Development of CRISPR-Cas tools towards rewiring 3D genome structure

Robin Kirkpatrick

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Jesse George Zalatan, Chair

Champak Chatterjee

Joshua Charles Vaughan

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Development of CRISPR-Cas tools towards rewiring 3D genome structure

Robin Kirkpatrick

Chair of the Supervisory Committee:

Jesse George Zalatan

Department of Chemistry

New genomic methods have revealed a high level of spatial organization in the nucleus, and the subnuclear positioning of genes correlates with transcriptional activity. It is thought that 3D genome organization and positioning directly influences transcriptional activity – perhaps by re-positioning genes near distal enhancers or to areas high in concentration of epigenetic modifiers. Limited studies in both human and yeast cells have revealed that gene repositioning affects transcription in some, but not all, genomic contexts. These studies raise the question: What rules govern transcriptional sensitivity to spatial re-positioning mechanisms? Such a question has remained elusive as conventional methodologies to rewire gene positioning rely on integration of a DNA binding site at every locus of interest. This approach stymies genomic screens involving multiple loci, particularly in human cells, where genome engineering remains difficult. Use of the CRISPR-Cas system enables researchers to target endogenous sites, bypassing the need for genome modification, which opens the door to reprogram the location of endogenous genes. In this work, I demonstrate two advances to the genome re-organization field. First, we show that a programmable CRISPR-Cas system can be used to localize genes to the nuclear periphery in budding yeast. Second, we demonstrate the development of a CRISPR-Cas allosteric sensor of DNA binding which might be fundamental to forming efficient long-range loops. These technological advancements may aid researchers to assess the functional consequences of gene repositioning.

Table of Contents

List of publications

Paper title	Journal	Chapter	Contribution	Status	Ref.
Conditional Recruitment to a DNA-Bound Complex Using a Colocalization-Dependent Protein Switch	ACS Synthetic Biology	Ch. 2	First author		1
CRISPR-Cas-mediated tethering recruits the yeast HMR locus to the nuclear periphery but fails to silence gene expression	In preparation	Ch. 3	First author		2

Reference

(1) Kirkpatrick, R. L., Lewis, K., Langan, R. A., Lajoie, M. J., Boyken, S. E., Eakman, M., et al. (2020). Conditional Recruitment to a DNA-Bound CRISPR-Cas Complex Using a Colocalization-Dependent Protein Switch. ACS Synthetic Biology, 9(9), 2316–2323. http://doi.org/10.1021/acssynbio.0c00012

(2) Kirkpatrick, R. L., Cliff, E., Cunningham-Bryant, D., Fernandez, B., Zalatan, J. (2020). CRISPR-Cas-mediated tethering recruits the yeast HMR locus to the nuclear periphery but fails to silence gene expression. In preparation

List of figures

Main figures

Supplementary figures

List of tables

Acknowledgments

First, I would like those involved in this work including my lab-mate and partner in tasty snacks Emily Cliff, the talented undergraduates Kieran Lewis and Maddie Eakman, and my skilled collaborators Robert Langan, Mark Lajoie, Scott Boyken and Ryan Kibler. Next, I would like to thank my scientific mentor Jesse Zalatan who been a better mentor than I could have asked for and taught me skills that include, but extend much beyond the science. To my coworkers who made lab more fun than grad school is designed for: Chloe, Erin, Hannah, Emily, and more. I would like to extend my deep gratitude to my friends and coworkers in the Zalatan lab for their extensive support in all areas of my life. I thank my friends that bring memories of coffee, science, space-team, hiking, play-dates, chocolate, puzzles, and more. I thank my friends and family for supporting me, including free child-care!, during my PhD and to my daughter who makes all of this the most fun in the world. I look forward to what is to come in the world of science and thank UW for being my academic home for all of these years.

Dedicated to my friends and family

Chapter 1 | Introduction

1.1 | Background

1.1.1 | Genomic positioning correlates with cellular function

Recent studies have revealed a high level of spatial organization of DNA in the nucleus[1] and genomic position correlates with gene function[2]. Genes can be physically positioned in 3-dimensional space near well-defined subnuclear structures such as the nuclear periphery[3,] the nuclear pore[4], or the nucleolus[5]. Genetic elements can also be positioned relative to other genetic elements such as transcriptional enhancers[6,7,8] or recombination templates[9] to form a 3-dimensional gene loops that range from kb to Mb scale[7-10]. 3-dimensional gene positioning often correlates with a variety of different genetic functions including genetic recombination[9], epigenetic state[3], and gene expression[3], suggesting that genome organization plays a crucial role in controlling cellular function.

Dynamic changes in 3-d gene positioning during cellular differentiation or extracellular signals further suggests a crucial regulatory role of gene positioning. For example, in fetal blood cells the LCR enhancer forms a 3-d loop with fetal beta globin genes to activate transcription. Later in differentiation, the fetal LCR beta globin loop gets reprogrammed to form a 3-d loop with adult beta globin genes resulting results in changes in gene expression[6,7,8]. Long-range looping also plays a role in genetic recombination. In budding yeast, yeast mating type is determined by the whether the a or alpha gene cassette is at the MAT locus. Interestingly, genetic templates contained at HML or HMR are copied into the MAT locus through formation of a long range DNA loop between one of the two genetic loci, suggesting that 3D genome organization can control cell type specification through promoting genetic recombination[9]. Finally, upon addition of extracellular signals, genes reposition from the nuclear interior to interact with nucleoporins to activate gene transcription in budding yeast[4].To determine the causative relationship between genome structure and gene function, tools to physically reposition genes in the nucleus and assess the impact on function are needed.

1.1.2 | Conventional 3D gene repositioning tools

To probe the role of genome structure on gene function, gene repositioning tools have been developed to re-localize genes to specific subnuclear sites. Early tools have focused on fusing hard coded DNA binding domains, such as Gal4, to localized proteins to direct genes to specific sites in the nucleus. In a fundamental yeast study, fusion of Gal4 to a localized nuclear membrane protein resulted in

repression of a reporter gene, likely as a result of re-localization to high concentrations of repressors at the nuclear periphery[11]. In human cells, fusion of LacI to nuclear membrane proteins re-localized genes to the periphery and was sufficient to establish gene silencing[12,13]. These fundamental studies using hard coded DBDs have demonstrated a causative role for gene positioning on gene function. However, these studies have been limited, in part, because genomic engineering of a DNA binding site is required at each site of interest.

1.1.3 | The utility of CRISPR-Cas gene repositioning tools

To overcome this technical challenge, CRISPR-Cas systems have been developed to re-localize genes in the nucleus. Because the CRISPR-Cas system is programmable, these systems can be targeted to endogenous sites and bypass the requirement to do site specific modification of a DNA binding domain target site. In yeast, indirect dCas9 fusion to a nuclear membrane protein allowed gene recruitment to the nuclear periphery and this localization was sufficient to bias plasmid segregation in daughter cells[14]. In human cells, fusion of dCas9 to chemically-inducible dimerization motifs was sufficient to form DNA loops or re-localize genes to well-defined subnuclear structures[15,16]. This repositioning was sufficient to cause changes in gene expression as well as cellular toxicity. These data suggest that CRISPR-Cas re-positioning tools may be useful towards dissecting the details of how 3D structural perturbations can influence gene function.

1.1.4 | Limitations on DNA re-localization technologies

CRISPR tools offer the promise to easily re-localize genes to either well defined subnuclear structures or to other DNA sites. Formation of DNA-DNA loops can be achieved by fusing CRISPR complexes to dimerization motifs which results in the bridging of two distant DNA sites. DNA looping experiments in particular are limited, in part, because the dimerizing DNA binding domain can also bind other free unbound DNA binding domains which results in inefficient DNA loops[17] (Figure 1.1). In the bacterial LacI looping system, this competition effect not only lowers the looping efficiency, but creates limitations of the length of DNA loops that can be formed[18]. At short DNA lengths, promotion of DNA-DNA interactions through random collision of the DNA loci overcomes the competition effect between non DNA-bound complexes. At long DNA lengths, free unbound DNA binding domains occupy each DNA target site, resulting in inefficient loop formation. The efficiency of loops depends on DNA binding

domain concentration, as high concentrations results in free unbound complexes outcompeting loop formation. Notably, low concentration of the lac repressor is crucial for efficient loop formation (<10 nM), and is most efficient at short range (~1-5kb). Interestingly, bacterial systems are able to control this low level of lac repressor, which is 20 copies per cell, through a feedback regulation circuit that precisely regulates expression levels[18].

Despite known issues with protein competition in looping systems, human cells, bacteria, and yeast cells are able to form long loop sizes ranging from kb several MB. There are several models that may explain these data. Natural systems may be able to drive long range or efficient loops by either controlling the concentration of looping proteins to low levels[18], mechanisms driving physical proximity of DNA elements[19], recruitment of loop stabilizing factors[16], or using cooperative or allosteric communication between DNA and protein binding[18]. To engineer long range DNA loops, engineering strategies modeled after natural systems should be considered. While difficult to engineer, advances in protein engineering may provide opportunities to engineer allosteric or cooperative communication between DNA binding and dimerization domain activity.

Current synthetic looping systems have been developed by linking DNA binding domains to dimerization motifs which results in the bridging of two DNA sites. These methods are limited, in part, by protein complexes competing for the DNA site. For instance, CRISPR looping using dCas9 fused to dimerization motifs in bacteria is limited to the kb range due to the competition effect of unbound CRISPR complexes[17]. In human cells, the length range of loops has been observed to be much higher. When dCas9 was linked to the Pyl1 ABI system, loops up to 40 kb were observed[16]. In a separate study, linking dCas9 to light activated protein complexes resulted in a 500 kb DNA loop[20]. The ability to form long range loops in human cells, despite known competition problems, may be a result of low protein concentrations, unknown cooperativity in the system, or mechanisms that promote DNA-DNA contacts human cells. Interestingly, in a separate experiment, telomere clustering was observed in human cells when nucleating phase separated droplets using dCas9 linked to nucleating proteins[21] (CasDrop). This observation in particular suggests that engineering cooperativity or allostery into looping systems may promote loop formation.

1.2 | Overview of thesis work

This thesis focuses on addressing two challenges in the gene re-localization field: development of

programmable gene recruitment tools and development of a protein motif that senses and responds to DNA binding. In parallel to the methods reported in the literature and outlined in Chapter 1, I developed an alternative method to re-localize genes to the nuclear periphery in budding yeast using CRISPR-Cas9 (Chapter 3). We benchmark this system against the Gal4 peripheral re-localization system and find that both Cas9 and Gal4 can recruit DNA to the periphery at equal efficiencies. However, Gal4 but not Cas9 mediated peripheral recruitment is sufficient to cause gene silencing, suggesting the precise structure of the recruitment complex is important for gene silencing. This study highlights potential challenges of using dCas9 to discover structure function rules in yeast. In a separate study (Chapter 2), we developed a CRISPR-Cas allosteric sensor of DNA binding where an effector is predominantly active in the DNA bound conformation. Future work will focus on testing if allosteric communication between dimerization motifs and DNA binding could make efficient DNA loops.

1.3 | References

(1) Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., and Dekker, J. (2009) Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 326, 289–293.

(2) Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., Yen, C.-A., Schmitt, A. D., Espinoza, C. A., and Ren, B. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503, 290–294.

(3) Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453, 948–951.

(4) Randise-Hinchliff, C., Coukos, R., Sood, V., Sumner, M. C., Zdraljevic, S., Meldi Sholl, L., Garvey Brickner, D., Ahmed, S., Watchmaker, L., and Brickner, J. H. (2016) Strategies to regulate transcription factor-mediated gene positioning and interchromosomal clustering at the nuclear periphery. J Cell Biol 212, 633–646.

(5) Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., Shendure, J., Fields, S., Blau, C. A., and Noble, W. S. (2010) A three-dimensional model of the yeast genome. Nature 465, 363–367.

(6) Carter, D., Chakalova, L., Osborne, C. S., Dai, Y.-F., and Fraser, P. (2002) Long-range chromatin regulatory interactions in vivo. Nature Genetics 32, 623–626.

(7) Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and de Laat, W. (2002) Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. Mol. Cell 10, 1453–1465.

(8) Palstra, R.-J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003) The β-globin nuclear compartment in development and erythroid differentiation. Nature Genetics 35, 190–194.

(9) Avşaroğlu, B., Bronk, G., Li, K., Haber, J. E., and Kondev, J. (2016) Chromosome-refolding model of mating-type switching in yeast. Proc. Natl. Acad. Sci. U.S.A. 113, E6929–E6938.

(10) Denker, A., and de Laat, W. (2016) The second decade of 3C technologies: detailed insights into nuclear organization. Genes Dev. 30, 1357–1382.

(11) Andrulis, E. D., Neiman, A. M., Zappulla, D. C., and Sternglanz, R. (1998) Perinuclear localization of chromatin facilitates transcriptional silencing. Nature 394, 592–595.

(12) Finlan, L. E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J. R., and Bickmore, W. A. (2008) Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. PLOS Genet (Reik, W., Ed.) 4, e1000039.

(13) Reddy, K. L., Zullo, J. M., Bertolino, E., and Singh, H. (2008) Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature 452, 243–247.

(14) Lin, J.-L., Ekas, H., Deaner, M., and Alper, H. S. (2019) CRISPR-PIN: Modifying gene position in the nucleus via dCas9-mediated tethering. Synth Syst Biotechnol 4, 73–78.

(15) Wang, H., Xu, X., Nguyen, C. M., Liu, Y., Gao, Y., Lin, X., Daley, T., Kipniss, N. H., La Russa, M., and Qi, L. S. (2018) CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization. Cell 175, 1405–1417.e14.

(16) Morgan, S. L., Mariano, N. C., Bermudez, A., Arruda, N. L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.-F., Hoffman, A. R., Huang, C.-C., Pitteri, S. J., and Wang, K. C. (2017) Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nature Communications 2017 8:1 8, 15993.

(17) Hao, N., Shearwin, K. E., and Dodd, I. B. (2017) Programmable DNA looping using engineered bivalent dCas9 complexes. Nature Communications 2017 8:1 8, 1628.

(18) Priest, D. G., Cui, L., Kumar, S., Dunlap, D. D., Dodd, I. B., and Shearwin, K. E. (2014) Quantitation of the DNA tethering effect in long-range DNA looping in vivo and in vitro using the Lac and λ repressors. PNAS 111, 349–354.

(19) Elf, J., Li, G.-W., and Xie, X. S. (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. Science 316, 1191–1194.

(20) Kim, J. H., Rege, M., Valeri, J., Dunagin, M. C., Metzger, A., Titus, K. R., Gilgenast, T. G., Gong, W., Beagan, J. A., Raj, A., and Phillips-Cremins, J. E. (2019) LADL: light-activated dynamic looping for endogenous gene expression control. Nature Methods 16, 633–639.

(21) Shin, Y., Chang, Y.-C., Lee, D. S. W., Berry, J., Sanders, D. W., Ronceray, P., Wingreen, N. S.,

Haataja, M., and Brangwynne, C. P. (2018) Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. Cell 175, 1481–1491.e13.

1.4 | Figures

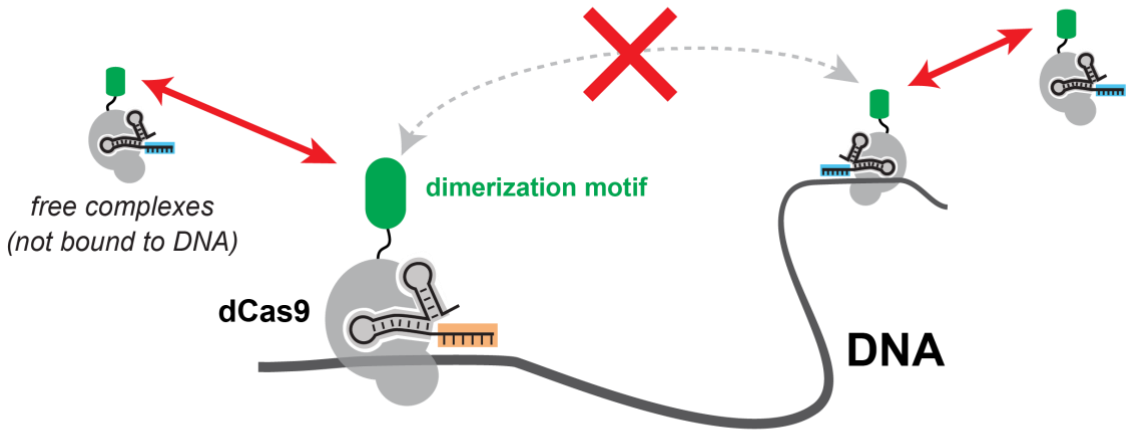


Figure 1.1 CRISPR-Cas DNA relocalization strategies suffer from competition of free unbound CRISPR-Cas complexes

DNA relocalization can be controlled by fusion of catalytically inactive dCas9 to a dimerization motif[17]. Looping efficiency can suffer from free, unbound complexes, outcompeting loop formation.

Chapter 2 | Conditional protein recruitment to a DNA-bound CRISPR-Cas complex using a proximity-sensitive Co-LOCKR switch

The work in this chapter was published and reprinted with permission from the journal of initial publication:

Kirkpatrick, R. L., Lewis, K., Langan, R. A., Lajoie, M. J., Boyken, S. E., Eakman, M., et al. (2020). Conditional Recruitment to a DNA-Bound CRISPR-Cas Complex Using a Colocalization-Dependent Protein Switch. ACS Synthetic Biology, 9(9), 2316–2323. http://doi.org/10.1021/acssynbio.0c00012

2.1 | Abstract

To spatially control biochemical functions at specific sites within a genome, we have engineered a synthetic switch that activates when bound to its DNA target site. The system uses two CRISPR-Cas complexes to colocalize components of a de novo designed protein switch (Co-LOCKR) to adjacent sites in the genome. Colocalization triggers a conformational change in the switch from an inactive, closed state to an active, open state with an exposed functional peptide. We prototype the system in yeast and demonstrate that DNA binding triggers switch activation, recruitment of a transcription factor, and expression of a downstream reporter gene. This DNA-triggered Co-LOCKR switch provides a platform to engineer sophisticated functions that should only be executed at a specific target site within the genome, with potential applications in a wide range of synthetic systems including epigenetic regulation, imaging, and genetic logic circuits.

2.2 | Introduction

Catalytically inactive CRISPR-Cas complexes, which bind but do not cleave DNA, can programmably recruit functional proteins to specific genomic target sites with applications for transcriptional control, epigenetic regulation, and imaging.[1,2] One outstanding challenge for the field is that these effector proteins can produce non-specific background effects. For example, epigenetic modifiers can mark off-target sites in the genome,[3] while unbound imaging probes can produce background signal that obscures the specific target site of interest.[4] Similarly, efforts to engineer long-range DNA loops are challenging in part because the desired interaction between two DNA-bound CRISPR-Cas complexes competes with the free, unbound CRISPR-Cas complexes in the cell.[5] In each of these cases, the problem is that the effector protein remains functional even when it is not bound at a specific DNA target site. As a first step to address this general problem, we have developed a conditional system where an effector protein can be activated only when the CRISPR-Cas complex is bound to its DNA target.

To engineer a conditional, DNA-triggered effector protein, we envisioned making the activity of the effector protein dependent on colocalization of two CRISPR-Cas complexes. Ideally, assembly of the functional effector would only occur when the two complexes are brought into proximity at adjacent genomic target sites. To achieve this behavior, we turned to a recently developed system called Co-LOCKR (colocalization-dependent latching orthogonal cage-key),[6] a designed protein switch that only activates when its two components, cage and key, are colocalized. The cage conformationally regulates a protein interaction module (the 'latch' peptide) that becomes activated only when key binding displaces the latch and exposes the interaction module (Figure 2.1).[7,8] By tuning the cage, latch, and key affinities appropriately, the conformational switch occurs only when key and cage are colocalized to the same subcellular location and in close physical proximity.[6]

To couple the conformational switch to DNA binding, we can use two orthogonal CRISPR-Cas complexes to recruit a Co-LOCKR cage and key to adjacent genomic target sites. When both cage and key are colocalized on DNA, the Co-LOCKR switches to the active state and exposes the protein interaction module. When the cage-tethered CRISPR-Cas complex is not bound to DNA, the cage adopts the inactive state (Figure 2.1). Thus, provided that the cage-key interaction is too weak to form without colocalization, the system should effectively function as a sensor for DNA binding: only CRISPR-Cas complexes that are bound to DNA and appropriately positioned will switch to the active state. An additional feature of this approach is that off-target binding events should not activate the switch, as

any off-target sites for the cage or key CRISPR-Cas complexes are unlikely to be in close proximity to each other. Conceptually similar approaches have been used for gene editing with paired Fok1 nucleases, nickases, and alternative DNA binding domains.[9-14]

We demonstrate here that Co-LOCKR switches can be coupled to CRISPR-Cas complexes to function as a proximity-sensitive sensor of DNA binding. We use a reporter gene assay to assess the switch function, and we systematically vary the stoichiometry, separation distance on DNA, linker lengths, expression levels, and protein interaction strengths to identify the key parameters that enable optimal colocalization-dependent switch activation on DNA. The use of tunable, designed proteins to implement a DNA-binding dependent protein switch lays the foundation for a variety of future applications, including synthetic epigenetic modifications, imaging tools, rewiring of genome structure, and genetic logic gates, that could benefit from coupling a biochemical function to DNA binding.

2.3 | Results

2.3.1 | CRISPR complexes recruit cage and key components to a transcriptional reporter

To engineer a DNA-dependent Co-LOCKR switch, we need one CRISPR-Cas complex that recruits the cage protein and another CRISPR-Cas complex programmed to an adjacent genomic site that recruits the key protein. To recruit different proteins to adjacent CRISPR-Cas complexes, we can use the catalytically inactive *S. pyogenes* dCas9 together with scaffold RNAs (scRNAs), which are modified sgRNAs with 3' hairpins that can recruit RNA binding proteins (RBPs).[15] By using orthogonal RNA hairpin-RBP pairs, we can program one target site to recruit the cage and the other target site to recruit the key (Figure 2.1). Alternatively, we can directly fuse cage and key to orthogonal CRISPR-Cas proteins.[1]

To test different DNA-dependent Co-LOCKR designs, we constructed a transcriptional reporter system in *S. cerevisiae* with multiple distinct CRISPR-Cas target sites upstream of a genomically-integrated fluorescent Venus reporter gene. These target sites can be used to colocalize the Co-LOCKR cage and key proteins in close proximity on DNA. Key binding to the cage exposes the protein interaction module, a Bim peptide that binds to the Bcl2 protein.[6] We fused Bcl2 to the transcriptional activator VP64, so that cage opening recruits Bcl2-VP64 to activate Venus reporter expression (Figure 2.1). Reporter activation thus serves as a proxy for the open state of the Co-LOCKR switch.

2.3.2 | Colocalization on genomic DNA can activate a Co-LOCKR switch

We initially prototyped the RNA recruitment strategy using scRNAs with PP7, MS2, and com hairpins together with their cognate RNA binding proteins PCP, MCP, and Com. We used a 2x PP7 scRNA to recruit the key-PCP fusion protein to the upstream target site (J5). PP7 binds PCP as a dimer, so a 2x PP7 scRNA recruits four key-PCP fusion proteins. To recruit the cage to the downstream, promoter-proximal target site (J4), we initially tested two strategies: either a 2x MS2 scRNA, which will recruit four MCP-cage fusion proteins, or a 1x com scRNA, which will recruit one Com-cage fusion protein (Figure 2.2A).[15] We engineered yeast reporter strains to express all protein components of the system: dCas9, Bcl2-VP64, key-PCP, and either MCP-cage or Com-cage. We then delivered a plasmid expressing two scRNAs to each strain: either J5 2x PP7 + J4 2x MS2 to the MCP-cage strain or J5 2x PP7 + J4 1x com to the Com-cage strain. In both cases, we observed Venus fluorescent reporter expression (Figure 2.2B).

To evaluate the significance of the reporter activation, we need to assess the background signal from colocalization-independent cage opening, which could occur with a recruited cage in the presence of co-expressed key or from a recruited key in the presence of co-expressed cage (Figure 2.2B). These contributions can be measured by delivering only one guide instead of both guides simultaneously. For the MCP-cage, the background from recruiting only the key is comparable to the Venus signal obtained when both cage and key are recruited, suggesting that most if not all of the observed expression arises from colocalization-independent effects where free cage binds the recruited key. In contrast, for the monomeric Com-cage, the background expression levels from recruiting only the key, or only the cage, are both significantly smaller than the Venus expression when both cage and key are recruited.

To confirm that the observed Venus expression with the Com-cage Co-LOCKR is truly colocalization-dependent, we need to consider whether the observed expression is larger than the combined effects of free key binding to the recruited cage and free cage binding to the recruited key (Figure S2.1). In control experiments, we demonstrated that VP64 recruitment to two target sites is additive rather than synergistic (Figure S2.1), suggesting that the additive sum of individual cage and key recruitment can be used to assess the contribution from colocalization-independent cage opening at both sites. The Com-cage colocalized activity is significantly larger than the sum of key only and cage only activity (Figure 2.2), suggesting that the Com-cage-mediated Co-LOCKR is a colocalization-dependent switch.

Finally, we assessed whether alternative RNA recruitment strategies could improve switch activation. We tested 1x PP7, 2x PP7, and 1x com scRNAs to recruit key fusion proteins in multiple combinations with 1x MS2, 2x MS2, and 1x com to recruit cage fusion proteins. For all MS2-mediated MCP-cage recruitment strategies, background reporter gene activation was relatively high and there was no significant switch activation above background (Figure S2.2). The dominant source of background activity is from the “key only” control, which presumably results from the DNA-tethered key recruiting free, partially open cages (Figure S2.1). This background is significantly reduced when the cage is fused to Com, a monomeric RNA binding protein. Using the Com-cage, both 1x PP7 and 2x PP7 produced colocalization-dependent switch activation above background, with more activation from 2x PP7 (Figure S2.2). Thus, it appears that providing a high stoichiometry of key to cage can help to promote Co-LOCKR activation, while delivering the cage on a multivalent protein is not effective.

2.3.3 | Direct protein fusions to orthogonal CRISPR-Cas complexes can activate a Co-LOCKR switch

We also tested an alternative recruitment strategy with direct fusions to orthogonal CRISPR-Cas proteins. We fused the cage to *S. py* dCas9 and the key to *Lachnospiraceae bacterium* dCpf1, which has previously been shown to be effective for transcriptional activation in yeast as a dCpf1-VPR fusion.[16] We observed colocalization-dependent activation that is significantly above the background (Figure S2.3), but the overall activation level was ~10X smaller than that observed with the RNA recruitment strategy (Figure 2.2). Unlike mammalian cells, in yeast the direct protein fusions to CRISPR-Cas complexes are often outperformed by RNA recruitment strategies (Figure S2.3),[15] and the relatively weak activation obtained with direct fusions of the Co-LOCKR switch is consistent with this behavior. We therefore proceeded to focus on the RNA recruitment strategy using the optimal 2x PP7 scRNA to recruit PCP-key and 1x com to recruit Com-cage (Figure 2.2 & Figure S2.2).

2.3.4 | Switch activation is sensitive to the distance between CRISPR-Cas complexes

Using the RNA recruitment strategy, we had initially targeted cage and key to two relatively close sites positioned 51 bases apart (Figure 2.2). The 51 base separation between PAM sites includes each 20-base target site, leaving 11 bases between each CRISPR-Cas complex. To explore the distance-dependence of the Co-LOCKR switch, we targeted the key to a range of sites further upstream. Shifting the key five bases upstream to a 56 base separation, which corresponds to a half-turn around the DNA helix, resulted in an almost complete loss in the colocalization-dependent activation, which was recovered with another half-turn shift to a 61 base separation (Figure 2.3). The same pattern continued with 66 and 71 base separations. Moving the key further upstream resulted in a complete loss of colocalization-dependent activation (Figure 2.3). A similar periodicity in colocalizing protein effectors to DNA was observed with assembly of the Fok1 dimer with two dCas9-Fok1 complexes.[9]

Colocalization-dependent activation could also be sensitive to linker lengths in the CRISPR-Cas complex, as linkers that are too short might not be able to reach their binding partners, while linkers that are too long might be too flexible to be effective. However, we varied the linker between key and PCP from 3 to 45 amino acids and saw little change in colocalization-dependent activation (Figure S2.4).

2.3.5 | Optimizing the Com-cage RNA-mediated Co-LOCKR switch

We explored two additional design parameters with the Com-cage Co-LOCKR switch: protein expression levels and cage-key interaction affinity. LOCKR switch activation is sensitive to protein concentration, with high key concentrations driving the switch towards the open state. [7,8] We therefore varied the expression level of the key-PCP fusion protein using a set of promoters with different expression levels (Figure 2.4); these promoters vary in strength over a >100-fold range (Table S2).[17] Our initial experiments were performed with the relatively strong *Adh1* promoter, which was effective for previous applications of RBP fusion proteins in yeast.[15] We expected that higher key levels might activate the switch without colocalization, while key levels that are too low would fail to bind the scRNA PP7 hairpin at the CRISPR-Cas complex. Consistent with this expectation, the strongest promoter tested, *pTdh3*, results in the highest level of background activation when only the cage is recruited (Figure 2.4B). *pAdh1* maintains a similar level of colocalization-dependent switch activation, but lower background compared to *pTdh3*. The weak promoters *pUra3* and *pCyc1* significantly reduced background activation, but also dramatically reduced colocalization-dependent activation.

Similarly, high expression levels of the Bcl2-VP64 fusion protein could drive the switch towards an open state. We therefore tested whether reducing Bcl2-VP64 levels could decrease background activation. When we switched from the strong *Adh1* promoter to the weaker *Ura3* promoter, we observed a significant decrease in background activation, but also a nearly complete loss in switch activation (Figure 2.4C).

To improve colocalization-dependent switch activation, we sought to reduce the background by tuning the interaction strength between cage and key. A key that interacts too strongly with the cage might activate the switch without colocalization, while a key that interacts too weakly might not be able to activate. The Co-LOCKR system has already been tuned to weaken the cage-key interface and minimize colocalization-independent activation.[6] To further weaken the cage-key interface, we truncated the key peptide over a range from 44 to 34 amino acids (Figure 2.5). Similar key truncations have been previously demonstrated to reduce cage key affinity in the Co-LOCKR system.[6,7] While most truncations had no significant effect, the 34 aa key significantly reduced background activation while maintaining the level of colocalization-dependent activation after correcting for background (Figure 2.5).

2.4 | Discussion

We show here that a colocalization-dependent protein switch, Co-LOCKR, can be adapted to act as a sensor for DNA binding. The switch undergoes a DNA-triggered conformational change when the Co-LOCKR cage and key modules are colocalized to CRISPR-Cas complexes at adjacent target sites in the genome. Multiple layers of modularity in the system provide powerful tools for precisely controlling biological functions: the Co-LOCKR switch can cage a diverse set of functional peptides,[6-8] and the CRISPR-Cas system enables programmable DNA targeting, which means that this switch can in principle be executed at a broad range of different sites in the genome. Further, the Co-LOCKR system could in principle be coupled with other programmable DNA binding systems, such as synthetic zinc fingers, that enable cooperative assembly of protein switch functions.[18]

The best-performing DNA-triggered Co-LOCKR switch utilized an RNA-recruitment strategy with the Com-cage (Figure 2.2 & Figure S2.3). Multiple CRISPR-Cas complex separation distances between 51-71 are effective, but within this window the relative orientation along the DNA helix appears to affect activity (Figure 2.3). While all of our designs produced some colocalization-independent activation, we found that this background could be reduced by weakening the cage-key interaction affinity. The most effective switch used a truncated 34 aa key and reduced the total background signal by 1.7-fold (Figure 2.5). Future work could potentially minimize this background further by strengthening the cage-latch interface together with more precise adjustments of expression levels and cage-key affinity.[6,7]

Engineering a DNA-triggered Co-LOCKR switch requires balancing interaction strengths and expression levels so that the switch only opens when cage and key are colocalized. Alternative approaches to obtain DNA-dependent protein activity face similar challenges in balancing affinities and concentrations. For example, we could minimize the unbound fraction of a DNA-binding effector protein by expressing it at only a few copies per cell and engineering sufficiently high affinity so that it binds its DNA target even at low concentrations. Practically, however, it is challenging to precisely control expression levels at only a few copies per cell, and at such low concentrations the affinities would need to be quite tight to achieve full occupancy at the DNA target. With the Co-LOCKR system, we can express the components at a high enough level to bind the DNA target, but the switch ensures that the concentration of unbound, functionally active protein is minimized.

There are multiple systems where minimizing the activity of unbound effector proteins could be advantageous. For example, with epigenetic modifiers, expression of effectors fused to DNA binding

domains can lead to non-specific modifications across the genome.[3] With a DNA-triggered Co-LOCKR, it should be possible to minimize the amount of active enzyme that is not bound to its DNA target. To achieve this goal, future work will need to develop a switch that triggers assembly of an enzyme. LOCKR switches that cage peptide fragments of split proteins have been engineered,[19] suggesting that a DNA-triggered epigenetic modifier should be achievable. A conceptually similar approach using split proteins fused to DNA binding domains has been shown to be effective for regulating the activity of the methyltransferase Dmnt3.[20,21] Because the affinity of a split protein interface may not be readily tunable, the LOCKR system could expand the toolbox of proteins that could be conditionally regulated as split proteins.

Another system where minimizing active, unbound protein would be advantageous is for engineering long-range loops in DNA to probe the relationship between genome structure and function. It is possible to engineer DNA loops using constitutively active protein interaction domains fused to DNA binding domains,[5,22,23] but free, unbound interaction domains can compete with loop formation.[5,24] With a DNA-triggered Co-LOCKR switch, it may be possible to minimize the concentration of free interaction domains, as any unbound complexes should remain caged and inactive. DNA-triggered switch opening would expose the interaction domain, which should promote interactions between DNA-bound complexes.

Finally, there are a broad range of potential applications of CRISPR-Cas guide RNAs as programmable elements in synthetic genetic circuits. CRISPR-Cas systems have been engineered to generate logic gates and transcriptional cascades.[25-30] The Co-LOCKR switch that we have implemented here acts as a simple two-input AND gate, where each input is an scRNA. While NOR gates have been previously described, and multiple NOR gates can be linked to construct AND gates,[29] there are practical limitations for delivering large numbers of circuit components to a biological system. A Co-LOCKR-based AND gate provides a potentially simpler alternative to achieve this function.

The DNA-triggered CRISPR Co-LOCKR switch combines DNA, RNA, and protein based logic to achieve sophisticated functional outcomes in a biological setting. The switch effectively functions as an allosteric sensor, with DNA binding triggering a conformational change at a distant site in the complex. These synthetic tools for allosteric control and spatial regulation of biochemical activity should provide new routes towards precise and tunable control of biological systems.

2.5 | Methods

2.5.1 | Yeast strain construction

Yeast (*S. cerevisiae*) transformations were performed with the standard lithium acetate method. The parent haploid yeast strain for reporter gene experiments was SO992 (W303; *MAT***a** *ura3 leu2 trp1 his3*). Complete descriptions of all yeast strains generated in this work are provided in Table S1. Reporter genes and protein expression constructs (Table S2) were integrated in single copy into the genome. Guide RNA expression constructs (Table S3) were delivered on CEN/ARS plasmids. Protein sequences and gRNA target site sequences are provided in the Supporting Information.

2.5.2 | Reporter gene design

The pJ1-Venus reporter gene is a modified version of a previously described 1XtetO-Venus reporter gene.[15] The upstream tetO target site was replaced with a short 163 base array of dCas9 target sites, spaced 10 bases apart, from a sequence originally designed for a bacterial reporter system.[31] pR4-Venus and pR5-Venus are derivatives of pJ1-Venus with 3 or 5 bases inserted upstream of the dCas9 target site (J4) used for cage targeting. These shifted promoters allowed an additional set of dCas9 target site spacing distances to be tested (Figure 2.3). pR6-Venus is a modified version of the 1XtetO-Venus reporter gene that retains the tetO target site and inserts a dCpf1 target site 47 bases upstream. pR6+5, pR6+10, pR6+15, and pR6+20 are derivatives of pR6 with 5, 10, 15, or 20 bases inserted upstream of the tetO site, which allows a range of dCpf1-dCas9 spacings to be tested (Figure S2.3). Complete sequences of the reporter genes are provided in the Supporting Information.

2.5.3 | Co-LOCKR fusion proteins

Cage and key protein designs for the Co-LOCKR switch have been described previously.[6] Cage proteins were fused to the C-terminus of RBPs and dCas9; C-terminal fusions of transcriptional regulators to RBPs and dCas9 are effective,[15] and LOCKR cages have been effective as C-terminal fusions *in vivo*.[7,8] Key proteins were fused to the N-terminus of RBPs and dCpf1, as this orientation was effective for key proteins in prior *in vivo* applications.[6-8]

2.5.4 | Flow cytometry

After transformation of guide RNA plasmids, yeast strains were grown overnight at 30 °C in minimal

media (SD complete, SD –Ura, SD –His, or SD –Ura –His). Overnight cultures were diluted 1:25 and grown for an additional 4-5 hours. Fluorescent protein expression levels were measured with an LSRII flow cytometer (BD Biosciences). To select single yeast cells, we applied a gate using the SSC-A vs. FSC-A plot. Median fluorescence values were recorded from the gated populations. Values reported in the plots are the mean ± s.d. of the median fluorescence values for at least three measurements (biological replicates).

2.6 | Acknowledgements

We thank Dustin Maly, David Shechner, Rhiju Das, Ryan Kibler, and members of the Zalatan and Baker groups for comments and discussion. We thank Jason Fontana for assistance with the design of the J1 reporter gene. R.L.K. was supported by a Genome Sciences NIH Training Grant T32 HG00035. R.A.L. was supported by Bruce and Jeannie Nordstrom, thanks to the Patty and Jimmy Barrier Gift for the Institute for Protein Design Directors Fund. M.J.L. was supported by a Washington Research Foundation Innovation Postdoctoral Fellowship and a Cancer Research Institute Irvington Fellowship from the Cancer Research Institute. S.E.B. was supported by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (#1017008). This work was supported by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (#1010814 to J.G.Z.), NIH R35 GM124773 (J.G.Z.), and by Open Philanthropy to D.B.

2.7 | Conflict of interest

M.J.L., S.E.B., R.A.L., and D.B. inventors on patents related to this work.

2.8 | References

(1) Barrangou, R., and Doudna, J. A. (2016) Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* 933–941.

(2) Wang, H., Xu, X., Nguyen, C. M., Liu, Y., Gao, Y., Lin, X., Daley, T., Kipniss, N. H., La Russa, M., and Qi, L. S. (2018) CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization. *Cell 175*, 1405–1417.e14.

(3) Galonska, C., Charlton, J., Mattei, A. L., Donaghey, J., Clement, K., Gu, H., Mohammad, A. W., Stamenova, E. K., Cacchiarelli, D., Klages, S., Timmermann, B., Cantz, T., Schöler, H. R., Gnirke, A., Ziller, M. J., and Meissner, A. (2018) Genome-wide tracking of dCas9-methyltransferase footprints. *Nat. Commun. 9*, 597.

(4) Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E. H., Weissman, J. S., Qi, L. S., and Huang, B. (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell 155*, 1479–1491.

(5) Hao, N., Shearwin, K. E., and Dodd, I. B. (2017) Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat. Commun. 8*, 1628.

(6) Lajoie, M. J., Boyken, S. E., Salter, A. I., Bruffey, J., Rajan, A., Langan, R. A., Olshefsky, A., Muhunthan, V., Gewe, M., Quijano-Rubio, A., Johnson, J., Lenz, G., Nguyen, A., Lau, K., King, N. P., Pun, S. H., Correnti, C. E., Riddell, S., and Baker, D. Designed protein logic to target cells with precise combinations of surface antigens. submitted.

(7) Langan, R. A., Boyken, S. E., Ng, A. H., Samson, J. A., Dods, G., Westbrook, A. M., Nguyen, T. H., Lajoie, M. J., Chen, Z., Berger, S., Mulligan, V. K., Dueber, J. E., Novak, W. R. P., El-Samad, H., and Baker, D. (2019) De novo design of bioactive protein switches. *Nature 572*, 205–210.

(8) Ng, A. H., Nguyen, T. H., Gómez-Schiavon, M., Dods, G., Langan, R. A., Boyken, S. E., Samson,

J. A., Waldburger, L. M., Dueber, J. E., Baker, D., and El-Samad, H. (2019) Modular and tunable biological feedback control using a de novo protein switch. *Nature 572*, 265–269.

(9) Guilinger, J. P., Thompson, D. B., and Liu, D. R. (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol. 32*, 577–582.

(10) Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., Goodwin, M. J., Aryee, M. J., and Joung, J. K. (2014) Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol. 32*, 569–576.

(11) Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., and Church, G. M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol. 31*, 833–838.

(12) Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell 154*, 1380–1389.

(13) Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., and Kim, J.-S. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res. 24*, 132–141.

(14) Bolukbasi, M. F., Gupta, A., Oikemus, S., Derr, A. G., Garber, M., Brodsky, M. H., Zhu, L. J., and Wolfe, S. A. (2015) DNA-binding-domain fusions enhance the targeting range and precision of Cas9. *Nat. Methods 12*, 1150–1156.

(15) Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., Tsai, J. C., Weissman, J. S., Dueber, J. E., Qi, L. S., and Lim, W. A. (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell 160*, 339–350.

(16) Lian, J., HamediRad, M., Hu, S., and Zhao, H. (2017) Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun. 8*, 1688.

(17) Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell 95*, 717–728.

(18) Bashor, C. J., Patel, N., Choubey, S., Beyzavi, A., Kondev, J., Collins, J. J., and Khalil, A. S. (2019) Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. *Science 364*, 593–597.

(19) Langan, R. A. (2019, June 15) *De Novo* Design of Bioactive Protein Switches, and Applications Thereof (Baker, D., Ed.).

(20) Chaikind, B., Kilambi, K. P., Gray, J. J., and Ostermeier, M. (2012) Targeted DNA methylation using an artificially bisected M.HhaI fused to zinc fingers. *PLoS ONE 7*, e44852.

(21) Xiong, T., Meister, G. E., Workman, R. E., Kato, N. C., Spellberg, M. J., Turker, F., Timp, W., Ostermeier, M., and Novina, C. D. (2017) Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. *Sci Rep 7*, 6732.

(22) Morgan, S. L., Mariano, N. C., Bermudez, A., Arruda, N. L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.-F., Hoffman, A. R., Huang, C.-C., Pitteri, S. J., and Wang, K. C. (2017) Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun. 8*, 15993.

(23) Kim, J.-H., Rege, M., Valeri, J., Dunagin, M. C., Metzger, A., Titus, K. R., Gilgenast, T. G., Gong, W., Beagan, J. A., Raj, A., and Phillips-Cremins, J. E. (2019) LADL: light-activated dynamic looping for endogenous gene expression control. *Nat. Methods 16*, 633–639.

(24) Priest, D. G., Cui, L., Kumar, S., Dunlap, D. D., Dodd, I. B., and Shearwin, K. E. (2014)

Quantitation of the DNA tethering effect in long-range DNA looping in vivo and in vitro using the Lac and λ repressors. *Proc. Natl. Acad. Sci. USA 111*, 349–354.

(25) Kiani, S., Beal, J., Ebrahimkhani, M. R., Huh, J., Hall, R. N., Xie, Z., Li, Y., and Weiss, R. (2014) CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods 11*, 723–726.

(26) Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P., and Lu, T. K. (2014) Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell 54*, 698–710.

(27) Didovyk, A., Borek, B., Hasty, J., and Tsimring, L. (2015) Orthogonal Modular Gene Repression in Escherichia coli Using Engineered CRISPR/Cas9. *ACS Synth. Biol.*

(28) Lebar, T., and Jerala, R. (2016) Benchmarking of TALE- and CRISPR/dCas9-Based Transcriptional Regulators in Mammalian Cells for the Construction of Synthetic Genetic Circuits. *ACS Synth. Biol. 5*, 1050–1058.

(29) Gander, M. W., Vrana, J. D., Voje, W. E., Carothers, J. M., and Klavins, E. (2017) Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nat. Commun. 8*, 15459.

(30) Hofmann, A., Falk, J., Prangemeier, T., Happel, D., Köber, A., Christmann, A., Koeppl, H., and Kolmar, H. (2019) A tightly regulated and adjustable CRISPR-dCas9 based AND gate in yeast. *Nucleic Acids Res. 47*, 509–520.

(31) Dong, C., Fontana, J., Patel, A., Carothers, J. M., and Zalatan, J. G. (2018) Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun. 9*, 2489.

2.9 | Figures

Figure 1: Colocalized CRISPR-Cas complexes can sense DNA binding

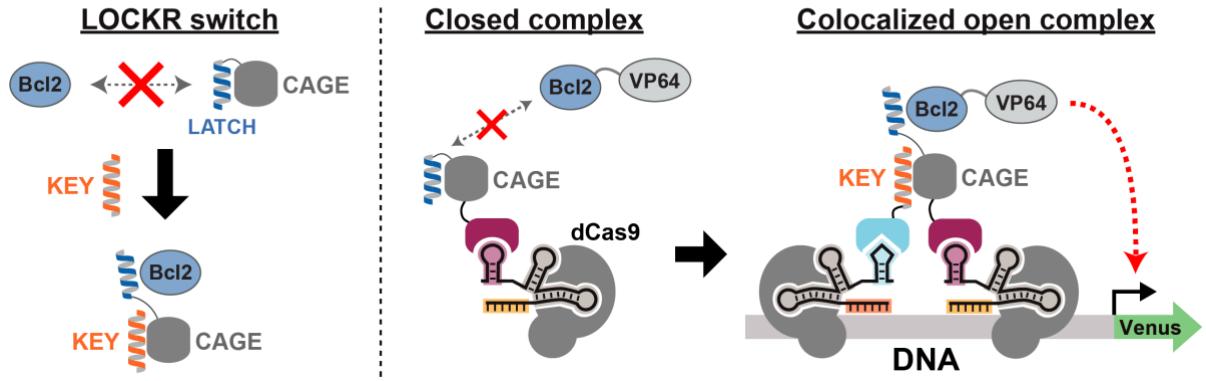


Figure 2.1 Colocalized CRISPR-Cas complexes can sense DNA binding

Colocalized CRISPR-Cas complexes can sense DNA binding. The LOCKR switch is a designed protein that can switch between two conformational states.[7] Colocalization of the key and cage with orthogonal CRISPR-Cas complexes to adjacent genomic target sites on DNA releases the latch, which allows recruitment of the Bcl2-VP64 transcriptional activator. CRISPR-Cas complexes are specified for either cage or key recruitment using orthogonal 3' RNA hairpins that recruit RNA binding proteins fused to the cage or key.

Figure 2: Colocalization-dependent activation of a transcriptional reporter

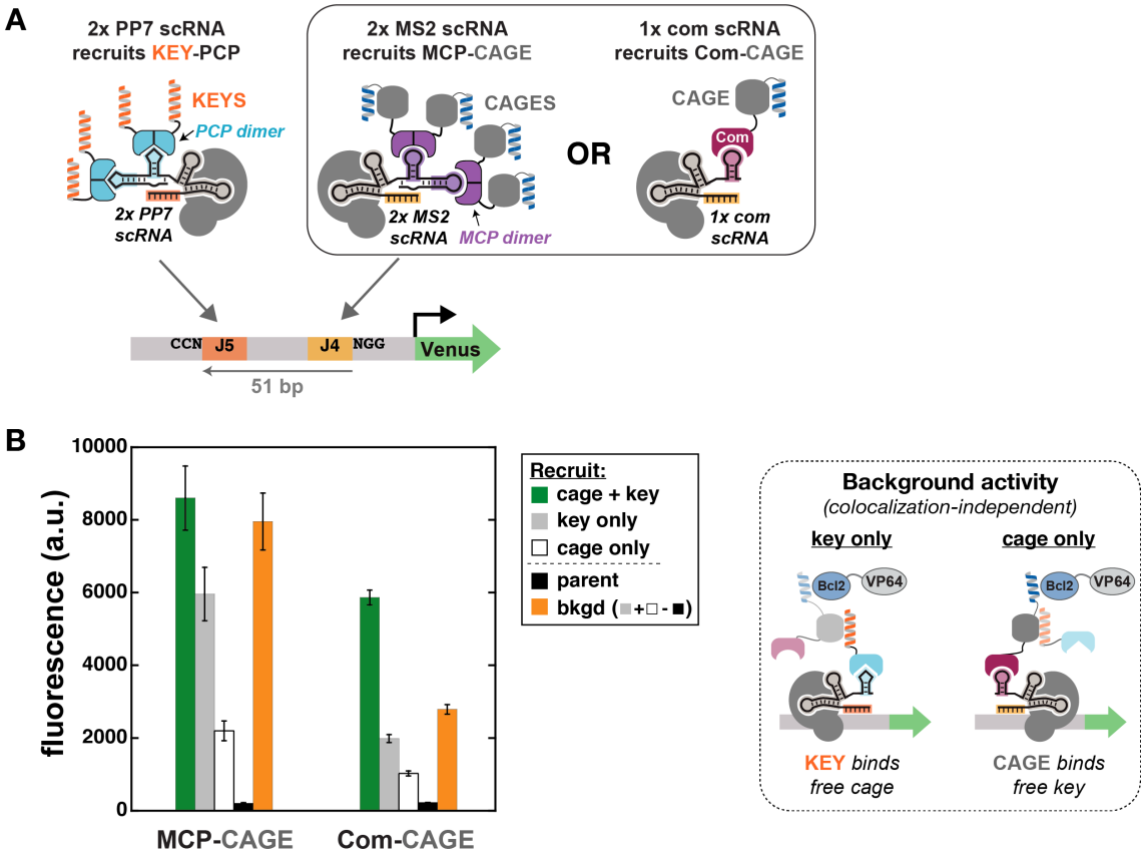


Figure 2.2 Colocalization-dependent activation of a transcriptional reporter

Colocalization-dependent activation of a transcriptional reporter. (A) A 2x PP7 scRNA targets the upstream site (J5) and recruits four key-PCP fusion proteins. At the downstream site (J4), a 2x MS2 scRNA recruits four MCP-cage fusion proteins. Alternatively, a 1x com scRNA recruits one Com-cage fusion protein. (B) Fluorescence reporter activity upon cage-key colocalization. For the MCP-cage, the observed “cage + key” signal appears to arise primarily from colocalization-independent opening of the Co-LOCKR switch. For the Com-cage, the observed “cage+key” signal is significantly larger than the background colocalization-independent activation. The background is the additive sum of the “key only” and “cage only” samples (Figure S2.1), and the errors are propagated by adding in quadrature. Data for parents were obtained with the unmodified parent strains yKL016 and yKL014. Fluorescence values are mean ± SD for at least three biological replicates.

Figure 3: CRISPR Co-LOCKR switch is sensitive to target site spacing

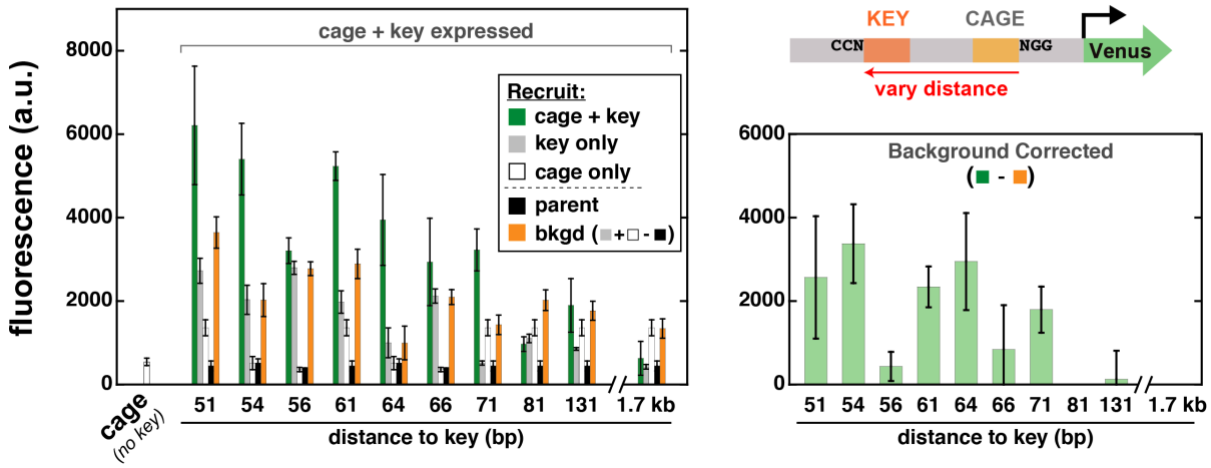


Figure 2.3 CRISPR Co-LOCKR

The CRISPR Co-LOCKR switch (2x PP7:key & 1x com:cage) is sensitive to target site spacing. The background colocalization-independent activation is calculated from the sum of “key only” and “cage only” samples as described in Figure S2.1. Subtracting the background from the observed “cage + key” activity gives the background corrected values shown. Data for parent strains were obtained by transforming the parental strains yKL014, yRK244, and yRK245 with empty vector (pRS316). Fluorescence values are mean ± SD for at least three biological replicates.

Figure 4: High key expression levels increase background activation

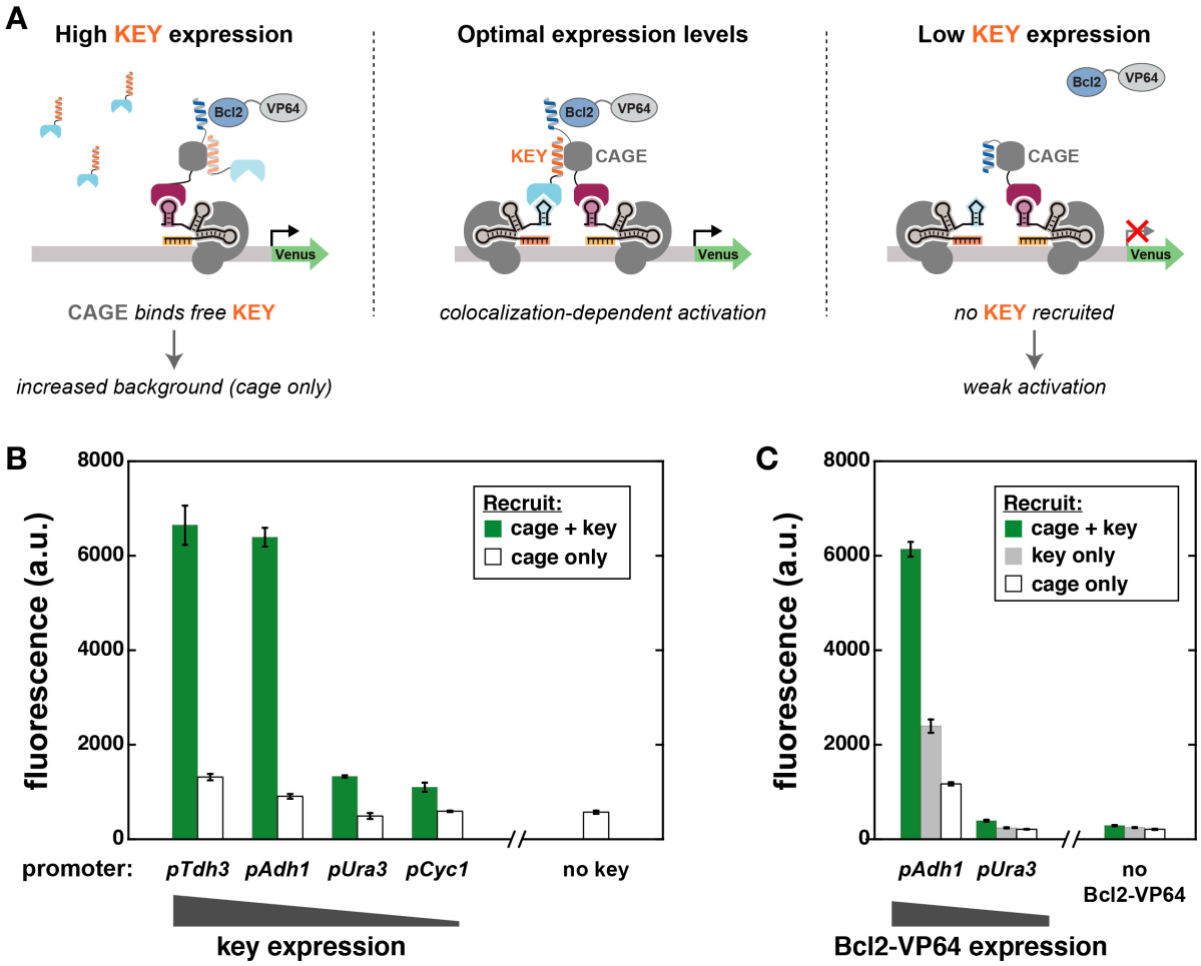


Figure 2.4 High key expression increases background activation

High key expression levels increase background activation of the CRISPR Co-LOCKR switch (2x PP7:key & 1x com:cage). (A) If key expression is too high, free key can bind and open the cage, which produces colocalization-independent background when only the cage is recruited. If key expression is too low, there will not be enough key-PCP fusion protein to bind the PP7 RNA hairpin, and there will be little to no colocalization-dependent activation. (B) When the key is expressed from a strong pTdh3 promoter, we observe an increase in background activation (cage only). When the key is expressed from weak pUra3 or pCyc1 promoters, there is little colocalization-dependent activation (cage + key). (C) Decreasing Bcl2-VP64 expression prevents switch activation. Fluorescence values in B & C are mean ± SD for at least three biological replicates.

Figure 5: Tuning key length reduces background activation

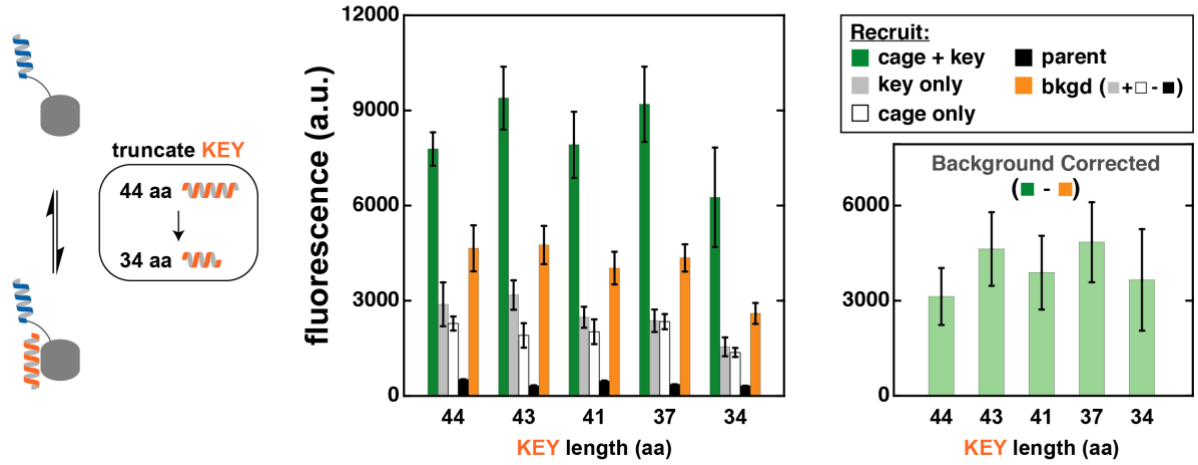


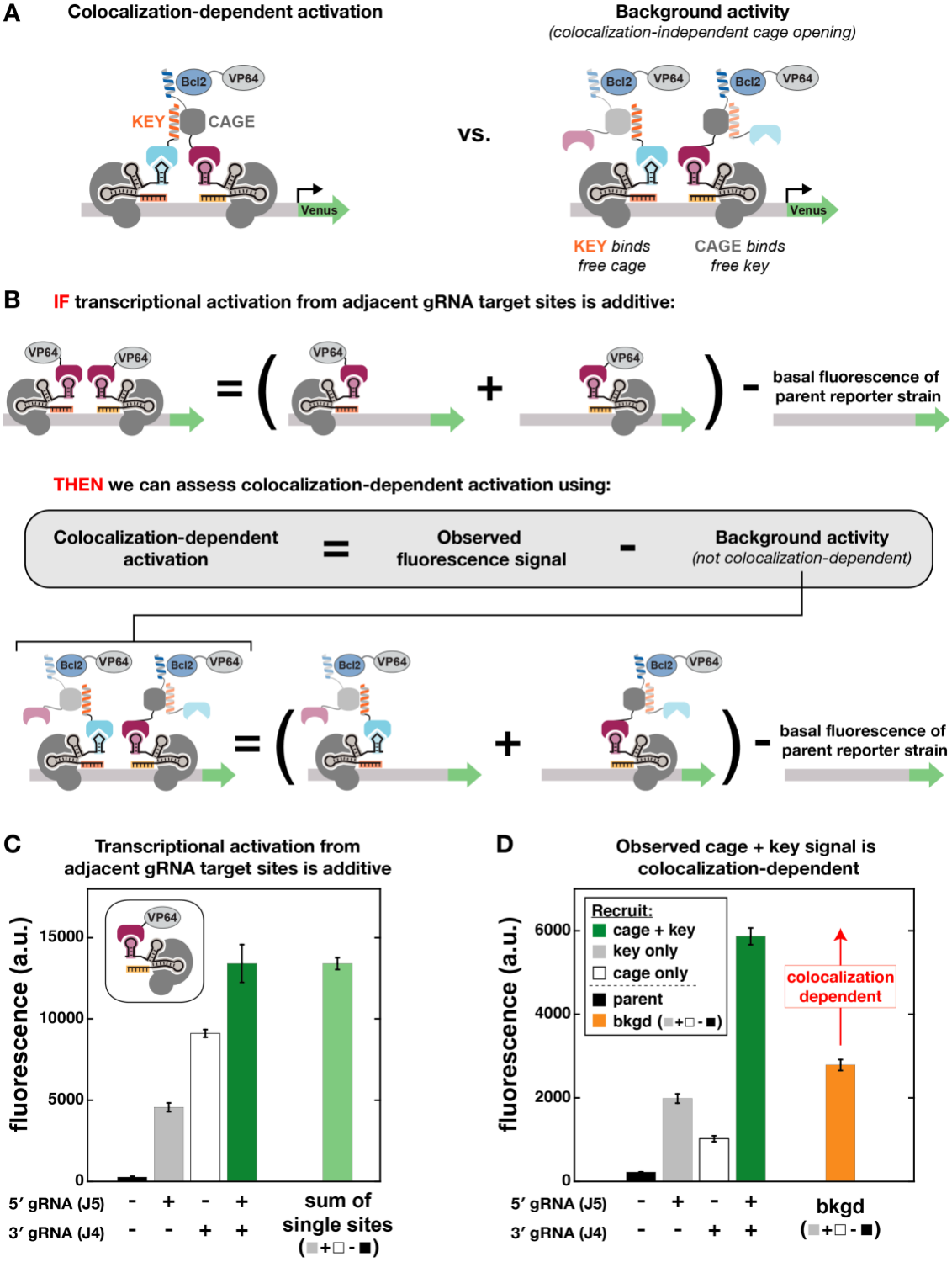
Figure 2.5 Gene activation is highly sensitive to CRISPR target site position.

Tuning key length reduces background activation of the CRISPR Co-LOCKR switch (2x PP7:key & 1x com:cage). The 44 aa key is the original key used in previous experiments. When the key is truncated to 34 aa, background activation decreases while the background corrected colocalization-dependent activation is not significantly affected. The background colocalization-independent activation is calculated from the sum of “key only” and “cage only” samples as described in Figure S2.1. Data for parent strains were obtained with the unmodified parent strains yKL014, yKL029, yKL030, yKL031, and yKL032. Fluorescence values are mean ± SD for at least three biological replicates.

2.10 | Supplemental information

2.10.1 | Supplementary figures

Figure S1: Reporter activation is colocalization-dependent



Supplementary Figure S2.1 Reporter gene activation is colocalization dependent

Reporter gene activation with the CRISPR-Cas Co-LOCKR switch is colocalization-dependent. (A)

Observed reporter gene activation could arise from colocalization-dependent switch opening, which allows the Bim peptide to recruit the Bcl-VP64 activator. Alternatively, reporter gene activation could arise from recruitment of Bcl-VP64 to each CRISPR-Cas complex. Although the co-LOCKR is designed to open only in when cage and key are colocalized, high concentrations of cage or key could drive the equilibrium towards the open state. (B) We can separately measure reporter gene activation from a recruited cage in the presence of co-expressed key, and recruited key in the presence of co-expressed cage. To predict the contribution from colocalization-independent cage opening when both cage and key are recruited, we need to assess whether VP64 recruitment to adjacent gRNA sites is additive. We can measure reporter gene activation when VP64 is recruited to both sites or each site individually. We can calculate an additive sum from each single target site for comparison to the observed activation when VP64 is recruited to both sites. The sum must be corrected for basal fluorescence from the parent reporter strain, which will be included twice if observed when the values for each single target site are added. This correction can be explicitly derived as:

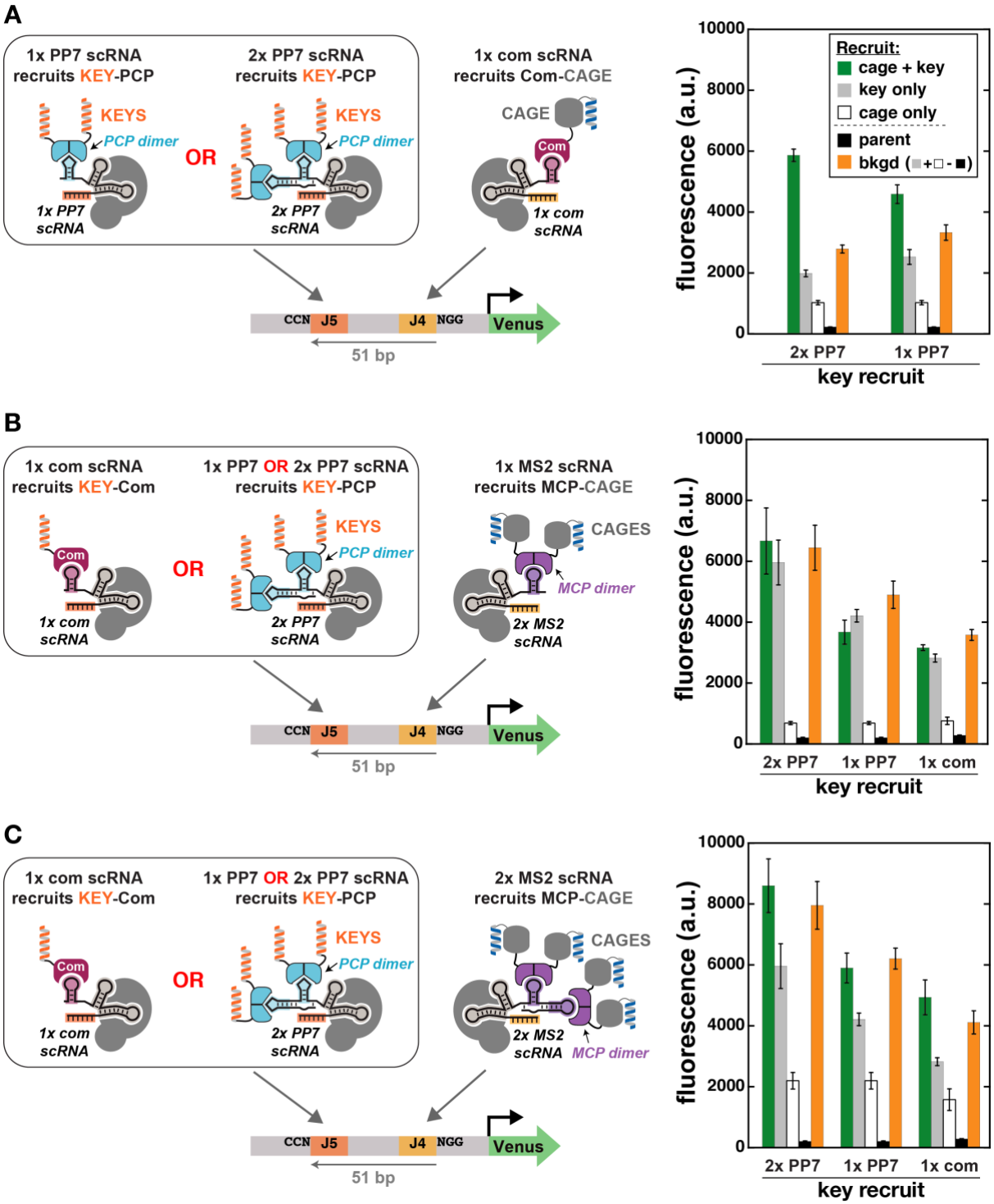
$$A_{\text{both sites}} = A_{\text{site1}} + A_{\text{site2}}$$

$$(\text{Obs} - \text{basal})_{\text{both sites}} = (\text{Obs} - \text{basal})_{\text{site1}} + (\text{Obs} - \text{basal})_{\text{site2}}$$

$$\text{Obs}_{\text{both,sites}} = \text{Obs}_{\text{site1}} + \text{Obs}_{\text{site2}} - \text{basal}$$

In practice, the correction for basal parental strain fluorescence is negligible. (C) Reporter gene activation from directly recruiting VP64 to both target sites is indistinguishable from the additive sum of VP64 recruitment to each individual target site. Data without either guide present was obtained by transforming the parental yRK266 strain with empty vector (pRS316). (D) Reporter gene activation from the co-LOCKR switch is larger than the background activation predicted from the sum of measured background activity from independent recruitment of cage and key. This co-LOCKR switch uses a 2x PP7 scRNA to recruit key-PCP and a 1x com scRNA to recruit Com-cage (Figure 2.2). Data for parent was obtained with the unmodified parent strain yKL014. Fluorescence values in (C) and (D) are mean ± SD for at least three biological replicates. The errors are propagated by adding in quadrature.

Figure S2: Alternate RNA recruitment modules affect switch activation

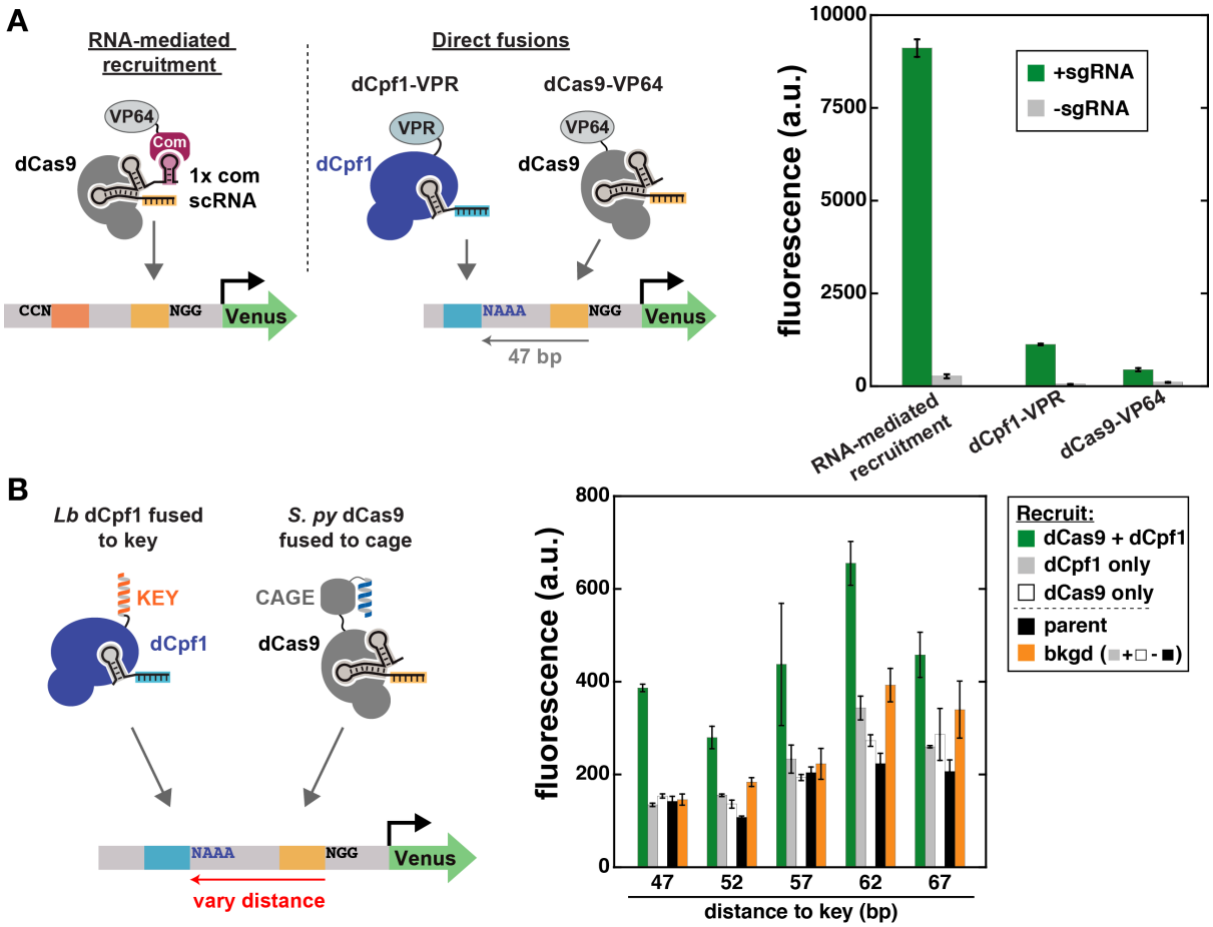


Supplementary Figure S2.2 Alternate RNA recruitment modules affect switch activation

Alternative RNA recruitment modules affect switch activation. (A) Either a 2x PP7 or a 1x PP7 scRNA targets the upstream site (J5) and recruits either four or two key-PCP fusion proteins. A 1x com scRNA recruits one Com-cage fusion protein to the downstream site (J4). Both key recruitment strategies

produce switch activation above background. The 2x PP7 switch produces more activation than 1x PP7. (B) Either a 2x PP7, 1x PP7, or 1x Com scRNA recruits key fusion proteins to the upstream site (J5). A 1x MS2 scRNA recruits two MCP-cage fusion proteins to the downstream site (J4). In each case, there is no significant switch activation above background. (C) Either a 2x PP7, 1x PP7, or 1x Com scRNA recruits key fusion proteins to the upstream site (J5). A 2x MS2 scRNA recruits four MCP-cage fusion proteins to the downstream site (J4). In each case, there is no significant switch activation above background. For A-C, fluorescence values are mean ± SD for at least three biological replicates. Data for the parent strains were obtained from the unmodified yKL014, yKL016, or yRK456 strains. Background values are the additive sums of the "key only" and "cage only" samples (Figure S2.1). We did not test every possible pairwise combination of MS2, PP7, and com scRNA recruitment strategies because the MCP and PCP proteins are structurally homologous and behave similarly in CRISPR gene activation assays.[1]

Figure S3: RNA-mediated recruitment is more effective than direct dCas9 fusions

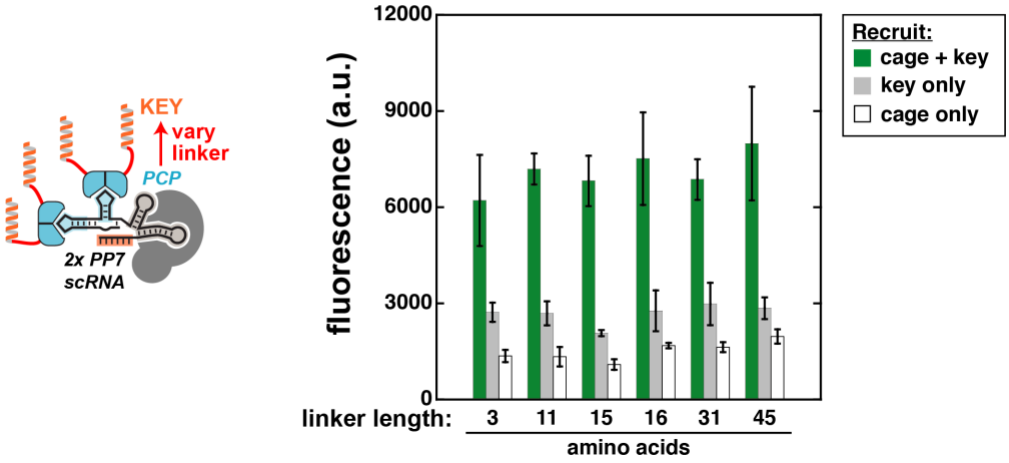


Supplementary Figure S2.3 RNA-mediated recruitment is more effective than direct dCas9 fusions

RNA recruitment is more effective than direct dCas9 fusions for co-LOCKR-mediated gene activation. (A) RNA recruitment is more effective than direct dCas9 fusions for reporter gene activation in yeast. Reporter gene activation with a 1x com scRNA and the Com-VP64 fusion protein was measured with the same reporter construct used for co-LOCKR-mediated activation. Reporter gene activation for dLbCpf1-VPR[2] and dCas9-VP64 was measured using a modified reporter gene construct where the upstream dCas9 PAM sites have been replaced with dCpf1-compatible PAM sites. Both dCpf1-VPR and dCas9-VP64 give significant reporter gene activation, although the values are substantially smaller than that observed with RNA-mediated recruitment of Com-VP64. This observation is consistent with prior results in yeast; in mammalian cells, dCas9 fusion proteins and RNA-mediated recruitment produce comparable reporter gene activation.[1] dCpf1-VPR is more effective than dCas9-VP64, likely because dCpf1 is fused to the tripartite VPR (VP64-p65-Rta) activator, which is more effective than VP64.[3] (B) Colocalization of dCpf1-Key and dCas9-CAGE results in significant gene activation, well

above the background predicted from the sum of colocalization-independent cage opening (Fig S4.1). The observed reporter gene activation is substantially smaller than that observed with RNA-mediated recruitment (Fig 4.2). Varying the distance between dCpf1 and dCas9 has modest effects. Data for parents were obtained by cotransforming the parental reporter strains yKL069, yKL073, yKL074, yKL075, and yKL076 with the off-target guide RNA plasmids pRK020 and pKL053. Fluorescence values are mean ± SD for at least three biological replicates.

Figure S4: Varing the key-PCP linker length has no significant effect



Supplementary Figure S2.4 Varying the key-PCP linker length has no significant effect

Varying the key-PCP linker length has no significant effect on reporter gene activation. The linker was varied between 3 and 45 amino acids (complete sequences appended below), but we observe no change in colocalization-dependent activation (cage + key) or background activation (key only, cage only). The 3 amino acid linker was used for all other experiments. Fluorescence values are mean ± SD for at least three biological replicates.

2.10.2 | Supplementary tables

Supplementary Table S2.1 Yeast strains

Strain[a]	**Genotype**	**Figure**
SO992	W303 *MAT**a** ura3 leu2 trp1 his3 can1R ade*	
yKL014	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage HO::HygB^R pAdh1-CLkey$_{44}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	2, 3, 4, 5 S1D, S2, S4
yKL016	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-MCP-CLcage HO::HygB^R pAdh1-CLkey$_{44}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	2, S2
yRK266	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pAdh1-Com-VP64*	S1C, S3A
yKL006	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage mfa2::Kan^R pAdh1-BCL2-VP64*	3, 4
yRK244	SO992 *trp1::pR4-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage mfa2::Kan^R pAdh1-BCL2-VP64 HO::HygB^R pAdh1-CLkey$_{44}$-PCP*	3
yRK245	SO992 *trp1::pR5-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage mfa2::Kan^R pAdh1-BCL2-VP64 HO::HygB^R pAdh1-CLkey$_{44}$-PCP*	3
yKL029	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage HO::HygB^R pAdh1-CLkey$_{43}$-PCPmfa2::Kan^R pAdh1-BCL2-VP64*	5
yKL030	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage HO::HygB^R pAdh1-CLkey$_{41}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	5
yKL031	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage HO::HygB^R pAdh1-CLkey$_{37}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	5
yKL032	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage HO::HygB^R pAdh1-CLkey$_{34}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	5
yRK456	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-MCP-CLcage HO::HygB^R pAdh1-CLkey$_{44}$-Com mfa2::Kan^R pAdh1-BCL2-VP64*	S2
yRK273	SO992 *trp1::R6-Venus leu2::pTdh3-dCas9-VP64*	S3A
yKL079	SO992 *trp1::R6-Venus leu2::pTdh3-dLbCpf1-VPR*	S3A
yKL069	SO992 *trp1::pR6-Venus leu2::pTdh3-CLkey$_{44}$-dLbCpf1*	S3B

	HO::HygB^R pUra3-dCas9-2xNLS-CLcage mfa2::Kan^R pAdh1-BCL2-VP64	
pKL073	SO992 *trp1::pR6+5-Venus leu2::pTdh3-CLkey$_{44}$-dLbCpf1* *HO::HygB^R pUra3-dCas9-2xNLS-CLcage mfa2::Kan^R pAdh1-BCL2-VP64*	S3B
yKL074	SO992 *trp1::pR6+10-Venus leu2::pTdh3-CLkey$_{44}$-dLbCpf1* *HO::HygB^R pUra3-dCas9-2xNLS-CLcage mfa2::Kan^R pAdh1-BCL2-VP64*	S3B
yKL075	SO992 *trp1::pR6+15-Venus leu2::pTdh3-CLkey$_{44}$-dLbCpf1* *HO::HygB^R pUra3-dCas9-2xNLS-CLcage mfa2::Kan^R pAdh1-BCL2-VP64*	S3B
yKL076	SO992 *trp1::pR6+20-Venus leu2::pTdh3-CLkey$_{44}$-dLbCpf1* *HO::HygB^R pUra3-dCas9-2xNLS-CLcage mfa2::Kan^R pAdh1-BCL2-VP64*	S3B
yKL021	*trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pCyc1-CLkey$_{44}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	4
yKL022	*trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pUra3-CLkey$_{44}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	4
yKL023	*trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pTdh3-CLkey$_{44}$-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	4
yRK325	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-PCP mfa2::Kan^R pUra3-BCL2-VP64*	4
yRK454	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-PCP*	4
yME007	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-(11aa)-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	S4
yME002	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-(15aa)-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	S4
yME008	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-(16aa)-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	S4
yME009	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-(31aa)-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	S4
yME003	SO992 *trp1::pJ1-Venus leu2::pTdh3-dCas9 his3::pTef1-Com-CLcage* *HO::HygB^R pAdh1-CLkey$_{44}$-(45aa)-PCP mfa2::Kan^R pAdh1-BCL2-VP64*	S4

[a] The SO992 parental strain has been described previously.[1] The Venus fluorescent reporter gene is a

derivative of GFP.[4]

Supplementary Table S2.2 Yeast protein expression plasmids

Plasmid[a]	**Parent Vector**[b]	**Marker**	**Promoter**[c]	**Gene**	**Used for yeast strain**
pJZC518	pNH605	*leu2*	*pTdh3*	dCas9-3xNLS	yKL006, yKL014, yKL016 yKL021, yKL022, yKL023, yKL029, yKL030, yKL031, yKL032, yRK244, yRK245, yRK266, yRK325, yRK454, yRK456
pRK201	pJW609	*kanMX*	*pAdh1*	Bcl2-VP64	yKL006, yKL014, yKL016, yKL021, yKL022, yKL023, yKL029, yKL030, yKL031, yKL032, yKL069, yKL073, yKL074, yKL075, yKL076, yRK244, yRK245, yRK456
pKL054	pJW609	*kanMX*	*pUra3*	Bcl2-VP64	yRK325
pKL001	pNH603	*his3*	*pTef1*	Com-CLcage	yKL006, yKL014, yKL021, yKL022, yKL023, yKL029, yKL030, yKL031, yKL032, yRK244, yRK245, yRK325, yRK454
pKL003	pNH603	*his3*	*pTef1*	MCP-CLcage	yKL016, yRK456

pRK442	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-Com	yRK456
pKL006	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-PCP	yKL014, yKL016, yRK244, yRK245, yRK325, yRK454
pKL024	pJW607	*hphMX*	*pCyc1*	CLkey$_{44}$-PCP	yKL021
pKL025	pJW607	*hphMX*	*pUra3*	CLkey$_{44}$-PCP	yKL022
pKL026	pJW607	*hphMX*	*pTdh3*	CLkey$_{44}$-PCP	yKL023
pKL028	pJW607	*hphMX*	*pAdh1*	CLkey$_{43}$-PCP	yKL029
pKL029	pJW607	*hphMX*	*pAdh1*	CLkey$_{41}$-PCP	yKL030
pKL030	pJW607	*hphMX*	*pAdh1*	CLkey$_{37}$-PCP	yKL031
pKL031	pJW607	*hphMX*	*pAdh1*	CLkey$_{34}$-PCP	yKL032
pME007	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-(11aa)-PCP	yME007
pME002	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-(15aa)-PCP	yME002
pME008	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-(16aa)-PCP	yME008
pME009	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-(31aa)-PCP	yME009
pME003	pJW607	*hphMX*	*pAdh1*	CLkey$_{44}$-(45aa)-PCP	yME003
pRK271	pNH604	*trp1*	*pJ1*	Venus	yKL006, yKL014, yKL016, yKL029, yKL030, yKL031, yKL032, yKL021, yKL022, yKL023, yRK266, yRK325, yRK454, yRK456
pRK339	pNH604	*trp1*	*pR4*	Venus	yRK244
pRK340	pNH604	*trp1*	*pR5*	Venus	yRK245
pRK369	pNH604	*trp1*	*pR6*	Venus	yKL069, yKL079, yRK273
pKL065	pNH604	*trp1*	*pR6+5*	Venus	yKL073
pKL066	pNH604	*trp1*	*pR6+10*	Venus	yKL074

pKL067	pNH604	*trp1*	*pR6+15*	Venus	yKL075
pKL068	pNH604	*trp1*	*pR6+20*	Venus	yKL076
pKL055	pJW607	*hphMX*	*pUra3*	dCas9-2xNLS-CLcage	yKL069, yKL073, yKL074, yKL075, yKL076
pKL058	pNH605	*leu2*	*pTdh3*	CLkey$_{44}$-dLbCpf1	yKL069, yKL073, yKL074, yKL075, yKL076
pJZC506	pNH603	*his3*	*pAdh1*	Com-VP64	yRK266
pJZC527	pNH605	*leu2*	*pTdh3*	dCas9-VP64	yRK273
pRK364	pNH605	*leu2*	*pTdh3*	dLbCpf1-VPR	yKL079

[a] pJZC518 and pJZC506 have been described previously.[1] pKL058 and pRK364 were cloned from addgene #104567;[5] dLbCpf1-VPR expression in yeast has been previously described.[2]

[b] The parent vectors used in this work are designed to integrate in single copy in the yeast genome. The pNH600 and pJW600 series of vectors have been described previously.[6,7] Vectors containing auxotrophic markers (*leu2*, *his3, trp1*) integrate at the corresponding endogenous locus. The pJW609 parent vector (kanMX marker) integrates at the *mfa2* locus and confers G418 resistance (*Kan*^{R}). The pJW607 parent vector (hphMX marker) integrates at the HO locus and confers hygromycin B resistance (*HygB*^{R}). Each of these vectors uses the same *C. alb. Adh1* terminator.[c] *S. cerevisiae* promoter pTdh3 is among the strongest yeast promoters. pAdh1 and pTef1 produce similar expression levels and are relatively strong, but somewhat weaker than pTdh3. pUra3 is a relatively weak promoter.[8,9] The minimal 247 bp pCyc1 promoter is the weakest promoter used here.[10] Based on microarray data, the endogenous pTdh3 and pUra3 promoters vary in expression level by >100-fold.[9] Complete reporter gene promoter sequences for pJ1, pR4, pR5, and pR6 derivatives are appended below.

Supplementary Table S2.3 Guide RNA expression plasmids[a]

Plasmid	**Parent Vector**	**sgRNA Target Site[b]**	**sgRNA/scRNA Design[c]**	**Figure**
pKL011	pRS316	1) J4 2) J5	1) 2x MS2 scRNA 2) 2x PP7 scRNA	2, S2

pKL013	pRS316	1) J4 2) J5	1) 1x com scRNA 2) 2x PP7 scRNA	2, 3, 4, 5, S1D, S2, S4
pKL014	pRS316	1) J4 2) J6	1) 1x com scRNA 2) 2x PP7 scRNA	3
pRK323	pRS316	1) J4 2) J7	1) 1x com scRNA 2) 2x PP7 scRNA	3
pRK324	pRS316	1) J4 2) J8	1) 1x com scRNA 2) 2x PP7 scRNA	3
pRK325	pRS316	1) J4 2) J9	1) 1x com scRNA 2) 2x PP7 scRNA	3
pRK326	pRS316	1) J4 2) J10	1) 1x com scRNA 2) 2x PP7 scRNA	3
pKL040	pRS316	1) J4 2) J5	1) 1x com scRNA 2) 1x PP7 scRNA	S2
pRK433	pRS316	1) J4 2) J5	1) 1x MS2 scRNA 2) 2x PP7 scRNA	S2
pRK440	pRS316	1) J4 2) J5	1) 2x MS2 scRNA 2) 1x PP7 scRNA	S2
pKL047	pRS316	1) J4 2) J5	1) 1x MS2 scRNA 2) 1x PP7 scRNA	S2
pRK441	pRS316	1) J4 2) J5	1) 1x MS2 scRNA 2) 1x com scRNA	S2
pRK290	pRS316	1) J4 2) J5	1) 2x MS2 scRNA 2) 1x com scRNA	S2
pRK275	pRS316	J4	2x MS2 scRNA	2, S2
pKL038	pRS316	J4	1x MS2 scRNA	S2
pKL016	pRS316	J4	1x com scRNA	2, 3, 4, 5, S1D, S2, S3, S4
pKL048	pRS316	J5	1x com scRNA	S1C, S2

pKL036	pRS316	1) J4 2) J5	1) 1x com scRNA 2) 1x com scRNA	S1C
pKL018	pRS316	J5	2x PP7 scRNA	2, 3, 4, 5, S1D, S2, S4
pKL019	pRS316	J6	2x PP7 scRNA	3
pME010	pRS316	J7	2x PP7 scRNA	3
pME011	pRS316	J8	2x PP7 scRNA	3
pME012	pRS316	J9	2x PP7 scRNA	3
pME013	pRS316	J10	2x PP7 scRNA	3
pKL039	pRS316	J5	1x PP7 scRNA	S2
pRK373	pRS313	TET	sgRNA	S3
pRK020	pRS313	W17 (off target control)	sgRNA	S3
pRK375	pRS316	R6	sgRNA (LbCpf1)	S3
pKL053	pRS316	TET_2 (off target control)	sgRNA (LbCpf1)	S3

[a] Guide RNA constructs were expressed from the pRS316 (*ura3* marker) or pRS313 (*his3* marker) CEN/ARS plasmid backbones with the SNR52 promoter and a SUP4 terminator.[1,11,12] sgRNA/scRNA designs and corresponding sequences have been described previously.[1,2,13] Guide RNA target sequences are provided in Table S4. All sgRNA and scRNA constructs (sgRNAs with 3' RNA recruitment domains)[1] are *S. py* Cas9 guide sequences unless otherwise noted.

Supplementary Table S2.4 Guide RNA target sequences

sgRNA	DNA Sequence	CRISPR system	Reporter[a]
J4	CGGTGTCCTGCGGTTACCAA	*S. py dCas9*	pJ1, pR4, pR5
J5	AGGTCGCCCGTGGTGGCCCA	*S. py dCas9*	pJ1, pR4, pR5
J6	TGGTGGCCCATGGTCACCAT	*S. py dCas9*	pJ1, pR4, pR5
J7	TGGTCACCATAGGTCACCCT	*S. py dCas9*	pJ1
J8	AGGTCACCCTTGGCAACCAA	*S. py dCas9*	pJ1
J9	AGGTCTCCGGTGGATACCGT	*S. py dCas9*	pJ1

J10	CCGGATCAAGATTGTACGTA	*S. py dCas9*	pJ1
W17	GAAGTCAGTTGACAGAGTCG	*S. py dCas9*	*off target control*
TET	ACTTTTCTCTATCACTGATA	*S. py* dCas9	pR6-Venus
R6	CGGTGCAAAGCAAAGTAAAG	*Lb dCpf1*	pR6-Venus
TET_2	CCACTCCCTATCAGTGATAGAGA	*Lb dCpf1*	*off target control*

[a] J4 is used to recruit the cage to the reporter. J5-10 are upstream target sites separated from J4 by 51, 61, 71, 81, 131, and 1671 bases respectively in the pJ1 reporter. The pR4 promoter has a 3 base insertion to access separation distances 54 and 64 bases (with J5 and J6). The pR5 promoter has a 5 base insertion to access separation distances 56 and 66 bases (with J5 and J6). Complete reporter gene promoter sequences for pJ1, pR4, pR5, and pR6 are appended below.

The target site separation distances are defined as the distance between PAM sites:

51 base separation for dCas9 target sites in pJ1:

CCN(N20)nnnnnnnnnn(N20)NGG

_____

(51 bases)

47 base separation for dCpf1 and dCas9 target sites in pR6:

(N20)TAAAnnnnnnnnnnnnnnnnnnnnn (N20)NGG

_____

(47 bases)

2.10.3 | Supplementary methods

Reporter gene promoters

All reporter gene promoters contain an array of gRNA target sites, a pCyc1 promoter fragment (bases -147 to -4, lowercase letters), an ATG start codon, and a Venus fluorescent reporter gene (annotated in green, only first 18 bases shown).

Promoters to vary distance between dCas9 target sites (Figure 2.2):

J4 (CGGTGTCCTGCGGTTACCAA), J5 (TGGGCCACCACGGGCGACCT), and J6 (**ATGGTGACCATGGGCCACCA**) sequences are annotated.

>pJ1

GCCTACGGTATCCACCGGAGACCTATGGCAGCCTCCGGCCGCCATAGGACACCTTTGGTTGCCA
AGGGTGACCT**ATGGTGACCATGGGCCACCA**CGGGCGACCTCAGGTATCCTGCGGTGTCCTGCG
GTTACCAAAGGCGTCCTTTGGGTTCCACCGGATACCTCCGGAAAGTGAAAGTCGAGCTCGGTAC
CCtatggcatgcatgtgctctgtatgtatataaaactcttgttttcttcttttctctaaatattctttccttatacattaggtcctttgtagcataaattactata
cttctatagacacgcaaacacaaatacacacactaaattaCCCGGATCAATTCGGGATGCTCGAGTCTAAAGGTGAA
GAATTA…

>pR4

GCCTACGGTATCCACCGGAGACCTATGGCAGCCTCCGGCCGCCATAGGACACCTTTGGTTGCCA
AGGGTGACCT**ATGGTGACCATGGGCCACCA**CGGGCGACCTGACCAGGTATCCTGCGGTGTCCT
GCGGTTACCAAAGGCGTCCTTTGGGTTCCACCGGATACCTCCGGAAAGTGAAAGTCGAGCTCGG
TACCCtatggcatgcatgtgctctgtatgtatataaaactcttgttttcttcttttctctaaatattctttccttatacattaggtcctttgtagcataaatta
ctatacttctatagacacgcaaacacaaatacacacactaaattaCCCGGATCAATTCGGGATGCTCGAGTCTAAAGGT
GAAGAATTA…

>pR5

GCCTACGGTATCCACCGGAGACCTATGGCAGCCTCCGGCCGCCATAGGACACCTTTGGTTGCCA
AGGGTGACCT**ATGGTGACCATGGGCCACCA**CGGGCGACCTCTGACCAGGTATCCTGCGGTGTC

CTGCGGTTACCAAAGGCGTCCTTTGGGTTCCACCGGATACCTCCGGAAAGTGAAAGTCGAGCTC
GGTACCCtatggcatgcatgtgctctgtatgtatataaaactcttgttttcttctttctctaaatattctttccttatacattaggtcctttgtagcataaa
ttactatacttctatagacacgcaaacacaaatacacacactaaattaCCCGGATCAATTCGGGATGCTCGAGTCTAAAG
GTGAAGAATTA…

Promoters to vary distance between dCpf1 and dCas9 target sites (Figure S2.3)

R6 (CTTTACTTTGCTTTGCACCG) and TET (ACTTTTCTCTATCACTGATA) gRNA target sequences are annotated.

>pR6

CTTTACTTTGCTTTGCACCGTAAACGTAAAGGAAACGCACTCAGGATACTTTTCTCTATCACTGATA
CGGAAAGTGAAAGTCGAGCTCGGTACCCtatggcatgcatgtgctctgtatgtatataaaactcttgttttcttctttctctaaatat
tctttccttatacattaggtcctttgtagcataaattactatacttctatagacacgcaaacacaaatacacacactaaattaCCCGGATCAAT
TCGGGATGCTCGAGTCTAAAGGTGAAGAATTA…

> pR6+5

CTTTACTTTGCTTTGCACCGTAAACGTAAAGGAAACGATTCGCACTCAGGATACTTTTCTCTATCAC
TGATACGGAAAGTGAAAGTCGAGCTCGGTACCCtatggcatgcatgtgctctgtatgtatataaaactcttgttttcttctttctct
ctaaatattctttccttatacattaggtcctttgtagcataaattactatacttctatagacacgcaaacacaaatacacacactaaattaCCCGG
ATCAATTCGGGATGCTCGAGTCTAAAGGTGAAGAATTA…

> pR6+10

CTTTACTTTGCTTTGCACCGTAAACGTAAAGGAAACGATTGCTACCGCACTCAGGATACTTTTCTCT
ATCACTGATACGGAAAGTGAAAGTCGAGCTCGGTACCCtatggcatgcatgtgctctgtatgtatataaaactcttgttttc
ttcttttctctaaatattctttccttatacattaggtcctttgtagcataaattactatacttctatagacacgcaaacacaaatacacacactaaattaC
CCGGATCAATTCGGGATGCTCGAGTCTAAAGGTGAAGAATTA…

> pR6+15

CTTTACTTTGCTTTGCACCGTAAACGTAAAGGAAACGATTGCTACTCCAGCGCACTCAGGATACTT
TTCTCTATCACTGATACGGAAAGTGAAAGTCGAGCTCGGTACCCtatggcatgcatgtgctctgtatgtatataaaa

ctcttgttttcttctttttctctaaatattctttccttatacattaggtcctttgtagcataaattactatacttctatagacacgcaaacacaaatacacacac

taaattaCCCGGATCAATTCGGGATGCTCGAGTCTAAAGGTGAAGAATTA…

> pR6+20

CTTTACTTTGCTTTGCACCGTAAACGTAAAGGAAACGATTGCTACTCCAGCCAAGCGCACTCAGG

ATACTTTTCTCTATCACTGATACGGAAAGTGAAAGTCGAGCTCGGTACCCtatggcatgcatgtgctctgtatgt

atataaaactcttgttttcttctttttctctaaatattctttccttatacattaggtcctttgtagcataaattactatacttctatagacacgcaaacacaaat

acacacactaaattaCCCGGATCAATTCGGGATGCTCGAGTCTAAAGGTGAAGAATTA…

Protein Sequences

Transcriptional activator

> NLS-Bcl2-linker-VP64

MPKKKRKVMAHAGRTGYDNREIVMKYIHYKLSQRGYEWDAGDVGAAPPGAAPAPGIFSSQPGHTPH

PAASRDPVARTSPLQTPAAPGAAAGPALSPVPPVVHLTLRQAGDDFSRRYRRDFAEMSSQLHLTPFT

ARGRFATVVEELFRDGVNWGRIVAFFEFGGVMCVESVNREMSPLVDNIALWMTEYLNRHLHTWIQDN

GGWDAFVELYGPSMRGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDF

DLDMLIN

Cage constructs fused to RNA binding proteins

>NLS-linker-MCP-linker-CLcage

MPKKKRKVGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQN

RKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGSGSELA

RKLLEASTKLQRLNIRLAEALLEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDD

AAKESEKILEEAREAISGSGSELAKLLLKAIAETQDLNLRAAKAFLEAAAKLQELNIRAVELLVKLTDPATI

REALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEKGSGSGSELARELLRAHAQLQRLNLELLR

ELLRALAQLQELNLDLLRLASELTDEIWIAQELRRIGDEFNAYYADAERLIREAAAASEKISREAERLIR

>NLS-linker-Com-linker-CLcage

MPKKKRKVGSMKSIRCKNCNKLLFKADSFDHIEIRCPRCKRHIIMLNACEHPTEKHCGKREKITHSDET

VRYGSGSGSELARKLLEASTKLQRLNIRLAEALLEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKS
KEIIRRAEKEIDDAAKESEKILEEAREAISGSGSELAKLLLKAIAETQDLNLRAAKAFLEAAAKLQELNIRA
VELLVKLTDPATIREALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEKGSGSGSELARELLRAH
AQLQRLNLELLRELLRALAQLQELNLDLLRLASELTDEIWIAQELRRIGDEFNAYYADAERLIREAAAAS
EKISREAERLIR

Key constructs fused to the PCP RNA binding protein

Key truncations are shown in Figure 2.4 and extended linkers are shown in Figure S2.4. The 15 and 45 amino acid linkers are derived from the disordered tether from the regulatory domain of PKA, which is well-characterized structurally and biochemically.[14] The 11, 16, and 31 amino acid linkers are derived from the XTEN linker, a flexible, extended polypeptide.[15,16]

Vary key length:

>NLS-linker-CLkey$_{44}$-linker-PCP *(44 aa key)*

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRGSGMSKTIVLSVG
EATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQV
WSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{43}$-linker-PCP *(43 aa key)*

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLKGSGMSKTIVLSVG
EATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQV
WSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{41}$-linker-PCP *(41 aa key)*

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERGSGMSKTIVLSVGEA
TRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWS
HDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{37}$-linker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISRGSGMSKTIVLSVGEATRTLT
EIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTI
VANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{34}$-linker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKGSGMSKTIVLSVGEATRTLTEIQ
STADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTIVAN
STEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

Vary key-PCP linker length:

> NLS-linker-CLkey$_{44}$-11aalinker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRGSETPGTSESAM
SKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLP
KVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{44}$-15aalinker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRRQESDTFIVSPTTF
HMSKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDS
GLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{44}$-16aalinker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRSGSETPGTSESAT
PESMSKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVV
DSGLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR

> NLS-linker-CLkey$_{44}$-31aalinker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRSGSETPGTSESAT
PESGSETPGTSESATPESMSKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAK

TAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPL
GR

> NLS-linker-CLkey$_{44}$-45aalinker-PCP

MPKKKRKVGSGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRRQESDTFIVSPTTF
HTQESSAVPVIEEDGESDSDSEDADLEVPVPMSKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGR
LRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVAT
SQVEDLVVNLVPLGR

<u>dCas9 fused to cage</u>

> dCas9-2xNLS-linker-CLcage

MLEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYH
LRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVD
AKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD
NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGA
SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT
NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF
EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFM
QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN
RLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK
DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS
KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFE
KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK

LKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL

TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDEGADPKKKRKVDPKKK

RKVGSGSGSELARKLLEASTKLQRLNIRLAEALLEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKS

KEIIRRAEKEIDDAAKESEKILEEAREAISGSGSELAKLLLKAIAETQDLNLRAAKAFLEAAAKLQELNIRA

VELLVKLTDPATIREALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEKGSGSGSELARELLRAH

AQLQRLNLELLRELLRALAQLQELNLDLLRLASELTDEIWIAQELRRIGDEFNAYYADAERLIREAAAAS

EKISREAERLIR

dCpf1 fused to key

> NLS-linker-CLkey$_{44}$-dLbCpf1-NLS

MLEPKKKRKVGSGPGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRGSGSGSKLE

KFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLN

NYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDKDEIALVNSFNGF

TTAFTGFFDNRENMFSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQEIKEKILNSDYDVED

FFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRESLS

FYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIR

DKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKEIIIQKVDEIYKVY

GSSEKLFDADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILL

KVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYAKCL

QKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCH

KLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYREVEEQGYKVSFESASKKEVDKLVEEGKLYM

FQIYNKDFSDKSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKN

PDNPKKTTTLSYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIARGERNLLY

IVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAGYISQVVHKI

CELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNYMVDKKSNPCATGGALKGYQITNK

FESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYVPEEDLFEFALDYK

NFSRTDADYIKKWKLYSYGNRIRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQ

SDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYN

IARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKHASKRPAATKKAGQAKKKK

2.10.4 | Supplementary references

(1) Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., Tsai, J. C., Weissman, J. S., Dueber, J. E., Qi, L. S., and Lim, W. A. (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell 160*, 339–350.

(2) Lian, J., HamediRad, M., Hu, S., and Zhao, H. (2017) Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun. 8*, 1688.

(3) Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C. D., Wiegand, D. J., Ter-Ovanesyan, D., Braff, J. L., Davidsohn, N., Housden, B. E., Perrimon, N., Weiss, R., Aach, J., Collins, J. J., and Church, G. M. (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods 12*, 326–328.

(4) Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2001) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol. 20*, 87–90.

(5) Tak, Y. E., Kleinstiver, B. P., Nuñez, J. K., Hsu, J. Y., Horng, J. E., Gong, J., Weissman, J. S., and Joung, J. K. (2017) Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat. Methods*.

(6) Zalatan, J. G., Coyle, S. M., Rajan, S., Sidhu, S. S., and Lim, W. A. (2012) Conformational control of the Ste5 scaffold protein insulates against MAP kinase misactivation. *Science 337*, 1218–1222.

(7) Chau, A. H., Walter, J. M., Gerardin, J., Tang, C., and Lim, W. A. (2012) Designing synthetic regulatory networks capable of self-organizing cell polarization. *Cell 151*, 320–332.

(8) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol. 4*, 975–986.

(9) Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell 95*, 717–728.

(10) Guarente, L., and Ptashne, M. (1981) Fusion of Escherichia coli lacZ to the cytochrome c gene of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA 78*, 2199–2203.

(11) DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013) Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res. 41*, 4336–4343.

(12) Sikorski, R. S., and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics 122*, 19–27.

(13) Zalatan, J. G. (2017) CRISPR-Cas RNA Scaffolds for Transcriptional Programming in Yeast. *Methods Mol. Biol. 1632*, 341–357.

(14) Smith, F. D., Reichow, S. L., Esseltine, J. L., Shi, D., Langeberg, L. K., Scott, J. D., and Gonen, T. (2013) Intrinsic disorder within an AKAP-protein kinase A complex guides local substrate phosphorylation. *eLife 2*, e01319.

(15) Guilinger, J. P., Thompson, D. B., and Liu, D. R. (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol. 32*, 577–582.

(16) Schellenberger, V., Wang, C.-W., Geething, N. C., Spink, B. J., Campbell, A., To, W., Scholle, M. D., Yin, Y., Yao, Y., Bogin, O., Cleland, J. L., Silverman, J., and Stemmer, W. P. C. (2009) A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. *Nat. Biotechnol. 27*, 1186–1190.

Chapter 3 | CRISPR-Cas-mediated tethering recruits the yeast HMR locus to the nuclear periphery but fails to silence gene expression

The work in this chapter was published and reprinted with permission from the journal of initial publication (License Number: 4574981073453):

Fontana, J., Dong, C., Ham, J. Y., Zalatan, J. G. & Carothers, J. M. Regulated Expression of sgRNAs Tunes CRISPRi in E. coli. *Biotechnology Journal* **13,** 1800069 (2018).

3.1 | Abstract

The physical organization of the genome plays a central role in biological processes ranging from cell division to gene regulation. To investigate the relationship between genome structure and function, we have developed a programmable CRISPR-Cas system for nuclear peripheral recruitment in yeast. We benchmarked this system at the HMR locus, a paradigm for recruitment-mediated silencing. Prior studies demonstrate that Gal4-mediated tethering to the nuclear periphery can silence a reporter gene at the HMR locus. Using microscopy and gene silencing assays, we demonstrate that CRISPR-Cas-mediated tethering can recruit the HMR locus but does not silence reporter gene expression. Additional control experiments with Gal4-mediated tethering shows that the silencing phenotype has an unexpected dependence on the structure of the protein tether. Next, we targeted the CRISPR-Cas recruitment system at different distances relative to the HMR locus and found that extent of recruitment depends on the distance between the gRNA target site and the target locus. Finally, we targeted the CRISPR-Cas recruitment system to the Gal2 locus, a model for inducible nuclear peripheral recruitment, and observed no detectable repositioning. These data suggest potential limitations for CRISPR-Cas-based peripheral recruitment and recruitment-mediated silencing in yeast.

3.2 | Introduction

An emerging body of data suggests that the 3D spatial organization of the genome plays an important role in eukaryotic gene regulation[1-4]. For example, genes positioned near the nuclear periphery tend to be repressed, and genes positioned in the nuclear interior tend to be active[3,4]. While genome organization is often conserved between cell types and organisms, there is significant variability that may reflect key functional differences[5]. Further support for the possibility that genome organization plays an important regulatory role comes from the observation that genes can be dynamically repositioned in response to extracellular signals[6-7]. To probe the biological function of genome structure, we have developed a programmable CRISPR-Cas system to relocalize genes to the nuclear periphery in yeast.

Prior methods to reposition genes have fused well characterized DNA binding domains (DBD), such as Gal4, to a recruitment domain protein that directs the tethered gene to specific sites in the nucleus[7-12]. In these studies, repositioning genes sometimes, but not always, leads to predictable changes in gene expression. The number of sites that have been studied with this approach has been relatively limited, in part because DBDs typically recognize specific DNA sequences and these motifs must be engineered into each genomic site of interest.

To address this challenge, CRISPR-Cas genome repositioning systems have been developed to target and spatially reposition genes within the nucleus. Because CRISPR-Cas targeting is programmable, such systems enable recruitment of endogenous genes and bypass the need for site specific gene modification of the recruitment target site[13,14]. In human cells, dCas9 fusion to a chemically-inducible dimerization domain allowed inducible recruitment to the nuclear envelope and other subnuclear sites. Reporter and endogenous gene expression could be perturbed by nuclear repositioning with this system, and localization of telomers to the nuclear periphery resulted in cellular toxicity[13]. In yeast cells, dCas9 fusion to a cohesin domain enabled targeting to a dockerin fused to a nuclear membrane protein. This system successfully recruited multiple endogenous loci to the nuclear periphery and was able to affect plasmid segregation to daughter cells[14]. These findings suggest that CRISPR gene relocalization systems will be useful to discover relationships between gene positioning and cellular behavior.

In parallel to the methods described above, we developed an alternative CRISPR-Cas repositioning system in yeast and tested it at the HMR locus, a well-characterized model system for

gene silencing. Gal4-mediated recruitment of HMR to the nuclear periphery can rescue silencing defects[8] and we assessed whether a CRISPR-Cas-mediated recruitment strategy would have the same functional effects. Using a nuclear membrane protein as the recruitment domain, we targeted either the CRISPR-Cas system or the Gal4 system to the HMR locus and measured nuclear peripheral relocalization by microscopy. We found that both systems produce similar, significant levels of recruitment. Next, we compared the ability of CRISPR-Cas and Gal4 recruitment systems to modulate gene expression, and found that Gal4, but not dCas9 based recruitment, is sufficient to silence the HMR locus. The Gal4-mediated silencing phenotype was sensitive to the structure of the Gal4-membrane protein fusion. Inserting a protein spacer between Gal4 and the nuclear periphery maintains recruitment but abrogates silencing, suggesting alternative tethering strategies can recruit genes to the periphery but silencing effects may depend on precise structural rules. We also tested the CRISPR-Cas recruitment system at the Gal2 locus, which is a model for dynamic repositioning in response to external stimuli[6,15,16]. We were unable to recruit Gal2 to the nuclear periphery with the CRISPR-Cas system, suggesting that not all endogenous target sites can be synthetically relocalized to the nuclear periphery.

3.3 | Results and discussion

3.3.1 | Gal4 can recruit the HMR locus to the nuclear periphery but fails to silence gene expression

The yeast *HMR* locus provided one of the earliest examples of a functional effect from gene repositioning, making it an ideal model system to prototype a new repositioning system. *HMR* contains a backup copy of yeast mating type sequences and is natively silenced[17]. Deletion of *HMR* regulatory regions results in locus de-repression[18], and silencing can be restored by synthetically tethering silencing factors using the Gal4 DNA binding domain[19]. Notably, Gal4-mediated tethering of nuclear membrane proteins to *HMR* can also restore gene silencing, likely due to the presence of high concentrations of silencing factors at the nuclear periphery[8].

To compare the ability of Gal4 or dCas9 to recruit *HMR* to the nuclear periphery in yeast, we constructed an *HMR* silencing reporter in *S. cerevisiae*, following previous designs[8,19]. We replaced the endogenous copy of *HMR* with a mutant in the *HMR-E* regulatory region containing a 2xUAS_G site upstream of a Trp1 reporter (Aeb::HMR) (see Supplemental Methods). We fused the Gal4 DBD to the nuclear membrane protein Yif1 and confirmed that $Gal4_{DBD}$-Yif1 expression silences the Trp1 reporter gene in a cell-spotting growth assay, as described previously[8] (Fig 3.2). To determine if $Gal4_{DBD}$-Yif1

physically repositions *HMR* to the nuclear periphery, which was not previously assessed, we further engineered the silencing reporter strain with a tetO array 2.4 kb downstream of the UAS_G site. Expression of tetR-GFP in this strain can be used to visualize the position of the *HMR* locus[20], and we confirmed that this modified reporter still silenced $Gal4_{DBD}$-Yif1 expression (Fig. S3.1). We expressed the mCherry-Heh2 fusion protein[21,22] to label the nuclear membrane, and used confocal microsopy to measure colocalization of the tetO array to the nuclear rim. We observed a significant change in peripheral localization when $Gal4_{DBD}$-Yif1 was expressed compared to the parental, increasing from 39% to 56% (Fig 3.2B).

3.3.2 | CRISPR-Cas can the recruit the HMR locus to the nuclear periphery

To determine if the CRISPR-Cas system can relocalize the *HMR* locus, we physically linked the CRISPR-Cas complex to Yif1, the same nuclear membrane protein used in the Gal4 tethering strategy. We linked Yif1 to the CRISPR-Cas complex through scaffold RNAs (scRNAs), which are sgRNAs engineered with additional hairpin motifs to recruit effectors fused to RNA binding proteins[23-26] (Fig 3.1). In this strategy, Yif1 is fused to an RNA binding protein, which engages directly with the scRNA. To target dCas9 to the *HMR* locus, we designed two guide RNA target sites adjacent to or overlapping the UAS_G site and cloned these sites into an scRNA construct with two MS2 RNA hairpins (Table S1). We transformed each individual scRNA construct separately into the reporter yeast strain expressing dCas9 and an MS2 coat protein (MCP) fusion to the N-terminus of Yif1 (MCP-Yif1), which binds the MS2 hairpins as a dimer. Using confocal microscopy, we observed a significant increase in peripheral localization with the scRNA-containing strains (Fig 3.2C). Guide RNAs with an off-target (OT) sequence or lacking the MS2 hairpins (–MS2) gave no significant peripheral recruitment, as expected.

Both dCas9 and Gal4 recruitment strategies resulted in similar recruitment phenotypes (Fig 3.2). For both the CRISPR-Cas and the Gal4 recruitment strategies, the 50-60% peripheral localization that we observe is comparable to the 50-80% range of values observed for endogenous and heterologous peripheral recruitment reported in the literature[7,22,27,28]. The background peripheral localization for the unrecruited *HMR* silencing reporter was ~40%, which is modestly larger than the 30% peripheral localization values reported for unrecruited genes in other systems. This larger background for the *HMR* reporter may be due to its proximity to the ChrIII telomere, as telomeres are endogenously localized to the periphery in yeast[29,30].

In addition to targeting sites relatively close the tetO array (2.4 kb), we also tested gRNA target sites at increasing distances. We designed MS2 scRNAs to target additional sites 9 kb, 15 kb, and 102 kb from the tetO array. For each of these sites, we did not detect statistically significant increases in peripheral localization (Fig 3.2C). These data suggest that the target site and microscopy reporter need to be in close physical proximity to observe gene relocalization. How these distances in base pairs translate to physical distances is uncertain. Using a base pair to distance conversion derived from microscopy data of 130 nm/bp[31], we can roughly estimate that these sites are 56 nm, 96 nm, and 450 nm respectively from the tetO array, but we lack any direct measurements of the distances for these specific sites to *HMR*. Regardless of the precise distance relationship, however, our data suggest that a CRISPR-Cas-based gene recruitment system can localize adjacent genomic regions to the periphery using a single scRNA targeting a unique site in the genome.

3.3.3 | Peripheral recruitment is not sufficient to silence reporter gene expression

To determine if CRISPR-Cas-mediated *HMR* recruitment produced the same silencing effect on gene expression as observed with Gal4-Yif1, we assessed Trp1 reporter gene expression using the cell-spotting growth assay. Although CRISPR-Cas-mediated tethering to Yif1 and $Gal4_{DBD}$-Yif1 was indistinguishable by microscopy (Fig 3.2), there was no detectable silencing at *HMR* with the CRISPR-Cas system (Fig 3.3). In addition to the two scRNA target sites used in microscopy assays (Fig 3.2), we tested 3 additional nearby target sites but saw no Trp silencing with any of these scRNAs. There are no immediately obvious distinguishing features of the CRISPR-Cas and Gal4 systems that could explain their distinct behaviors. The affinities of the recruitment interactions are all relatively similar, in the range of ~10^{-9} M, for Gal4 dimer binding to the UAS_G[32], dCas9-gRNA binding to cognate DNA[33,34], and the MS2 RNA hairpin for the MCP dimer (the $MCP_{V29I\Delta FG}$ variant)[35].

$Gal4_{DBD}$-Yif1-mediated silencing at *HMR* is known to require the presence of endogenous cis-regulatory sites[8], and it is possible that the precise structural arrangement of the $Gal4_{DBD}$-Yif1 fusion protein relative to these sites or other associated regulatory factors might be important for silencing. To test this possibility, we inserted maltose binding protein (MBP) between $Gal4_{DBD}$ and Yif1 ($Gal4_{DBD}$-M-Yif1), which we estimate extends the distance between $Gal4_{DBD}$ and Yif1 by ~40 Å based on the crystal structure of MBP[36]. By microscopy, $Gal4_{DBD}$-M-Yif1 resulted in peripheral localization that was indistinguishable from $Gal4_{DBD}$-Yif1 or CRISPR mediated recruitment (Fig 3.2B, 3.2C). Unlike $Gal4_{DBD}$-

Yif1, however, the $Gal4_{DBD}$-M-Yif1 construct produced only a partial silencing phenotype (Fig 3.3). This observation suggests that, at least at the *HMR* locus in yeast, peripheral gene silencing may depend on the precise structure of the recruitment machinery.

Reporter gene expression can be silenced indirect by Gal4 fusion to Yif1 and directly by Gal4 fusions to silencing factor proteins[8,19]. To determine if this behavior is an idiosyncratic feature of Yif1 mediated perinuclear silencing or a more general property, we assessed the ability of dCas9 to silence Aeb::HMR by recruiting the Sir1 silencing protein. First, we tethered Sir1 to the Aeb::HMR Trp1 reporter by expressing $Gal4_{DBD}$-Sir1 and found significant gene silencing, consistent with previous observations (Fig S3.2). To assess if CRISPR-Cas mediated Sir1 is able to silence the Trp1 reporter, we targeted Sir1 to Aeb::HMR Trp1 reporter using the CRISPR-Cas complex and observed no gene silencing (Fig S3.2). Thus, Gal4, but not dCas9, mediated recruitment of both Sir1 and Yif1 is sufficient to confer gene silencing at the Aeb::HMR reporter.

3.3.4 | The CRISPR-Cas system does not recruit the Gal2 locus to the nuclear periphery

To determine if CRISPR-Cas-mediated recruitment is effective at other sites in the genome, we targeted the *GAL2* locus. In response to galactose, yeast cells localize *GAL2* to the nuclear pore complex and activate *GAL2* expression by ~20-fold[6,15,37]. This behavior indicates that it is possible to reposition *GAL2* and provides a useful positive control for comparison to synthetic recruitment strategies. We constructed a reporter strain with a tetO array at the *GAL2* locus to visualize its position and confirmed that galactose induction recruits *GAL2* to the periphery. The extent of peripheral localization increased from 32% in glucose to 47% in galactose, comparable to previously reported results[15] (Fig 3.4B). When we used a Yif1-tethered scRNA to target the CRISPR-Cas system to *GAL2*, however, we did not detect significant repositioning to the nuclear periphery (Fig 3.4C). We also tested a system with simultaneous expression of four scRNAs targeting adjacent sites at *GAL2*, but again observed no detectable repositioning (Fig 3.4B). Thus, although our CRISPR-Cas recruitment system was effective at *HMR*, we were unable to detect repositioning at a different site in the yeast genome. This behavior is in contrast with an alternative recruitment system, CRISPR-PIN, which was able to effectively recruit multiple distinct endogenous loci in yeast[14].

3.4 | Conclusions

Recently, it has been reported that dCas9 can be used form DNA loops and relocalize genes to the

nuclear periphery in a variety of cell types. Prior results in yeast suggest that dCas9 can recruit an engineered target site to the nuclear periphery in budding yeast using single sgRNAs. In principal, such tools could offer the ability to use both targeted hypothesis-driven probes at individual loci and unbiased screens that span a region of interest or the entire genome. In this study, we highlight the challenges of discovery-based CRISPR-mediated re-localization assays in yeast. First, we find that we can recruit the silencing-defective HMR locus, but not the endogenous Gal2 gene to the nuclear periphery in yeast. Second, HMR mediated silencing is obtainable when Gal4-Yif1, but not Gal4-MBP-Yif1 or CRISPR-Cas Yif1 is used, suggesting significant context rules for peripheral-mediated silencing. These observations suggest potential limitations for CRISPR-Cas based re-localization experiments in yeast.

3.5 | Acknowledgements

The authors thank Susan Gasser, Sue Biggins, Dan Gottschling, Joshua Vaughan, Tyler Chozinski, Dustin Maly, Jack Rose, Jay Shendure, Bill Noble, David Shechner, James Carothers, Jason Brickner, Wendell Lim, Geeta Narlikar, Wai Chan, the Biology Imaging Facility at the University of Washington, and members of the Zalatan group for technical assistance, advice, and helpful discussions. This work was supported by the University of Washington (J.G.Z.), a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (J.G.Z.), a Genome Sciences NIH Training Grant T32 HG00035 (R.L.K.), and NIH R35 GM124773 (J.G.Z.).

3.8 | References

(1) Avşaroğlu B, Bronk G, Li K, Haber JE, Kondev J. Chromosome-refolding model of mating-type switching in yeast. Proc Natl Acad Sci USA. 2016;113: E6929–E6938. doi:10.1073/pnas.1607103113

(2) Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326: 289–293. doi:10.1126/science.1181369

(3) Egecioglu D, Brickner JH. Gene positioning and expression. Curr Opin Cell Biol. 2011;23: 338–345. doi:10.1016/j.ceb.2011.01.001

(4) Ramani V, Shendure J, Duan Z. Understanding Spatial Genome Organization: Methods and Insights. Genomics Proteomics Bioinformatics. 2016;14: 7–20. doi:10.1016/j.gpb.2016.01.002

(5) Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485: 376–380. doi:10.1038/nature11082

(6) Casolari JM, Brown CR, Komili S, West J, Hieronymus H, Silver PA. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell. 2004;117: 427–439. doi:10.1016/s0092-8674(04)00448-9

(7) Brickner JH, Walter P. Gene recruitment of the activated INO1 locus to the nuclear membrane. PLoS Biol. 2004;2: e342. doi:10.1371/journal.pbio.0020342

(8) Andrulis ED, Neiman AM, Zappulla DC, Sternglanz R. Perinuclear localization of chromatin facilitates transcriptional silencing. Nature. 1998;394: 592–595. doi:10.1038/29100

(9) Taddei A, Hediger F, Neumann FR, Bauer C, Gasser SM. Separation of silencing from perinuclear anchoring functions in yeast Ku80, Sir4 and Esc1 proteins. EMBO J. 2004;23: 1301–1312. doi:10.1038/sj.emboj.7600144

(10) Menon BB, Sarma NJ, Pasula S, Deminoff SJ, Willis KA, Barbara KE, et al. Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. Proc Natl Acad Sci USA. 2005;102: 5749–5754. doi:10.1073/pnas.0501768102

(11) Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, et al. Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS Genet. 2008;4: e1000039. doi:10.1371/journal.pgen.1000039

(12) Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature. 2008;452: 243–247. doi:10.1038/nature06727

(13) Wang H, Xu X, Nguyen CM, Liu Y, Gao Y, Lin X, et al. CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization. Cell. 2018;175: 1405–1417.e14. doi:10.1016/j.cell.2018.09.013

(14) Lin J-L, Ekas H, Deaner M, Alper HS. CRISPR-PIN: Modifying gene position in the nucleus via dCas9-mediated tethering. Synth Syst Biotechnol. 2019;4: 73–78. doi:10.1016/j.synbio.2019.02.001

(15) Dieppois G, Iglesias N, Stutz F. Cotranscriptional recruitment to the mRNA export receptor Mex67p contributes to nuclear pore anchoring of activated genes. Mol Cell Biol. 2006;26: 7858–7870. doi:10.1128/MCB.00870-06

(16) Sood V, Brickner JH. Nuclear pore interactions with the genome. Curr Opin Genet Dev. 2014;25: 43–49. doi:10.1016/j.gde.2013.11.018

(17) Haber JE. Mating-type genes and MAT switching in Saccharomyces cerevisiae. Genetics. 2012;191: 33–64. doi:10.1534/genetics.111.134577

(18) Brand AH, Micklem G, Nasmyth K. A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation. Cell. 1987;51: 709–719. doi:10.1016/0092-8674(87)90094-8

(19) Chien CT, Buck S, Sternglanz R, Shore D. Targeting of SIR1 protein establishes transcriptional silencing at HM loci and telomeres in yeast. Cell. 1993;75: 531–541. doi:10.1016/0092-8674(93)90387-6

(20) Rohner S, Gasser SM, Meister P. Modules for cloning-free chromatin tagging in Saccharomyces cerevisae. Yeast. 2008;25: 235–239. doi:10.1002/yea.1580

(21) Meinema AC, Laba JK, Hapsari RA, Otten R, Mulder FAA, Kralt A, et al. Long unfolded linkers facilitate membrane protein import through the nuclear pore complex. Science. 2011;333: 90–93. doi:10.1126/science.1205741

(22) Egecioglu DE, D'Urso A, Brickner DG, Light WH, Brickner JH. Approaches to studying subnuclear organization and gene-nuclear pore interactions. Methods Cell Biol. 2014;122: 463–485. doi:10.1016/B978-0-12-417160-2.00021-7

(23) Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013;31: 833–838. doi:10.1038/nbt.2675

(24) Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013;154: 442–451.

doi:10.1016/j.cell.2013.06.044

(25) Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell. 2015;160: 339–350. doi:10.1016/j.cell.2014.11.052

(26) Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature. 2015;517: 583–588. doi:10.1038/nature14136

(27) Ahmed S, Brickner DG, Light WH, Cajigas I, McDonough M, Froyshteter AB, et al. DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery. Nat Cell Biol. 2010;12: 111–118. doi:10.1038/ncb2011

(28) Brickner DG, Sood V, Tutucci E, Coukos R, Viets K, Singer RH, et al. Subnuclear positioning and interchromosomal clustering of the GAL1-10 locus are controlled by separable, interdependent mechanisms. Mol Biol Cell. 2016;27: 2980–2993. doi:10.1091/mbc.E16-03-0174

(29) Palladino F, Laroche T, Gilson E, Axelrod A, Pillus L, Gasser SM. SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. Cell. 1993;75: 543–555. doi:10.1016/0092-8674(93)90388-7

(30) Gotta M, Laroche T, Formenton A, Maillet L, Scherthan H, Gasser SM. The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type Saccharomyces cerevisiae. J Cell Biol. 1996;134: 1349–1363. doi:10.1083/jcb.134.6.1349

(31) Bystricky K, Heun P, Gehlen L, Langowski J, Gasser SM. Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. Proc Natl Acad Sci USA. 2004;101: 16495–16500. doi:10.1073/pnas.0402766101

(32) Reece RJ, Ptashne M. Determinants of binding-site specificity among yeast C6 zinc cluster proteins. Science. 1993;261: 909–911. doi:10.1126/science.8346441

(33) Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014;507: 62–67. doi:10.1038/nature13011

(34) Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol. 2016;34: 339–344. doi:10.1038/nbt.3481

(35) Lim F, Peabody DS. Mutations that increase the affinity of a translational repressor for RNA. Nucleic Acids Res. 1994;22: 3748–3752. doi:10.1093/nar/22.18.3748

(36) Sharff AJ, Rodseth LE, Spurlino JC, Quiocho FA. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. Biochemistry. 1992;31: 10657–10663. doi:10.1021/bi00159a003

(37) Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, et al. Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc Natl Acad Sci USA. 1997;94: 13057–13062. doi:10.1073/pnas.94.24.13057

3.9 | Figures

Figure 1: CRISPR-Cas Recruits DNA Sites to Subnuclear Compartments

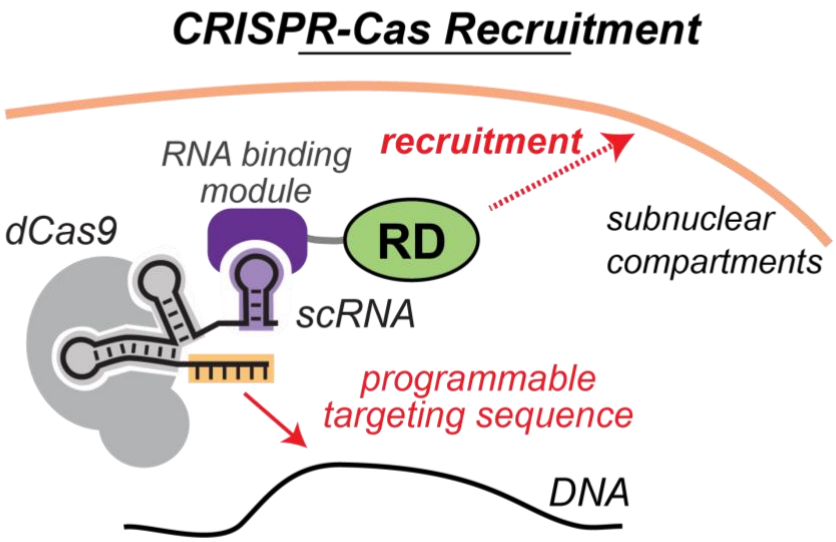


Figure 3.1 CRISPR-Cas based gene relocalization

The CRISPR-Cas system recruits DNA sites to subnuclear compartments. A programmable CRISPR-Cas DNA binding complex is physically linked to an effector protein recruitment domain (RD) via a scaffold RNA (scRNA), a modified gRNA that includes additional RNA hairpins to recruit RNA binding modules. RDs can be nuclear envelope proteins, lamins, nucleoporins, or other proteins that localize to specific sites in the nucleus.

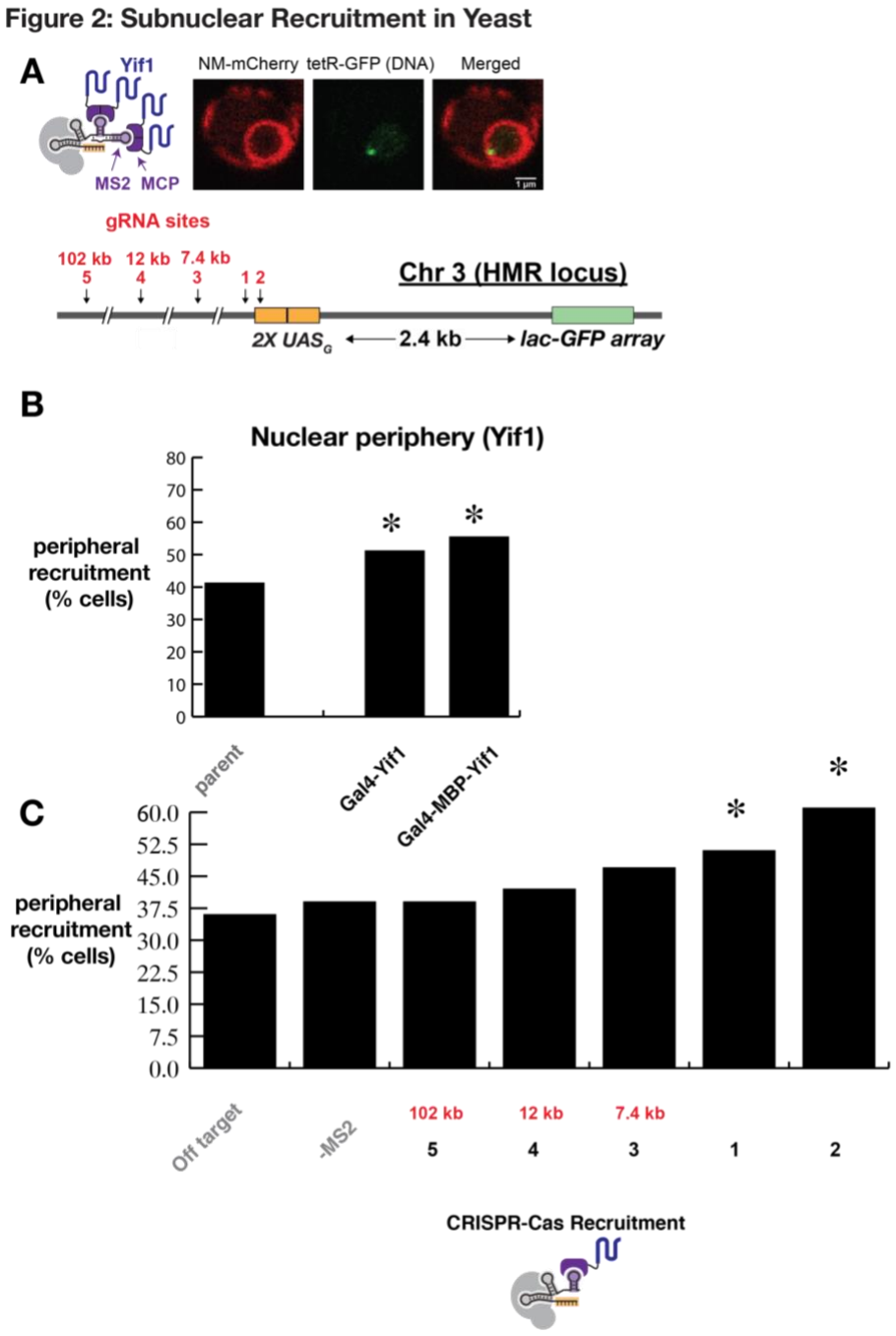


Figure 3.2 Subnuclear recruitment in yeast

A) A 2x MS2 scRNA recruits the MCP dimer. Each MCP monomer is fused to Yif1, a nuclear membrane protein. The CRISPR-Cas complex targets sites at the HMR locus, which has been engineered to include a tetO-GFP array. TetR-GFP binds to the array and defines the position of the HMR locus, and an mCherry-Heh2 fusion protein defines the nuclear envelope. Representative confocal microscopy images show the nuclear envelope, defined by an mCherry-Heh2 fusion protein, and the HMR locus,

defined by a tetR-GFP protein that binds the tetO-GFP array.

B) Peripheral recruitment (scored as described in the methods) for yeast strains with and without recruitment systems. CRISPR-Cas results in a significant increase in peripheral localization when targeted to the *HMR* locus with a single scRNA target site. The parent yeast strain (yRK119) is indistinguishable from negative control strains with an off-target scRNA (OT, see Table S1) or a gRNA lacking the MS2 recruitment hairpins (–MS2). The *p* values for targets 1-5, relative to the –MS2 control, are 0.03, 0.002, 0.05, 0.006, and 0.06, respectively. $Gal4_{DBD}$-Yif1 and $Gal4_{DBD}$-M-Yif1 (containing MBP between $Gal4_{DBD}$ and Yif1) also show significant peripheral recruitment relative to a strain containing only $Gal4_{DBD}$ (*p* values 0.01 and 0.0009, respectively). Values reported are the mean ± SD for three measurements (technical replicates). Each measurement was performed with 30 cells. Statistical significance for increased recruitment relative to a corresponding negative control (indicated by *) was evaluated using a one-sided unpaired *t*-test with a cutoff $p \leq 0.05$.

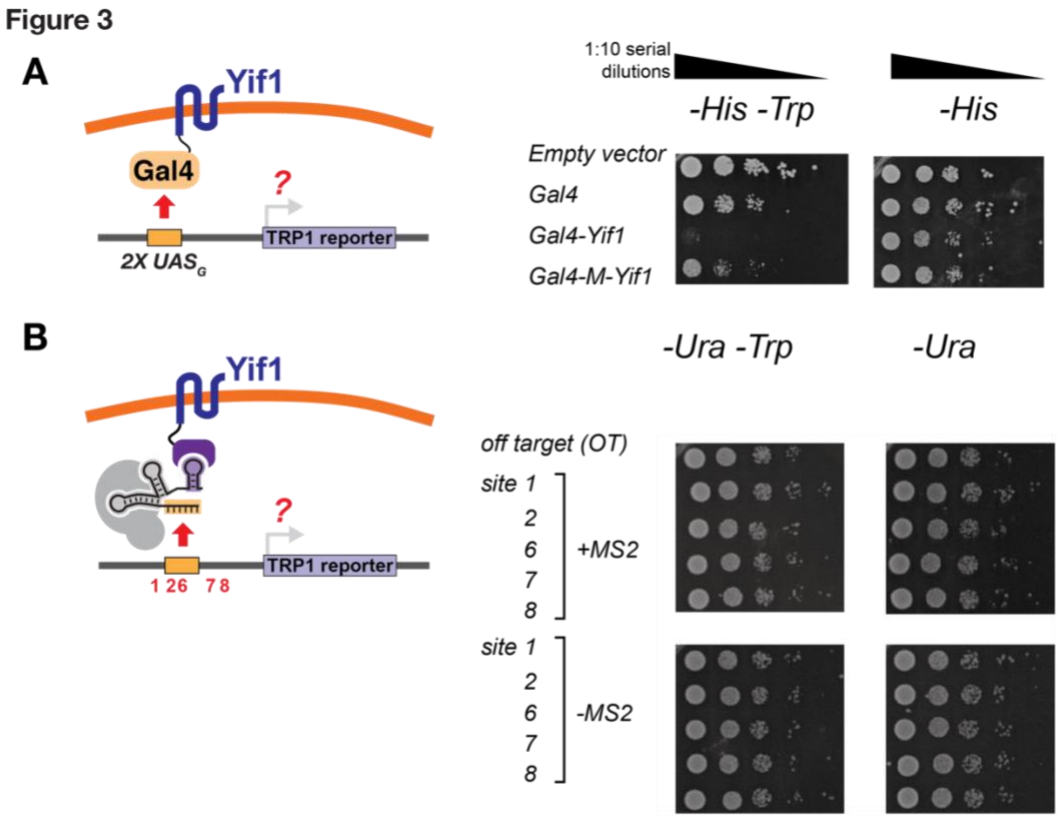


Figure 3.3 Trp reporter assay for gene silencing at HMR in yeast

A) Peripheral recruitment of the HMR locus silences a Trp1 reporter gene, leading to a growth defect on media lacking Trp. Consistent with published reports[8], $Gal4_{DBD}$-Yif1 expression results in a growth defect on –His –Trp media. Cells expressing $Gal4_{DBD}$-M-Yif1, which has MBP inserted between $Gal4_{DBD}$ and Yif1, produces in a partial silencing phenotype. The parent strain is yRK036 and Gal4 constructs are delivered on p423 (His selection).

B) The CRISPR-Cas recruitment system, which we verified recruits HMR to the periphery using sites 1 and 2, does not silence the Trp1 reporter gene. There is no detectable growth defect on –Ura –Trp media. The parent strain is yRK045, and 2x MS2 scRNA or –MS2 sgRNA constructs are delivered on pRS316 (Ura selection).

For A) and B), images are representative of three independent experiments (biological replicates).

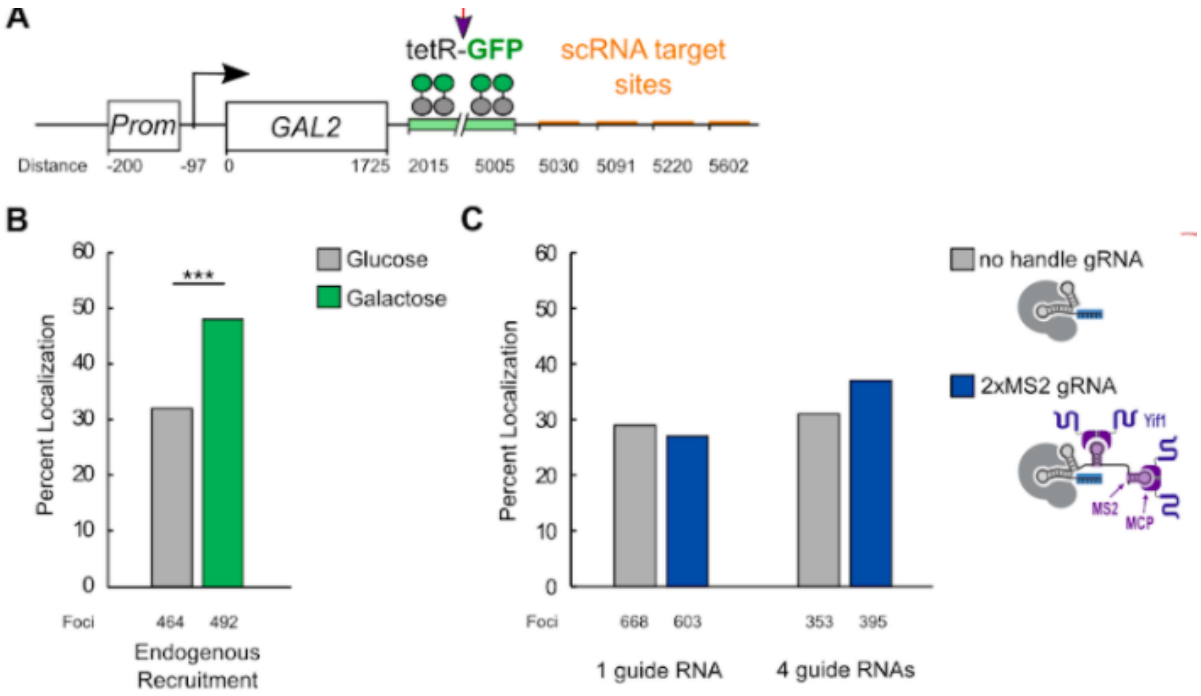


Figure 3.4 CRISPR-Cas targeting to the Gal2 locus results in no gene recruitment

A) Schematic of localization reporter at the Gal2 locus. B) Growth of yeast in different carbon sources results in differential recruitment to the nuclear periphery. C) Targeting of the CRISPR-Cas system to the Gal2 locus using either 1 sgRNA, or 4 sgRNAs is insufficient to recruit the Gal2 locus to the nuclear periphery.

3.10 | Supplementary information

3.10.1 | Supplementary tables

Supplementary Table S3.1 gRNA target sites in yeast

sgRNA target[a]	DNA Sequence	Location[b]
HMR_1	ATTATATTGCAAAAACTCGA	ChrIII, -2496 (partially overlaps UAS_G)
HMR_2	GGACAGTCCTCCGTCGACGG	ChrIII, -2472 (within UAS_G)
HMR_6	GGAGGACTGTCCTCCGTCGA	ChrIII, -2439 (within UAS_G)
HMR_7	TCATGTACTAAACTAAAATC	ChrIII, -2414
HMR_8	AGAATAAGCGCAGGTACTCC	ChrIII, -2341
HMR_3	TATTCTCCAAAACAATAATA	ChrIII, 286640 - 286621

HMR_4	AATTCGAATAAGATAAACAG	ChrIII, 280020 - 280039
HMR_5	AAAGGGTTTATATCCGAAGG	ChrIII, 193185 - 193204
W17	GAAGTCAGTTGACAGAGTCG	Synthetic target site, used for off-target(OT) control in Fig 3.2 and 3.3
GAL2_1	CACATCACCAGACTTATCTC	ChrXII, +45, 292251-292271
GAL2_2	TGAGAATGTTCGAACGATCC	ChrXII, +149, 292355-292375
GAL2_3	AAATAAGTCAGGTACTTGCC	ChrXII, +81, 292287-292307
GAL2_4	CGTCATTTAGGTCTAAAGTC	ChrXII, +402, 292608-292628

[a] For HMR target sites, see complete HMR reporter sequence in Supplemental Methods. For HMR3 there are two repeats of the 20 base sequence in the UAS_G, but only one of these repeats has an appropriately positioned PAM.

[b] Location numbers given for HMR and GAL2 sites are the distance from the 5' end (PAM distal) of the guide sequence to the integration sites for their respective tetO arrays (see Supplemental Methods).

Supplementary Table S3.2 Yeast strains

Strain	**Description**	**Genotype**
SO992	W303 derivative	*MAT***a** *ura3 leu2 trp1 his3 can1R ade*
yRK036	HMR reporter	*SO992 HMR::2x* UAS_G *Aeb Trp1*
yRK045	HMR reporter/dCas9/MCP-Yif1	*yRK036 Leu2::pTdh dCas9 His3::pAdh MCP-Yif1*
yRK113[a]	yRK036/tetO-array(HMR)/ tetR-GFP/mCherry-Heh2	*yRK036 HMR:: pRS14 TetO LEU2 (60x) HO::pUra3 TetR-GFP_hph*^{R} *mfa2::pTef1 mCherry-Heh2_KanMX*
yRK119[b]	yRK036/tetO-array(HMR)/ tetR-GFP/mCherry-Heh2	*yRK036 HMR::pRS8 TetO HIS3 (60x) HO::pUra3 TetR-GFP_hph*^{R} *mfa2::pTef1 mCherry-Heh2_KanMX*
yRK124	yRK119/dCas9/MCP-Yif1	*yRK119 Leu2::pTdh dCas9 pAdh MCP-Yif1*

yEC059[d]	mCherry-Heh2/tetR-GFP/tetO-array(GAL2)/dCas9/MCP-Yif1	*S0992 Mfa2:: pTef1 mCherry Heh2 Kan HO:: pUra3 TetR-GFP HygB, HIS3:: TetO at GAL2, LEU2::pTdh dCas9 pAdh MCP Yif1*

[a] yRK113 used Leu selection to deliver the tetO array and is the parent for all Gal4-derivative strains in Fig 3.2.

[b] yRK119 used His selection to deliver the tetO array and is the parent for all CRISPR recruitment strains at HMR in Fig 3.2.

[c] cSLQ.Sc003 has been described previously.[4,6]

[d] yEC059 used His selection to deliver the tetO array and is the parent strain for all GAL2 recruitment strains

Supplementary Table S3.3 Yeast protein expression plasmids

Plasmid[a]	**Parent Vector**[b]	**Marker**	**Promoter**	**Gene**	**Terminator**
pJZC518	pNH605	*leu2*	*pTdh3*	dCas9	*C. alb. Adh1*
pRK076	pNH603	*His3*	*pAdh*	MCP-Yif1$_{55-314}$	*C. alb. Adh1*
pRK067	p423	*his3*	*pAdh*	Gal4$_{DBD}$-Yif1$_{55-314}$	*cyc*
pRK069	p423	*his3*	*pAdh*	Gal4$_{DBD}$	*cyc*
pRK144	p423	*his3*	*pAdh*	Gal4$_{DBD}$-MBP-Yif1$_{55-314}$	*cyc*
pRK149	pJW607	*hph^{R}*	*pUra3*	TetR-GFP	*C. alb. Adh1*
pRK160	pJW609	*KanMX*	*pTef1*	mCherry-Heh2$_{138-378}$	*C. alb. Adh1*
pRK159	pNH605	*leu2*	*1) pTdh3* *2) pAdh*	1) dCas9 2) MCP-Yif1$_{55-314}$	*1) C. alb. Adh1* *2) C. alb. Adh1*

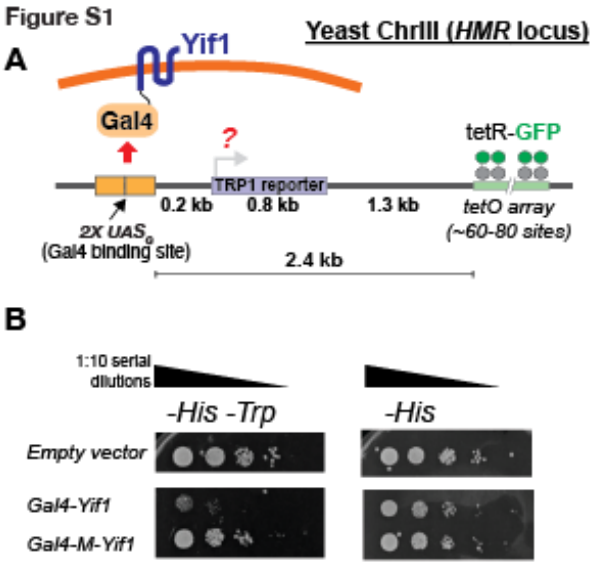
[a] pJZC518 for dCas9 expression and the general strategy for delivering MCP-effector proteins, either on separate integrating plasmids or together with dCas9 on a single integration cassette, have been described previously.[4,7]

[b] The pNH600 series of yeast single copy integration vectors has been described previously.[8-10]

Supplementary Table S3.4 PCR primers used in this study

Location	Forward Oligo	Reverse Oligo
HMR	CAAAAGGCGGCCTGGAATGTAAAGAAAA TCATAAGCCGTAAGAGATCTCCGAATAA CGGTAGCGGATGCCGGGAGCAGAC	ACCTTGTGCCCCATAAAGGTTCCATGTTCCAT AAGTCTTCAATAATACTTTTGTATATTAGTGA GCTGATACCGCTCGCC
Gal2	AACGATAACATGCTCTGCCATCCTTTGT TCACCGAGCAAAATTAAAAACGCAAAAT GAATAGCGGATGCCGGGAGCAGAC	CTAGAGGACCCCCCAGAGATAAGTCTGGTGAT GTGGTCCTTTAATAATTTCATAGGGACAGTGA GCTGATACCGCTCGCC

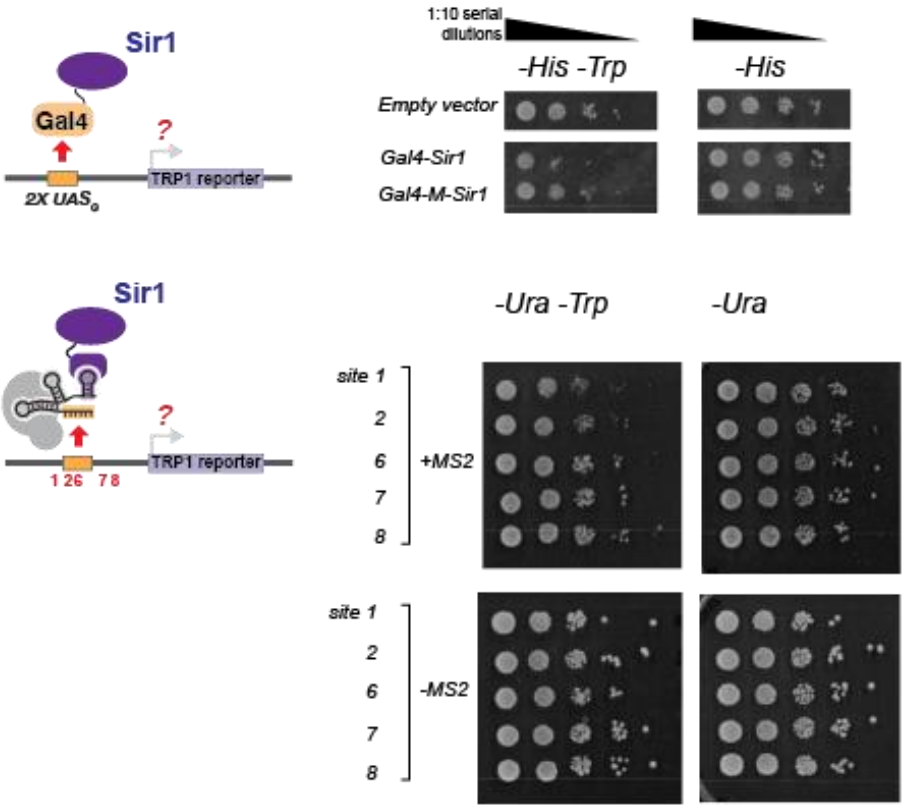
3.10.2 | Supplementary figures



Supplementary Figure S3.1 Lac array insertion does not perturb recruitment mediated gene silencing

Peripheral recruitment of the HMR locus silences a Trp1 reporter gene in a strain containing the lac array, leading to a growth defect on media lacking Trp. Consistent with the same reporter without an integrated array, $Gal4_{DBD}$-Yif1 expression results in a growth defect on –His –Trp media. Cells expressing $Gal4_{DBD}$-M-Yif1, which has MBP inserted between $Gal4_{DBD}$ and Yif1, produces in a partial silencing phenotype. The parent strain is yRK113 and Gal4 constructs are delivered on p423 (His selection).

Figure S2



Supplementary Figure S3.2 CRISPR mediated recruitment of Sir1 results in no gene silencing

A) Recruitment of Gal4-Sir1 to the Aeb::HMR Trp reporter leads to a growth defect on media lacking Trp, indicating gene silencing. B) Recruitment of Sir1 using the CRISPR system to a Trp reporter does not result in a growth defect on media lacking Trp, indicating no gene silencing. The parent strain is yRK036 and Gal4 constructs are delivered on p423 (His selection). For CRISPR experiments, the parent strain is yRK046 and guide constructs are delivered on p316 (Ura selection).

3.10.3 | Supplemental methods

Yeast strain construction and manipulation

Yeast (*S. cerevisiae*) transformations were performed with the standard lithium acetate method. The parent haploid yeast strain for reporter gene experiments was SO992 (W303; *MAT***a** *ura3 leu2 trp1 his3*). Complete descriptions of all yeast strains generated in this work are provided in Table S3, and descriptions of the plasmids are in Table S4. dCas9 and MCP fusion proteins were expressed as described previously from constructs integrated in single copy into the yeast genome.[4] Yif1 and Nup84 were cloned from yeast genomic DNA. Yif1 fusion proteins used Yif1(55-314),[1] fused to the C-terminus of $Gal4_{DBD}$, dCas9, or MCP. For Nup84, the C-terminus was fused to the N-terminus of MCP or TetR,

following the design of the Nup84-LexA fusion.[2] Guide RNA constructs were expressed as described previously from the pRS316 CEN/ARS plasmid (*ura3* marker) with the SNR52 promoter and SUP4 terminator.[4] All $Gal4_{DBD}$ constructs and derivatives were expressed from the p423 2μ plasmid. The HMR Trp reporter sequence at the endogenous HMR locus was constructed by transforming a linear DNA fragment containing *HMR*-E_UAS_G_Trp1 with >280 bases of flanking homology and selecting on SD – Trp plates. Integration of the full reporter cassette was verified by colony PCR. TetO arrays were integrated at ChrIII (HMR) and ChrXII using either pRS8 (*his3* marker) or pRS14 (*leu2* marker) (gifts from Susan Gasser), using previously described methods (Table S5).[6] Based on plasmid sizes, we estimate that our TetO arrays contained ~60-80 tetO repeats. The tetR-GFP and mCherry-$Heh2_{138-378}$ fusion proteins were integrated in single copy in the yeast genome. The expression cassette for the tetR-GFP protein was derived from pGVH29 (gift from Susan Gasser).[7] The mCherry-$Heh2_{138-378}$ construct was designed following previous reports.[8,10]

Trp Silencing Assay

After transformations, yeast strains were grown overnight at 30 °C in selective media (SD –Ura or SD –His as appropriate). Cultures were diluted to OD_{600} 0.2 in selective media lacking Trp and serially diluted (1:10). 10 μL of each dilution was spotted on selective SD plates with or without Trp. Plates were incubated at 30 °C and evaluated after 2 days.

Flow Cytometry

After transformations, yeast strains were grown overnight at 30 °C in selective media (SD –Ura or SD –His). Overnight cultures were diluted 1:25 and grown for an additional 4 hours. Fluorescent protein expression levels were measured with a MACSQuant flow cytometer (Miltenyi Biotec).

Yeast Microscopy

After transformations, yeast strains were grown overnight at 30 °C on selective plates (SD –Ura or SD –His) or YPD (parent strain). Cells were resuspended in YPD at a starting OD_{600} ~0.15 and grown to OD_{600} 0.3-0.5. 5 mL cultures were pelleted, washed in synthetic complete media (SCM), and resuspended in 100 μL SCM. 10 μL of these cells were pipetted onto agarose pads in 13 x 1 mm silicone isolator wells (Electron Microscopy Sciences) and covered with a No. 1.5 coverslip. Imaging was

performed using a Leica TCS SP5 II laser scanning confocal microscope with a 63x oil immersion objective. The pixel size was 90.1 nm and the z-step size was 0.21 µm. The optical thickness of each slice is 0.98 µm. For each cell, the tetR-GFP spot was assigned to a particular z-plane based on its maximum intensity. In that z-plane, we defined the nuclear periphery as the pixel corresponding to the center of the Heh2-mCherry peak along the radial axis. GFP spots were classified as “peripheral” if the center of spot was within two pixels of the nuclear periphery (i.e. separated by no more than one pixel). Cells in which the GFP spot was assigned to the bottom or top slice of the nucleus were excluded from analysis.[11] >150 cells were analyzed per experiment.

3.10.4 | gRNA, effector protein, and HMR reporter sequence

Guide RNA Sequence Designs

sgRNA and scRNA sequences were constructed as described previously.[4]Alternative target sites for yeast (Table S1) or human (Table S2) cells were cloned with standard PCR methods._

Yeast sequences

Parent sgRNA

ATTATATTGCAAAAACTCGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTAT

CAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTTTGTTTTTTATGTCT

2x MS2 scRNA

ATTATATTGCAAAAACTCGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTAT

CAACTTGAAAAAGTGGCACCGAGTCGGTGCgggagcACATGAGGATCACCCATGTgccacgagcgA

CATGAGGATCACCCATGTcgctcgtgttcccTTTTTTTGTTTTTTATGTCT

Annotations: 20 base target site (HMR1), 2x MS2, SUP4 terminator

Effector Protein Sequences

>MCP$_{V29I\Delta FG}$-Yif1$_{55-314}$

MPKKKRKVGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT

IKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSAGMGGFPFQDP
RGSMAFQLGQSAFSNFIGQDNFNQFQETVNKATANAAGSQQISTYFQVSTRYVINKLKLILVPFLNGTKN
WQRIMDSGNFLPPRDDVNSPDMYMPIMGLVTYILIWNTQQGLKGSFNPEDLYYKLSSTLAFVCLDLLILK
LGLYLLIDSKIPSFSLVELLCYVGYKFVPLILAQLLTNVTMPFNLNILIKFYLFIAFGVFLLRSVKFNLL
SRSGAEDDDIHVSISKSTVKKCNYFLFVYGFIWQNVLMWLMG

> Gal4$_{DBD}$-Yif1$_{55-314}$

MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVESRLERLEQLFLL
IFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDMPLTLRQHRISATSSSEESSNKG
QRQLTVSAAPEFGGFPFQDPRGSMAFQLGQSAFSNFIGQDNFNQFQETVNKATANAAGSQQISTYFQVST
RYVINKLKLILVPFLNGTKNWQRIMDSGNFLPPRDDVNSPDMYMPIMGLVTYILIWNTQQGLKGSFNPED
LYYKLSSTLAFVCLDLLILKLGLYLLIDSKIPSFSLVELLCYVGYKFVPLILAQLLTNVTMPFNLNILIK
FYLFIAFGVFLLRSVKFNLLSRSGAEDDDIHVSISKSTVKKCNYFLFVYGFIWQNVLMWLMG

> Gal4$_{DBD}$-M-Yif1$_{55-314}$ *(M = maltose binding protein, MBP, 3x HA)*

MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVESRLERLEQLFLL
IFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDMPLTLRQHRISATSSSEESSNKG
QRQLTVSAAPEFMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGP
DIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPK
TWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD
LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGIN
AASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMS
AFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRISTSQFGSGYPYDVPDYAGSG
YPYDVPDYAGSGYPYDVPDYAGSGKFGGFPFQDPRGSMAFQLGQSAFSNFIGQDNFNQFQETVNKATANA
AGSQQISTYFQVSTRYVINKLKLILVPFLNGTKNWQRIMDSGNFLPPRDDVNSPDMYMPIMGLVTYILIW
NTQQGLKGSFNPEDLYYKLSSTLAFVCLDLLILKLGLYLLIDSKIPSFSLVELLCYVGYKFVPLILAQLL
TNVTMPFNLNILIKFYLFIAFGVFLLRSVKFNLLSRSGAEDDDIHVSISKSTVKKCNYFLFVYGFIWQNV
LMWLMG

HMR reporter sequence (pRK105 integration plasmid)

The *HMR* reporter sequence (Fig. 3.2A & S3.2) was designed following a previously described silencing

reporter (Aeb)[1,12] that contains a binding site for Gal4 (the UAS_G or upstream activating sequence)[13] within the *HMR*-E region and a downstream Trp1 reporter gene integrated at the endogenous *HMR* locus in *S. cerevisiae* chromosome 3. Genes at the *HMR* locus are normally silenced, but the UAS_G insertion disrupts endogenous regulatory sites to allow gene expression. The $Gal4_{DBD}$ binds at the UAS_G, and $Gal4_{DBD}$ fusion proteins can rescue the silencing phenotype by directly recruiting silencing factors or by recruiting the *HMR* locus to the nuclear periphery.[1,12]

In our reporter, we integrated a lacO array approximately 2.4 kb downstream of the UAS_G site, which allows direct visualization of the locus with lacI-GFP. The original reporter described in the literature contains a 3X UAS_G repeat, while our reporter contains a 2X UAS_G repeat. In functional silencing assays, the 2X UAS_G reporter construct was effectively silenced by $Gal4_{DBD}$-Yif1 expression (Fig S3.3), similar to that described for the 3X UAS_G reporter.[1]

Annotations[12,14,15]: *HMR*-E (Aeb), UAS_G, Trp1 (5'utr-ORF-utr-3'), *HMR*-I

ctagtacttaaaaaaactgtagtttcagtgcaaaaaagttttaacattacgtatcttgtacccttttttattgcatatagaaag
gtcaaataatccttcacatcatgaaatataagctaaatcgcatttcttttcgtccacatttgcaaacaaaacttttcaataat
aattttataaatagtatcaatatatatatatatatttatttgtttacttttttctatcagtgttttcaatttttttattaaacaa
tgtttgatttttttcaatcgcaatttaatacctaaatataaaaaatgttattatattgcaaaaaCTCGACGGAGGACAGTCCTC
CGTCGACGGAGGACAGTCCTCCGTCGAGaatatttgaaagcaatagatcatgtactaaactaaaatcagggaaattaagactc
cttttgaagtaatacctattacttactaatacgtttgagaataagcgcaggtactcctggtttttgttaaaattacaaatttta
tacttagcattacgaagattctcgattccgaaaaacaaaaatttttatcgtcatatacaaatctagggtcgaaaaaagaaaagg
agagggccaagagggagggcattggtgactattgagcacgtgagtatacgtgattaagcacacaaaggcagcttggagtaTGT
CTGTTATTAATTTCACAGGTAGTTCTGGTCCATTGGTGAAAGTTTGCGGCTTGCAGAGCACAGAGGCCGCAGAATGTGCTCTA
GATTCCGATGCTGACTTGCTGGGTATTATATGTGTGCCCAATAGAAAGAGAACAATTGACCCGGTTATTGCAAGGAAAATTTC
AAGTCTTGTAAAAGCATATAAAAATAGTTCAGGCACTCCGAAATACTTGGTTGGCGTGTTTCGTAATCAACCTAAGGAGGATG
TTTTGGCTCTGGTCAATGATTACGGCATTGATATCGTCCAACTGCATGGAGATGAGTCGTGGCAAGAATACCAAGAGTTCCTC
GGTTTGCCAGTTATTAAAAGACTCGTATTTCCAAAAGACTGCAACATACTACTCAGTGCAGCTTCACAGAAACCTCATTCGTT
TATTCCCTTGTTTGATTCAGAAGCAGGTGGGACAGGTGAACTTTTGGATTGGAACTCGATTTCTGACTGGGTTGGAAGGCAAG
AGAGCCCCGAAAGCTTACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTGGTGATGCGCTTAGATTAAATGGCGTT
ATTGGTGTTGATGTAAGCGGAGGTGTGGAGACAAATGGTGTAAAAGACTCTAACAAAATAGCAAATTTCGTCAAAAATGCTAA
GAAATAGgttattactgagtagtatttatttaagtattgtttgtgcacttgcctgcaggccttttgaaaagcaagcataaaaa
ataaattcgttttcaatgattaaaatagcatagtcgggtttttcttttagtttcagctttccgcaacagtataattttataaa

ccctggttttggttttgtagagtggttgacgaataattatgctgaagtacgtggtgacggatattgggaagatgtgtttgtac

atttggccttatagagtgtggtcgtggcggaggttgtttatctttcgagtactgaatgttgtcagtatagctatcctatttga

aactccccatcgtcttgctcttgttctcaatgtttgtttatatactcatatttctatgtgtttatacaattgctattgtttat

ataatgtagtgacattttctcttaatcttatactaatttctatgacatttatataagaagagacttatgatcaacataatttt

gcaaactttgagagaaatatgtctttctactgcgataaagttattatttagattacatgtcaccaacattttcgtatatggcg

atataaatttatcatgttttggtatgataatttaatttttaaaaaaacaaatttaattgacctcattaattaatatttattaa

taccttttaatgttgaggtaaatagctattttctctcttctttttccttttagttggaatttgcacaagaaaatgtttttccacac

actttagcgtttttttcctaaatgttggaataaaaaacaactatcatctatcaaCTAGTAGTCACACTACCAATGTGTTATCAT

TATACTGTGTTAAACAATGACATAAGGTATGAAAATTTGTCAACGAAGTTAGAGAAAGCTGGATGCAAGGATTGATAATGTGG

TAGGAAAATGAAACATATAACGGAATGAGGAATAATCGTAATATCAGTATATAGAAATATAGATTCCCTTTTGAGGATTCCTA

TATCCTCGAGGAGAACTTCTAGTATATTCTATATACCTAATATTATTACTTTTATCTACAATGCAACCCCACAATAATATAAA

AATTCACCAATTCCGCATCTGCAGATTACTTTCCTAAATTTGCATATAGAATTGTCAAGCGCAAATCCGACGTCGATTCCGCG

GCGGATGGGTCATTCTAGGTCATTCTACCAATTTTATTTGAGACCAGGTTTATTCAACCGGTAACATAGAAATATTCATACAA

TTAAGCTTCTATGGCCAAGTTGGTAAGGCGCCACACTAGTAATGTGGAGATCATCGGTTCAAATCCGATTGGAAGCATTTTTT

ATCACGTTATTCGGTGACACCCAGGTTGCCGCCGCGTTCGCGTCCATCGTCATCTGAAAAATAATGAATATTAATGGACCTTG

TGCCCCATAAAGGTTCCATGTTCCATAAGTCTTCAATAATACTTTTGTATATTA**[tetO_array_integration_site]**

ACCGTTATTCGGAGATCTCTTACGGCTTATGATTTTCTTTACATTCCAGGCCGCCTTTTG

HMR-E (Aeb), W17, Trp1 (5'utr-ORF-utr-3'), *HMR*-I

ctagtacttaaaaaaactgtagtttcagtgcaaaaaagttttaacattacgtatcttgtacccttttttattgcatatagaaag

gtcaaataatccttcacatcatgaaatataagctaaatcgcatttcttttcgtccacatttgcaaacaaaacttttcaataat

aattttataaatagtatcaatatatatatatatatttatttgtttacttttttctatcagtgttttcaatttttttattaaacaa

tgtttgatttttttcaatcgcaatttaatacctaaatataaaaaatgttattatattgcaaaaaCTCGACGGAGGACAGTCCTC

CGTCGACGGAGGACAGTCCTCCGTCGAGaatatttgaaagcaatagatcatgtactaaactaaaatcagggaaattaagactc

cttttgaagtaatacctattacttactaatacgtttgagaataagcgcaggtactcctggtttttgttaaaattacaaattta

tacttagcattacgaagattctcgattccgaaaaacaaaaatttttatcgtcatatacaaatctagggtcgaaaaaagaaaagg

agagggccaagagggagggcattggtgactattgagcacgtgagtatacgtgattaagcacacaaaggcagcttggagtaTGT

CTGTTATTAATTTCACAGGTAGTTCTGGTCCATTGGTGAAAGTTTGCGGCTTGCAGAGCACAGAGGCCGCAGAATGTGCTCTA

GATTCCGATGCTGACTTGCTGGGTATTATATGTGTGCCCAATAGAAAGAGAACAATTGACCCGGTTATTGCAAGGAAAATTTC

AAGTCTTGTAAAAGCATATAAAAATAGTTCAGGCACTCCGAAATACTTGGTTGGCGTGTTTCGTAATCAACCTAAGGAGGATG

TTTTGGCTCTGGTCAATGATTACGGCATTGATATCGTCCAACTGCATGGAGATGAGTCGTGGCAAGAATACCAAGAGTTCCTC

GGTTTGCCAGTTATTAAAAGACTCGTATTTCCAAAAGACTGCAACATACTACTCAGTGCAGCTTCACAGAAACCTCATTCGTT

TATTCCCTTGTTTGATTCAGAAGCAGGTGGGACAGGTGAACTTTTGGATTGGAACTCGATTTCTGACTGGGTTGGAAGGCAAG

AGAGCCCCGAAAGCTTACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTGGTGATGCGCTTAGATTAAATGGCGTT
ATTGGTGTTGATGTAAGCGGAGGTGTGGAGACAAATGGTGTAAAAGACTCTAACAAAATAGCAAATTTCGTCAAAAATGCTAA
GAAATAGgttattactgagtagtatttatttaagtattgtttgtgcacttgcctgcaggccttttgaaaagcaagcataaaaa
ataaattcgttttcaatgattaaaatagcatagtcgggtttttcttttagtttcagctttccgcaacagtataatttttataaa
ccctggttttggttttgtagagtggttgacgaataattatgctgaagtacgtggtgacggatattgggaagatgtgtttgtac
atttggccttatagagtgtggtcgtggcggaggttgtttatctttcgagtactgaatgttgtcagtatagctatcctatttga
aactccccatcgtcttgctcttgttctcaatgtttgtttatatactcatatttctatgtgtttatacaattgctattgtttat
ataatgtagtgacattttctcttaatcttatactaatttctatgacatttatataagaagagacttatgatcaacataatttt
gcaaactttgagagaaatatgtctttctactgcgataaagttattatttagattacatgtcaccaacattttcgtatatggcg
atataaatttatcatgttttggtatgataatttaatttttaaaaaaacaaatttaattgacctcattaattaatatttattaa
tacctttaatgttgaggtaaatagctattttctctcttctttttccttttagttggaatttgcacaagaaaatgtttttccacac
actttagcgttttttcctaaatgttggaataaaaaacaactatcatctatcaaCTAGTAGTCACACTACCAATGTGTTATCAT
TATACTGTGTTAAACAATGACATAAGGTATGAAAATTTGTCAACGAAGTTAGAGAAAGCTGGATGCAAGGATTGATAATGTGG
TAGGAAAATGAAACATATAACGGAATGAGGAATAATCGTAATATCAGTATATAGAAATATAGATTCCCTTTTGAGGATTCCTA
TATCCTCGAGGAGAACTTCTAGTATATTCTATATACCTAATATTATTACTTTTATCTACAATGCAACCCCACAATAATATAAA
AATTCACCAATTCCGCATCTGCAGATTACTTTCCTAAATTTGCATATAGAATTGTCAAGCGCAAATCCGACGTCGATTCCGCG
GCGGATGGGTCATTCTAGGTCATTCTACCAATTTTATTTGAGACCAGGTTTATTCAACCGGTAACATAGAAATATTCATACAA
TTAAGCTTCTATGGCCAAGTTGGTAAGGCGCCACACTAGTAATGTGGAGATCATCGGTTCAAATCCGATTGGAAGCATTTTTT
ATCACGTTATTCGGTGACACCCAGGTTGCCGCCGCGTTCGCGTCCATCGTCATCTGAAAAATAATGAATATTAATGGACCTTG
TGCCCCATAAAGGTTCCATGTTCCATAAGTCTTCAATAATACTTTTGTATATTA**[tetO_array_integration_site]**
ACCGTTATTCGGAGATCTCTTACGGCTTATGATTTTCTTTACATTCCAGGCCGCCTTTTG

Annotation: GAL2 (ORF)

AATAGTAATAGTTAAGTAAACACAAGATTAACATAATAAAAAAAATAATTCTTTCATAATGGCAGTTGAGGAGAACAATATGC
CTGTTGTTTCACAGCAACCCCAAGCTGGTGAAGACGTGATCTCTTCACTCAGTAAAGATTCCCATTTAAGCGCACAATCTCAA
AAGTATTCTAATGATGAATTGAAAGCCGGTGAGTCAGGGTCTGAAGGCTCCCAAAGTGTTCCTATAGAGATACCCAAGAAGCC
CATGTCTGAATATGTTACCGTTTCCTTGCTTTGTTTGTGTGTTGCCTTCGGCGGCTTCATGTTTGGCTGGGATACCGGTACTA
TTTCTGGGTTTGTTGTCCAAACAGACTTTTTGAGAAGGTTTGGTATGAAACATAAGGATGGTACCCACTATTTGTCAAACGTC
AGAACAGGTTTAATCGTCGCCATTTTCAATATTGGCTGTGCCTTTGGTGGTATTATACTTTCCAAAGGTGGAGATATGTATGG
CCGTAAAAAGGGTCTTTCGATTGTCGTCTCGGTTTATATAGTTGGTATTATCATTCAAATTGCCTCTATCAACAAGTGGTACC
AATATTTCATTGGTAGAATCATATCTGGTTTGGGTGTCGGCGGCATCGCCGTCTTATGTCCTATGTTGATCTCTGAAATTGCT
CCAAAGCACTTGAGAGGCACACTAGTTTCTTGTTATCAGCTGATGATTACTGCAGGTATCTTTTTGGGCTACTGTACTAATTA
CGGTACAAAGAGCTATTCGAACTCAGTTCAATGGAGAGTTCCATTAGGGCTATGTTTCGCTTGGTCATTATTTATGATTGGCG

CTTTGACGTTAGTTCCTGAATCCCCACGTTATTTATGTGAGGTGAATAAGGTAGAAGACGCCAAGCGTTCCATTGCTAAGTCT

AACAAGGTGTCACCAGAGGATCCTGCCGTCCAGGCAGAGTTAGATCTGATCATGGCCGGTATAGAAGCTGAAAAACTGGCTGG

CAATGCGTCCTGGGGGGAATTATTTTCCACCAAGACCAAAGTATTTCAACGTTTGTTGATGGGTGTGTTTGTTCAAATGTTCC

AACAATTAACCGGTAACAATTATTTTTTCTACTACGGTACCGTTATTTTCAAGTCAGTTGGCCTGGATGATTCCTTTGAAACA

TCCATTGTCATTGGTGTAGTCAACTTTGCCTCCACTTTCTTTAGTTTGTGGACTGTCGAAAACTTGGGACATCGTAAATGTTT

ACTTTTGGGCGCTGCCACTATGATGGCTTGTATGGTCATCTACGCCTCTGTTGGTGTTACTAGATTATATCCTCACGGTAAAA

GCCAGCCATCTTCTAAAGGTGCCGGTAACTGTATGATTGTCTTTACCTGTTTTTATATTTTCTGTTATGCCACAACCTGGGCG

CCAGTTGCCTGGGTCATCACAGCAGAATCATTCCCACTGAGAGTCAAGTCGAAATGTATGGCGTTGGCCTCTGCTTCCAATTG

GGTATGGGGGTTCTTGATTGCATTTTTCACCCCATTCATCACATCTGCCATTAACTTCTACTACGGTTATGTCTTCATGGGCT

GTTTGGTTGCCATGTTTTTTTTATGTCTTTTTCTTTGTTCCAGAAACTAAAGGCCTATCGTTAGAAGAAATTCAAGAATTATGG

GAAGAAGGTGTTTTACCTTGGAAATCTGAAGGCTGGATTCCTTCATCCAGAAGAGGTAATAATTACGATTTAGAGGATTTACA

ACATGACGACAAACCGTGGTACAAGGCCATGCTAGAATAATGCGTTTGAAGTGAGACGCTCCATCATCTCTCTTAATTTTTCA

TGACTGACGTTTTTTCTTCATTTTAATTATCATAGTATTTGTTTGAAAAAAAAAAAAAAAAAATTTCCCTTATCAATGATATCC

TTACGATTATATAAATTCCTTACCTAAACCTATTATTTGTGTACATATATCAGAGTATTATTACATATATAACCTTTTTCTCT

AAAACAGGAAAAAAAAAAGAAAACGATAACATGCTCTGCCATCCTTTGTTCACCGAGCAAAATTAAAAACGCAAAATGAAT**[t**

etO_array_integration_site]TGTCCCTATGAAATTATTAAAGGACCACATCACCAGACTTATCTCTGGGGGGTCCT

CTAGAAAATAAGTCAGGTACTTGCCTGGACTTTCTTCCAGTTGAATTCCTGAGCTAACATACAATTAATGGAGTGAGA

3.10.5 | Supplementary references

(1) Andrulis, E. D., Neiman, A. M., Zappulla, D. C. & Sternglanz, R. Perinuclear localization of chromatin facilitates transcriptional silencing. Nature 394, 592–595 (1998).

(2) Menon, B. B. et al. Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. Proc. Natl. Acad. Sci. USA 102, 5749–5754 (2005).

(3) Cherry, J. M. et al. Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. 40, D700–5 (2012).

(4) Zalatan, J. G. et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 160, 339–350 (2015).

(5) Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–451 (2013).

(6) Rohner, S., Gasser, S. M. & Meister, P. Modules for cloning-free chromatin tagging in Saccharomyces cerevisae. Yeast 25, 235–239 (2008).

(7) Bystricky, K., Heun, P., Gehlen, L., Langowski, J. & Gasser, S. M. Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. Proc. Natl. Acad. Sci. USA 101, 16495–16500 (2004).

(8) Meinema, A. C. et al. Long unfolded linkers facilitate membrane protein import through the nuclear pore complex. Science 333, 90–93 (2011).

(9) Zalatan, J. G., Coyle, S. M., Rajan, S., Sidhu, S. S. & Lim, W. A. Conformational control of the Ste5 scaffold protein insulates against MAP kinase misactivation. Science 337, 1218–1222 (2012).

(10) Egecioglu, D. E., D'Urso, A., Brickner, D. G., Light, W. H. & Brickner, J. H. Approaches to studying subnuclear organization and gene-nuclear pore interactions. Methods Cell Biol. 122, 463–485 (2014).

(11) Meister, P., Gehlen, L. R., Varela, E., Kalck, V. & Gasser, S. M. Visualizing yeast chromosomes and nuclear architecture. Methods Enzymol 470, 535–567 (2010).

(12) Chien, C. T., Buck, S., Sternglanz, R. & Shore, D. Targeting of SIR1 protein establishes transcriptional silencing at HM loci and telomeres in yeast. Cell 75, 531–541 (1993).

(13) Bram, R. J., Lue, N. F. & Kornberg, R. D. A GAL family of upstream activating sequences in yeast: roles in both induction and repression of transcription. EMBO J. 5, 603–608 (1986).

(14) Abraham, J., Nasmyth, K. A., Strathern, J. N., Klar, A. J. & Hicks, J. B. Regulation of mating-type information in yeast. Negative control requiring sequences both 5‘ and 3’ to the regulated region. J. Mol. Biol. 176, 307–331 (1984).

(15) Brand, A. H., Micklem, G. & Nasmyth, K. A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation. Cell 51, 709–719 (1987).