ± s.e.m.; $n = 3$ biological replicates). **c**, General schematic of constitutive TALE transcriptional repressor packaged into AAV. Effector domain SID4X is highlighted. hSyn, human synapsin promoter; 2A, *Thosea asigna* virus 2A self-cleaving peptide[36]; WPRE, woodchuck hepatitis post-transcriptional response element; bGH pA, bovine growth hormone poly-A signal. phiLOV2.1[37] (330 bp) was chosen as a shorter fluorescent marker to ensure efficient AAV packaging. **d**, A TALE targeting either the endogenous mouse locus *Grm5* or *Grm2* was fused to SID4X and virally transduced into primary neurons. SID4X-

Figure 13 (Continued)

mediated target gene downregulation is shown for each TALE relative to levels in control neurons expressing GFP only (mean ± s.e.m.; n = 3–4 biological replicates).

We further hypothesized that TALE-mediated targeting of histone effectors, such as SID, to endogenous loci could induce specific epigenetic modifications, which would enable the interrogation of epigenetic as well as transcriptional dynamics (**Fig. 14a**). We fused CRY2PHR with SID4X (**Fig. 14b**) and observed light-mediated transcription repression of *Grm2* in neurons (**Fig. 14c**) accompanied by ~2-fold reduction in H3K9 acetylation at the targeted *Grm2* promoter (**Fig. 14d**).

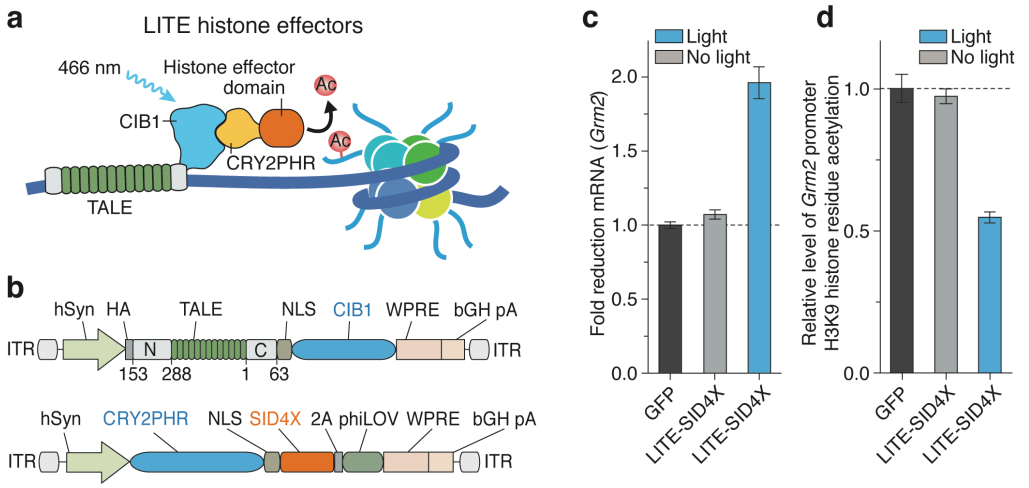

Figure 14 | LITE mediated epigenetic modifications. a, LITE epigenetic modifiers (epiLITE). **b**, epiLITE AAV vectors. **c**, epiLITE-mediated repression of endogenous *Grm2* in neurons (mean ± s.e.m; n = 4 biological replicates). **d**, epiLITE-mediated decrease in H3K9 histone acetylation at the *Grm2* promoter (mean ± s.e.m; n = 4 biological replicates).

In order to expand the diversity of histone residue targets for locus-specific histone modification, we next derived a set of 32 repressive histone effector domains (**Supplementary Tables 1-5**). Selected from across a wide phylogenetic spectrum, the

domains include histone deacetylases (HDACs), methyltransferases (HMTs), acetyltransferase (HAT) inhibitors, as well as HDAC and HMT recruiting proteins. Preference was given to proteins and functional truncations of small size to facilitate efficient AAV packaging. The resulting epigenetic mark-modifying TALE-histone effector fusion constructs (epiTALEs) were evaluated in primary neurons and Neuro 2a cells for their ability to repress *Grm2* and *Neurog2* transcription, respectively (**Fig. 15** and **Fig. 16**). In primary neurons, 23 out of 24 epiTALEs successfully repressed transcription of *Grm2* ($p < 0.05$). Similarly, epiTALE expression in Neuro 2a cells led to decreased *Neurog2* expression for 20 of the 32 histone effector domains tested (**Fig. 16**; $p < 0.05$). We then expressed a subset of promising epiTALEs in primary neurons and Neuro 2a cells and quantified the relative histone residue mark levels at the target locus using ChIP-RT-qPCR (**Fig. 15** and **Fig. 17**). The effector domains characterized by ChIP-RT-qPCR are listed with their corresponding histone marks and species of origin (**Supplementary Tables 1-5**). These domains provide a ready source of epigenetic effectors for LITE-mediated control of specific epigenetic modifications.

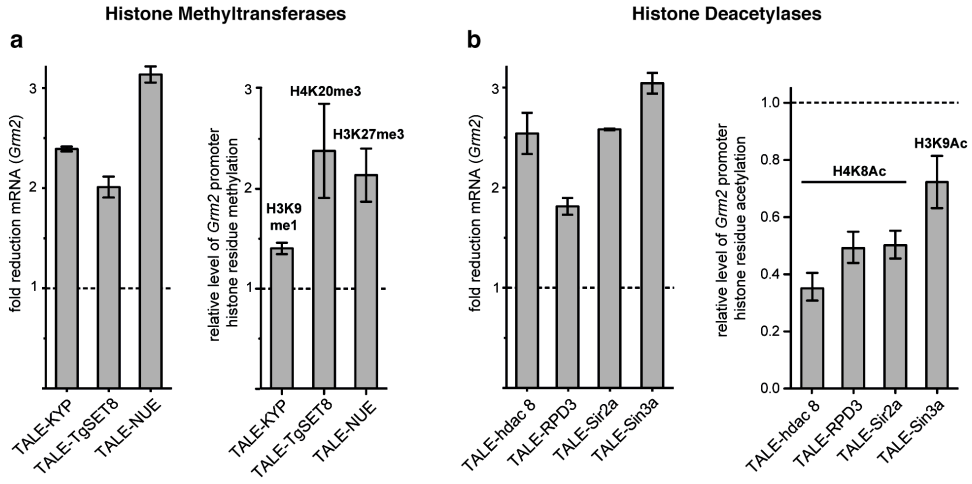

Figure 15 | TALE-mediated epigenetic modifications in primary neurons. a, **b**, epiTALE methyltransferases mediated decrease in *Grm2* mRNA and corresponding enrichment of

Figure 15 (Continued)

H3K9me1, H4K20me3 and H3K27me3 at the *Grm2* promoter (n = 3 biological replicates). **c**, **d**, epiTALE histone deacetylases mediated repression of *Grm2* and corresponding decreases in H4K8Ac and H3K9Ac marks at the *Grm2* promoter (n=2–5 biological replicates). Mean ± s.e.m. in all panels.

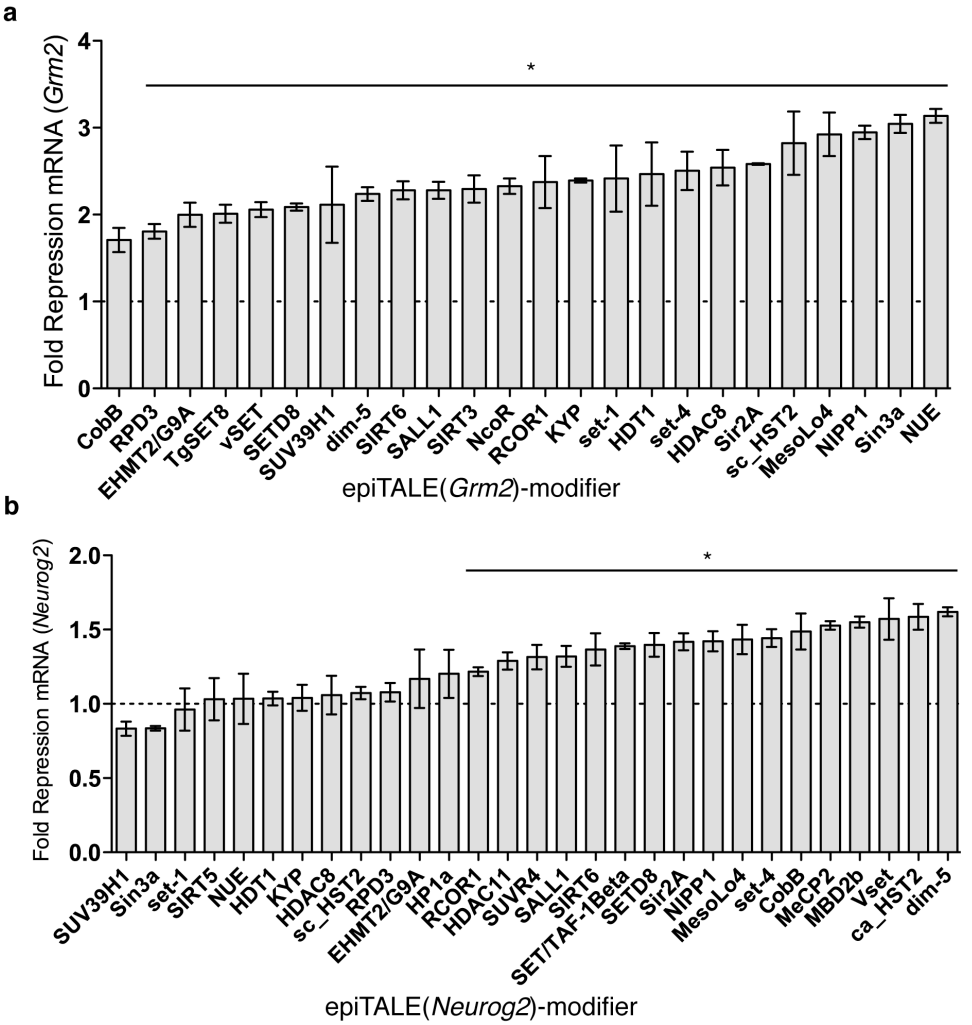

Figure 16 | Transcriptional repression by a diverse set of epiTALEs. a, 24 different histone effector domains were each fused to a *Grm2* targeting TALE. TALE–effector fusions were expressed in primary cortical mouse neurons using AAV transduction. *Grm2* mRNA levels were measured using qRT–PCR relative to neurons transduced with GFP only. (*$p < 0.05$; mean ± s.e.m.; n = 2–3 biological replicates.) **b**, A total of 32 epiTALEs were transfected into Neuro 2a

Figure 16 (Continued)

cells. 20 of them mediated significant repression of the targeted *Neurog2* locus (*p < 0.05; mean ± s.e.m.; *n* = 2–3 biological replicates).

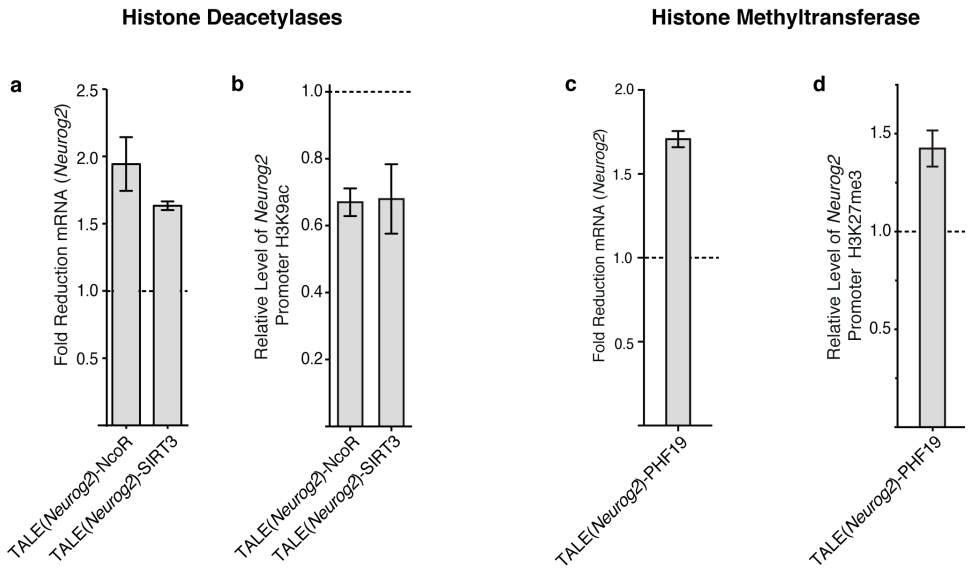

Figure 17 | EpiTALE mediated transcriptional repression and histone modifications in Neuro 2a cells. a, TALEs fused to histone-deacetylating epigenetic effectors NcoR and SIRT3 targeting the murine *Neurog2* locus in Neuro 2a cells were assayed for repressive activity on *Neurog2* transcript levels (mean ± s.e.m.; *n* = 2–3 biological replicates). **b**, ChIP qRT–PCR showing a reduction in H3K9 acetylation at the *Neurog2* promoter for NcoR and SIRT3 epiTALEs (mean ± s.e.m.; *n* = 2–3 biological replicates). **c**, The epigenetic effector PHF19 with known histone methyltransferase binding activity was fused to a TALE targeting *Neurog2*. Repression of *Neurog2* mRNA levels was observed (mean ± s.e.m.; *n* = 2–3 biological replicates). **d**, ChIP qRT–PCR showing an increase in H3K27me3 levels at the *Neurog2* promoter for the PHF19 epiTALE (mean ± s.e.m.; *n* = 2–3 biological replicates)

Section VI: CRISPR/Cas9 transcriptional effectors and CasLITEs

The CRISPR/Cas9 RNA-guided nuclease system has facilitated major increases in the efficiency[22, 38, 39], speed[40], and scalability[41] of mammalian genome engineering. The capacity for Cas9 to be targeted to a specific genetic locus using a short, typically 20 bp, RNA guide sequence fused to a <100 bp chimeric backbone, together called a synthetic guide RNA (sgRNA) enables the rapid generation of distinct targeting sequences. Panels of greater than 800 guide sequences have been tested in a high-throughput arrayed format[38], while mixed libraries of up to 90,000 guides have been applied to pooled knockout screens for fundamental cellular processes[42] and chemotherapeutic resistance[41]. Further, the small size of the sgRNA allows for convenient multiplexing of guides targeted to a single[22, 23, 43] or several genetic targets[22, 23, 44]. The ease of synthesis and potential for multiplexing of sgRNAs have made CRISPR/Cas9 systems an attractive successor to TALE based DNA-editing and transcriptional modulation.

The discovery of a nucleolytically inactive Cas9 mutant[27] (mutations D10A and H840A) inspired us, and our colleagues in the field[26, 45-48], to explore the use of dead Cas9 (dCas9) as a transcriptional modulator. In principle, dCas9 is fused to an effector domain and expressed in concert with an sgRNA targeted to the promoter of a gene of interest. When the complex is expressed in a cell, the dCas9 facilitates the interaction of the sgRNA with chromosomal DNA, the guide sequence pairs with its target site, stabilizes the interaction, and allows the effector domain to act (**Fig. 18a**). To test this concept, we

fused dCas9 to either the VP64 activator domain or the SID4X repressor domain and selected guide sequences to target human genetic loci (**Fig. 18b**–**d**). Expression of dCas9-VP64 with an sgRNA targeted to the KLF4 promoter resulted in up to 2.5 fold activation in 293FT cells, depending on the particular guide sequence of the sgRNA (**Fig. 18f**). Similarly, expression of SID4X-dCas9 with an sgRNA targeted to the promoter or 5' UTR of SOX2 resulted in up to 2 fold repression, again depending on the sgRNA guide sequence (**Fig. 18g**).

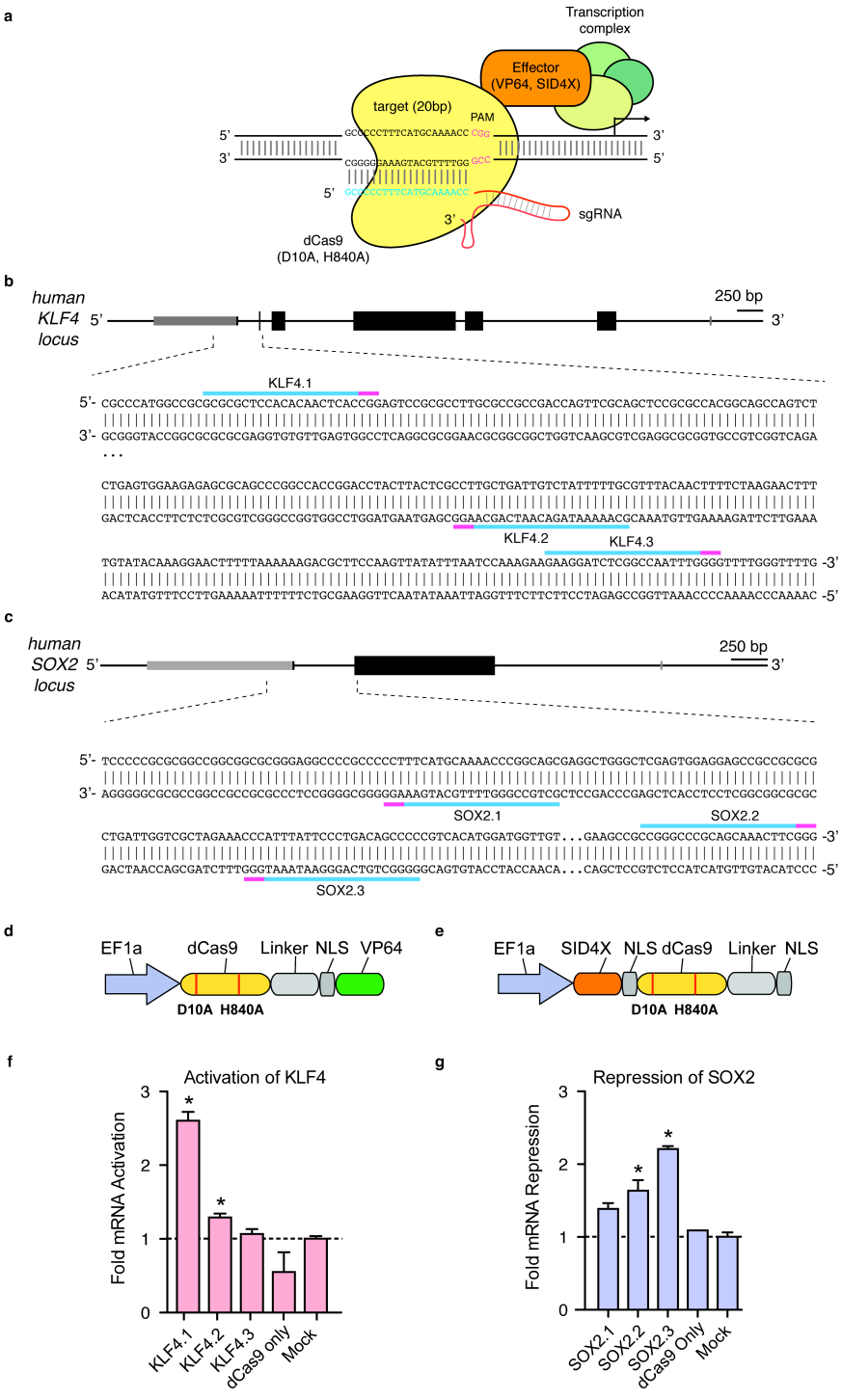

Figure 18 | RNA-guided DNA binding protein Cas9 targeting of transcription effector domains to specific genomic loci. a, The RNA-guided nuclease Cas9 from the type II Streptococcus pyogenes CRISPR/Cas system can be converted into a nucleolytically inactive RNA-guided DNA binding protein (dCas9) by introducing two alanine substitutions (D10A and H840A). Schematic showing that a synthetic guide RNA (sgRNA) can direct dCas9-effector

Figure 18 (Continued)

fusion to a specific locus in the human genome. The sgRNA contains a 20-bp guide sequence at the 5′ end, which specifies the target sequence. On the target genomic DNA, the 20-bp target site needs to be followed by a 5′-NGG PAM motif. **b**, **c**, Schematics showing the sgRNA target sites in the human KLF4 and SOX2 loci, respectively. Blue bars indicate target sites. Magenta bars indicate PAM sequences. **d**, **e**, Schematics of the dCas9–VP64 transcription activator and SID4X–dCas9 transcription repressor constructs. **f**, **g**, dCas9–VP64 and SID4X–dCas9 mediated activation of KLF4 and repression of SOX2, respectively. All mRNA levels were measured relative to GFP transfected 293FT cells (mean ± s.e.m.; n = 3 biological replicates).

To further validate our CRISPR/Cas9 activator approach, we expressed dCas9-VP64 in mouse Neuro 2a cells in combination with one of several sgRNAs targeted to the mouse *Neurog2* locus (**Supplementary Table 11**). We found that 6 out of 7 guides facilitated significant upregulation in *Neurog2* expression with the most potent guide producing ~20 fold expression (**Fig. 19**). We also verified the activity of a set of previously published guides targeted to several human genes and found that our dCas9-VP64 system produced results similar to those published by our colleagues[43, 47] (**Fig. 20**, **Supplementary Table 11**).

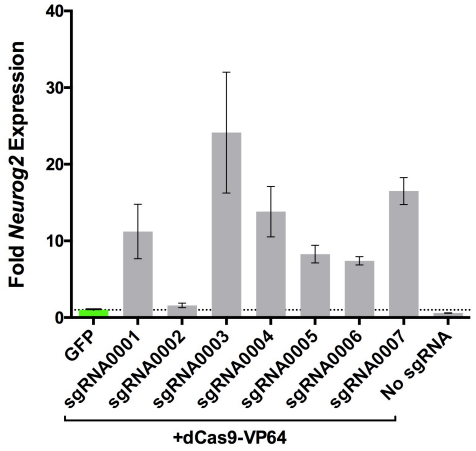

Figure 19 | dCas9 activation of murine *Neurog2*. sgRNA0001-0007 were designed to target the proximal regions of the mouse *Neurog2* locus. sgRNA0001 overlapped the transcription start site

Figure 19 (Continued)

(TSS) *Neurog2*. sgRNA0002-0006 were distributed throughout the 300 bp region upstream of the TSS. sgRNA0007 was located at bp 25 to 44 downstream of the TSS, within the *Neurog2* 5' untranslated region (UTR). sgRNA0002 overlapped the TATA-box site of the *Neurog2* promoter. Each sgRNA was expressed in combination with dCas9-VP64 in Neuro 2a cells. "No sgRNA" condition was transfected with an identical mass of dCas9-VP64 plasmid, as in sgRNA conditions, as well as GFP plasmid in place of an sgRNA plasmid (mean ± s.e.m.; n=3 biological replicates).

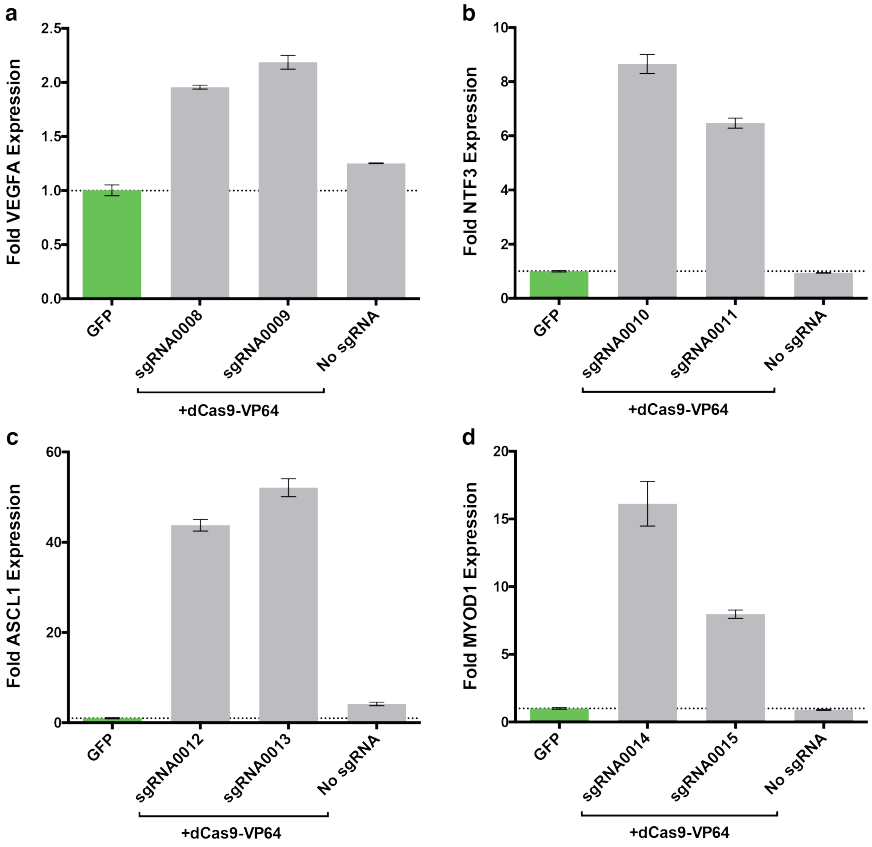

Figure 20 | CRISPR/dCas9 mediated human gene activation using validated sgRNA sequences. a-d, All sgRNAs were contransfected with dCas9-VP64 in 293FT cells. Fold expression was calculated relative to cells transfected with GFP only (mean ± s.e.m.; n=3 biological replicates). **a**, Guide sequences sgRNA0008-0009, known to increase VEGF concentration as previously measured by ELISA assay[47], were verified as expression activators of human VEGFA **b-d**, Guide sequences sgRNA0010-0011, sgRNA0012-0013, and sgRNA0014-

Figure 20 (Continued)

0015, known to increase NTF3[47], ASCL1[43], and MYOD1[43] expression, respectively, were verified as expression activators of their target genes.

In order to outfit the CRISPR/Cas9 activator system with the inducible properties of our TALE based LITE system, we constructed a cryptochrome heterodimer and dCas9 based light-inducible transcriptional effector called CasLITE. The CasLITE system comprises three components: a dCas9-CIB1 fusion protein, a CRY2PHR-effector fusion protein, and an sgRNA (**Fig. 21**). The dCas9-CIB1 forms a complex with the sgRNA to provide targeting to the gene of interest. Upon photostimulation, the CRY2PHR/CIB1 binding event localizes the effector domain to the target locus, inducing transcriptional modulation.

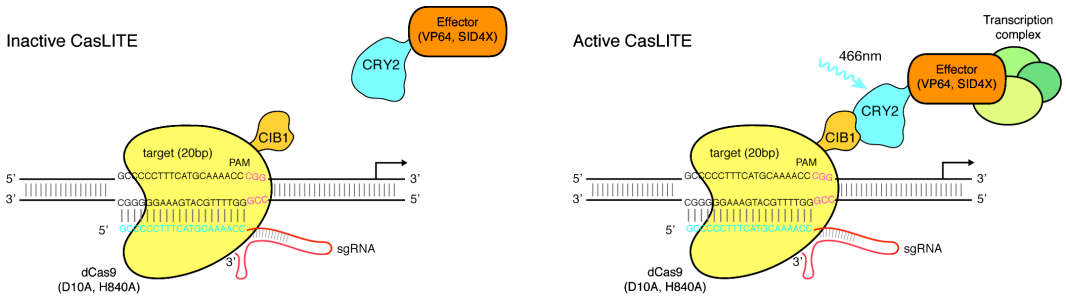

Figure 21 | Overview of CasLITE activation. dCas9-CIB1, in complex with an sgRNA, guides the system to the target locus via base-pairing of the sgRNA guide sequence with DNA at the gene of interest. Blue light stimulation induces CRY2/CIB1 dimerization, recruiting the effector domain to the target promoter.

We tested several versions of a VP64 activator CasLITE system with guides targeted, independently, to human ASCL1, MYOD1, and IL1RN (**Fig. 22**). We found that one iteration, CasLITE1.0 2X NLS, designed from our initial LITE CIB1/CRY2PHR system but using a double NLS linker, produced the highest absolute activation. Another version

of the system, CasLITE2.0, based on the inducible nuclear import LITE 2.0 system, resulted in both a lower activation level and a lower background signal. Fold induction was observed to be highly variable between gene targets, which future iterations of the system will aim to improve.

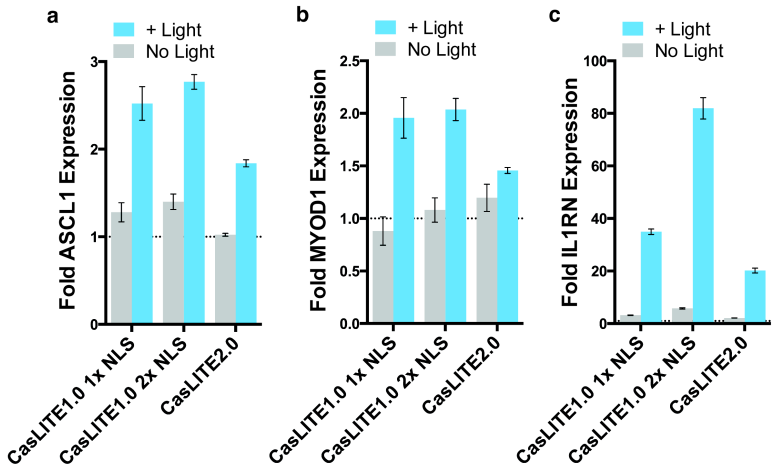

Figure 22 | Light-inducible RNA-guided transcriptional activation using CasLITEs. Several versions of the CasLITE system were tested for their capacity to provide light-mediated activation of human genes in 293FT cells. CasLITE1.0 was modeled after the basic CRY2PHR/CIB1 LITE architecture, with a single NLS between the CRY2PHR and VP64 domains and either a single or double NLS between the dCas9 and CIB1 domains. CasLITE2.0 was modeled after the LITE2.0 architecture for light-inducible nuclear import, with an NLS at both the N-term and C-term of CRY2PHR and a GS linker between the dCas9 and CIB1 domains. All fold expression levels were calculated by comparison to light-stimulated cells expressing GFP only (mean ± s.e.m.; n=3 biological replicates). **a**, sgRNA0012, which targets the ASCL1 locus, was co-expressed with each version of the CasLITE system. **b**, sgRNA0014, which targets the MYOD1 locus, was co-expressed with each version of the CasLITE system. **c**, 4 sgRNAs designed to target the IL1RN locus were co-expressed with each version of the CasLITE system.

To expand the applicability of the constitutive CRISPR/Cas9 activator approach to a wide range of biological applications, such as genome-wide activation screens, the potency of individual sgRNA/dCas9-activator combinations must be enhanced. We hypothesized

that the activation potency of current CRISPR/Cas9 activators may be hindered by the placement of the activator domain at either the N- or C-term of the dCas9. In fact, observation of the Cas9/DNA/RNA crystal structure[49] suggested that a VP64 domain fused to either terminus would be poorly localized to the target DNA.

Further observation of the crystal structure suggested that the sgRNA stemloops are not occluded by Cas9 and are in close proximity to the target DNA. By modifying these stemloops to become transcriptional activator recruitment sites, we hoped to improve the potency of the CRISPR/Cas9 actvators. To do so, we employed the MS2 phage coat protein, which binds directly to a particular RNA sequence[50-53]. We designed modified sgRNA backbones, with an MS2 binding site inserted into the native sgRNA stemloop sites (**Fig. 23a**). An MS2 binding site replaced either the sgRNA tetraloop or stemloop 2 or both. The first iteration of the system consisted of 3 components: dCas9, an sgRNA with an inserted MS2 site(s), and an MS2-VP64 fusion protein. The dCas9-VP64/sgRNA complex localizes to the gene of interest while the MS2-VP64 binds to its site on the sgRNA and induces transcription. When targeted to the mouse *Neurog2* locus in Neuro 2a cells, a more potent activation, compared to the basic dCas9-VP64 system, was observed in all conditions that included both the MS2-VP64 and an MS2 binding site in the sgRNA (**Fig. 23b**). The combination of MS2-VP64, dCas9-VP64, and an sgRNA with MS2 binding sites at both the tetraloop and stemloop 2 resulted in an ~8 fold improvement in target gene expression.

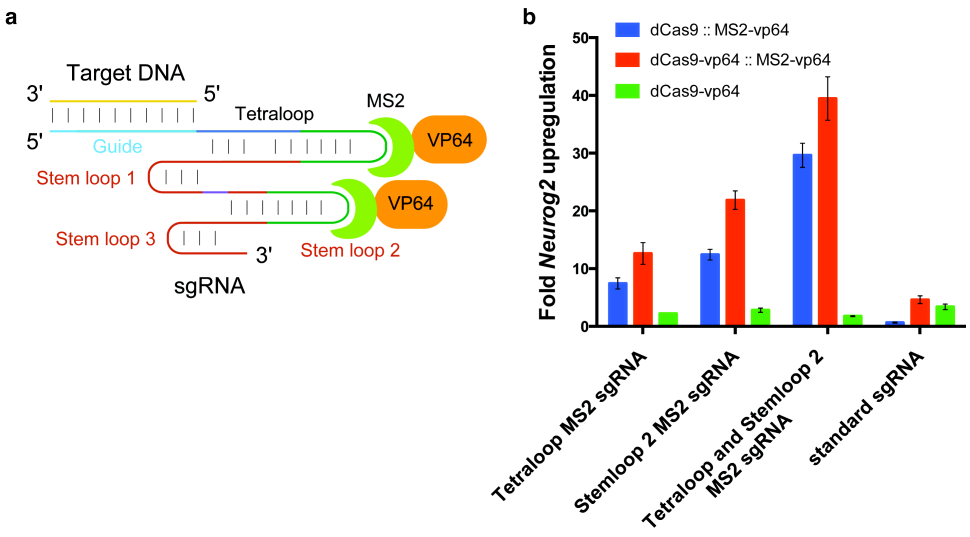

Figure 23 | MS2-VP64 mediated transcriptional activation. a, Schematic of the MS2-effector and MS2 sgRNA interaction. The tetraloop and/or stem-loop 2 of the standard chimeric sgRNA were replaced with the sequence encoding an MS2 binding site. These binding sites promote the recruitment of an MS2 phage coat protein, which was fused to the VP64 activation domain. **b**, Three component combinations of sgRNA, MS2-VP64, and dCas9 were tested for their ability to activate expression from the mouse *Neurog2* locus. The target sequence sgRNA0003 was inserted into several sgRNA backbones: an sgRNA backbone containing an MS2 binding site at the tetraloop and/or stem-loop 2, as well as the standard sgRNA backbone, which contained no MS2 binding site. Each of these sgRNAs was co-expressed with MS2-VP64 and either dCas9 with no effector domain or dCas9-VP64. Fold *Neurog2* expression was calculated relative to cells expressing GFP only (mean ± s.e.m.; n=3 biological replicates).

The design of the MS2/dCas9 activator system opens the door for the use of independent effector domains on the dCas9-activator and the MS2-activator. We hypothesized that greater activation levels could be achieved by placing distinct, complimentary activation domains on the dCas9 and MS2 proteins. We created proteins with MS2 and dCas9 fused to either VP64 or the p65 activation domain. Interestingly, we found that, when used in combination, the p65 and VP64 effectors produced greater levels of activation than versions of the system consisting of an equivalent number of a single type of activator

domain (**Fig. 24**). All told, the combination of MS2-p65, dCas9-VP64, and an sgRNA with two MS2 binding sites produced ~500 fold activation when targeted to human ASCL1 in 293FT cells, compared to 15 fold activation using our standard dCas9-VP64 and standard sgRNA. Comparatively, activation of human MYOD1 increased from ~40 fold under the standard system to ~170 fold when targeted by the MS2-p65/dCas9-VP64 combination. Given the increased potency of the MS2/dCas9 activator and its capacity for synergistic complimentary effector domain recruitment, we believe this architecture offers an attractive alternative to previous CRISPR/Cas9 transcriptional effector approaches.

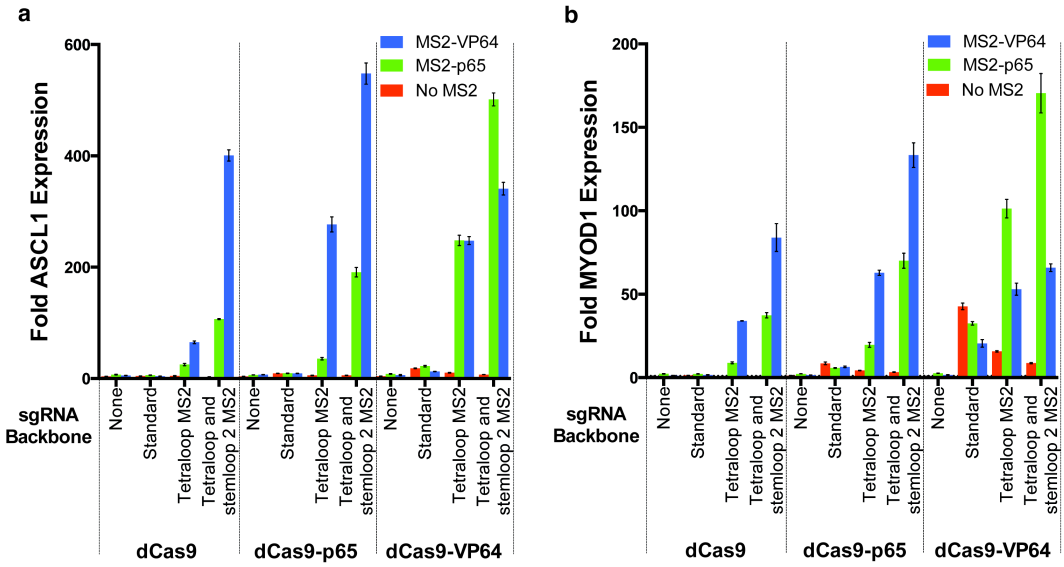

Figure 24 | Dual recruitment of p65 and VP64 activation domains by the MS2/CRISPR/dCas9 activator system. Three component combinations of an sgRNA, MS2-activator, and dCas9 were tested for their ability to activate human genes of interest in 293FT cells. Modified MS2 sgRNA backbones contained an MS2 binding site at the tetraloop and/or stem-loop 2 of the sgRNA. The standard sgRNA contained no MS2 binding site. Cells received either the MS2-VP64 or MS2-p65 activator protein. Cells also received either dCas9 with no effector domain, dCas9-VP64, or dCas9-p65. Conditions marked with "None" or "No MS2", received an empty pUC19 expression plasmid instead of an sgRNA plasmid or MS2-activator plasmid, respectively. Fold gene expression was determined relative to cells expressing GFP only

Figure 24 (Continued)

(mean ± s.e.m.; n=2–3 biological replicates). **a**, Guide sequence sgRNA0012, targeted to the human ASCL1 locus, was tested with each combination of the MS2/CRISPR/dCas9 activator system. **b**, Guide sequence sgRNA0014, targeted to the human MYOD1 locus, was tested with each combination of the MS2/CRISPR/dCas9 activator system.

Section VII: Considerations and Continuations

In considering this work, it is worth noting some of the open questions and caveats of the systems we have developed. These perspectives should also shed light on our continuing progress towards new transcriptional engineering tools.

Although we have characterized our TALE and CRISPR/Cas9 based transcriptional modulation tools from an engineering perspective, a number of questions remain regarding the basic biological responses to such systems. For instance, expression levels vary widely from gene to gene when activated by TALE-TFs (**Fig. 5c**) or dCas9-activators (**Fig. 20**). Basal expression level, locus-specific epigenetic states, local chromatin conformation, positive and negative transcriptional feedback, amongst other mechanisms, may contribute to the diversity of gene-specific responses to engineered transcriptional modulators. A systematic study of a large panel of activation targets, integrating transcriptional dynamics and epigenetic data, will be necessary to elucidate the particular mechanisms which attenuate or amplify gene expression engineering. An understanding of how artificial transcription factors interact with endogenous transcriptional regulation processes would prove invaluable for classifying locus-specific regulation mechanisms and interrogating gene expression in a systematic manner.

In addition to deepening our understanding of targeted activation processes, progress in the field will require a study of the genome-wide off-target effects of engineered transcriptional modulators. Significant effort has been dedicated to studying indel

formation by TALE[54, 55] and CRISPR/Cas9[38] nucleases at predicted off-target sites. Preliminary RNA-seq data sets have also been generated for dCas9 repressors targeted to fluorescent reporters[45] and dCas9 activators with sgRNAs targeted to the IL1RN locus[43]. However, a thorough genome-wide transcriptomic characterization is yet to be performed for either TALE or dCas9 based activators. For our continuing work, it will be of primary importance to understand any off-target effects of the MS2/dCas9 activator system. Wide adoption of either TALE or dCas9 transcriptional modulators will depend on a clear picture of off-target expression changes, as well as cytotoxicity and other cellular side-effects, in order to ensure clean experimental results or safe therapeutic applications.

As was highlighted at several points in the preceding sections (**Fig. 2c** and **d**, **Fig. 7**), light intensity is a critical parameter for the LITE system. We have observed that differences in tissue contexts can affect the light dosage tolerance of stimulated cells. It is likely that each individual cell line and primary tissue type will possess unique light sensitivity characteristics. Cell types could also exhibit markedly different light sensitivities within *in vitro* and *in vivo* contexts. Moving beyond viability studies, subtler analyses of photosensitivity, including gene expression analysis, would expand our understanding of the cell-physiological effects of the LITE system.

Complementary to elucidating the critical biological questions discussed above, we hope that our continuing work will yield further engineering improvements to increase the utility of the systems we have developed. For instance, the incorporation of photodimers that can be activated by wavelengths outside of the blue spectrum would open the door to

multimodal expression modulation of multiple genes. Capitalizing on the modularity of our TALE and dCas9 based architectures, we aim to continue to expand the toolbox of domains that we can apply to transcriptional and epigenetic engineering by testing effectors from across the phylogenetic spectrum.

There are a number of characterizations and improvements necessary to continue the development of dCas9 transcriptional modulators. Decoupling locus-specific regulation mechanisms from the variability in activator efficiency defined by targeting sequence content and location will facilitate a more rational approach to transcriptional modulator design. Testing our MS2/dCas9 system at a greater diversity of targets will allow us to evaluate the robustness this new iteration. Utilizing the MS2/dCas9 architecture to push the boundaries of gene activation multiplexing would help realize one of the main advantages supposedly offered by RNA-guided transcriptional effector systems. Assessing whether the high levels of transcription that we have observed using the MS2/dCas9 system are translated to the protein level is a key pending question in our current work. Looking ahead, porting the MS2/dCas9 system to lentiviral and AAV delivery systems will expand the diversity of contexts in which we can apply this new set of tools. Finally, the development of a virally delivered MS2/dCas9 activator system that can robustly activate expression using single sgRNA sequences will allow us to pursue genome-wide transcriptional activation screening applications. We believe that such a transcriptional modulation system, capable of interrogating large pathways or whole transcriptomes, will enable previously unachievable systems-level approaches to understanding gene expression dynamics.

Section VIII: Materials and Methods

Design and construction of LITEs

All LITE constructs sequences can be found in **Supplementary Sequences**. We evaluated full-length CRY2 as well as a truncation consisting of the photolyase homology region alone (CRY2PHR, amino acids 1-498)[6]. For CIB1, we tested the full-length protein as well as an N-terminal domain-only fragment (CIBN, amino acids 1-170)[6]. The efficacy of each design was determined based on the level of light-dependent up-regulation of the endogenous target *Neurog2* mRNA (**Fig. 1b**). In order to use AAV as a vector for the delivery of LITE components, we needed to ensure that the total viral genome size of each recombinant AAV, with the LITE transgenes included, did not exceed the packaging limit of 4.8 kb[56]. We shortened the TALE N- and C-termini (keeping 136 aa in the N-terminus and 63 aa in the C-terminus) and exchanged the CRY2PHR (1.5kb) and CIB1 (1kb) domains (TALE-CIB1 and CRY2PHR-VP64; **Fig. 6a**). TALE binding sequences were selected based on Dnase I-sensitive regions in the promoter of each target gene. TALE targeting sequences are listed in **Supplementary Table 6.**

Neuro 2a culture and experiments

Neuro 2a cells (Sigma-Aldrich) were grown in media containing a 1:1 ratio of OptiMEM (Life Technologies) to high-glucose DMEM with GlutaMax and Sodium Pyruvate (Life Technologies) supplemented with 5% HyClone heat-inactivated FBS (Thermo Scientific), 1% penicillin/streptomycin (Life Technologies), and passaged at 1:5 every 2

days. 120,000 cells were plated in each well of a 24-well plate 18-20 h prior to transfection. 1 h before transfection, media was changed to DMEM supplemented with 5% HyClone heat-inactivated FBS and 1% penicillin/streptomycin. Cells were transfected with 1.0 μg total of construct DNA (at equimolar ratios) per well with 1.5 μL of GenJet (SignaGen Laboratories) transfection reagent according to the manufacturer's instructions. Media was exchanged 24 h and 44 h post-transfection and light stimulation was started at 48 h. Stimulation parameters were: 5 mW cm^{-2}, 466 nm, 7 % duty cycle (1 s light pulse 0.067 Hz) for 24 h unless indicated otherwise in figure legends. RNA was extracted using the Rneasy kit (Qiagen) according to manufacturer's instructions and 1 μg of RNA per sample was reverse-transcribed using qScript (Quanta Biosystems). Relative mRNA levels were measured by quantitative real-time PCR (qRT-PCR) using TaqMan probes specific for the targeted gene as well as GAPDH as an endogenous control (Life Technologies, see **Supplementary Table 7** for Taqman probe IDs). ΔΔCt analysis was used to obtain fold-changes relative to negative controls transduced with GFP only and subjected to light stimulation. Toxicity experiments were conducted using the LIVE/DEAD assay kit (Life Technologies) according to manufacturer's protocol.

AAV vector production

The ssDNA-based genome of AAV is less susceptible to recombination, thus providing an advantage over RNA-based lentiviral vectors for the packaging and delivery of highly repetitive TALE sequences[33]. 293FT cells (Life Technologies) were grown in antibiotic-free D10 media (DMEM high glucose with GlutaMax and Sodium Pyruvate, 10% heat-inactivated Hyclone FBS, and 1% 1M HEPES) and passaged daily at 1:2-2.5. The total

number of passages was kept below 10 and cells were never grown beyond 85% confluence. The day before transfection, $1x10^6$ cells in 21.5 mL of D10 media were plated onto 15 cm dishes and incubated for 18-22 hours or until ~80% confluence. For use as a transfection reagent, 1 mg/mL of PEI "Max" (Polysciences) was dissolved in water and the pH of the solution was adjusted to 7.1. For AAV production, 10.4 µg of pDF6 helper plasmid, 8.7 µg of pAAV1 serotype packaging vector, and 5.2 µg of pAAV vector carrying the gene of interest were added to 434 µL of serum-free DMEM and 130 µL of PEI "Max" solution was added to the DMEM-diluted DNA mixture. The DNA/DMEM/PEI cocktail was vortexed briefly and incubated at room temperature for 15 min. After incubation, the transfection mixture was added to 22 mL of complete media, vortexed briefly, and used to replace the media for a 15 cm dish of 293FT cells. For supernatant production, transfection supernatant was harvested at 48 h, filtered through a 0.45 µm PVDF filter (Millipore), distributed into aliquots, and frozen for storage at -80°C.

Primary cortical neuron culture

Dissociated cortical neurons were prepared from C57BL/6N mouse embryos on E16 (Charles River Labs). Cortical tissue was dissected in ice-cold HBSS – (50 mL 10x HBSS, 435 mL dH_2O, 0.3 M HEPES pH 7.3, and 1% penicillin/streptomycin). Cortical tissue was washed 3X with 20 mL of ice-cold HBSS and then digested at 37°C for 20 min in 8 mL of HBSS with 240 µL of 2.5% trypsin (Life Technologies). Cortices were then washed 3 times with 20 mL of warm HBSS containing 1 mL FBS. Cortices were gently triturated in 2 ml of HBSS and plated at 150,000 cells/well in poly-D-lysine coated

24-well plates (BD Biosciences). Neurons were maintained in Neurobasal media (Life Technologies), supplemented with 1X B27 (Life Technologies), GlutaMax (Life Technologies) and 1% penicillin/streptomycin.

Primary neuron transduction and light stimulation experiments

Primary cortical neurons were transduced with 250 µL of AAV1 supernatant on DIV 5. The media and supernatant were replaced with regular complete neurobasal the following day. Neurobasal was exchanged with Minimal Essential Medium (Life Technologies) containing 1X B27, GlutaMax (Life Technologies) and 1% penicillin/streptomycin 6 days after AAV transduction to prevent formation of phototoxic products from HEPES and riboflavin contained in Neurobasal during light stimulation. For co-transduction of primary neurons with two AAV vectors, the co-delivery efficiency was >80%, with individual components having transduction efficiencies of 83-92%.

Light stimulation was started 6 days after AAV transduction (DIV 11) with an intensity of 5 mW/cm^2, duty cycle of 0.8% (250 ms pulses at 0.033Hz or 500 ms pulses at 0.016Hz), 466 nm blue light for 24 h unless indicated otherwise in figure legends. RNA extraction and reverse transcription were performed using the Cells-to-Ct kit according to the manufacturers instructions (Life Technologies). Relative mRNA levels were measured by quantitative real-time PCR (qRT-PCR) using TaqMan probes as described above for Neuro 2a cells.

Immunohistochemistry of primary neurons

For immunohistochemistry of primary neurons, cells were plated on poly-D-lysine/laminin coated coverslips (BD Biosciences) after harvesting. AAV1 transductions were performed as described above. Neurons were fixed 7 days post-transduction with 4% paraformaldehyde (Sigma Aldrich) for 15 min at room temperature. Blocking and permeabilization were performed with 10% normal goat serum (Life Technologies) and 0.5% Triton-X100 (Sigma-Aldrich) in DPBS (Life Technologies) for 1 h at room temperature. Neurons were incubated with primary antibodies overnight at 4°C, washed 3X with DPBS and incubated with secondary antibodies for 90 min at RT. For antibody providers and concentrations used, see **Supplementary Table 8**. Coverslips were finally mounted using Prolong Gold Antifade Reagent with DAPI (Life Technologies) and imaged on an Axio Scope A.1 (Zeiss) with an X-Cite 120Q light source (Lumen Dynamics). Images were acquired using an AxioCam MRm camera and AxioVision 4.8.2.

Western Blots

For preparation of total protein lysates, primary cortical neurons were harvested after light stimulation in ice-cold lysis buffer (RIPA, Cell Signaling; 0.1% SDS, Sigma-Aldrich; and cOmplete ULTRA protease inhibitor mix, Roche Applied Science). Cell lysates were sonicated for 5 min at 'M' setting with the Bioruptor water bath sonicator (Diagenode) and centrifuged at 21,000 x g for 10 min at 4°C. Protein concentration was determined using the RC DC protein assay (Bio-Rad). 30-40 μg of total protein per lane was separated under non-reducing conditions on 4-15% Tris-HCl gels (Bio-Rad) along with Precision Plus Protein Dual Color Standard (Bio-Rad). After wet electrotransfer to

polyvinylidene difluoride membranes (Millipore) and membrane blocking for 45 min in 5% BLOT-QuickBlocker (Millipore) in Tris-buffered saline (TBS, Bio-Rad), western blots were probed with anti-mGluR2 (Abcam, 1:1.000) and anti-α-tubulin (Sigma-Aldrich 1:20,000) overnight at 4°C, followed by washing and anti-mouse-IgG HRP antibody incubation (Sigma-Aldrich, 1:5,000 – 1:10,000). For further antibody details see **Supplementary Table 8**. Detection was performed via ECL Western blot substrate (SuperSignal West Femto Kit, Thermo Scientific). Blots were imaged with an AlphaImager system (Innotech), and quantified using ImageJ software 1.46r.

Production of concentrated and purified AAV1/2 vectors

Production of concentrated and purified AAV for stereotactic injection *in vivo* was performed using the same initial steps outlined above for production of AAV1 supernatant. However, for transfection, equal ratios of AAV1 and AAV2 serotype plasmids were used instead of AAV1 alone. Five 15cm plates were transfected per construct and cells were harvested with a cell-scraper 48 h post transfection. Purification of AAV1/2 particles was performed using HiTrap heparin affinity columns (GE Healthcare)[57]. We added a second concentration step down to a final volume of 100 µl per construct using an Amicon 500 µl concentration column (100 kDa cutoff, Millipore) to achieve higher viral titers. Titration of AAV was performed by qRT-PCR using a custom Taqman probe for WPRE (Life Technologies). Prior to qRT-PCR, concentrated AAV was treated with DnaseI (New England Biolabs) to achieve a measurement of DnaseI-resistant particles only. Following DnaseI heat-inactivation, the viral envelope

was degraded by Proteinase K digestion (New England Biolabs). Viral titer was calculated based on a standard curve with known WPRE copy numbers.

Stereotactic injection of AAV1/2 and optical implant

All animal procedures were approved by the MIT Committee on Animal Care. Adult (10-14 weeks old) male C57BL/6N mice were anaesthetized by intraperitoneal (i.p.) injection of Ketamine/Xylazine (100 mg/kg Ketamine and 10 mg/kg Xylazine) and pre-emptive analgesia was applied (Buprenex, 1 mg/kg, i.p.). Craniotomy was performed according to approved procedures and 1µl of AAV1/2 was injected into ILC at 0.35/1.94/-2.94 (lateral, anterior and inferior coordinates in mm relative to bregma). During the same surgical procedure, an optical cannula with fiber (Doric Lenses) was implanted into ILC unilaterally with the end of the optical fiber located at 0.35/1.94/-2.64 relative to bregma. The cannula was affixed to the skull using Metabond dental cement (Parkell Inc) and Jet denture repair (Lang Dental) to build a stable, supporting cone. The incision was sutured and proper post-operative analgesics were administered for three days following surgery.

Immunohistochemistry on ILC brain sections

Mice were injected with a lethal dose of Ketamine/Xylazine anaesthetic and transcardially perfused with PBS and 4% paraformaldehyde (PFA). Brains were additionally fixed in 4% PFA at 4°C overnight and then transferred to 30% sucrose for cryoprotection overnight at room temperature. Brains were then transferred into Tissue-Tek Optimal Cutting Temperature (OCT) Compound (Sakura Finetek) and frozen at -80°C. 18 µm sections were cut on a cryostat (Leica Biosystems) and mounted on

Superfrost Plus glass slides (Thermo Fischer). Sections were post-fixed with 4% PFA for 15 min, and immunohistochemistry was performed as described for primary neurons above.

Light stimulation and mRNA level analysis in ILC

Neurons at the injection site were efficiently co-transduced by both viruses, with >80% of transduced cells expressing both TALE(*Grm2*)-CIB1 and CRY2PHR-VP64 (**Fig. 8b** and **c**). 8 days post-surgery, awake and freely moving mice were stimulated using a 473 nm laser source (OEM Laser Systems) connected to the optical implant via fiber patch cables and a rotary joint. Stimulation parameters were the same as used on primary neurons: 5 mW (total output), 0.8% duty cycle (500 ms light pulses at 0.016 Hz) for a total of 12 h. Brain tissue from the fiber optic cannula implantation site was analyzed (**Fig. 3h**) for changes in *Grm2* mRNA. Experimental conditions, including transduced constructs and light stimulation are listed in **Supplementary Table 9**.

After the end of light stimulations, mice were euthanized using CO_2 and the prefrontal cortices (PFC) were quickly dissected on ice and incubated in RNA later (Qiagen) at 4°C overnight. 200 µm sections were cut in RNA later at 4°C on a vibratome (Leica Biosystems). Sections were then frozen on a glass coverslide on dry ice and virally transduced ILC was identified under a fluorescent stereomicroscope (Leica M165 FC). A 0.35 mm diameter punch of ILC, located directly ventrally to the termination of the optical fiber tract, was extracted (Harris uni-core, Ted Pella). The brain punch sample was then homogenized using an Rnase-free pellet-pestle grinder (Kimble Chase) in 50 µl

Cells-to-Ct RNA lysis buffer and RNA extraction, reverse transcription and qRT-PCR was performed as described for primary neuron samples.

Chromatin Immunoprecipitation

Neurons or Neuro 2a cells were cultured and transduced or transfected as described above. ChIP samples were prepared as previously described[58] with minor adjustments for the cell number and cell type. Cells were harvested in 24-well format, washed in 96-well format, and transferred to microcentrifuge tubes for lysis. Sample cells were directly lysed by water bath sonication with the Biorupter sonication device for 21 minutes using 30s on/off cycles (Diagenode). qPCR was used to assess enrichment of histone marks at the targeted locus. qPCR primer sequences are listed in **Supplementary Table 10.**

Statistical analysis

All experiments were performed with a minimum of two independent biological replicates. Statistical analysis was performed with Prism (GraphPad) using Student's two-tailed t-test when comparing two conditions, ANOVA with Tukey's post-hoc analysis when comparing multiple samples with each other, and ANOVA with Dunnett's post-hoc analysis when comparing multiple samples to the negative control.

Photostimulation Hardware – *in vitro*

In vitro light stimulation experiments were performed using a custom built LED photostimulation device. All electronic elements were mounted on a custom printed circuit board (ExpressPCB). Blue LEDs with peaks 466 nm (model #: YSL-R542B5C-

A11, China Young Sun LED Technology; distributed by SparkFun Electronics as 'LED – Super Bright Blue' COM-00529), were arrayed in groups of three aligned with the wells of a Corning 24-well plate. LED current flow was regulated by a 25 mA DynaOhm driver (LEDdymanics #4006-025). Columns of the LED array were addressed by TTL control (Fairchild Semiconductor PN2222BU-ND) via an Arduino UNO microcontroller board. Light output was modulated via pulse width modulation. Light output was measured from a distance of 80 mm above the array utilizing a Thorlabs PM100D power meter and S120VC photodiode detector. In order to provide space for ventilation and to maximize light field uniformity, an 80 mm tall ventilation spacer was placed between the LED array and the 24-well sample plate. Fans (Evercool EC5015M12CA) were mounted along one wall of the spacer unit, while the opposite wall was fabricated with gaps to allow for increased airflow.

Ambient light exposure

All cells were cultured at low light levels (<0.01 mW/cm^2) at all times except during stimulation. These precautions were taken as ambient light in the room (0.1-0.2 mW/cm^2) was found to significantly activate the LITE system (**Fig. 2c**). No special precautions were taken to shield animals from light during *in vivo* experiments – even assuming ideal propagation within the implanted optical fiber, an estimation of light transmission at the fiber terminal due to ambient light was <0.01 mW (based on 200 μm fiber core diameter and 0.22 numerical aperture).

Optimization of light stimulation parameters in Neuro2A cells

To minimize near-UV induced cytotoxicity, we selected 466 nm blue LEDs to activate TALE-CRY2, a wavelength slightly red-shifted from the CRY2 absorption maxima of 450 nm but still maintaining over 80% activity[29] (**Fig. 2a**). To minimize light exposure, we selected a mild stimulation protocol (1 s light pulses at 0.067 Hz, ~7% duty cycle). This was based on our finding that light duty cycle had no significant effect on LITE-mediated transcriptional activation over a wide range of duty cycle parameters (1.7% to 100% duty cycles, **Fig. 2b**). Illumination with a range of light intensities from 0 to 10 mW/cm^2 revealed that *Ngn2* mRNA levels increased as a function of intensity up to 5 mW/cm^2.

However, increases in *Ngn2* mRNA levels declined at 10 mW/cm^2 (**Fig. 2c**), suggesting that higher intensity light may have detrimental effects on either LITE function or on cell physiology. To better characterize this observation, we performed an ethidium homodimer-1 cytotoxicity assay with a calcein counterstain for living cells and found a significantly higher percentage of ethidium-positive cells at the higher stimulation intensity of 10 mW/cm^2. Conversely, the ethidium-positive cell count from 5 mW/cm^2 stimulation was indistinguishable from unstimulated controls (**Fig. 2d**). Thus 5 mW/cm^2 appeared to be optimal for achieving robust LITE activation while maintaining low cytotoxicity.

Quantification of LIVE/DEAD® assay using ImageJ software.

Images of LIVE/DEAD (Life Technologies) stained cells were captured by fluorescence microscopy and processed as follows: Background was subtracted (*Process* → *Subtract*

Background). A threshold based on fluorescence area was set to ensure accurate identification of cell state (*Image* → *Adjust* → *Threshold*). A segmentation analysis was performed to enable automated counting of individual cells (*Process* → *Binary* → *Watershed*). Finally, debris signals were filtered and cells were counted (*Analyze* → *Analyze Particles*). Toxicity was determined as the percentage of dead cells.

Reduction of light-induced toxicity in primary neurons

Initial application of LITEs in neurons revealed that cultured neurons were much more sensitive to blue light than Neuro 2a cells. Stimulation parameters previously optimized for Neuro 2a cells (466 nm, 5 mW/cm^2 intensity, 7% duty cycle with 1 s light pulse at 0.067 Hz for a total of 24 h) caused >50% toxicity in primary neurons. We therefore tested survival with a lower duty cycle, as we had previously observed that a wide range of duty cycles had little effect on LITE-mediated transcriptional activation (**Fig. 2b**). A reduced duty cycle of 0.8% (0.5 s light pulses at 0.0167 Hz) at the same light intensity (5 mW/cm^2) was sufficient to maintain a high survival rate that was indistinguishable from that of unstimulated cultures (**Fig. 7**).

Light propagation and toxicity in *in vivo* experiments

Previous studies have investigated the propagation efficiency of different wavelengths of light in brain tissue. For 473 nm light (wavelength used in this study), there was a >90% attenuation after passing through 0.35 mm of tissue[59]. An estimated 5 mW/cm^2 light power density was estimated based on a tissue depth of 0.35 mm of tissue (the diameter of brain punch used in this study) and a total power output of 5 mW. The light

stimulation duty cycle used *in vivo* was the same (0.8%, 0.5 s at 0.0167 Hz) as that used for primary neurons.

Chemically-inducible TALEs

Neuro2A cells were grown in a medium containing a 1:1 ratio of OptiMEM (Life Technologies) to high-glucose DMEM with GlutaMax and Sodium Pyruvate (Life Technologies) supplemented with 5% HyClone heat-inactivated FBS (Thermo Scientific), 1% penicillin/streptomycin (Life Technologies) and 25mM HEPES (Sigma Aldrich). 150,000 cells were plated in each well of a 24-well plate 18-24 hours prior to transfection. Cells were transfected with 1 μg total of construct DNA (at equimolar ratios) per well and 2 μL of Lipofectamine 2000 (Life Technologies) according to the manufacturer's recommended protocols. Media was exchanged 12 hours post-transfection. For the kinetics test, chemical induction was started 24 hours post-transfection, when abscisic acid (ABA, Sigma Aldrich) was added to fresh media to a final concentration of 250 μM. RNA was extracted using the Rneasy kit (Qiagen) according to manufacturer's instructions and 1 μg of RNA per sample was reverse-transcribed using qScript (Quanta Biosystems). Relative mRNA levels were measured by quantitative real-time PCR (qRT-PCR) using Taqman probes specific for the targeted gene as well as mouse GAPDH as an endogenous control (Life Technologies, see **Supplementary Table 7** for Taqman probe IDs). ΔΔCt analysis was used to obtain fold-changes relative to negative controls where cells were subjected to mock transfection with GFP.

Development of AAV1 supernatant process

Traditional AAV particle generation requires laborious production and purification processes[60], making the testing of many constructs in parallel impractical. In this study, a simple yet highly effective process of AAV production using filtered supernatant from transfected 293FT cells was developed (**Fig. 4**). A previous study indicated that AAV particles produced in 293FT cells could be found not only within the cytoplasm but also in considerable amounts within the culture media[61]. The ratio of viral particles between the supernatant and cytosol of host cells varied depending on the AAV serotype, and secretion was enhanced if polyethylenimine (PEI) was used to transfect the viral packaging plasmids[61].

In the current study, we demonstrated that $2x10^5$ 293FT cells transfected with AAV vectors carrying TALEs (**Fig. 5a**) and packaged using AAV1 serotype were capable of producing 250 µl of AAV1 at a concentration of $5.6 \pm 0.24 \times 10^{10}$ DnaseI resistant genome copies (GC) per mL. 250 µl of filtered supernatant was able to transduce 150,000 primary cortical neurons at efficiencies of 80-90% (**Fig. 2b** and **Fig. 8b** and **c**). This process was also successfully adapted to a 96-well format, enabling the production of 125 µl AAV1 supernatant from up to 96 different constructs in parallel. 35µl of supernatant can then be used to transduce one well of primary neurons cultured in 96-well format, enabling transductions in biological triplicates from a single well.

Cas9 transcriptional effectors: SOX2 and KLF4 experiments

HEK 293FT cells were co-transfected with mutant Cas9 fusion protein and a synthetic guide RNA (sgRNA) using Lipofectamine 2000 (Life Technologies) 24 hours after

seeding into a 24 well dish. 72 hours post-transfection, total RNA was purified (Rneasy Plus, Qiagen). 1 μg of RNA was reverse transcribed into cDNA (qScript, Quanta BioSciences). Quantitative real-time PCR was done according to the manufacturer's protocol (Life Technologies) and performed in triplicate using TaqMan Assays for hKlf4 (Hs00358836_m1), hSox2 (Hs01053049_s1), and the endogenous control GAPDH (Hs02758991_g1).

The hSpCas9 activator plasmid was cloned into a lentiviral vector under the expression of the hEF1a promoter (pLenti-EF1a-Cas9-NLS-VP64). The hSpCas9 repressor plasmid was cloned into the same vector (pLenti-EF1a-SID4x-NLS-Cas9-NLS). Guide sequences (20bp) targeted to the KLF4 locus are: 5'-GCGCGCTCCACACAACTCAC, 5'-GCAAAAATAGACAATCAGCA, GAAGGATCTCGGCCAATTTG. Spacer sequences for guide RNAs targeted to the SOX2 locus are: 5'-GCTGCCGGGTTTTGCATGAA, 5'-CCGGGCCCGCAGCAAACTTC, 5'-GGGGCTGTCAGGGAATAAAT

***Neurog2* activation with dCas9-VP64 and standard sgRNA backbone**

The locus for the mouse gene *Neurog2* was selected for testing dCas9-VP64 transcriptional activation in Neuro 2A cells. 7 sgRNA targeting sequences were selected from region 300 bp upstream to 100 bp downstream of the annotated *Neurog2* transcriptional start site (**Supplementary Table 11**). These guide sequences were cloned by the golden gate method into the px362 plasmid containing the U6 promoter upstream of a chimeric sgRNA cloning backbone.

Neuro 2A cells were cultured in medium consisting of a 1:1 ratio of D5 (100:5:1 DMEM high glucose w/Glutamax and sodium pyruvate: fetal bovine serum: penecilin/streptomycin). For transfection, 120,000 cells in 0.5 mL culture medium were plated in each well of a 24 well plate. After 22 h in culture, each well of Neuro 2A cells was transfected with 1 μg of DNA using 5 μL of Lipofectamine following the recommended steps of the manufacturer (Life Technologies). For all dCas9-VP64 + sgRNA samples, 0.5 μg dCas9-VP64 plasmid and 0.5 μg sgRNA expression plasmid were used. For control samples, plasmid masses were replaced with GFP expressing plasmid as necessary. 4 h after transfection, the medium was replaced with 1 mL culture medium. 48 h after transfection, RNA was purified using the Macherey Nagel Nucleospin 96 RNA kit. Reverse transcription was performed with qScript cDNA supermix according to the manufacturer's protocol. qPCR was performed using Taqman probes and Taqman Fast Advanced Master Mix from Life Technologies on a Roche LightCycler 480 II real-time PCR machine.

Validation of human gene activation targets with dCas9-VP64 standard sgRNA backbone

8 previously published sgRNA sequences[43, 47] (**Supplementary Table 11**) were selected for testing in 293FT cells using our pXRP057 dCas9-VP64 activator construct. Guide sequences were cloned by the golden gate method into the px362 plasmid containing the U6 promoter upstream of a chimeric sgRNA cloning backbone.

293FT cells were cultured in D10 medium (100:10:1 ratio of DMEM high glucose w/Glutamax and sodium pyruvate: fetal bovine serum: 100x HEPES solution). For transfection, 100,000 cells in 0.5 mL culture medium were plated in each well of a 24 well plate. After 22 h in culture, each well of 293FT cells was transfected with 1 μg of DNA using 5 μL of Lipofectamine following the recommended steps of the manufacturer (Life Technologies). For all dCas9-VP64 + sgRNA samples, 0.5 μg dCas9-VP64 plasmid and 0.5 ug sgRNA expression plasmid were used. For control samples plasmid masses were replaced by GFP expressing plasmid as necessary. 4 h after transfection, the medium was replaced with 1 mL culture medium. 48 h after transfection, RNA was purified using the Macherey Nagel Nucleospin 96 RNA kit. Reverse transcription was performed with qScript cDNA supermix according to the manufacturer's protocol. qPCR was performed using Taqman probes and Taqman Fast Advanced Master Mix from Life Technologies on a Roche LightCycler 480 II real-time PCR machine.

CasLITEs

The CasLITE system consists of 3 components: a dCas9-CIB1 fusion protein, the CRY2PHR-VP64, and an sgRNA guide sequence. For the experiments described here, 3 different versions of dCas9-CIB1 were synthesized: (1) dCas9-GS-NLS-CIB1 comprising the double mutant (D10A H840A) dCas9 from our pXRP057 plasmid, a glycine serine linker, SV40 nuclear localization signal, and CIB1 from LITE1.0. (2) dCas9-GS-NLS-NLS-CIB1 comprising the double mutant dCas9, glycine serine linker, 2 SV40 NLS sequences, and CIB1 from LITE1.0. (3) dCas9-GS-CIB1(mNLS d318-334) comprising the double mutant dCas9, glycine serine linker from LITE2.0, and CIB1 from LITE2.0

sgRNA sequences targeted to ASCL1 and MYOD1, previously validated with a constitutive dCas9-VP64 activator, were selected for testing the CasLITE systems in human cells. Guide sequences were cloned by the golden gate method into the px362 plasmid containing the U6 promoter upstream of a chimeric sgRNA cloning backbone.

293FT cells were cultured in D10 medium without HEPES (10:1 ratio of DMEM high glucose w/Glutamax and sodium pyruvate: fetal bovine serum). For transfection, 100,000 cells in 0.5 mL culture medium were plated in each well of a 24 well plate. After 22 h in culture, each well of 293FT cells was transfected with 1.5 µg of DNA using 7.5 uL of Lipofectamine following the recommended steps of the manufacturer (Life Technologies). For all CasLITE + sgRNA samples, 0.5 µg dCas9-CIB1 plasmid, 0.5 µg of CRY2PHR-VP64 plasmid (LITE1.0 or LITE2.0 to match CIB1), and 0.5 µg sgRNA expression plasmid were used. Experiments with 4 sgRNAs used equal masses of each of 4 sgRNAs totaling 0.5 µg. For control samples, plasmid masses were replaced by GFP expressing plasmid as necessary. 4 h after transfection, the medium was replaced with 1 mL culture medium. 48 h after transfection excitation was started on light stimulated samples using 5 mW/cm^2 475 nm blue light. 12 h after starting stimulation, all samples, including all controls, were harvested for RNA using the Macherey Nagel Nucleospin 96 RNA kit. Reverse transcription was performed with qScript cDNA supermix according to the manufacturer's protocol. qPCR was performed using Taqman probes and Taqman Fast Advanced Master Mix from Life Technologies on a Roche LightCycler 480 II real-time PCR machine.

Design and validation of MS2/dCas9 activators: *Neurog2* locus

The MS2 domain RNA binding sequence, shown in detail in Supplementary Sequences, was inserted into the nucleotide sequence of the standard chimeric sgRNA backbone. The MS2 stemloop was inserted into either the tetraloop of the sgRNA, stemloop 2 of the sgRNA, or into both sites[49]. These 3 sgRNA backbones, as well as the standard sgRNA backbone, were synthesized as ultramers (Integrated DNA Technologies) with the sgRNA0003 sequence, ATACGATGAAAAGAATAAGC, as a guide sequence targeted to the mouse *Neurog2* promoter. The ultramers were used in combination with a U6 promoter primer to generate PCR product[38] which included, from 5' to 3': the U6 promoter, the sgRNA0003 sequence, and each respective MS2 or control backbone.

Neuro 2a cells (Sigma-Aldrich) were grown in media containing a 1:1 ratio of OptiMEM (Life Technologies) to high-glucose DMEM with GlutaMax and sodium pyruvate (Life Technologies) supplemented with 5% HyClone heat-inactivated FBS (Thermo Scientific), 1% penicillin/streptomycin (Life Technologies), and passaged at 1:5 every 2 days. 120,000 cells were plated in each well of a 24-well plate 18–20 h before transfection. Cells were transfected with Lipofectamine transfection reagent (Life Technologies) according to the manufacturer's instructions. Plasmid DNA was used for transfection of MS2-vp62 and Cas9 constructs, while PCR product was transfected for the guide RNA expression cassette. 48 h after transfection, RNA was extracted using the Rneasy kit (Qiagen) according to manufacturer's instructions and 1 μg of RNA per sample was reverse-transcribed using qScript (Quanta Biosystems). Relative mRNA

levels were measured by reverse transcription and quantitative PCR (qRT–PCR) using TaqMan probes specific for the targeted gene as well as GAPDH as an endogenous control (Life Technologies). ΔΔCt analysis was used to obtain fold-changes relative to negative controls transfected with GFP only.

MS2 1.0 Optimization: ASCL1 and MYOD1 Targets

Human HEK293FT cells were maintained in high-glucose DMEM with GlutaMax and sodium pyruvate (Life Technologies) supplemented with 10% heat-inactivated characterized HyClone fetal bovine serum (Thermo Scientific) and 1% penicillin/streptomycin (Life Technologies). Cells were passaged daily at a ratio 1 to 2 or 1 to 2.5. 20,000 HEK293FT cells were plated in 100 μL of culture medium in poly-D-lysine coated 96-well plates (BD biosciences). 24 h after plating, cells were transfected with a 1:1:1 mass ratio of:

- sgRNA backbone plasmid with gene specific targeting sequence or pUC19 control plasmid
- MS2-VP64 plasmid or MS2-p65 plasmid or pUC19 control plasmid
- dCas9 plasmid or dCas9-VP64 plasmid or dCas9-p65 plasmid or pUC19 control plasmid

Total plasmid mass per well was 0.3 μg. Transfection was performed with 1.5 μL Lipfectamine 2000 (Life Technologies), according to the manufacturer's instructions. Culture medium was changed 5 h after transfection.

48 h after transfection, cell lysis and reverse transcription were performed using a Cells-to-Ct kit (Life Technologies). Gene expression levels were quantified using Taqman qPCR probes (Life technologies) and Fast Advanced Master Mix (Life Technologies). ASCL1 and MYOD1 expression levels were calculated relative to GAPDH expression level. Fold gene expression levels were determined by comparison to samples transfected with GFP plasmid only.

Appendix A: Supplementary Tables

Supplementary Table 1 | HDAC Recruiter Effector Domains

Subtype/ Complex	Name	Substrate (if known)	Effect (if known)	Organism	Full size (aa)	Selected truncation (aa)	Final size (aa)	Catalytic domain
Sin3a	**MeCP2**	-	-	*R. norvegicus*	492	207-492[62]	286	-
Sin3a	**MBD2b**	-	-	*H. sapiens*	262	45-262[63]	218	-
Sin3a	**Sin3a**	-	-	*H. sapiens*	1273	524-851[64]	328	627-829: HDAC1 interaction
NcoR	**NcoR**	-	-	*H. sapiens*	2440	420-488[65]	69	-
NuRD	**SALL1**	-	-	*M. musculus*	1322	1-93[66]	93	-
CoREST	**RCOR1**	-	-	*H. sapiens*	482	81-300[67, 68]	220	-

Supplementary Table 2 | HDAC Effector Domains

Subtype/ Complex	Name	Substrate (if known)	Effect (if known)	Organism	Full size (aa)	Selected truncation (aa)	Final size (aa)	Catalytic domain
HDAC I	**HDAC8**	-	-	*X. laevis*	325	1-325	325	1-272: HDAC
HDAC I	**RPD3**	-	-	*S. cerevisiae*	433	19-340[69]	322	19-331: HDAC
HDAC IV	**MesoLo4**	-	-	*M. loti*	300	1-300[70]	300	-
HDAC IV	**HDAC11**	-	-	*H. sapiens*	347	1-347[71]	347	14-326: HDAC
HD2	**HDT1**	-	-	*A. thaliana*	245	1-211[72]	211	-
SIRT I	**SIRT3**	H3K9Ac H4K16Ac H3K56Ac	-	*H. sapiens*	399	143-399[73]	257	126-382: SIRT
SIRT I	**HST2**	-	-	*C. albicans*	331	1-331[74]	331	-
SIRT I	**CobB**	-	-	*E. coli (K12)*	242	1-242[75]	242	-
SIRT I	**HST2**	-	-	*S. cerevisiae*	357	8-298[76]	291	-
SIRT III	**SIRT5**	H4K8Ac H4K16Ac	-	*H. sapiens*	310	37-310[77]	274	41-309: SIRT
SIRT III	**Sir2A**	-	-	*P. falciparum*	273	1-273[78]	273	19-273: SIRT
SIRT IV	**SIRT6**	H3K9Ac H3K56Ac	-	*H. sapiens*	355	1-289[79]	289	35-274: SIRT

Supplementary Table 3 | Histone Methyltransferase (HMT) Effector Domains

Subtype/ Complex	Name	Substrate (if known)	Effect (if known)	Organism	Full size (aa)	Selected truncation (aa)	Final size (aa)	Catalytic domain
SET	**NUE**	H2B, H3, H4	-	*C. trachomatis*	219	1-219[80]	219	-
SET	**vSET**	-	H3K27me3	*P. bursaria chlorella virus*	119	1-119[81]	119	4-112: SET2
SUV39 family	**EHMT2/ G9A**	H1.4K2, H3K9, H3K27	H3K9me1/ 2, H1K25me1	*M. musculus*	1263	969-1263[82]	295	1025-1233: preSET, SET, postSET
SUV39	**SUV39H1**	-	H3K9me2/ 3	*H. sapiens*	412	79-412[83]	334	172-412: preSET, SET, postSET
Suvar3-9	**dim-5**	-	H3K9me3	*N. crassa*	331	1-331[84]	331	77-331: preSET, SET, postSET
Suvar3-9 (SUVH subfamily)	**KYP**	-	H3K9me1/ 2	*A. thaliana*	624	335-601[85]	267	-
Suvar3-9 (SUVR subfamily)	**SUVR4**	H3K9me1	H3K9me2/ 3	*A. thaliana*	492	180-492[86]	313	192-462: preSET, SET, postSET
Suvar4-20	**SET4**	-	H4K20me3	*C. elegans*	288	1-288[87]	288	-
SET8	**SET1**	-	H4K20me1	*C. elegans*	242	1-242[87]	242	-
SET8	**SETD8**	-	H4K20me1	*H. sapiens*	393	185-393[88]	209	256-382: SET
SET8	**TgSET8**	-	H4K20me1 /2/3	*T. gondii*	1893	1590-1893[89]	304	1749-1884: SET

Supplementary Table 4 | Histone Methyltransferase (HMT) Recruiter Effector Domains

Subtype/ Complex	Name	Substrate (if known)	Effect (if known)	Organism	Full size (aa)	Selected truncation (aa)	Final size (aa)	Catalytic domain
-	**Hp1a**	-	H3K9me3	*M. musculus*	191	73-191[90]	119	121-179: chromo-shadow
-	**PHF19**	-	H3K27me3	*H. sapiens*	580	(1-250) + GGSG linker + (500-580)	335	163-250[91]: PHD2
-	**NIPP1**	-	H3K27me3	*H. sapiens*	351	1-329[92]	329	310-329: EED

Supplementary Table 5 | Histone Acetyltransferase Inhibitor Effector Domains

Subtype/ Complex	Name	Substrate (if known)	Effect (if known)	Organism	Full size (aa)	Selected truncation (aa)	Final size (aa)	Catalytic domain
-	**SET/ TAF-1β**	-	-	*M. musculus*	289	1-289[93]	289	-

Supplementary Table 6 | genomic sequences targeted by TALEs (5' to 3')

5-HT1B	TATCTGAACTCTCC
5-HTT	TGTCTGTCTTGCAT
Arc	TGGCTGTTGCCAGG
BDNF	TACCTGGAGCTAGC
c-Fos	TACACAGGATGTCC
DNMT3a	TTGGCCCTGTGCAG
DNMT3b	TAGCGCAGCGATCG
gad65	TATTGCCAAGAGAG
gad67	TGACTGGAACATAC
GR (GCR, NR3C1)	TGATGGACTTGTAT
HAT1	TGGACCTTCTCCCT
HCRTR1	TAGGTCTCCTGGAG
HCRTR2	TGGCTCAGGAACTT
HDAC1	TTCTCTAAGCTGCC
HDAC2	TGAGCCCTGGAGGA
HDAC4	TGCCTAAGATGGAG
JMJD2A	TGTAGTGAGTGTTC
MCH-R1	TGTCTAGGTGATGT
NET	TCTCTGCTAGAAGG
Scn1a	TCTAGGTCAAGTGT
SIRT1	TCCTCTGCTCCGCT
tet1	TCTAGGAGTGTAGC
tet3	TGCCTGGCTGCTGG
5-HT1B	TATCTGAACTCTCC
Grm2	TCAGAGCTGTCCTC
Grm5	TGCAAGAGTAGGAG
5-HT2A	TAGTGACTGATTCC
Grin2a	TTGGAGGAGCACCA
Neurog2	TGAATGATGATAATAC

Supplementary Table 7 | Taqman probes (Life Technologies)

Target	Species	Probe #
Neurog2	mouse	Mm00437603_g1
Grm5 (mGluR5)	mouse	Mm00690332_m1
Grm2 (mGluR2)	mouse	Mm01235831_m1
Grin2a (NMDAR2A)	mouse	Mm00433802_m1
GAPD (GAPDH)	mouse	4352932E
KLF4	human	Hs00358836_m1
GAPD (GAPDH)	human	4352934E
WPRE	custom	
5-HT1A	mouse	Mm00434106_s1
5-HT1B	mouse	Mm00439377_s1
5-HTT	mouse	Mm00439391_m1
Arc	mouse	Mm00479619_g1
BDNF	mouse	Mm04230607_s1
c-Fos	mouse	Mm00487425_m1
CBP/P300	mouse	Mm01342452_m1
CREB	mouse	Mm00501607_m1
CRHR1	mouse	Mm00432670_m1
DNMT1	mouse	Mm01151063_m1
DNMT3a	mouse	Mm00432881_m1
DNMT3b	mouse	Mm01240113_m1
egr-1 (zif-268)	mouse	Mm00656724_m1
Gad65	mouse	Mm00484623_m1
Gad67	mouse	Mm00725661_s1
GR (GCR, NR3C1)	mouse	Mm00433832_m1
HAT1	mouse	Mm00509140_m1
HCRTR1	mouse	Mm01185776_m1
HCRTR2	mouse	Mm01179312_m1
HDAC1	mouse	Mm02391771_g1
HDAC2	mouse	Mm00515108_m1
HDAC4	mouse	Mm01299557_m1
JMJD2A	mouse	Mm00805000_m1
M1 (CHRM1)	mouse	Mm00432509_s1
MCH-R1	mouse	Mm00653044_m1
NET (SLC6A2)	mouse	Mm00436661_m1
NR2B subunit	mouse	Mm00433820_m1
OXTR	mouse	Mm01182684_m1
Scn1a	mouse	Mm00450580_m1
SIRT1	mouse	Mm00490758_m1
Tet1	mouse	Mm01169087_m1
Tet2	mouse	Mm00524395_m1
Tet3	mouse	Mm00805756_m1

Supplementary Table 8 | Antibodies

Primary Antibodies

Target	Host	Clone #	Manufacturer	Product #	IsoType	Concentration
mGluR2	mouse	mG2Na-s	Abcam	Ab15672	IgG	1:1000
α-tubulin	mouse	B-5-1-2	Sigma-Aldrich	T5168	IgG1	1:20000
NeuN	mouse	A60	Millipore	MAB377	IgG1	1:200
HA (Alexa Fluor 594)	mouse	6E2	Cell Signaling	3444	IgG1	1:100
GFP	chicken	polyclonal	Aves Labs	GFP-1020	IgY	1:500

Secondary Antibodies

Target	Host	Conjugate	Manufacturer	Product #		Concentration
mouse IgG	goat	HRP	Sigma-Aldrich	A9917		1:5000-10000
mouse IgG	goat	Alexa Fluor 594	Life Technologies	A11005		1:1000
chicken IgG	Goat	Alexa Fluor 488	Life Technologies	A11039		1:1000

ChIP Antibodies

Target	Host	Epitope	Manufacturer	Product #	IsoType	Concentration
H3K9me1	mouse	1-18	Millipore	17-680	IgG	2 µl/IP
H3K9me2	mouse	1-18	Millipore	17-681	IgG	4 µl/IP
H3K9Ac	rabbit	polyclonal	Millipore	17-658	IgG	3 µg/IP
H4K20me1	rabbit	15-24	Millipore	17-651	IgG	4 µg/IP
H4K8Ac	rabbit	polyclonal	Millipore	17-10099	IgG	1.5 µl/IP
H4K20me3	rabbit	18-22	Millipore	17-671	IgG	7 µl/IP
H3K27me3	rabbit	polyclonal	Millipore	17-622	IgG	4 µg/IP

Supplementary Table 9 | Viral transduction and light stimulation parameters for *in vivo* LITE-mediated activation of *Grm2* in the mouse infralimbic cortex (ILC). *Grm2* mRNA levels in the ipsilateral LITE-expressing hemisphere are compared with the contralateral mCherry-expressing control hemisphere for all three experimental conditions shown in **Fig. 8d**.

Experimental condition	ILC Hemisphere (ipsilateral)		ILC Hemisphere (contralateral)
	AAV vector	**Light stimulation**	**AAV vector**
GFP	GFP	yes	mCherry
LITEs / no Light	TALE-CIB1::CRY2PHR-VP64	no	mCherry
LITEs / + Light	TALE-CIB1::CRY2PHR-VP64	yes	mCherry

Supplementary Table 10 | qPCR primers used for CHIP-qPCR

target	**Primers**
Grm2 promoter	Forward: CTGTGCTGAAGGATCTGGGG Reverse: ATGCTGCAGGCATAGGACAA
Neurog2 promoter	Forward: GAGGGGGAGAGGGACTAAAGA Reverse: GCTCTCCCTCCCCAGCTTA
Myt-1 promoter control	Cell Signaling Technologies SimpleChIP® Mouse MYT-1 Promoter Primers #8985
RPL30 Intron 2 control	Cell Signaling Technologies SimpleChIP® Mouse RPL30 Intron 2 Primers #7015

Supplementary Table 11 | sgRNA guide sequences

sgRNA ID	Guide sequence	Gene	Organism
sgRNA0001	TGGTTCAGTGGCTGCGTGTC	*Neurog2*	*Mus musculus*
sgRNA0002	TGTTTTCTTGGTGGTATATA	*Neurog2*	*Mus musculus*
sgRNA0003	ATACGATGAAAAGAATAAGC	*Neurog2*	*Mus musculus*
sgRNA0004	GGGGGAGAGGGACTAAAGAA	*Neurog2*	*Mus musculus*
sgRNA0005	GGGCGGGGGAAGGGTAGGTG	*Neurog2*	*Mus musculus*
sgRNA0006	ATTAGATAAAGGGGGGACGG	*Neurog2*	*Mus musculus*
sgRNA0007	CGGCTTTAACTGGAGTGCCT	*Neurog2*	*Mus musculus*
sgRNA0008	GAGCAGCGTCTTCGAGAGTG	VEGFA	*Homo sapiens*
sgRNA0009	GGTGAGTGAGTGTGTGCGTG	VEGFA	*Homo sapiens*
sgRNA0010	GCGCGGCGCGGAAGGGGTTA	NTF3	*Homo sapiens*
sgRNA0011	GCGGCGCGGCGCGGGCCGGC	NTF3	*Homo sapiens*
sgRNA0012	GCAGCCGCTCGCTGCAGCAG	ASCL1	*Homo sapiens*
sgRNA0013	ATGGAGAGTTTGCAAGGAGC	ASCL1	*Homo sapiens*
sgRNA0014	GGGCCCCTGCGGCCACCCCG	MYOD1	*Homo sapiens*
sgRNA0015	GAGGTTTGGAAAGGGCGTGC	MYOD1	*Homo sapiens*
sgRNA0021	TTGTACTCTCTGAGGTGCTC	IL1RN	*Homo sapiens*
sgRNA0022	TACGCAGATAAGAACCAGTT	IL1RN	*Homo sapiens*
sgRNA0023	GCATCAAGTCAGCCATCAGC	IL1RN	*Homo sapiens*
sgRNA0024	TGAGTCACCCTCCTGGAAAC	IL1RN	*Homo sapiens*

Appendix B: Supplementary Sequences

> TALE(*Ngn2*)-NLS-CRY2

MSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAAD
APPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHH
EALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTV
AGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALE
TVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQ
ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGG
KQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASN
NGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLT
PEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQ
AHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPH
APALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVG
VTELEARSGTLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEG
DQTRASASPKKKRKVEASKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYP
GRASRWWMKQSLAHLSQSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDH
TVKEKLVERGISVQSYNGDLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITA
AAEAIWACSIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTS
LLSPYLHFGEISVRHVFQCARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLL
SHLRFFPWDADVDKFKAWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWK
WGMKYFWDTLLDADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELAR
LPTEWIHHPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPDEIVADS
FEALGANTIKEPGLCPSVSSNDQQVPSAVRYNGSKRVKPEEEEERDMKKSRGFDERELFSTAESSSS
SSVFFVSQSCSLASEGKNLEGIQDSSDQITTSLGKNG

> TALE(*Ngn2*)-NLS-CRY2PHR

MSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAAD
APPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHH
EALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTV
AGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALE
TVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQ
ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGG
KQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASN
NGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLT
PEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQ
AHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPH
APALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVG
VTELEARSGTLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEG
DQTRASASPKKKRKVEASKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYP
GRASRWWMKQSLAHLSQSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDH
TVKEKLVERGISVQSYNGDLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITA
AAEAIWACSIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTS
LLSPYLHFGEISVRHVFQCARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLL
SHLRFFPWDADVDKFKAWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWK
WGMKYFWDTLLDADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELAR
LPTEWIHHPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAP

> CIB1-NLS-VP64_2A_GFP

MNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPM
MYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEEEKSKITE
QNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISE
RMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTP
MTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSH
VQNLYGNLGVASPKKKRKVEASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDM
LGSDALDDFDLDMLINSRGSGEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKF
SVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGY
VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQ
KNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEF
VTAAGITLGMDELYK

> CIBN-NLS-VP64_2A_GFP

MNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPM
MYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEEEKSKITE
QNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIASPKKKRKVEASGSGRADALDDFDL
DMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINSRGSGEGRGSLLTCGDVEEN
PGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTT
LTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGI
DFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPV
LLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

> CIB1-NLS-VP16_2A_GFP

MNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPM
MYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEEEKSKITE
QNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISE
RMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTP
MTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSH
VQNLYGNLGVASPKKKRKVEASAPPTDVSLGDELHLDGEDVAMAHADALDDFDLDMLGDGDSP
GPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGGEFPGIRRSRGSGEGRGSLLTCGDVEENPG
PVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLT
YGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

> CIB1-NLS-p65_2A_GFP

MNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPM
MYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEEEKSKITE
QNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISE
RMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTP
MTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSH
VQNLYGNLGVASPKKKRKVEASPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLAQPPAPAPVL
TPGPPQSLSAPVPKSTQAGEGTLSEALLHLQFDADEDLGALLGNSTDPGVFTDLASVDNSEFQQLL
NQGVSMSHSTAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSAL
LSQISSSGQSRGSGEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG
DATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFF
KDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNF
KIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITL
GMDELYK

.> **HA-TALE(12mer)-NLS-VP64_2A_GFP**

MYPYDVPDYAVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTV
AVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVT
AVEAVHAWRNALTGAPLNLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASX
XGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAI

ASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGRPALESIVAQL
SRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASPKKKRK
VEASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINSR
GSGEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL
KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTR
AEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

> **HA-TALE(12mer)-NLS-SID4X_2A_phiLOV2.1**
MYPYDVPDYAVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTV
AVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVT
AVEAVHAWRNALTGAPLNLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASX
XGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGRPALESIVAQL
SRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASPKKKRK
VEASPKKKRKVEASGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLER
REREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLER
REREAEHGYASMLPSRSRGSGEGRGSLLTCGDVEENPGPIEKSFVITDPRLPDYPIIFASDGFLELTE
YSREEIMGRNARFLQGPETDQATVQKIRDAIRDQRETTVQLINYTKSGKKFWNLLHLQPVRDRKG
GLQYFIGVQLVGSDHV

> **HA-TALE(12mer)-NLS-CIB1**
MYPYDVPDYAVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTV
AVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVT
AVEAVHAWRNALTGAPLNLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASX
XGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGRPALESIVAQL
SRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASPKKKRK
VEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNL
PMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEEEKSKI
TEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKI
SERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAS
TPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDS
HVQNLYGNLGV

> **CRY2PHR-NLS-VP64_2A_GFP**
MKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPGRASRWWMKQSLAHLS
QSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTVKEKLVERGISVQSYNG
DLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITAAAEAIWACSIEELGLENE
AEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLLSPYLHFGEISVRHVFQ
CARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLLSHLRFFPWDADVDKFK
AWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLLDADLE
CDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIHHPWDAPLTVLK
ASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKVEASGSGRADALDDFD

LDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINSRGSGEGRGSLLTCGDVEE
NPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVT
TLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKG
IDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGP
VLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKV

> **CRY2PHR-NLS-SID4X_2A_phiLOV2.1**
MKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPGRASRWWMKQSLAHLS
QSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTVKEKLVERGISVQSYNG
DLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITAAAEAIWACSIEELGLENE
AEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLLSPYLHFGEISVRHVFQ
CARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLLSHLRFFPWDADVDKFK
AWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLLDADLE
CDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIHHPWDAPLTVLK
ASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKVEASGSGMNIQMLLEA
ADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEA
ADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPSRSRGSGEGRGS
LLTCGDVEENPGPIEKSFVITDPRLPDYPIIFASDGFLELTEYSREEIMGRNARFLQGPETDQATVQKI
RDAIRDQRETTVQLINYTKSGKKFWNLLHLQPVRDRKGGLQYFIGVQLVGSDHV

> **TALE(*KLF4*)-NLS-CRY2PHR**
MSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAAD
APPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHH
EALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTV
AGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNGGGKQALE
TVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQ
ALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIG
GKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLT
PEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIK
RTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELE
ARSGTLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEGDQTRA
SASPKKKRKVEASKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPGRASR
WWMKQSLAHLSQSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTVKEK
LVERGISVQSYNGDLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITAAAEAI
WACSIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLLSPY
LHFGEISVRHVFQCARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLLSHLRF
FPWDADVDKFKAWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKY
FWDTLLDADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIH
HPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAP

> **HA-NLS-TALE(*p11*, N136)-SID**
MYPYDVPDYASPKKKRKVEASVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVA
LSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQ
LLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGK
QALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHD
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQV

VAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTN
RRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARS
GTLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEGDQTRASAS
GSGMNIQMLLEAADYLERREREAEHGYASMLP.

> **HA-NLS-TALE(*p11*, N136)-SID4X**
MYPYDVPDYASPKKKRKVEASVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVA
LSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQ
LLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGK
QALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHD
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQV
VAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTN
RRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARS
GTLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEGDQTRASAS
GSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLP
GSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLP
SR

>HA-TALE(*ng2*, C63)-GS-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKE
VASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSM
WDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-wNES-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA

HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLYPER
LRRILTNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR
EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEV
ASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMW
DSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-mNES-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLQLPP
LERLTLNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR
EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEV
ASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMW
DSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-ptk2NES-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLDLAS
LILNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLP
MMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEKSKIT
EQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKIS
ERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAST
PMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDS
HVQNLYGNLGV

> HA-TALE(*ng2*, C63)-mapkkNES-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV

RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLQKK
LEELELNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR
EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEV
ASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMW
DSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-cib1Δ3-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKE
VASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSS

> HA-TALE(*ng2*, C63)-wNLS-cib1Δ3-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLYPER
LRRILTNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR

EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEV
ASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSS

> HA-TALE(*ng2*, C63)-mNLS-cib1Δ3-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASPKKK
RKVEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKE
VASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSS

> HA-TALE(*ng2*, C63)-GS-cib1-mutNLS-mutbHLH

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARAAQATDSHSIAEAV
AREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAK
EVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPS
MWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-wNES-cib1-mutNLS-mutbHLH

SRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADA
PPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHE
ALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVA
GELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNGGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQA
LETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGG
KQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASN
IGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQV
VAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTP
EQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG

LTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPAL
AALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLYPERLRRILTNGA
IGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGE
TTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEKSKITEQNNG
STKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARAAQATDSHSIAEAVAREKISERMK
FLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTPMTV
VPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSHVQN
LYGNLGV

> HA-TALE(*ng2*, C63)-GS-cib1Δ1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPSPEMVLSGYSHE
MVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-wNLS-cib1Δ1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLYPER
LRRILTNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR
EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSGEEEKSKITEQNNGSTKSIKKMKH
KAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKFLQDLVPGCD
KITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVN
PMQQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-cib1Δ2-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG

GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKE
VASTPMTVVPSPEMVLSGYGGSPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-wNES-cib1Δ2-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLYPER
LRRILTNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR
EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEV
ASTPMTVVPSPEMVLSGYGGSPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-NLS-cib1-mutNLS-mutbHLH

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASPKKK
RKVEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARAAQATDSHSIAEAV
AREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAK
EVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPS
MWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-NLS-cib1Δ1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASPKKK
RKVEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPSPEMVLSGYSHE
MVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-NLS-cib1Δ2-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASLYPER
LRRILTNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGL
NLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEK
SKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRR
EKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEV
ASTPMTVVPSPEMVLSGYGGSPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-iNES1-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLLYPERLRRILTNPTFDSPLAGFFADSSMITGGE
MDSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRD

DLVEEGEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQAT
DSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPD
FDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLS
CFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-iNES2-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDLYPERLR
RILTSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNR
DDLVEEGEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQA
TDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRP
DFDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPL
SCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-iNES3-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGLYPERLRRILTDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNR
DDLVEEGEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQA
TDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRP
DFDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPL
SCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-iNES4-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL

TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAAAKKMTMNRDDLVEEGL
YPERLRRILTEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRG
QATDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNP
RPDFDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSD
PLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-iNES5-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKLYPERLRRILTMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQA
TDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRP
DFDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPL
SCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-GS-iNES6-cib1-mutNLS
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASGGGG
SGGGGSNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTLYPERLRRILTKELEKTDYIHVRARRGQA
TDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRP
DFDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPL
SCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-NLS-cib1Δ1

MYPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDE
VQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPK
VRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGA
RALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASN
GGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQV
VAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP
EQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQ
AHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRL
LPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIV
AQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASSPKK
KRKVEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTA
GLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEE
EKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERV
RREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPSPEMVLSGYSHE
MVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

> HA-TALE(*ng2*, C63)-NLS-cib1Δ2
YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKV
RSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGAR
ALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQA
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVA
QLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAASSPKKK
RKVEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAG
LNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDLVEEGEEE
KSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR
REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKE
VASTPMTVVPSPEMVLSGYGGSPLSCFNNGEAPSMWDSHVQNLYGNLGV

> **alpha-importin-NLS-CRY2PHR-NLS-VP64_2A_GFP**
MKRPAATKKAGQAKKKKKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYP
GRASRWWMKQSLAHLSQSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDH
TVKEKLVERGISVQSYNGDLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITA
AAEAIWACSIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTS
LLSPYLHFGEISVRHVFQCARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLL
SHLRFFPWDADVDKFKAWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWK
WGMKYFWDTLLDADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELAR
LPTEWIHHPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKK
RKVEASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN
SRGSGEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKL
TLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIED
GSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK
V

> **mutNES-CRY2PHR-NLS-VP64_2A_GFP**

MEQKLISEEDLKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPGRASRWW
MKQSLAHLSQSLKAAGSDATLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTVKEKLVE
RGISVQSYNGDLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITAAAEAIWAC
SIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLLSPYLHF
GEISVRHVFQCARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLLSHLRFFP
WDADVDKFKAWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYF
WDTLLDADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIH
HPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKVEASG
SGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINSRGSGEG
RGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD
HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKV

> **CRY2PHR-NLS-VP64-NLS_2A_GFP**
MKMDKKTIVWFRRDLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPGRASRWWMKQSLAHLS
QSLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTVKEKLVERGISVQSYNG
DLLYEPWEIYCEKGKPFTSFNSYWKKCLDMSIESVMLPPPWRLMPITAAAEAIWACSIEELGLENE
AEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLLSPYLHFGEISVRHVFQ
CARMKQIIWARDKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSLLSHLRFFPWDADVDKFK
AWRQGRTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLLDADLE
CDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIHHPWDAPLTVLK
ASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKVEASGSGRADALDDFD
LDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINSPKKKRKVEASSRGSGEGR
GSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTG
KLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE
GDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD
HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKV

> Neurog2-TALE(N240,C63)-PYL

MSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAAD
APPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHH
EALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTV
AGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHNGGKQALE
TVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQA
LETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHNGG
KQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASN
GGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHNGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQV
VAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKR
TNRRIPERTSHRVAASMANSESSSSPVNEEENSQRISTLHHQTMPSDLTQDEFTQLSQSIAEFHTYQL
GNGRCSSLLAQRIHAPPETVWSVVRRFDRPQIYKHFIKSCNVSEDFEMRVGCTRDVNVISGLPANT
SRERLDLLDDDRRVTGFSITGGEHRLRNYKSVTTVHRFEKEEEEERIWTVVLESYVVDVPEGNSEE
DTRLFADTVIRLNLQKLASITEAMNRNNNNNNSSQVR

>ABI-NLS-VP64

MVPLYGFTSICGRRPEMEAAVSTIPRFLQSSSGSMLDGRFDPQSAAHFFGVYDGHGGSQVANYCR
ERMHLALAEEIAKEKPMLCDGDTWLEKWKKALFNSFLRVDSEIESVAPETVGSTSVVAVVFPSHIF
VANCGDSRAVLCRGKTALPLSVDHKPDREDEAARIEAAGGKVIQWNGARVFGVLAMSRSIGDRY
LKPSIIPDPEVTAVKRVKEDDCLILASDGVWDVMTDEEACEMARKRILLWHKKNAVAGDASLLA
DERRKEGKDPAAMSAAEYLSKLAIQRGSKDNISVVVVDLKPRRKLKSKPLNASPKKKRKVEASGS
GRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN

>hSpCas9(D10A,H840A)-Linker-NLS-VP64
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSAGGGGSGGGGSGGGGS
GPKKKRKVAAAGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDL
DMLIN

>SID4X-NLS-FLAG-Linker-hSpCas9(D10A,H840A)-NLS
MGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASM
LPGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASM
LPSRSAGPKKKRKVAAADYKDDDDKSRLEPGEKPYKCPECGKSFSQSGALTRHQRTHTRDKKYSI
GLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK
KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV
DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYD
DDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL
VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEET
ITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP
AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK
DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI
LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE
NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGK
SDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA
LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD
LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFV
EQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF
DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK

Epigenetic effector domain sequences

>hs_NCoR

ASSPKKKRKVEASMNGLMEDPMKVYKDRQFMNVWTDHEKEIFKDKFIQHPKNFGLIASYLERKS
VPDCVLYYYLTKKNENYKEF

>pf_Sir2A
ASSPKKKRKVEASMGNLMISFLKKDTQSITLEELAKIIKKCKHVVALTGSGTSAESNIPSFRGSSNSI
WSKYDPRIYGTIWGFWKYPEKIWEVIRDISSDYEIEINNGHVALSTLESLGYLKSVVTQNVDGLHE
ASGNTKVISLHGNVFEAVCCTCNKIVKLNKIMLQKTSHFMHQLPPECPCGGIFKPNIILFGEVVSSD
LLKEAEEEIAKCDLLLVIGTSSTVSTATNLCHFACKKKKKIVEINISKTYITNKMSDYHVCAKFSELT
KVANILKGSSEKNKKIMEF

>nc_DIM5
ASSPKKKRKVEASMEKAFRPHFFNHGKPDANPKEKKNCHWCQIRSFATHAQLPISIVNREDDAFL
NPNFRFIDHSIIGKNVPVADQSFRVGCSCASDEECMYSTCQCLDEMAPDSDEEADPYTRKKRFAYY
SQGAKKGLLRDRVLQSQEPIYECHQGCACSKDCPNRVVERGRTVPLQIFRTKDRGWGVKCPVNIK
RGQFVDRYLGEIITSEEADRRRAESTIARRKDVYLFALDKFSDPDSLDPLLAGQPLEVDGEYMSGP
TRFINHSCDPNMAIFARVGDHADKHIHDLALFAIKDIPKGTELTFDYVNGLTGLESDAHDPSKISEM
TKCLCGTAKCRGYLWEF

>sc_HST2
ASSPKKKRKVEASTEMSVRKIAAHMKSNPNAKVIFMVGAGISTSCGIPDFRSPGTGLYHNLARLKL
PYPEAVFDVDFFQSDPLPFYTLAKELYPGNFRPSKFHYLLKLFQDKDVLKRVYTQNIDTLERQAGV
KDDLIIEAHGSFAHCHCIGCGKVYPPQVFKSKLAEHPIKDFVKCDVCGELVKPAIVFFGEDLPDSFS
ETWLNDSEWLREKITTSGKHPQQPLVIVVGTSLAVYPFASLPEEIPRKVKRVLCNLETVGDFKANK
RPTDLIVHQYSDEFAEQLVEELGWQEDFEKILTAQGGMGEF

>hs_SIRT3
ASSPKKKRKVEASMVGAGISTPSGIPDFRSPGSGLYSNLQQYDLPYPEAIFELPFFFHNPKPFFTLAK
ELYPGNYKPNVTHYFLRLLHDKGLLLRLYTQNIDGLERVSGIPASKLVEAHGTFASATCTVCQRPF
PGEDIRADVMADRVPRCPVCTGVVKPDIVFFGEPLPQRFLLHVVDFPMADLLLILGTSLEVEPFASL
TEAVRSSVPRLLINRDLVGPLAWHPRSRDVAQLGDVVHGVESLVELLGWTEEMRDLVQRETGKL
DGPDKEF

>hs_NIPP1
ASSPKKKRKVEASMAAAANSGSSLPLFDCPTWAGKPPPGLHLDVVKGDKLIEKLIIDEKKYYLFGR
NPDLCDFTIDHQSCSRVHAALVYHKHLKRVFLIDLNSTHGTFLGHIRLEPHKPQQIPIDSTVSFGAST
RAYTLREKPQTLPSAVKGDEKMGGEDDELKGLLGLPEEETELDNLTEFNTAHNKRISTLTIEEGNL
DIQRPKRKRKNSRVTFSEDDEIINPEDVDPSVGRFRNMVQTAVVPVKKKRVEGPGSLGLEESGSRR
MQNFAFSGGLYGGLPPTHSEAGSQPHGIHGTALIGGLPMPYPNLAPDVDLTPVVPSAVNMNPAPN
PAVYNPEAVNEEF

>ct_NUE
ASSPKKKRKVEASMTTNSTQDTLYLSLHGGIDSAIPYPVRRVEQLLQFSFLPELQFQNAAVKQRIQ
RLCYREEKRLAVSSLAKWLGQLHKQRLRAPKNPPVAICWINSYVGYGVFARESIPAWSYIGEYTGI
LRRRQALWLDENDYCFRYPVPRYSFRYFTIDSGMQGNVTRFINHSDNPNLEAIGAFENGIFHIIIRAI
KDILPGEELCYHYGPLYWKHRKKREEFVPQEEEF

>hs_MBD2b
ASSPKKKRKVEASARYLGNTVDLSSFDFRTGKMMPSKLQKNKQRLRNDPLNQNKGKPDLNTTLPI
RQTASIFKQPVTKVTNHPSNKVKSDPQRMNEQPRQLFWEKRLQGLSASDVTEQIIKTMELPKGLQ
GVGPGSNDETLLSAVASALHTSSAPITGQVSAAVEKNPAVWLNTSQPLCKAFIVTDEDIRKQEERV
QQVRKILEDALMADILSRAADTEEMDIEMDSGDEAEF

>ca_HST2
ASSPKKKRKVEASMPSLDDILKPVAEAVKNGKKVTFFNGAGISTGAGIPDFRSPDTGLYANLAKLN
LPFAEAVFDIDFFKEDPKPFYTLAEELYPGNFAPTKFHHFIKLLQDQGSLKRVYTQNIDTLERLAGV
EDKYIVEAHGSFASNHCVDCHKEMTTETLKTYMKDKKIPSCQHCEGYVKPDIVFFGEGLPVKFFD

LWEDDCEDVEVAIVAGTSLTVFPFASLPGEVNKKCLRVLVNKEKVGTFKHEPRKSDIIALHDCDIV
AERLCTLLGLDDKLNEVYEKEKIKYSKAETKEIKMHEIEDKLKEEAHLKEDKHTTKVDKKEKQND
ANDKELEQLIDKAKAEF

>hs_PHF19
ASSPKKKRKVEASMENRALDPGTRDSYGATSHLPNKGALAKVKNNFKDLMSKLTEGQYVLCRW
TDGLYYLGKIKRVSSSKQSCLVTFEDNSKYWVLWKDIQHAGVPGEEPKCNICLGKTSGPLNEILIC
GKCGLGYHQQCHIPIAGSADQPLLTPWFCRRCIFALAVRKGGALKKGAIARTLQAVKMVLSYQPE
ELEWDSPHRTNQQQCYCYCGGPGEWYLRMLQCYRCRQWFHEACTQCLNEPMMFGDRFYLFFCS
VCNQGPGGSGSDSSAEGASVPERPDEGIDSHTFESISEDDSSLSHLKSSITNYFGAAGRLACGEKYQ
VLARRVTPEGKVQYLVEWEGTTPYEF

>hs_HDAC11
ASSPKKKRKVEASMLHTTQLYQHVPETRWPIVYSPRYNITFMGLEKLHPFDAGKWGKVINFLKEE
KLLSDSMLVEAREASEEDLLVVHTRRYLNELKWSFAVATITEIPPVIFLPNFLVQRKVLRPLRTQTG
GTIMAGKLAVERGWAINVGGGFHHCSSDRGGGFCAYADITLAIKFLFERVEGISRATIIDLDAHQG
NGHERDFMDDKRVYIMDVYNRHIYPGDRFAKQAIRRKVELEWGTEDDEYLDKVERNIKKSLQEH
LPDVVVYNAGTDILEGDRLGGLSISPAGIVKRDELVFRMVRGRRVPILMVTSGGYQKRTARIIADSI
LNLFGLGLIGPESPSVSAQNSDTPLLPPAVPEF

>ml_MesoLo4
ASSPKKKRKVEASMPLQIVHHPDYDAGFATNHRFPMSKYPLLMEALRARGLASPDALNTTEPAPA
SWLKLAHAADYVDQVISCSVPEKIEREIGFPVGPRVSLRAQLATGGTILAARLALRHGIACNTAGG
SHHARRAQGAGFCTFNDVAVASLVLLDEGAAQNILVVDLDVHQGDGTADILSDEPGVFTFSMHG
ERNYPVRKIASDLDIALPDGTGDAAYLRRLATILPELSARARWDIVFYNAGVDVHAEDRLGRLALS
NGGLRARDEMVIGHFRALGIPVCGVIGGGYSTDVPALASRHAILFEVASTYAEF

>pbcv1_vSET
ASSPKKKRKVEASMFNDRVIVKKSPLGGYGVFARKSFEKGELVEECLCIVRHNDDWGTALEDYLF
SRKNMSAMALGFGAIFNHSKDPNARHELTAGLKRMRIFTIKPIAIGEEITISYGDDYWLSRPRLTQN
EF

>at_KYP
ASSPKKKRKVEASDISGGLEFKGIPATNRVDDSPVSPTSGFTYIKSLIIEPNVIIPKSSTGCNCRGSCT
DSKKCACAKLNGGNFPYVDLNDGRLIESRDVVFECGPHCGCGPKCVNRTSQKRLRFNLEVFRSAK
KGWAVRSWEYIPAGSPVCEYIGVVRRTADVDTISDNEYIFEIDCQQTMQGLGGRQRRLRDVAVPM
NNGVSQSSEDENAPEFCIDAGSTGNFARFINHSCEPNLFVQCVLSSHQDIRLARVVLFAADNISPMQ
ELTYDYGYALDSVHEF

>tg_TgSET8
ASSPKKKRKVEASASRRTGEFLRDAQAPSRWLKRSKTGQDDGAFCLETWLAGAGDDAAGGERG
RDREGAADKAKQREERRQKELEERFEEMKVEFEEKAQRMIARRAALTGEIYSDGKGSKKPRVPSL
PENDDDALIEIIIDPEQGILKWPLSVMSIRQRTVIYQECLRRDLTACIHLTKVPGKGRAVFAADTILK
DDFVVEYKGELCSEREARERERQRYNRSKVPMGSFMFYFKNGSRMMAIDATDEKQDFGPARLINH
SRRNPNMTPRAITLGDFNSEPRLIFVARRNIEKGEELLVDYGERDPDVIKEHPWLNSEF

>hs_SIRT6
ASSPKKKRKVEASMSVNYAAGLSPYADKGKCGLPEIFDPPEELERKVWELARLVWQSSSVVFHTG
AGISTASGIPDFRGPHGVWTMEERGLAPKFDTTFESARPTQTHMALVQLERVGLLRFLVSQNVDG
LHVRSGFPRDKLAELHGNMFVEECAKCKTQYVRDTVVGTMGLKATGRLCTVAKARGLRACRGE
LRDTILDWEDSLPDRDLALADEASRNADLSITLGTSLQIRPSGNLPLATKRRGGRLVIVNLQPTKHD
RHADLRIHGYVDEVMTRLMKHLGLEIPAWDGPRVLERALPPLEF

>ce_Set1
ASSPKKKRKVEASMKVAAKKLATSRMRKDRAAAASPSSDIENSENPSSLASHSSSSGRMTPSKNT
RSRKGVSVKDVSNHKITEFFQVRRSNRKTSKQISDEAKHALRDTVLKGTNERLLEVYKDVVKGRG

IRTKVNFEKGDFVVEYRGVMMEYSEAKVIEEQYSNDEEIGSYMYFFEHNNKKWCIDATKESPWK
GRLINHSVLRPNLKTKVVEIDGSHHLILVARRQIAQGEELLYDYGDRSAETIAKNPWLVNTEF

>mm_G9a
ASSPKKKRKVEASVRTEKIICRDVARGYENVPIPCVNGVDGEPCPEDYKYISENCETSTMNIDRNIT
HLQHCTCVDDCSSSNCLCGQLSIRCWYDKDGRLLQEFNKIEPPLIFECNQACSCWRSCKNRVVQS
GIKVRLQLYRTAKMGWGVRALQTIPQGTFICEYVGELISDAEADVREDDSYLFDLDNKDGEVYCI
DARYYGNISRFINHLCDPNIIPVRVFMLHQDLRFPRIAFFSSRDIRTGEELGFDYGDRFWDIKSKYFT
CQCGSEKCKHSAEAIALEQSRLARLDPHPELLPDLSSLPPINTEF

>hs_SIRT5
ASSPKKKRKVEASSSSMADFRKFFAKAKHIVIISGAGVSAESGVPTFRGAGGYWRKWQAQDLATP
LAFAHNPSRVWEFYHYRREVMGSKEPNAGHRAIAECETRLGKQGRRVVVITQNIDELHRKAGTK
NLLEIHGSLFKTRCTSCGVVAENYKSPICPALSGKGAPEPGTQDASIPVEKLPRCEEAGCGGLLRPH
VVWFGENLDPAILEEVDRELAHCDLCLVVGTSSVVYPAAMFAPQVAARGVPVAEFNTETTPATN
RFRFHFQGPCGTTLPEALACHENETVSEF

>xl_HDAC8
ASSPKKKRKVEASMSRVVKPKVASMEEMAAFHTDAYLQHLHKVSEEGDNDDPETLEYGLGYDC
PITEGIYDYAAAVGGATLTAAEQLIEGKTRIAVNWPGGWHHAKKDEASGFCYLNDAVLGILKLRE
KFDRVLYVDMDLHHGDGVEDAFSFTSKVMTVSLHKFSPGFFPGTGDVSDIGLGKGRYYSINVPLQ
DGIQDDKYYQICEGVLKEVFTTFNPEAVVLQLGADTIAGDPMCSFNMTPEGIGKCLKYVLQWQLP
TLILGGGGYHLPNTARCWTYLTALIVGRTLSSEIPDHEFFTEYGPDYVLEITPSCRPDRNDTQKVQEI
LQSIKGNLKRVVEF

>mm_HP1a
ASSPKKKRKVEASMKEGENNKPREKSEGNKRKSSFSNSADDIKSKKKREQSNDIARGFERGLEPEK
IIGATDSCGDLMFLMKWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKESA
KSEF

>at_HDT1
ASSPKKKRKVEASMEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQN
LVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNA
AKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTPKAP
VSAKKAKVAVTPQKTDEKKKGGKAANQSEF

>mm_SAll
ASSPKKKRKVEASMSRRKQAKPQHFQSDPEVASLPRRDGDTEKGQPSRPTKSKDAHVCGRCCAEF
FELSDLLLHKKSCTKNQLVLIVNESPASPAKTFPPGPSLNDEF

>hs_SETD8
ASSPKKKRKVEASSCDSTNAAIAKQALKKPIKGKQAPRKKAQGKTQQNRKLTDFYPVRRSSRKSK
AELQSEERKRIDELIESGKEEGMKIDLIDGKGRGVIATKQFSRGDFVVEYHGDLIEITDAKKREALY
AQDPSTGCYMYYFQYLSKTYCVDATRETNRLGRLINHSKCGNCQTKLHDIDGVPHLILIASRDIAA
GEELLYDYGDRSKASIEAFPWLKHEF

>sc_RPD3
ASSPKKKRKVEASRRVAYFYDADVGNYAYGAGHPMKPHRIRMAHSLIMNYGLYKKMEIYRAKP
ATKQEMCQFHTDEYIDFLSRVTPDNLEMFKRESVKFNVGDDCPVFDGLYEYCSISGGGSMEGAAR
LNRGKCDVAVNYAGGLHHAKKSEASGFCYLNDIVLGIIELLRYHPRVLYIDIDVHHGDGVEEAFY
TTDRVMTCSFHKYGEFFPGTGELRDIGVGAGKNYAVNVPLRDGIDDATYRSVFEPVIKKIMEWYQ
PSAVVLQCGGDSLSGDRLGCFNLSMEGHANCVNYVKSFGIPMMVVGGGGYTMRNVARTWCFET
GLLNNVVLDKDLPYEF

>ec_CobB

ASSPKKKRKVEASMEKPRVLVLTGAGISAESGIRTFRAADGLWEEHRVEDVATPEGFDRDPELVQ
AFYNARRRQLQQPEIQPNAAHLALAKLQDALGDRFLLVTQNIDNLHERAGNTNVIHMHGELLKV
RCSQSGQVLDWTGDVTPEDKCHCCQFPAPLRPHVVWFGEMPLGMDEIYMALSMADIFIAIGTSGH
VYPAAGFVHEAKLHGAHTVELNLEPSQVGNEFAEKYYGPASQVVPEFVEKLLKGLKAGSIAEF
>hs_SUV39H1
ASSPKKKRKVEASNLKCVRILKQFHKDLERELLRRHHRSKTPRHLDPSLANYLVQKAKQRRALRR
WEQELNAKRSHLGRITVENEVDLDGPPRAFVYINEYRVGEGITLNQVAVGCECQDCLWAPTGGC
CPGASLHKFAYNDQGQVRLRAGLPIYECNSRCRCGYDCPNRVVQKGIRYDLCIFRTDDGRGWGV
RTLEKIRKNSFVMEYVGEIITSEEAERRGQIYDRQGATYLFDLDYVEDVYTVDAAYYGNISHFVNH
SCDPNLQVYNVFIDNLDERLPRIAFFATRTIRAGEELTFDYNMQVDPVDMESTRMDSNFGLAGLPG
SPKKRVRIECKCGTESCRKYLFEF

>hs_RCOR1
ASSPKKKRKVEASSNSWEEGSSGSSSDEEHGGGGMRVGPQYQAVVPDFDPAKLARRSQERDNLG
MLVWSPNQNLSEAKLDEYIAIAKEKHGYNMEQALGMLFWHKHNIEKSLADLPNFTPFPDEWTVE
DKVLFEQAFSFHGKTFHRIQQMLPDKSIASLVKFYYSWKKTRTKTSVMDRHARKQKREREESEDE
LEEANGNNPIDIEVDQNKESKKEVPPTETVPQVKKEKHSTEF

>hs_sin3a
ASSPKKKRKVEASYKESVHLETYPKERATEGIAMEIDYASCKRLGSSYRALPKSYQQPKCTGRTPL
CKEVLNDTWVSFPSWSEDSTFVSSKKTQYEEHIYRCEDERFELDVVLETNLATIRVLEAIQKKLSRL
SAEEQAKFRLDNTLGGTSEVIHRKALQRIYADKAADIIDGLRKNPSIAVPIVLKRLKMKEEEWREA
QRGFNKVWREQNEKYYLKSLDHQGINFKQNDTKVLRSKSLLNEIESIYDERQEQATEENAGVPVG
PHLSLAYEDKQILEDAAALIIHHVKRQTGIQKEDKYKIKQIMHHFIPDLLFAQRGDLSDVEEEEEEE
MDVDEATGAVEF

>at_SUVR4
ASSPKKKRKVEASQSAYLHVSLARISDEDCCANCKGNCLSADFPCTCARETSGEYAYTKEGLLKE
KFLDTCLKMKKEPDSFPKVYCKDCPLERDHDKGTYGKCDGHLIRKFIKECWRKCGCDMQCGNRV
VQRGIRCQLQVYFTQEGKGWGLRTLQDLPKGTFICEYIGEILTNTELYDRNVRSSSERHTYPVTLD
ADWGSEKDLKDEEALCLDATICGNVARFINHRCEDANMIDIPIEIETPDRHYYHIAFFTLRDVKAM
DELTWDYMIDFNDKSHPVKAFRCCCGSESCRDRKIKGSQGKSIERRKIVSAKKQQGSKEVSKKRK
EF

>rn_MeCP2_NLS
ASSPKKKRKVEASVQVKRVLEKSPGKLLVKMPFQASPGGKGEGGGATTSAQVMVIKRPGRKRKA
EADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPL
LVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHAESPKAPMPLLPPPPPPE
PQSSEDPISPPEPQDLSSSICKEEKMPRAGSLESDGCPKEPAKTQPMVAAAATTTTTTTTTVAEKYK
HRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVSEF

>mm_SET-TAF1B
ASSPKKKRKVEASMAPKRQSAILPQPKKPRPAAAPKLEDKSASPGLPKGEKEQQEAIEHIDEVQNEI
DRLNEQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHYLT
RVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDLTKRSSQTQN
KASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMDDEEGEAEDDDDDD
EEEEGLEDIDEEGDEDEGEEDDDEDEGEEGEEDEGEDDEF

>ce_Set4
ASSPKKKRKVEASMQLHEQIANISVTFNDIPRSDHSMTPTELCYFDDFATTLVVDSVLNFTTHKMS
KKRRYLYQDEYRTARTVMKTFREQRDWTNAIYGLLTLRSVSHFLSKLPPNKLFEFRDHIVRFLNM
FILDSGYTIQECKRYSQEGHQGAKLVSTGVWSRGDKIERLSGVVCLLSSEDEDSILAQEGSDFSVM
YSTRKRCSTLWLGPGAYINHDCRPTCEFVSHGSTAHIRVLRDMVPGDEITCFYGSEFFGPNNIDCEC
CTCEKNMNGAFSYLRGNENAEPIISEKKTKYELRSRSEF

CasLITE Sequences

> pSAMCL001 CasLITE1.0 2X NLS: dCas9-GS-NLS-NLS-CIB1 DNA sequence

ATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGAT
CACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACA
GCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACC
CGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCA
AGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGT
CCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGAC
GAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAG
CACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGG
CCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCA
GCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACG
CCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAG
CTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACC
CCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACAC
CTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCT
GGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGA
TCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTG
ACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC
CAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAA
GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA
GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAAC
CGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGG
GGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTT
CGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCG
ATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACC
GTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCT
GAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCG
TGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCC
GGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAG
GACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCT
GACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCG
ACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCG
GAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGT
CCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAG
AGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCT
CGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC
AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACG
AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGAC
ATCAACCGGCTGTCCGACTACGATGTGGACgcTATCGTGCCTCAGAGCTTTCTGAAGGACGACT
CCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC
CTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGA
TTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACA
GATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAG
TGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACC
GCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACT
TCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCC
GGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCG
GGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCG
AGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTG

ATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGC
CTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGA
AAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT
CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTC
CCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG
GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA
AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCAC
TACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGC
TAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT
TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGA
TCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACa
gcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaagaaaaagagga
aggtggcggccgctGctagcAGCCCCAAGAAGAAGAGAAAGGTGGAGGCCAGCAACGGCGCGATTGGT
GGGGATTTGCTGCTTAACTTTCCCGACATGTCCGTGTTGGAACGTCAGCGCGCACATTTGAAG
TATCTTAACCCCACCTTCGACTCCCCGTTGGCCGGGTTCTTTGCGGACTCATCTATGATTACGG
GAGGGGAAATGGACAGCTACCTCTCAACGGCCGGATTGAATCTTCCGATGATGTATGGAGAA
ACCACTGTAGAAGGCGACTCGCGACTCTCGATTTCGCCTGAAACGACGCTGGGAACAGGGAA
CTTCAAGAAACGGAAATTCGACACGGAGACAAAAGATTGCAACGAAAAGAAGAAGAAAATG
ACCATGAATCGCGATGATCTGGTAGAGGAGGGAGAGGAGGAAAAGTCGAAGATTACTGAAC
AGAACAATGGGTCTACCAAAAGTATCAAAAAGATGAAGCACAAAGCTAAGAAAGAAGAGAA
CAATTTCAGCAATGACAGCAGTAAAGTCACAAAAGAACTGGAGAAAACGGATTACATTCACG
TGAGGGCGCGACGAGGGCAGGCTACAGATTCACATTCAATTGCGGAGAGAGTACGGAGAGAG
AAAATCTCAGAAAGGATGAAGTTCCTCCAAGACCTTGTGCCAGGTTGTGACAAGATCACAGG
CAAAGCAGGAATGCTGGATGAGATCATCAACTACGTCCAATCGTTGCAAAGACAAATTGAGT
TTCTCTCGATGAAACTGGCCATCGTGAATCCTAGACCGGATTTCGACATGGATGACATCTTTG
CGAAAGAAGTGGCATCCACTCCCATGACGGTTGTGCCCTCACCGGAGATGGTCTTGTCTGGTT
ACAGCCACGAAATGGTGCATTCGGGTTATTCAAGCGAGATGGTCAATTCGGGATACCTTCACG
TCAATCCCATGCAGCAGGTGAATACTTCCAGTGATCCACTCTCCTGCTTTAACAACGGCGAGG
CCCCTTCGATGTGGGACTCCCACGTACAGAATCTCTATGGAAATCTCGGAGTC

> pSAMCL001 CasLITE1.0 2X NLS: dCas9-GS-NLS-NLS-CIB1 amino acid sequence
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSAGGGGSGGGGSGGGGS
GPKKKRKVAAAASSPKKKRKVEASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFA
DSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKK
KKMTMNRDDLVEEGEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHV

RARRGQATDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMK
LAIVNPRPDFDMDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQ
VNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

>pSAMCL002 CasLITE1.0 1X NLS: dCas9-GS-NLS-CIB1 DNA sequence
ATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGAT
CACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACA
GCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACC
CGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCA
AGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGT
CCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGAC
GAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAG
CACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGG
CCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCA
GCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACG
CCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAG
CTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACC
CCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACAC
CTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCT
GGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGA
TCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTG
ACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC
CAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAA
GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA
GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAAC
CGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGG
GGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTT
CGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCG
ATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACC
GTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCT
GAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCG
TGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCC
GGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAG
GACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCT
GACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCG
ACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCG
GAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGT
CCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAG
AGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCT
CGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC
AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACG
AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGAC
ATCAACCGGCTGTCCGACTACGATGTGGACgcTATCGTGCCTCAGAGCTTTCTGAAGGACGACT
CCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC
CTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGA
TTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACA
GATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAG
TGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACC
GCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACT
TCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCC

GGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCG
GGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCG
AGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTG
ATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGC
CTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGA
AAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT
CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTC
CCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG
GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA
AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCAC
TACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGC
TAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT
TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGA
TCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACa
gcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaagaaaaagagga
aggtggcggccgctGctagcAACGGCGCGATTGGTGGGGATTTGCTGCTTAACTTTCCCGACATGTCCGT
GTTGGAACGTCAGCGCGCACATTTGAAGTATCTTAACCCCACCTTCGACTCCCCGTTGGCCGG
GTTCTTTGCGGACTCATCTATGATTACGGGAGGGGAAATGGACAGCTACCTCTCAACGGCCGG
ATTGAATCTTCCGATGATGTATGGAGAAACCACTGTAGAAGGCGACTCGCGACTCTCGATTTC
GCCTGAAACGACGCTGGGAACAGGGAACTTCAAGAAACGGAAATTCGACACGGAGACAAAA
GATTGCAACGAAAAGAAGAAGAAAATGACCATGAATCGCGATGATCTGGTAGAGGAGGGAG
AGGAGGAAAAGTCGAAGATTACTGAACAGAACAATGGGTCTACCAAAAGTATCAAAAAGAT
GAAGCACAAAGCTAAGAAAGAAGAGAACAATTTCAGCAATGACAGCAGTAAAGTCACAAAA
GAACTGGAGAAAACGGATTACATTCACGTGAGGGCGCGACGAGGGCAGGCTACAGATTCACA
TTCAATTGCGGAGAGAGTACGGAGAGAGAAAATCTCAGAAAGGATGAAGTTCCTCCAAGACC
TTGTGCCAGGTTGTGACAAGATCACAGGCAAAGCAGGAATGCTGGATGAGATCATCAACTAC
GTCCAATCGTTGCAAAGACAAATTGAGTTTCTCTCGATGAAACTGGCCATCGTGAATCCTAGA
CCGGATTTCGACATGGATGACATCTTTGCGAAAGAAGTGGCATCCACTCCCATGACGGTTGTG
CCCTCACCGGAGATGGTCTTGTCTGGTTACAGCCACGAAATGGTGCATTCGGGTTATTCAAGC
GAGATGGTCAATTCGGGATACCTTCACGTCAATCCCATGCAGCAGGTGAATACTTCCAGTGAT
CCACTCTCCTGCTTTAACAACGGCGAGGCCCCTTCGATGTGGGACTCCCACGTACAGAATCTC
TATGGAAATCTCGGAGTC

>pSAMCL002 CasLITE1.0 1X NLS: dCas9-GS-NLS-CIB1 amino acid sequence
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSAGGGGSGGGGSGGGGS

GPKKKRKVAAAASNGAIGGDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMD
SYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTMNRDDL
VEEGEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDS
HSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFD
MDDIFAKEVASTPMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCF
NNGEAPSMWDSHVQNLYGNLGV

>pSAMCL003 CasLITE2.0: dCas9-GS-cib1(mNLS Δ318-334) DNA sequence
ATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGAT
CACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACA
GCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACC
CGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCA
AGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGT
CCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGAC
GAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAG
CACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGG
CCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCA
GCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACG
CCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAG
CTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACC
CCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACAC
CTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCT
GGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGA
TCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTG
ACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC
CAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAA
GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA
GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAAC
CGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGG
GGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTT
CGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCG
ATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACC
GTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCT
GAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCG
TGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCC
GGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAG
GACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCT
GACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCG
ACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCG
GAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGT
CCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAG
AGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCT
CGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC
AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACG
AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGAC
ATCAACCGGCTGTCCGACTACGATGTGGACgcTATCGTGCCTCAGAGCTTTCTGAAGGACGACT
CCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC
CTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGA
TTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACA
GATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAG
TGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACC

GCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACT
TCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCC
GGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCG
GGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCG
AGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTG
ATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGC
CTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGA
AAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT
CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTC
CCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG
GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA
AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCAC
TACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGC
TAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT
TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGA
TCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGAC
GctagcggcggcggcggcagcggcggcggcggcagcAACGGCGCGATTGGTGGGGATTTGCTGCTTAACTTTCC
CGACATGTCCGTGTTGGAACGTCAGCGCGCACATTTGAAGTATCTTAACCCCACCTTCGACTC
CCCGTTGGCCGGGTTCTTTGCGGACTCATCTATGATTACGGGAGGGGAAATGGACAGCTACCT
CTCAACGGCCGGATTGAATCTTCCGATGATGTATGGAGAAACCACTGTAGAAGGCGACTCGC
GACTCTCGATTTCGCCTGAAACGACGCTGGGAACAGGGAACTTCAAGGCGGCGAAATTCGAC
ACGGAGACAAAAGATTGCAACGAAGCGGCGAAGAAAATGACCATGAATCGCGATGATCTGGT
AGAGGAGGGAGAGGAGGAAAAGTCGAAGATTACTGAACAGAACAATGGGTCTACCAAAAGT
ATCAAAAAGATGAAGCACAAAGCTAAGAAAGAAGAGAACAATTTCAGCAATGACAGCAGTA
AAGTCACAAAAGAACTGGAGAAAACGGATTACATTCACGTGAGGGCGCGACGAGGGCAGGC
TACAGATTCACATTCAATTGCGGAGAGAGTACGGAGAGAGAAAATCTCAGAAAGGATGAAGT
TCCTCCAAGACCTTGTGCCAGGTTGTGACAAGATCACAGGCAAAGCAGGAATGCTGGATGAG
ATCATCAACTACGTCCAATCGTTGCAAAGACAAATTGAGTTTCTCTCGATGAAACTGGCCATC
GTGAATCCTAGACCGGATTTCGACATGGATGACATCTTTGCGAAAGAAGTGGCATCCACTCCC
ATGACGGTTGTGCCCTCACCGGAGATGGTCTTGTCTGGTTACAGCCACGAAATGGTGCATTCG
GGTTATTCAAGCGAGATGGTCAATTCGGGATACCTTCACGTCAATCCCATGCAGCAGGTGAAT
ACTTCCAGT

>pSAMCL003 CasLITE2.0: dCas9-GS-cib1(mNLS Δ318-334) amino acid sequence
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG

APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDASGGGGSGGGGSNGAIG
GDLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETT
VEGDSRLSISPETTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEKSKITEQNNGST
KSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKFL
QDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTPMTVVP
SPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSS

MS2 activator sequences

In all sgRNA backbone sequences, NNNNNNNNNNNNNNNNNNN represents the locus-specific guide sequence of each sgRNA.

>Tetraloop MS2 sgRNA backbone
NNNNNNNNNNNNNNNNNNNgttttagagctaggccAACATGAGGATCACCCATGTCTGCAGggcctagca
agttaaaataaggctagtccgttatcaCGCCGAAAGGCGggcaccgAGTcggtgcTTTTT

>Stem-loop 2 MS2 sgRNA backbone
NNNNNNNNNNNNNNNNNNNgttttagagctaGAAAtagcaagttaaaataaggctagtccgttatcaacttggccAACATG
AGGATCACCCATGTCTGCAGggccaagtggcaccgAGTcggtgcTTTTT

>Tetraloop and stem-loop 2 MS2 sgRNA backbone
NNNNNNNNNNNNNNNNNNNgttttagagctaggccAACATGAGGATCACCCATGTCTGCAGggcctagca
agttaaaataaggctagtccgttatcaacttggccAACATGAGGATCACCCATGTCTGCAGggccaagtggcaccgAGTcggt
gcTTTTT

>Standard sgRNA backbone
NNNNNNNNNNNNNNNNNNNgttttagagctaGAAAtagcaagttaaaataaggctagtccgttatcaacttGAAAaagtggca
ccgAGTcggtgcTTTTT

>pSAMca014: MS2-VP64 DNA sequence
ATGGCTTCAAACTTTACTCAGTTCGTGCTCGTGGACAATGGTGGGACAGGGGATGTGACAGTG
GCTCCTTCTAATTTCGCTAATGGGGTGGCAGAGTGGATCAGCTCCAACTCACGGAGCCAGGCC
TACAAGGTGACATGCAGCGTCAGGCAGTCTAGTGCCCAGAAgAGAAAGTATACCATCAAGGT
GGAGGTCCCCAAAGTGGCTACCCAGACAGTGGGCGGAGTCGAACTGCCTGTCGCCGCTTGGA
GGTCCTACCTGAACATGGAGCTCACTATCCCAATTTTCGCTACCAATTCTGACTGTGAACTCAT
CGTGAAGGCAATGCAGGGGCTCCTCAAAGACGGTAATCCTATCCCTTCCGCCATCGCCGCTAA
CTCAGGTATCTACagcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAG
CggacctaagaaaaagaggaaggtggcggccgctggatccGGACGGGCTGACGCATTGGACGATTTTGATCTGGAT
ATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCCTTGAT
GACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATT
AAC

>pSAMca014: MS2-VP64 amino acid sequence
MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVE
VPKVATQTVGGVELPVAAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY
SAGGGGSGGGGSGGGGSGPKKKRKVAAAGSGRADALDDFDLDMLGSDALDDFDLDMLGSDAL
DDFDLDMLGSDALDDFDLDMLIN

Description: The MS2-VP64 activator protein consists of the following domains from N-term to C-term: the N55K mutant of the MS2 bacteriophage coat protein 3X GGGGS linker, SV40 nuclear localization signal, and VP64 activation domain. Functionally, the MS2 domain binds to its specific RNA aptamer, the 3XGGGGS linker provides mechanical flexibility between the MS2 and VP64 domains, the SV40 nuclear localization signal facilitates transport of the protein into the nucleus, and the VP64 activation domain promotes transcriptional activation.

>pSAMca006: standard sgRNA backbone

NNNNNNNNNNNNNNNNNNNgttttagagctaGAAAtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcacc
gagtcggtgcTTTTTTT

Description: +83 nucleotide chimeric backbone used in Zhang Lab CRISPR/Cas9 publications

>pSAMca009: Tetraloop and stem-loop 2 MS2 sgRNA backbone
NNNNNNNNNNNNNNNNNNNgttttagagctaggccAACATGAGGATCACCCATGTCTGCAGggcctagca
agttaaaataaggctagtccgttatcaacttggccAACATGAGGATCACCCATGTCTGCAGggccaagtggcaccgagtcggtgc
TTTTTTT

Description: MS2-binding loop ggccAACATGAGGATCACCCATGTCTGCAGggcc replaces nucleotides +13 to +16 and nucleotides +53 to +56 of the standard sgRNA backbone. The resulting structure is an sgRNA scaffold in which the tetraloop and stemloop 2 sequences have been replaced by an MS2 binding loop. The tetraloop and stemloop 2 were selected for replacement based on information obtained from the Cas9/RNA/DNA crystal structure. Specifically, the tetraloop and stemloop 2 were found to protrude from the Cas9 protein in such a way which suggested that adding an MS2 binding loop would not interfere with any Cas9 residues. Additionally, the proximity of the tetraloop and stemloop 2 sites to the DNA suggested that localization to these locations would result in a high degree of interaction between the DNA and any recruited protein, such as a transcriptional activator.

> pSAMca003: Tetraloop MS2 sgRNA backbone
NNNNNNNNNNNNNNNNNNNgttttagagctaggccAACATGAGGATCACCCATGTCTGCAGggcctagca
agttaaaataaggctagtccgttatcaCGCCgaaaGGCGggcaccgagtcggtgcTTTTTTT

Description: The MS2-binding loop sequence ggccAACATGAGGATCACCCATGTCTGCAGggcc replaces nucleotides +13 to +16 of the standard sgRNA backbone. The sequence CGCC replaces nucleotides +49 to +52 of the standard sgRNA backbone. The sequence GGCG replaces nucleotides +57 to +60 of the standard sgRNA backbone. The tetraloop MS2-binding loop insertion was designed with the same rationale as described for pSAMca009 above. The CGCC and GGCG sequences replace the stem portion of stemloop 2. The increased base-pairing strength of the CGCC-GGCG stem compared to the original ACTT-AAGT stem was hypothesized to provide additional stability to the stemloop 2 structure, thereby increasing sgRNA performance or longevity.

>pSAMca001: MS2-p65 DNA sequence
ATGGCTTCAAACTTTACTCAGTTCGTGCTCGTGGACAATGGTGGGACAGGGGATGTGACAGTG
GCTCCTTCTAATTTCGCTAATGGGGTGGCAGAGTGGATCAGCTCCAACTCACGGAGCCAGGCC
TACAAGGTGACATGCAGCGTCAGGCAGTCTAGTGCCCAGAAgAGAAAGTATACCATCAAGGT
GGAGGTCCCCAAAGTGGCTACCCAGACAGTGGGCGGAGTCGAACTGCCTGTCGCCGCTTGGA
GGTCCTACCTGAACATGGAGCTCACTATCCCAATTTTCGCTACCAATTCTGACTGTGAACTCAT
CGTGAAGGCAATGCAGGGGCTCCTCAAAGACGGTAATCCTATCCCTTCCGCCATCGCCGCTAA
CTCAGGTATCTACagcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAG
CggacctaagaaaaagaggaaggtggcggccgctggatccCCTTCAGGGCAGATCAGCAACCAGGCCCTGGCTCTG
GCCCCTAGCTCCGCTCCAGTGCTGGCCCAGACTATGGTGCCCTCTAGTGCTATGGTGCCTCTGG
CCCAGCCACCTGCTCCAGCCCCTGTGCTGACCCCAGGACCACCCCAGTCACTGAGCGCTCCAG
TGCCCAAGTCTACACAGGCCGGCGAGGGGACTCTGAGTGAAGCTCTGCTGCACCTGCAGTTCG
ACGCTGATGAGGACCTGGGAGCTCTGCTGGGGAACAGCACCGATCCCGGAGTGTTCACAGAT
CTGGCCTCCGTGGACAACTCTGAGTTTCAGCAGCTGCTGAATCAGGGCGTGTCCATGTCTCAT
AGTACAGCCGAACCAATGCTGATGGAGTACCCCGAAGCCATTACCCGGCTGGTGACCGGCAG
CCAGCGGCCCCCCGACCCCGCTCCAACTCCCCTGGGAACCAGCGGCCTGCCTAATGGGCTGTC
CGGAGATGAAGACTTCTCAAGCATCGCTGATATGGACTTTAGTGCCCTGCTGTCACAGATTTC
CTCTAGTGGGCAG

>pSAMca001: MS2-p65 amino acid sequence
MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVE
VPKVATQTVGGVELPVAAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY
SAGGGGSGGGGSGGGGSGPKKKRKVAAAGSPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLA

QPPAPAPVLTPGPPQSLSAPVPKSTQAGEGTLSEALLHLQFDADEDLGALLGNSTDPGVFTDLASV
DNSEFQQLLNQGVSMSHSTAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSI
ADMDFSALLSQISSSGQ

Description: The MS2-p65 activator protein consists of the following domains from N-term to C-term: the N55K mutant of the MS2 bacteriophage coat, 3X GGGGS linker, SV40 nuclear localization signal, and p65 activation domain. Functionally, the MS2 domain binds to its specific RNA aptamer, the 3XGGGGS linker provides mechanical flexibility between the MS2 and p65 domains, the SV40 nuclear localization signal facilitates transport of the protein into the nucleus, and the p65 activation domain promotes transcriptional activation.

>pXRP057B: dCas9-GS-NLS-VP64 DNA sequence
atgGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA
CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGC
ATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCG
GCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAG
AGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCC
TTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGA
GGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCA
CCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCC
ACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAG
CTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGC
CAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGC
TGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCC
CCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACC
TACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG
GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT
CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA
CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACC
AGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAG
TTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAG
AGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACC
TGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACC
GGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGG
GAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTC
GAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGA
TAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGT
GTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA
GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTG
AAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGG
CGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGA
CAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGA
CACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGAC
GACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGA
AGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCC
GACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGA
GGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATC
TGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTC
GTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC
AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACG
AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGAC
ATCAACCGGCTGTCCGACTACGATGTGGACgcTATCGTGCCTCAGAGCTTTCTGAAGGACGACT
CCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC
CTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGA
TTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT

AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACA
GATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAG
TGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACC
GCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACT
TCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCC
GGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCG
GGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCG
AGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTG
ATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGC
CTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGA
AAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT
CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTC
CCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG
GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA
AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCAC
TACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGC
TAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT
TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGA
TCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACa
gcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaagaaaaagagga
aggtggcggccgctggatccGGACGGGCTGACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGA
CGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCCTTGATGACTTTGACCTCGAC
ATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAAC

>pXRP057B: dCas9-GS-NLS-VP64 amino acid sequence
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSAGGGGSGGGGSGGGGS
GPKKKRKVAAAGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDL
DMLIN

>pSAMca005: dCas9-GS-NLS-p65 DNA sequence
atgGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA
CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGC
ATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCG

GCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAG
AGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCC
TTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGA
GGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCA
CCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCC
ACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAG
CTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGC
CAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGC
TGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCC
CCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACC
TACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG
GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT
CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA
CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACC
AGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAG
TTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAG
AGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACC
TGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACC
GGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGG
GAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTC
GAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGA
TAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGT
GTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA
GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTG
AAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGG
CGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGA
CAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGA
CACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGAC
GACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGA
AGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCC
GACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGA
GGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATC
TGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTC
GTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC
AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACG
AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGAC
ATCAACCGGCTGTCCGACTACGATGTGGACgcTATCGTGCCTCAGAGCTTTCTGAAGGACGACT
CCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC
CTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGA
TTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACA
GATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAG
TGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACC
GCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACT
TCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCC
GGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCG
GGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCG
AGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTG
ATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGC
CTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGA
AAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT
CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTC
CCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG

GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA
AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCAC
TACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGC
TAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT
TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGA
TCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACa
gcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaagaaaaagagga
aggtggcggccgctggatccCCTTCAGGGCAGATCAGCAACCAGGCCCTGGCTCTGGCCCCTAGCTCCGC
TCCAGTGCTGGCCCAGACTATGGTGCCCTCTAGTGCTATGGTGCCTCTGGCCCAGCCACCTGCT
CCAGCCCCTGTGCTGACCCCAGGACCACCCCAGTCACTGAGCGCTCCAGTGCCCAAGTCTACA
CAGGCCGGCGAGGGGACTCTGAGTGAAGCTCTGCTGCACCTGCAGTTCGACGCTGATGAGGA
CCTGGGAGCTCTGCTGGGGAACAGCACCGATCCCGGAGTGTTCACAGATCTGGCCTCCGTGGA
CAACTCTGAGTTTCAGCAGCTGCTGAATCAGGGCGTGTCCATGTCTCATAGTACAGCCGAACC
AATGCTGATGGAGTACCCCGAAGCCATTACCCGGCTGGTGACCGGCAGCCAGCGGCCCCCCG
ACCCCGCTCCAACTCCCCTGGGAACCAGCGGCCTGCCTAATGGGCTGTCCGGAGATGAAGACT
TCTCAAGCATCGCTGATATGGACTTTAGTGCCCTGCTGTCACAGATTTCCTCTAGTGGGCAG

>pSAMca005: dCas9-GS-NLS-p65 amino acid sequence
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSAGGGGSGGGGSGGGGS
GPKKKRKVAAAGSPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLAQPPAPAPVLTPGPPQSLSA
PVPKSTQAGEGTLSEALLHLQFDADEDLGALLGNSTDPGVFTDLASVDNSEFQQLLNQGVSMSHS
TAEPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLSQISSSGQ

>pXRP057A: dCas9-GS-NLS DNA sequence
atgGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA
CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGC
ATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCG
GCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAG
AGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCC
TTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGA
GGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCA
CCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCC
ACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAG
CTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGC
CAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGC

TGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCC
CCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACC
TACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG
GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT
CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA
CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACC
AGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAG
TTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAG
AGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACC
TGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACC
GGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGG
GAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTC
GAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGA
TAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGT
GTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA
GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTG
AAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGG
CGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGA
CAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGA
CACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGAC
GACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGA
AGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCC
GACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGA
GGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATC
TGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTC
GTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC
AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACG
AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGAC
ATCAACCGGCTGTCCGACTACGATGTGGACgcTATCGTGCCTCAGAGCTTTCTGAAGGACGACT
CCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC
CTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGA
TTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACA
GATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAG
TGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACC
GCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACT
TCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCC
GGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCG
GGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCG
AGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTG
ATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGC
CTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGA
AAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT
CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTC
CCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG
GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA
AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCAC
TACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGC
TAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT
TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGA
TCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACa

gcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaagaaaaagagga
aggtggcggccgct

>pXRP057A: dCas9-GS-NLS amino acid sequence
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY
PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED
LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY
LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG
YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSAGGGGSGGGGSGGGGS
GPKKKRKVAAA

Bibliography

1. Deisseroth, K. Optogenetics. *Nat Methods* **8**, 26-29 (2011).

2. Zhang, F. et al. The Microbial Opsin Family of Optogenetic Tools. *Cell* **147**, 1446-1457 (2011).

3. Levskaya, A., Weiner, O.D., Lim, W.A. & Voigt, C.A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**, 997-1001 (2009).

4. Yazawa, M., Sadaghiani, A.M., Hsueh, B. & Dolmetsch, R.E. Induction of protein-protein interactions in live cells using light. *Nat Biotech* **27**, 941-945 (2009).

5. Strickland, D. et al. TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat Meth* **9**, 379-384 (2012).

6. Kennedy, M.J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat Meth* **7**, 973-975 (2010).

7. Shimizu-Sato, S., Huq, E., Tepperman, J.M. & Quail, P.H. A light-switchable gene promoter system. *Nat Biotech* **20**, 1041-1044 (2002).

8. Ye, H., Baba, M.D.-E., Peng, R.-W. & Fussenegger, M. A Synthetic Optogenetic Transcription Device Enhances Blood-Glucose Homeostasis in Mice. *Science* **332**, 1565-1568 (2011).

9. Polstein, L.R. & Gersbach, C.A. Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. *J Am Chem Soc* **134**, 16480-16483 (2012).

10. Bugaj, L.J., Choksi, A.T., Mesuda, C.K., Kane, R.S. & Schaffer, D.V. Optogenetic protein clustering and signaling activation in mammalian cells. *Nat Methods* **10**, 249-252 (2013).

11. Zhang, F. et al. Multimodal fast optical interrogation of neural circuitry. *Nature* **446**, 633-639 (2007).

12. Boch, J. et al. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **326**, 1509-1512 (2009).

13. Moscou, M.J. & Bogdanove, A.J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **326**, 1501 (2009).

14. Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotech* **29**, 149-153 (2011).

15. Liu, H. et al. Photoexcited CRY2 Interacts with CIB1 to Regulate Transcription and Floral Initiation in Arabidopsis. *Science* **322**, 1535-1539 (2008).

16. Yizhar, O., Fenno, L.E., Davidson, T.J., Mogri, M. & Deisseroth, K. Optogenetics in neural systems. *Neuron* **71**, 9-34 (2011).

17. Airan, R.D., Thompson, K.R., Fenno, L.E., Bernstein, H. & Deisseroth, K. Temporally precise in vivo control of intracellular signalling. *Nature* **458**, 1025-1029 (2009).

18. Wang, X., Chen, X. & Yang, Y. Spatiotemporal control of gene expression by a light-switchable transgene system. *Nat Meth* **9**, 266-269 (2012).

19. Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology* **29**, 143-148 (2011).

20. Geissler, R. et al. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**, e19509 (2011).

21. Cong, L., Zhou, R., Kuo, Y.-c., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun* **3**, 968 (2012).

22. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).

23. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).

24. Beerli, R.R., Segal, D.J., Dreier, B. & Barbas, C.F., 3rd 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. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14628-14633 (1998).

25. Bikard, D. et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Research* **41**, 7429-7437 (2013).

26. Qi, L.S. et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173-1183 (2013).

27. Jinek, M. et al. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).

28. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2579-E2586 (2012).

29. Banerjee, R. et al. The Signaling State of Arabidopsis Cryptochrome 2 Contains Flavin Semiquinone. *Journal of Biological Chemistry* **282**, 14916-14922 (2007).

30. Moore, M.J. & Proudfoot, N.J. Pre-mRNA Processing Reaches Back toTranscription and Ahead to Translation. *Cell* **136**, 688-700 (2009).

31. Proudfoot, N.J., Furger, A. & Dye, M.J. Integrating mRNA Processing with Transcription. *Cell* **108**, 501-512 (2002).

32. Liang, F.S., Ho, W.Q. & Crabtree, G.R. Engineering the ABA Plant Stress Pathway for Regulation of Induced Proximity. *Sci Signal* **4** (2011).

33. Holkers, M. et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Research* (2012).

34. Zhang, F. et al. Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. *Nature protocols* **5**, 439-456 (2010).

35. de Groote, M.L., Verschure, P.J. & Rots, M.G. Epigenetic Editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. *Nucleic Acids Research* **40**, 10596-10613 (2012).

36. Szymczak, A.L. et al. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nature Biotechnology* **22**, 589-594 (2004).

37. Christie, J.M. et al. Structural tuning of the fluorescent protein iLOV for improved photostability. *J Biol Chem* **287**, 22295-22304 (2012).

38. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* **31**, 827-832 (2013).

39. Ran, F.A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389 (2013).

40. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918 (2013).

41. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).

42. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84 (2014).

43. Perez-Pinera, P. et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* **10**, 973-976 (2013).

44. Cheng, A.W. et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res* **23**, 1163-1171 (2013).

45. Gilbert, L.A. et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154**, 442-451 (2013).

46. Kearns, N.A. et al. Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development* **141**, 219-223 (2014).

47. Maeder, M.L. et al. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* **10**, 977-979 (2013).

48. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology* **31**, 833-838 (2013).

49. Nishimasu, H. et al. Crystal structure of cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935-949 (2014).

50. Peabody, D.S. The RNA binding site of bacteriophage MS2 coat protein. *Embo J* **12**, 595-600 (1993).

51. Peabody, D.S. & Lim, F. Complementation of RNA binding site mutations in MS2 coat protein heterodimers. *Nucleic Acids Research* **24**, 2352-2359 (1996).

52. Johansson, H.E. et al. A thermodynamic analysis of the sequence-specific binding of RNA by bacteriophage MS2 coat protein. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 9244-9249 (1998).

53. Lim, F., Spingola, M. & Peabody, D.S. Altering the RNA binding specificity of a translational repressor. *J Biol Chem* **269**, 9006-9010 (1994).

54. Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nature Biotechnology* **29**, 731-734 (2011).

55. Li, T. et al. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Research* **39**, 6315-6325 (2011).

56. Wu, Z., Yang, H. & Colosi, P. Effect of Genome Size on AAV Vector Packaging. *Mol Ther* **18**, 80-86 (2009).

57. McClure, C., Cole, K.L., Wulff, P., Klugmann, M. & Murray, A.J. Production and titering of recombinant adeno-associated viral vectors. *J Vis Exp*, e3348 (2011).

58. Blecher-Gonen, R. et al. High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. *Nature protocols* **8**, 539-554 (2013).

59. Witten, Ilana B. et al. Recombinase-Driver Rat Lines: Tools, Techniques, and Optogenetic Application to Dopamine-Mediated Reinforcement. *Neuron* **72**, 721-733 (2011).

60. Grieger, J.C., Choi, V.W. & Samulski, R.J. Production and characterization of adeno-associated viral vectors. *Nat. Protocols* **1**, 1412-1428 (2006).

61. Lock M, A.M., Vandenberghe LH, Samanta A, Toelen J, Debyser Z, Wilson JM Rapid, Simple, and Versatile Manufacturing of Recombinant Adeno-Associated Viral Vectors at Scale. *Human Gene Therapy* **21**, 1259-1271 (2010).

62. Nan, X. et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389 (1998).

63. Boeke, J., Ammerpohl, O., Kegel, S., Moehren, U. & Renkawitz, R. The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. *J Biol Chem* **275**, 34963-34967 (2000).

64. Laherty, C.D. et al. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* **89**, 349-356 (1997).

65. Zhang, J., Kalkum, M., Chait, B.T. & Roeder, R.G. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Molecular Cell* **9**, 611-623 (2002).

66. Lauberth, S.M. & Rauchman, M. A conserved 12-amino acid motif in Sall1 recruits the nucleosome remodeling and deacetylase corepressor complex. *J Biol Chem* **281**, 23922-23931 (2006).

67. Gu, H. & Roizman, B. Herpes simplex virus-infected cell protein 0 blocks the silencing of viral DNA by dissociating histone deacetylases from the CoREST-REST complex. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17134-17139 (2007).

68. Ouyang, J., Shi, Y., Valin, A., Xuan, Y. & Gill, G. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. *Molecular Cell* **34**, 145-154 (2009).

69. Vannier, D., Balderes, D. & Shore, D. Evidence that the transcriptional regulators SIN3 and RPD3, and a novel gene (SDS3) with similar functions, are involved in transcriptional silencing in S. cerevisiae. *Genetics* **144**, 1343-1353 (1996).

70. Gregoretti, I.V., Lee, Y.M. & Goodson, H.V. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *J Mol Biol* **338**, 17-31 (2004).

71. Gao, L., Cueto, M.A., Asselbergs, F. & Atadja, P. Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. *J Biol Chem* **277**, 25748-25755 (2002).

72. Wu, K., Tian, L., Malik, K., Brown, D. & Miki, B. Functional analysis of HD2 histone deacetylase homologues in Arabidopsis thaliana. *Plant J* **22**, 19-27 (2000).

73. Scher, M.B., Vaquero, A. & Reinberg, D. SirT3 is a nuclear NAD+-dependent histone deacetylase that translocates to the mitochondria upon cellular stress. *Genes & Development* **21**, 920-928 (2007).

74. Hnisz, D., Schwarzmuller, T. & Kuchler, K. Transcriptional loops meet chromatin: a dual-layer network controls white-opaque switching in Candida albicans. *Mol Microbiol* **74**, 1-15 (2009).

75. Landry, J. et al. The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5807-5811 (2000).

76. Wilson, J.M., Le, V.Q., Zimmerman, C., Marmorstein, R. & Pillus, L. Nuclear export modulates the cytoplasmic Sir2 homologue Hst2. *EMBO Rep* **7**, 1247-1251 (2006).

77. Gertz, M. & Steegborn, C. Function and regulation of the mitochondrial sirtuin isoform Sirt5 in Mammalia. *Biochim Biophys Acta* **1804**, 1658-1665 (2010).

78. Zhu, A.Y. et al. Plasmodium falciparum Sir2A preferentially hydrolyzes medium and long chain fatty acyl lysine. *ACS Chem Biol* **7**, 155-159 (2012).

79. Tennen, R.I., Berber, E. & Chua, K.F. Functional dissection of SIRT6: identification of domains that regulate histone deacetylase activity and chromatin localization. *Mech Ageing Dev* **131**, 185-192 (2010).

80. Pennini, M.E., Perrinet, S., Dautry-Varsat, A. & Subtil, A. Histone methylation by NUE, a novel nuclear effector of the intracellular pathogen Chlamydia trachomatis. *PLoS Pathog* **6**, e1000995 (2010).

81. Mujtaba, S. et al. Epigenetic transcriptional repression of cellular genes by a viral SET protein. *Nat Cell Biol* **10**, 1114-1122 (2008).

82. Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H. & Shinkai, Y. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *Embo J* **27**, 2681-2690 (2008).

83. Snowden, A.W., Gregory, P.D., Case, C.C. & Pabo, C.O. Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr Biol* **12**, 2159-2166 (2002).

84. Rathert, P., Zhang, X., Freund, C., Cheng, X. & Jeltsch, A. Analysis of the substrate specificity of the Dim-5 histone lysine methyltransferase using peptide arrays. *Chem Biol* **15**, 5-11 (2008).

85. Jackson, J.P. et al. Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. *Chromosoma* **112**, 308-315 (2004).

86. Thorstensen, T. et al. The Arabidopsis SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. *Nucleic Acids Research* **34**, 5461-5470 (2006).

87. Vielle, A. et al. H4K20me1 contributes to downregulation of X-linked genes for C. elegans dosage compensation. *PLoS genetics* **8**, e1002933 (2012).

88. Couture, J.F., Collazo, E., Brunzelle, J.S. & Trievel, R.C. Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. *Genes & Development* **19**, 1455-1465 (2005).

89. Sautel, C.F. et al. SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. *Molecular and Cellular Biology* **27**, 5711-5724 (2007).

90. Hathaway, N.A. et al. Dynamics and memory of heterochromatin in living cells. *Cell* **149**, 1447-1460 (2012).

91. Ballare, C. et al. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat Struct Mol Biol* **19**, 1257-1265 (2012).

92. Jin, Q. et al. The protein phosphatase-1 (PP1) regulator, nuclear inhibitor of PP1 (NIPP1), interacts with the polycomb group protein, embryonic ectoderm development (EED), and functions as a transcriptional repressor. *J Biol Chem* **278**, 30677-30685 (2003).

93. Cervoni, N., Detich, N., Seo, S.B., Chakravarti, D. & Szyf, M. The oncoprotein Set/TAF-1beta, an inhibitor of histone acetyltransferase, inhibits active demethylation of DNA, integrating DNA methylation and transcriptional silencing. *J Biol Chem* **277**, 25026-25031 (2002).