UCSF
UC San Francisco Previously Published Works

Title
Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds

Permalink
https://escholarship.org/uc/item/6kg004rt

Journal
Cell, 160(1-2)

ISSN
0092-8674

Authors
Zalatan, Jesse G
Lee, Michael E
Almeida, Ricardo
et al.

Publication Date
2015

DOI
10.1016/j.cell.2014.11.052

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

NIH Public Access
Author Manuscript
Cell. Author manuscript; available in PMC 2016 January 15.

Published in final edited form as:
Cell. 2015 January 15; 160(0): 339–350. doi:10.1016/j.cell.2014.11.052.

Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds

Jesse G. Zalatan[1,2,9], **Michael E. Lee**[5,6], **Ricardo Almeida**[1,2], **Luke A. Gilbert**[1,2,7,8], **Evan H. Whitehead**[1,3,8,10], **Marie La Russa**[1,3,4,8,10], **Jordan C. Tsai**[1,2,7,8], **Jonathan S. Weissman**[1,2,3,7,8], **John E. Dueber**[5,6], **Lei S. Qi**[1,3,8,10,*], and **Wendell A. Lim**[1,2,3,8,*]

[1]Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, CA 94158, USA

[2]Howard Hughes Medical Institute, University of California San Francisco, San Francisco, CA 94158, USA

[3]UCSF Center for Systems and Synthetic Biology, University of California San Francisco, San Francisco, CA 94158, USA

[4]Biomedical Sciences Graduate Program, University of California San Francisco, San Francisco, CA 94158, USA

[5]Department of Bioengineering, Berkeley, CA 94720, USA

[6]Energy Biosciences Institute, Berkeley, CA 94720, USA

[7]Center for RNA Systems Biology, Berkeley, CA 94720, USA

[8]California Institute for Quantitative Biomedical Research, San Francisco, CA 94158, USA

Summary

Eukaryotic cells execute complex transcriptional programs in which specific loci throughout the genome are regulated in distinct ways by targeted regulatory assemblies. We have applied this principle to generate synthetic CRISPR-based transcriptional programs in yeast and human cells. By extending guide RNAs to include effector protein recruitment sites, we construct modular scaffold RNAs that encode both target locus and regulatory action. Sets of scaffold RNAs can be used to generate synthetic multi-gene transcriptional programs in which some genes are activated and others are repressed. We apply this approach to flexibly redirect flux through a complex branched metabolic pathway in yeast. Moreover, these programs can be executed by inducing

© 2014 Elsevier Inc. All rights reserved.

*Correspondence: wendell.lim@ucsf.edu, slqi@stanford.edu.
[9]Current address: Department of Chemistry, University of Washington, Seattle, WA 98195, USA
[10]Current address: Department of Bioengineering, Department of Chemical and Systems Biology, and ChEM-H, Stanford University, Stanford, CA 94305, USA

Supplemental Information
Supplemental Information includes Extended Experimental Procedures, four figures, and eight tables and can be found with this article online.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

expression of the dCas9 protein, which acts as a single master regulatory control point. CRISPR-associated RNA scaffolds provide a powerful way to construct synthetic gene expression programs for a wide range of applications including rewiring cell fates or engineering metabolic pathways.

Introduction

Eukaryotic cells achieve many different states by executing complex transcriptional programs that allow a single genome to be interpreted in numerous, distinct ways. In these programs, specific loci throughout the genome must be regulated independently. For example, during development, it is often critical to activate sets of genes associated with a new cell fate while simultaneously repressing sets of genes associated with a prior or alternative fate. Similarly, environmental conditions often trigger shifts in metabolic state, which requires activating a new set of enzymes and repressing other previously expressed enzymes, leading to new metabolic fluxes. These complex multi-locus, multi-directional expression programs are encoded largely by the pattern of transcriptional activators, repressors, or other regulators that assemble at distinct sites in the genome. Reprogramming these instructions to produce a different cell type or state thus requires precisely targeted changes in gene expression over a broad set of genes.

How might we engineer novel gene expression programs that match the sophistication of natural programs? Such capabilities would provide powerful tools to probe how changes in gene expression programs lead to diverse cell types. These tools would also provide the ability to engineer more sophisticated designer cell types for therapeutic or biotechnological applications. Although a number of transcriptional engineering platforms have been developed, there are major constraints for constructing complex transcriptional programs. For example synthetic transcription factors (such as designed zinc fingers or transcription activator-like (TAL) effectors) can target a specific regulatory action to a key genomic locus, but it is challenging to simultaneously target many loci in parallel because each DNA-binding protein must be individually designed and tested (Gaj et al., 2013). The bacterial type II CRISPR (*c*lustered *r*egularly *i*nterspaced *s*hort *p*alindromic *r*epeats) interference system (CRISPRi) provides an alternative suite of tools for genome regulation (Qi et al., 2013). In particular, a catalytically inactive Cas9 (dCas9) protein which lacks endonuclease activity can be used to flexibly target many loci in parallel, by using Cas9 binding guide RNAs that recognize target DNA sequences based only on predictable Watson-Crick base pairing. CRISPRi regulation can be used to achieve activation or repression by fusing dCas9 to activator or repressor modules (Gilbert et al., 2013; Mali et al., 2013a), but these direct protein fusions are constrained to only one direction of regulation. Thus it remains challenging to engineer regulatory programs in which many loci are targeted simultaneously, but with distinct types of regulation at each locus.

To develop a platform for synthetic genome regulation that allows locus-specific action, we took inspiration from natural regulatory systems that encode both target specificity and regulatory function in the same molecule. In cell signaling pathways, scaffold proteins act to physically assemble interacting components so that functional outcomes can be precisely

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

controlled in time and space (Good et al., 2011). Similar scaffolding principles apply in genome organization, where, for example, long non-coding RNA (lncRNA) molecules are proposed to act as assembly scaffolds that recruit key epigenetic modifiers to specific genomic loci (Figure 1A) (Rinn and Chang, 2012; Spitale et al., 2011). The idea that RNA can be used to coordinate biological assemblies has important implications for engineering. RNA is inherently modular and programmable: DNA targets can be recognized by base pairing, and modular RNA-protein interaction domains can be used to recruit specific proteins (Figure 1A). The ability of engineered RNA scaffolds to coordinate functional protein assemblies has already been elegantly demonstrated (Delebecque et al., 2011).

To implement a synthetic, modular RNA-based system for locus-specific transcriptional programming, we can extend the CRISPR single-guide RNA (sgRNA) sequence with modular RNA domains that recruit RNA-binding proteins. This approach converts the sgRNA into a scaffold RNA (scRNA) that physically links DNA binding and protein recruitment activities (Figure 1B). Critically, a single scRNA molecule encodes both information about the target locus and instructions about what regulatory function to execute at that locus. This approach allows multidirectional regulation (i.e. simultaneous activation and repression) of different target genes as part of the same regulatory program. Engineering multivalent RNA recruitment sites on each scRNA offers the further possibility of independently tuning the strength of activation or repression at each target site. The potential viability of this approach is supported by a recent report showing that a sgRNA extended with MS2 hairpins can recruit activators to a reporter gene in human cells (Mali et al., 2013a).

Here we demonstrate that CRISPR sgRNAs can be repurposed as scaffolding molecules to recruit transcriptional activators or repressors, thus enabling flexible and parallel programmable locus-specific regulation. We use the budding yeast *S. cerevisiae* as a testbed to identify 3 orthogonal RNA-protein binding modules and to optimize scRNA designs for single and multivalent recruitment sites. We show that the system developed in yeast also functions efficiently in human cells to regulate reporter and endogenous target genes, and we extend its scope to include recruitment of chromatin modifiers for gene repression. We then demonstrate the use of CRISPR scaffold RNA molecules to construct synthetic multi-gene expression programs. Specifically, we are able to regulate multiple genes in a highly-branched biosynthetic pathway in yeast to express key enzymes in alternative combinations. These synthetic transcriptional programs, by combinatorially altering metabolic organization, allow us to flexibly redirect the pathway between five distinct possible product output states. Finally, we show that dCas9 can act as a master regulator of these gene expression programs, receiving input signals and acting as a single control point to execute a multi-gene response encompassing simultaneous activation and repression of downstream target genes.

Results

CRISPR RNA Scaffolds Efficiently Activate Gene Expression in Yeast

The minimal sgRNA previously used in CRISPR engineering consists of several modular domains: a 20 nucleotide variable DNA targeting sequence and two structured RNA

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

domains – dCas9-binding and 3′ tracrRNA – which are necessary for proper structure formation and binding to Cas9 (Jinek et al., 2012; 2014; Nishimasu et al., 2014). Here, to generate scaffold RNA (scRNA) constructs with additional protein recruitment capabilities, we first introduced a single RNA hairpin domain to the 3′ end of the sgRNA, connected by a two base linker. For these recruitment RNA modules, we used the well-characterized viral RNA sequences MS2, PP7, and com, which are recognized by the MCP, PCP, and Com RNA binding proteins respectively. We fused the transcriptional activation domain VP64 to each of the corresponding RNA binding proteins.

We first tested the CRISPR scRNA platform in yeast. A strain with a tet-promoter driving a fluorescent protein reporter was transformed to express dCas9, scRNAs targeting the tet operator, and the corresponding VP64 fusion proteins. We observed significant reporter gene expression using each of the three RNA binding recruitment modules (Figure 2A). scRNA constructs with recruitment hairpin domains connected to the sgRNA by linkers longer than two bases (up to 20 bases) gave weaker reporter gene expression (Figure S1A). scRNA designs with recruitment sequences attached to the 5′ end of the sgRNA gave no significant activation (Figure S1B). Northern blot analysis indicated that the 5′ extended sequence was degraded (Figure S2).

Gene activation mediated by scRNA-recruitment of VP64 was substantially greater than that for the direct dCas9-VP64 fusion protein. Both MCP and PCP bind to their RNA targets as dimers (Chao et al., 2008), which may account for some of the difference. The oligomerization state of the Com protein has not been directly determined but functional data consistent with a monomer has been reported (Wulczyn and Kahmann, 1991).

Three RNA-Protein Recruitment Modules Act in an Orthogonal Manner

To determine if there is crosstalk between RNA hairpins and non-cognate binding proteins (e.g. MS2 RNA recruiting the PCP protein), we expressed all three RNA hairpin designs (MS2, PP7, and com) in yeast strains containing either the MCP, PCP, or Com fusion proteins. We used a 7X tetO reporter to maximize sensitivity for detecting any weak cross-activation. No significant crosstalk was detected between mismatched pairs of scRNA sequences and binding proteins (Figure 2B). The strong activation of reporter gene expression only with cognate scRNA and RNA binding protein pairs demonstrates the potential for simultaneous, independent regulation of multiple target genes.

Multivalent Recruitment to scRNAs

To tune the valency of effectors recruited to each gene target, we introduced one, two, or three MS2 RNA hairpins to the 3′ end of the sgRNA. Surprisingly, reporter gene expression decreased with increasing numbers of MS2 hairpins (Figure S1C). Northern blot analysis indicated that steady state RNA levels decreased with two or three MS2 hairpins, suggesting that RNA expression or stability is limiting for these constructs (Figure S2A).

To address the apparent stability problem of multi-hairpin scRNAs, we constructed an alternative RNA design in which double-stranded linkers were inserted between the two repeats of the recruitment hairpins to enforce stable, local hairpin formation. These alternative designs produced stronger reporter gene activation for both MS2 and PP7

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

modules relative to the analogous single hairpin scRNAs (Figure 2C). Northern blot analysis of the 2x constructs with double-stranded linkers indicated steady state RNA levels comparable to single hairpin scRNA and unmodified sgRNA constructs (Figure S2A).

The strongest activation for a single scRNA construct was obtained with a mixed hairpin construct containing two different recruitment motifs for the MCP-VP64 effector protein (2x MS2 (wt+f6)) – this construct contained one MS2 hairpin and a second aptamer hairpin (f6) that had been selected to bind to the MCP protein (Hirao et al., 1998). Attempts to design 2x constructs with double-stranded linkers using the com RNA module were unsuccessful, possibly because the cognate Com protein binds to single stranded RNA at the base of the com hairpin (Hattman, 1999). RNA constructs with three MS2 hairpins connected by double-stranded linkers did not improve reporter gene expression beyond that obtained with the 2x MS2 scRNA. Northern blot analysis suggests that these constructs are stably expressed, so the lack of increased expression may be a result of misfolding or steric constraints.

To develop a platform for recruitment of more complex protein assemblies, we designed a heterologous MS2-PP7 scRNA sequence using the 2x double-stranded linker structure. Reporter gene activation was substantially stronger in yeast cells with both MCP-VP64 and PCP-VP64 effector proteins compared to cells with only a single type of effector protein, indicating that distinct RNA binding proteins can be recruited to the same target site (Figure 2D). This provides an effective approach to combinatorially recruit multiple effectors for the logical control of target genes.

scRNAs Can Mediate Activation of Reporter and Endogenous Genes in Human Cells

To test the efficacy of scRNA-based protein effector recruitment in human cells, we ported the system from yeast to HEK293T cells. The dCas9-binding hairpin of the sgRNA was modified as described previously to improve activity in human cells (see Supplementary Information) (Chen et al., 2013). In HEK293T cells expressing dCas9, expression of an scRNA with the corresponding VP64 fusion protein effector produced substantial activation of a 7x tet-driven GFP reporter gene for all three RNA binding modules (Figure 3A), although there are some quantitative differences from the activity trends observed in yeast. GFP activation with 1x MS2 and 1x PP7 scRNA constructs was relatively weak compared to both corresponding multivalent 2x scRNA constructs and the dCas9-VP64 fusion protein.

To determine if endogenous genes could be activated by targeting a single site upstream of the coding sequence, we designed 10 target sequences for the C-X-C chemokine receptor type 4 (CXCR4) (Table S3). CXCR4 expression is low in HEK293T cells, and changes in gene expression can be quantified at the single cell level by antibody staining. CXCR4 has previously been a target for CRISPR-based gene silencing in cell types with high basal expression levels (Gilbert et al., 2013). We used the divalent 2x MS2 (wt+f6) scRNA design to recruit the MCP-VP64 protein, and we observed increases in CXCR4 expression for nine of the ten target sites (Figure S3). For the three strongest target sites, we compared CXCR4 activation mediated by scRNA to that with dCas9-VP64 and observed consistently stronger output with scRNA (Figure 3B).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

scRNAs Recruit Chromatin Modifiers to Enhance Gene Silencing in Human Cells

In human cells, CRISPRi-mediated repression is relatively modest but can be enhanced by fusing dCas9 to the KRAB domain (Gilbert et al., 2013), a potent transcriptional repressor that recruits chromatin modifiers to silence target genes (Groner et al., 2010). To determine if scRNAs could recruit KRAB to enhance CRISPR-based gene silencing, we fused KRAB to RNA binding domains and designed scRNA constructs to target an SV40 promoter driving GFP expression. We targeted one site (P1) upstream of the transcriptional start site (TSS) and another site (NT1) that overlaps the TSS. Recruitment of a Com-KRAB fusion protein to either site by a com scRNA represses the GFP reporter beyond that obtained by CRISPRi alone (there is no significant CRISPRi effect at the P1 site upstream of the TSS) (Figure 3C). The behavior of the KRAB domain recruited by scRNA was similar to that obtained with a direct dCas9-KRAB fusion protein. MCP-KRAB and PCP-KRAB fusion proteins were ineffective at mediating repression, potentially because MCP and PCP form dimers (Chao et al., 2008), which could interfere with KRAB function.

Simultaneous ON/OFF Gene Regulation in Human Cells

The successful application of scRNA-mediated transcriptional control in human cells opens the way towards simultaneous ON/OFF gene regulatory switches mediated by orthogonal RNA binding proteins fused to transcriptional activators (VP64) or repressors (KRAB). To test this possibility, we targeted endogenous CXCR4 for activation with MCP-VP64 while simultaneously targeting an additional endogenous gene for repression with COM-KRAB in HEK293T cells. We selected the β-1,4-N-acetyl-galactosaminyl transferase (B4GALNT1) gene from a set of target sites previously validated for repression with the dCas9-KRAB fusion protein (Gilbert et al., 2014). We observe simultaneous activation of CXCR4 and repression of B4GALNT1 measured by RT-qPCR, and these changes in gene expression are similar to that observed when single genes were targeted (Figure 3D). Importantly, activation and repression are mediated by a single scRNA for each target gene. Thus, this platform could in principle be used for large-scale screening of pairwise combinations of genes that yield a target phenotype when one gene is activated and the other is repressed.

Harnessing scRNA Multi-Gene ON/OFF Transcriptional Programs to Redirect Metabolic Pathway Output in Yeast

The complex multi-gene transcriptional programs that can be generated using scRNAs and dCas9 have the potential to rewire and control diverse cellular networks. One particularly interesting application is metabolic control. In biotechnology production strains, there is often competition between pathways required for cell growth versus production of the desired product. In these cases, being able to facilely control the expression of sets of metabolic enzymes, especially with bidirectional (ON/OFF) control, is essential to optimizing new flux patterns and, thereby, production of the desired product (Paddon et al., 2013; Ro et al., 2006). There is a notable lack of approaches to flexibly increase the expression of enzymes in a desired pathway branch while simultaneously downregulating the expression of enzymes in a competing branch.

To test the ability of our scRNA programs to redirect metabolic pathway outputs, we turned to the highly-branched bacterial violacein biosynthetic pathway (Hoshino, 2011). The

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

complete five-gene pathway (*VioABEDC*) produces the violet pigment violacein, and branch points at the last two enzymatic steps (*VioD* and *VioC*) can direct pathway output among four distinctly-colored products (Figure 4A). The five-gene pathway can be reconstituted in yeast, and tuning the promoter strength for expression of *VioD* and *VioC* redirects pathway output to different products in a predictable manner (Lee et al., 2013). The four product states are visually distinguishable in yeast colonies and easily quantified by HPLC, making this pathway an ideal model system to simultaneously tune expression levels of multiple independent target genes to control functional output states.

We designed a yeast reporter strain with two key control points: the first (*VioA*) regulates total precursor flux into the pathway and the second regulates flow at the *VioC*/*VioD* branch point. The *VioBED* genes are expressed by strong promoters and *VioAC* genes are under the control of weak promoters (Figure 4B and Table S4), so that turning *VioA* ON will drive flux into the pathway, and flipping the ON/OFF expression states of *VioC* and *VioD* will redirect the product output. The eight possible pairwise ON/OFF combinations of these three genes leads to five distinct output states: one state with complete pathway output off and four alternative product states when the pathway is on. To access all five states, we designed an scRNA program to target *VioA* and *VioC* with independent activators (2x PP7 and 1x MS2, respectively) and to target *VioD* with CRISPRi-mediated repression (Figure 4B and Table S2). Activation of *VioA* routes pathway flux to the proviolacein product (PV) (Figure 4C). Once *VioA* is activated, activation of *VioC* or repression of *VioD* reroutes flux in a predictable manner. Expressing all three scRNA constructs simultaneously activates *VioA* and *VioC* and represses *VioD* to route flux into the pathway and towards the deoxyviolacein (DV) product. The scRNA/dCas9 platform flexibly and efficiently generates each of the multi-gene transcriptional states necessary to yield all possible metabolic outputs of the violacein pathway.

Importantly, competition for a fixed pool of dCas9 is not limiting when multiple scRNA constructs are expressed; we observe no significant differences in scRNA-mediated effects at individual gene targets when up to four scRNA constructs are expressed (Figure S4). This result suggests that a large-scale scRNA program can be implemented, as intracellular dCas9 concentration is not a limiting factor, although it remains possible that effects from limiting dCas9 will be observed with different expression levels, or if greater numbers of scRNAs are expressed.

dCas9 Acts as a Master Regulator to Execute Expression Programs

The dCas9 protein is a central regulatory node in the execution of scRNA-mediated gene expression programs, raising the possibility that it could act as a single synthetic master regulator, controlling expression levels for multiple downstream genes (Figure 5A). We designed a system in which expression of dCas9 controls a switch from a cell type that produces the PV metabolic product to one that produces DV. Expression of dCas9 was controlled by an inducible pGal10-dCas9 construct. The starting yeast strain contained the *VioABED* genes under the control of strong promoters, and *VioC* under the control of a weak promoter (Table S4). We introduced a two-scRNA program to switch *VioC*/*VioD* from OFF/ON to ON/OFF, redirecting output from PV to DV. When all components are present

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

in yeast, but Gal inducer is absent, PV is the dominant product. When the Gal inducer is present, dCas9 is expressed to execute the simultaneous switch of *VioC* ON and *VioD* OFF such that pathway output is routed to DV (Figure 5B). Thus, multiple scRNAs can be regulated using expression of the dCas9 protein as a single control point.

Discussion

CRISPR Toolkit Enables Construction of Complex Regulatory Circuits

A wide range of CRISPR-related technologies have recently emerged for editing and manipulating target genomes (Mali et al., 2013b; Sander and Joung, 2014). A key advantage of these tools is that they interface with core biological mechanisms, thus allowing the system to be easily ported between different organisms. Watson-Crick base-pairing rules specify target site selection, and synthetic effector proteins interface with conserved features of the transcriptional machinery to control gene expression. Here we have expanded the scope of the CRISPR toolkit further by adding another basic feature of biological systems, spatial organization mediated by scaffolding molecules, to link functional effector domains to specific genomic target sites. A modular scaffold RNA encodes, within a single molecule, the information specifying the target site in the genome and the particular regulatory function to be executed at that site. scRNAs encode this information using a 5′ 20 base targeting sequence, a common dCas9-binding domain, and a 3′ protein recruitment domain. Expression of multiple RNA scaffolds simultaneously permits independent, programmable control of multiple genes in parallel. Most simply, this approach provides a straightforward method to implement simultaneous multi-gene ON/OFF regulatory switching programs.

scRNAs also allow straightforward fine-tuning of output levels in a more analog fashion by altering the valency of effector proteins recruited to an individual target site. Although not explored here, an additional layer of expression control could come from the choice of scRNA target site. In this work we screened several candidate target sites to identify those that produced maximal output for further analysis (Figure S3, Table S2 & S3). To access a range of intermediate output levels, target sites that are less effective could also be selected. More systematic screening approaches will provide general rules to select target sites for varying output levels (Gilbert et al., 2014).

Finally, there are many different classes of protein effectors and epigenetic modifiers that could be recruited via scRNAs to produce different levels and types of gene activation or repression. Qualitatively different regulatory strategies could be implemented, such as regulators that can produce stable, long-lived chromatin states that persist well after an input stimulus is removed. Recent progress towards recruiting a library of epigenetic modifiers with zinc finger proteins (Keung et al., 2014) suggests that a similar range of functionality could be achieved by recruitment via scRNAs. Thus it may be possible to construct even more nuanced and sophisticated gene expression programs by using a variety of regulators with CRISPR scRNAs, and by recruiting these regulators in a combinatorial fashion.

These scRNA-encoded transcriptional programs have several advantages that are lacking in other platforms. First, they are easily programmable using Watson-Crick base pairing to target desired endogenous loci in the genome. TAL effectors can be used to generate

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

complex programs, and these effectors can produce larger effects on gene expression than CRISPR-based approaches, but this requires the custom design of many distinct TAL specificities (Kabadi and Gersbach, 2014). Second, scRNA programs allow for distinct regulatory actions to take place at each target locus. While CRISPRi programs can be targeted to many distinct sites in the genome, fusing or tethering a regulatory effector directly to the Cas9 protein allows only one type of regulatory event (e.g. activation or repression) to take place at all targets. By tethering effectors to binding motifs in the scRNA, which also encodes the target, we have created single RNA molecules that modularly specify both a target loci and regulatory outcome in their sequence. Third, although multiple scRNAs can be expressed to target many genes, they can still be controlled by a single master regulatory event, the expression of the dCas9 protein, allowing temporal control over the entire multi-gene program.

Orthogonal dCas9 proteins from other species (besides *S. pyogenes*) can recognize guide RNAs with different dCas9 binding modules (Esvelt et al., 2013) and thus can provide another potential layer for modular control in CRISPR engineered transcriptional circuits that is complementary to the scaffold RNAs explored here (Figure 6). For example, one could create, in a single cell, alternative sets of scRNA programs, each executed by an orthogonal dCas9 ortholog. In such a case, one could switch between distinct programs by controlling the expression of the dCas9 master regulators.

Applications: Reprogramming Complex Networks Controlling Cell Function and Fate

scRNA encoded transcriptional programs provide powerful tools for manipulating complex cellular behaviors, such as differentiation or metabolism. In metabolic engineering, microorganisms can be engineered for biosynthesis by heterologous expression of the desired metabolic pathway. Designing these microbial production factories requires careful engineering to prevent detrimental effects on host growth and metabolism, to avoid buildup of toxic intermediates, and to coordinate the expression of multiple genes to switch from growth to production phase (Keasling, 2012). Often optimizing production requires a coordinated increase in the expression of enzymes that convert precursors into the desired product, as well as simultaneous repression of enzymes that direct these precursors towards alternative products. Since these alternative products are often necessary for growth, optimized production requires precise and coordinated temporal control. It is difficult to construct complex programs of this type with only a handful of well-characterized inducible promoters.

A CRISPR RNA-encoded gene expression program can address these challenges by activating multiple target pathway genes while simultaneously repressing branch points that divert metabolites to cell growth. Execution of the program can be controlled by a dCas9 master regulator that is induced at the appropriate time. To avoid toxic intermediate buildup, expression levels of target pathway genes can be tuned to different levels, using differential multivalent recruitment of activators, to prevent bottlenecks. One potential limitation of the CRISPR-mediated approach, however, is that metabolic flux is often regulated by mechanisms that act post-transcriptionally (Daran-Lapujade et al., 2007). Nevertheless, there are many situations in which selectively altering transcription levels of pathway enzymes

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

can optimize metabolic flux (Ajikumar et al., 2010; Du et al., 2012; Latimer et al., 2014; Yuan and Zhao, 2013).

To improve metabolite production, CRISPR RNA-based scaffolds could also be used as a rapid prototyping strategy to screen gene expression programs that simultaneously alter the expression levels of multiple metabolic enzymes. The regions of expression space that are then identified by such screens could then be custom constructed with specific promoters to achieve finer control. CRISPR tools can also be combined with other approaches to optimize metabolic networks. Global transcription machinery engineering (gTME) screens mutations in general transcription factors or coactivators to modify the expression of many genes simultaneously (Alper et al., 2006). gTME could be used to identify potential target genes for control by scRNA-encoded programs and a dCas9 master regulator. Alternatively, a dCas9 master regulator could be used to switch between global transcription programs by activating and repressing modified general transcription factors that elicit global changes in gene expression.

Finally scRNA/CRISPR programs are easily transferable to different hosts. Most metabolic engineering efforts use well-characterized and genetically tractable hosts like *E. coli* or *S. cerevisiae*, which offer many desirable industrial characteristics. CRISPR-based tools to modify and regulate host genomes may dramatically expand the space of microorganisms that can be engineered for biosynthesis. Microbial strains or plants that have desirable industrial characteristics or metabolic precursors but lack good tools for genome manipulation may now be accessible for engineering. CRISPR-based tools could also be used to optimize target molecule production in the native host organism for a desired pathway rather than in a heterologous host.

Another broad area of applications for customized expression programs is in controlling cell fate decisions. During development, master regulators specify cell fates by directly or indirectly regulating multiple downstream target genes, and their presence or absence can determine the outcome of a developmental lineage (Chan and Kyba, 2013). A CRISPR-based multidirectional ON/OFF switch program could provide a straightforward method for genetic reprogramming by synthetically mimicking the behavior of master regulators. scRNA programs could be used to simultaneously activate and repress different master regulators, or to bypass master regulators and directly engage the next layer of target genes to specify cell fates. scRNA programs could also be used to create customized hybrid cell fate states that are not generated by natural master regulators, but that might still be useful in a therapeutic or research context. In either scenario, the ability of dCas9 itself to act as a synthetic master regulator will be a useful tool for controlling the timing of differentiation. Synthetic control of cell fate reprogramming could provide powerful new tools for regenerative medicine or other cell-based therapeutics.

CRISPR scRNAs as Screening Tools for Biology

High-throughput synthetic lethal screens have proven extremely powerful in analyzing complex biological systems and shedding light on strategies for treating disease networks. Such screens, however, whether they utilize siRNAs or CRISPRi sgRNAs, rely on perturbing the expression of multiple genes in one direction (usually repression). It is

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

equally likely that we can learn new features of networks by simultaneously activating and repressing different combinations of genes. This is particularly true in cases in which a particular cellular outcome requires both activation of that response, but also simultaneous inactivation of genes involved in driving competing, alternative responses (Rais et al., 2013). The multi-directional, but high-throughput, regulation that can be achieved with the scRNA/CRISPR platform is ideal for this type of exploration.

Experimental Procedures

scRNA Sequence Design

sgRNA sequences were extended to include hairpin sequences for MS2 (C5 variant) (Lowary and Uhlenbeck, 1987), PP7 (Lim et al., 2001), or com (Hattman, 1999). Sequences for linkers to the guide RNA and between hairpins were generated by RNA Designer (Andronescu et al., 2004). Candidate sequences were linked to the complete sgRNA sequence and evaluated in NUPACK (Zadeh et al., 2011) to confirm that the extended hairpins were compatible with sgRNA folding. Successful candidates were then evaluated for function in yeast as described below. The 2x MS2 (wt+f6) scRNA design uses the SELEX f6 aptamer, which was selected to bind the MCP protein (Hirao et al., 1998). Sequences of the minimal sgRNA, extended scRNAs, and RNA-binding modules are described in the Extended Experimental Procedures and Table S1.

Construct Design for CRISPR in Yeast

Mammalian codon-optimized *S. pyogenes* dCas9 (Qi et al., 2013) with three C-terminal SV40 NLSs was expressed from a constitutive Tdh3 or inducible Gal10 promoter. The dCas9-VP64 fusion protein was constructed with two C-terminal SV40 NLSs, the VP64 domain (Beerli et al., 1998), and an additional SV40 NLS. RNA-binding proteins MCP (ΔFG/V29I mutant) (Lim and Peabody, 1994), PCP (ΔFG mutant) (Chao et al., 2008), and Com (Hattman, 1999) were expressed with an N-terminal SV40 NLS and a C-terminal VP64 fusion domain. All protein expression constructs were integrated in single copy into the yeast genome. Complete descriptions of these constructs are provided in Table S5. sgRNA constructs were expressed from the pRS316 CEN/ARS plasmid (*ura3* marker) with the SNR52 promoter and SUP4 terminator (DiCarlo et al., 2013). sgRNA target sites are listed in Table S2. 20 base guide sequences upstream of an appropriate PAM motif for *S. pyogenes* dCas9 (Qi et al., 2013) were selected. For genes that had not been previously targeted for CRISPR-based transcriptional regulation, we screened 8 candidate sites upstream of the gene and tested each site independently for the desired output (Table S2). The target site with the strongest effect on output was used for subsequent experiments. Methods for manipulation and analysis of yeast strains are described in the Extended Experimental Procedures.

Construct Design for CRISPR in Human Cells

Plasmids for expression of *S. pyogenes* dCas9, dCas9 fusion proteins, and sgRNA constructs were described previously (Gilbert et al., 2013). dCas9 constructs were expressed from an SFFV promoter with two C-terminal SV40 NLSs and a tagBFP. The dCas9-KRAB fusion protein was constructed with a KRAB domain (Margolin et al., 1994) fused to the C-

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

terminus of the tagBFP. The dCas9-VP64 fusion protein was constructed with two C-terminal SV40 NLSs, the VP64 domain, an additional SV40 NLS, and a tagBFP. sgRNA sequences were modified as described previously for expression in human cells (see Extended Experimental Procedures) (Chen et al., 2013). sgRNAs were expressed using a lentiviral U6-based expression vector derived from pSico that expresses mCherry from a CMV promoter. To simultaneously express sgRNAs and RNA-binding protein effectors, the mCherry cassette was modified to express the protein effector followed by an IRES and mCherry. RNA-binding proteins (MCP, PCP, and Com) were expressed with an N-terminal SV40 NLS and a C-terminal VP64 or KRAB fusion domain. Complete descriptions of these constructs are provided in Table S7. sgRNA target site sequences are listed in Table S3. To simultaneously target two genes in human cells, we designed a pSico-derived construct with a U6 promoter driving expression of sgRNAs and a CMV promoter driving expression of a protein effector followed by a p2A sequence and tagBFP (Table S7 & Figure S3C). Methods for manipulation and analysis of human cells are described in the Extended Experimental Procedures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Jason Park, Krista McNally, Tammy Hsu, Matthew Larson, Sandra Torres, Max Horlbeck, William DeLoache and members of the Lim lab for technical assistance, advice, and helpful discussions. L.S.Q. acknowledges support from the UCSF Center for Systems and Synthetic Biology, NIH Office of the Director (OD), and National Institute of Dental & Craniofacial Research (NIDCR). This work was supported by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (J.G.Z.), NIH P50 GM081879 (J.G.Z., R.A., L.S.Q., and W.A.L.), NIH R01 DA036858 (J.S.W., L.S.Q., and W.A.L.), Howard Hughes Medical Institute (J.G.Z., L.A.G., J.C.T., R.A., J.S.W., and W.A.L.), Leukemia and Lymphoma Society Postdoctoral Fellowship (L.A.G.), NSF SynBERC EEC-0540879 (R.A. and W.A.L.), Energy Biosciences Institute (M.E.L. and J.E.D.), National Defense Science & Engineering Graduate Fellowship Program (M.E.L.), Genentech Predoctoral Fellowship (M.L.), and NIH DP5 OD017887 (E.H.W., M.L., and L.S.Q.).

References

Ajikumar PK, Xiao WH, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science. 2010; 330:70–74. [PubMed: 20929806]

Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G. Engineering yeast transcription machinery for improved ethanol tolerance and production. Science. 2006; 314:1565–1568. [PubMed: 17158319]

Andronescu M, Fejes AP, Hutter F, Hoos HH, Condon A. A new algorithm for RNA secondary structure design. J Mol Biol. 2004; 336:607–624. [PubMed: 15095976]

Beerli RR, Segal DJ, Dreier B, Barbas CF. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. P Natl Acad Sci Usa. 1998; 95:14628–14633.

Chan SS-K, Kyba M. What is a Master Regulator? J Stem Cell Res Ther. 2013; 3

Chao JA, Patskovsky Y, Almo SC, Singer RH. Structural basis for the coevolution of a viral RNA-protein complex. Nat Struct Mol Biol. 2008; 15:103–105. [PubMed: 18066080]

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013; 155:1479–1491. [PubMed: 24360272]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Daran-Lapujade P, Rossell S, van Gulik WM, Luttik MAH, de Groot MJL, Slijper M, Heck AJR, Daran JM, de Winde JH, Westerhoff HV, et al. The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels. P Natl Acad Sci Usa. 2007; 104:15753–15758.

Delebecque CJ, Lindner AB, Silver PA, Aldaye FA. Organization of intracellular reactions with rationally designed RNA assemblies. Science. 2011; 333:470–474. [PubMed: 21700839]

DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Research. 2013; 41:4336–4343. [PubMed: 23460208]

Du J, Yuan Y, Si T, Lian J, Zhao H. Customized optimization of metabolic pathways by combinatorial transcriptional engineering. Nucleic Acids Research. 2012; 40:e142. [PubMed: 22718979]

Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat Methods. 2013; 10:1116–1121. [PubMed: 24076762]

Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013; 31:397–405. [PubMed: 23664777]

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell. 2014; 159:647–661. [PubMed: 25307932]

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013; 154:442–451. [PubMed: 23849981]

Good MC, Zalatan JG, Lim WA. Scaffold proteins: hubs for controlling the flow of cellular information. Science. 2011; 332:680–686. [PubMed: 21551057]

Groner AC, Meylan S, Ciuffi A, Zangger N, Ambrosini G, Dénervaud N, Bucher P, Trono D. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS Genet. 2010; 6:e1000869. [PubMed: 20221260]

Hattman S. Unusual transcriptional and translational regulation of the bacteriophage Mu mom operon. Pharmacol Ther. 1999; 84:367–388. [PubMed: 10665835]

Hirao I, Spingola M, Peabody D, Ellington AD. The limits of specificity: an experimental analysis with RNA aptamers to MS2 coat protein variants. Mol Divers. 1998; 4:75–89. [PubMed: 10425631]

Hoshino T. Violacein and related tryptophan metabolites produced by Chromobacterium violaceum: biosynthetic mechanism and pathway for construction of violacein core. Appl Microbiol Biotechnol. 2011; 91:1463–1475. [PubMed: 21779844]

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337:816–821. [PubMed: 22745249]

Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014; 343:1247997. [PubMed: 24505130]

Kabadi AM, Gersbach CA. Engineering synthetic TALE and CRISPR/Cas9 transcription factors for regulating gene expression. Methods. 2014; 69:188–197. [PubMed: 25010559]

Keasling JD. Synthetic biology and the development of tools for metabolic engineering. Metab Eng. 2012; 14:189–195. [PubMed: 22314049]

Keung AJ, Bashor CJ, Kiriakov S, Collins JJ, Khalil AS. Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. Cell. 2014; 158:110–120. [PubMed: 24995982]

Latimer LN, Lee ME, Medina-Cleghorn D, Kohnz RA, Nomura DK, Dueber JE. Employing a combinatorial expression approach to characterize xylose utilization in Saccharomyces cerevisiae. Metab Eng. 2014; 25:20–29. [PubMed: 24930894]

Lee ME, Aswani A, Han AS, Tomlin CJ, Dueber JE. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. Nucleic Acids Research. 2013; 41:10668–10678. [PubMed: 24038353]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Lim F, Peabody DS. Mutations that increase the affinity of a translational repressor for RNA. Nucleic Acids Research. 1994; 22:3748–3752. [PubMed: 7937087]

Lim F, Downey TP, Peabody DS. Translational repression and specific RNA binding by the coat protein of the Pseudomonas phage PP7. J Biol Chem. 2001; 276:22507–22513. [PubMed: 11306589]

Lowary PT, Uhlenbeck OC. An RNA mutation that increases the affinity of an RNA-protein interaction. Nucleic Acids Research. 1987; 15:10483–10493. [PubMed: 3697094]

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013a; 31:833–838. [PubMed: 23907171]

Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. Nat Methods. 2013b; 10:957–963. [PubMed: 24076990]

Margolin JF, Friedman JR, Meyer WK, Vissing H, Thiesen HJ, Rauscher FJ. Krüppel-associated boxes are potent transcriptional repression domains. P Natl Acad Sci Usa. 1994; 91:4509–4513.

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014; 156:935–949. [PubMed: 24529477]

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 2013; 496:528–532. [PubMed: 23575629]

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013; 152:1173–1183. [PubMed: 23452860]

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M, et al. Deterministic direct reprogramming of somatic cells to pluripotency. Nature. 2013; 502:65–70. [PubMed: 24048479]

Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–166. [PubMed: 22663078]

Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature. 2006; 440:940–943. [PubMed: 16612385]

Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014; 32:347–355. [PubMed: 24584096]

Spitale RC, Tsai MC, Chang HY. RNA templating the epigenome: long noncoding RNAs as molecular scaffolds. Epigenetics. 2011; 6:539–543. [PubMed: 21393997]

Wulczyn FG, Kahmann R. Translational stimulation: RNA sequence and structure requirements for binding of Com protein. Cell. 1991; 65:259–269. [PubMed: 1826635]

Yuan Y, Zhao H. Directed evolution of a highly efficient cellobiose utilizing pathway in an industrial Saccharomyces cerevisiae strain. Biotechnol Bioeng. 2013; 110:2874–2881. [PubMed: 23616289]

Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA. NUPACK: Analysis and design of nucleic acid systems. J Comput Chem. 2011; 32:170–173. [PubMed: 20645303]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

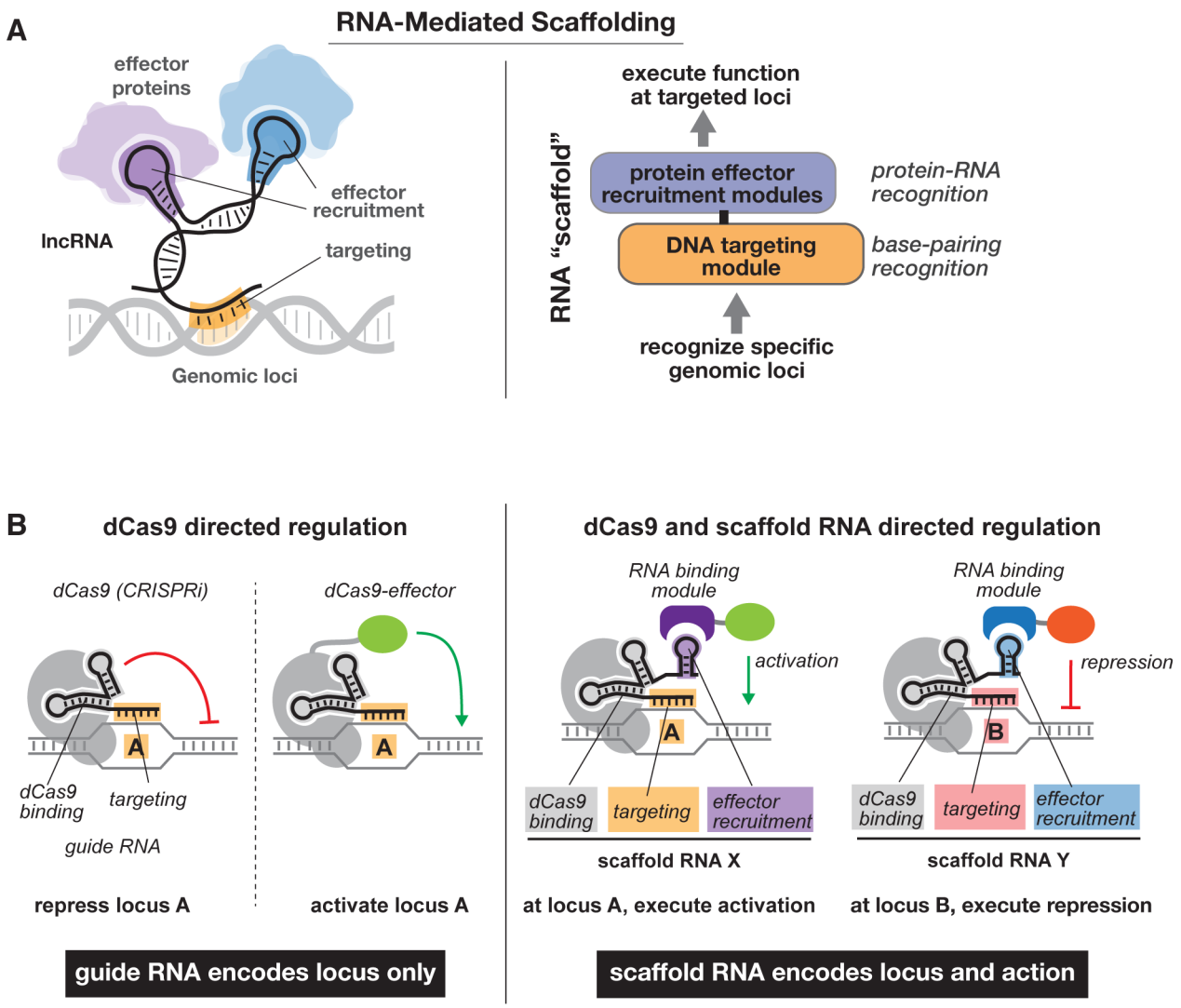

Figure 1. Genomic Regulatory Programming Using CRISPR and Multi-Domain Scaffolding RNAs

(A) lncRNA molecules may act as scaffolds to physically assemble epigenetic modifiers at their genomic targets. Modular RNA structures can encode domains for protein binding and DNA targeting to co-localize proteins to genomic loci.

(B) CRISPR RNA scaffold-based recruitment allows simultaneous regulation of independent gene targets. The minimal CRISPRi system silences target genes when dCas9 and an sgRNA assemble to physically block transcription. Fusing dCas9 to transcriptional activators or repressors provides additional functionality. When function is encoded in dCas9 (CRISPRi) or dCas9-effector fusion proteins, the sgRNA recruits the same function to every target site. To encode both target and function in a scaffold RNA, sgRNA molecules are extended with additional domains to recruit RNA binding proteins that are fused to functional effectors. This approach allows distinct types of regulation to be executed at individual target loci, thus allowing simultaneous activation and repression.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

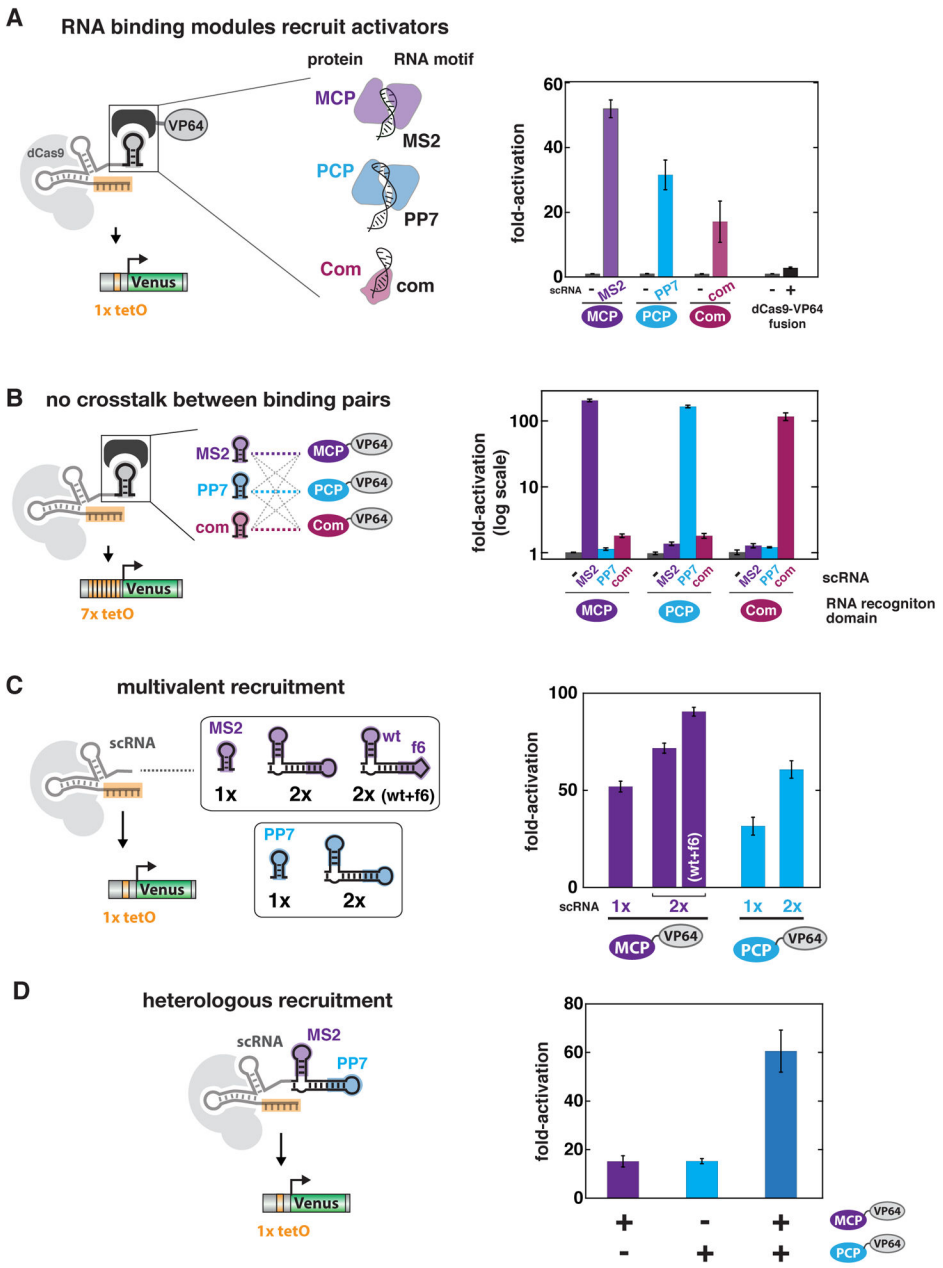

Figure 2. Multiple Orthogonal RNA Binding Modules Can Be Used to Construct CRISPR Scaffolding RNAs

(A) scRNA constructs with MS2, PP7, or com RNA hairpins recruit their cognate RNA-binding proteins fused to VP64 to activate reporter gene expression in yeast. A yeast strain with an unmodified sgRNA and the dCas9-VP64 fusion protein gives comparatively weaker reporter gene activation. The MS2 and PP7 RNA hairpins bind at a dimer interface on their corresponding MCP and PCP binding partner proteins (Chao et al., 2008), potentially recruiting two VP64 effectors to each RNA hairpin. The structure of the com RNA hairpin in complex with its binding protein has not been reported, but functional data suggest that a single Com monomer protein binds at the base of the com RNA hairpin (Wulczyn and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Kahmann, 1991). scRNA constructs and corresponding RNA-binding proteins were expressed in yeast with dCas9 and a 1x tetO-VENUS reporter gene. Representative flow cytometry data are presented in Figure S1.

(B) There is no significant crosstalk between mismatched pairs of scRNA sequences and non-cognate binding proteins. scRNA constructs and RNA-binding proteins were expressed in yeast with dCas9, using a 7x tetO-VENUS reporter gene to detect any potential weak crosstalk between mismatched pairs. The y-axis is on a log-scale, and activity with cognate scRNA-binding protein pairs is significantly greater with the 7x tet reporter compared to the 1x reporter.

(C) Multivalent recruitment with two RNA hairpins connected by a double-stranded linker produces stronger reporter gene activation compared to single RNA hairpin recruitment domains. The 2x MS2 (wt+f6) construct was designed with an aptamer sequence (f6) selected to bind to the MCP protein (Hirao et al., 1998). This construct has two distinct sequences to recruit the same protein, which may help to prevent misfolding between hairpin domains that can occur when two identical hairpins are linked on the same RNA.

(D) A mixed MS2-PP7 scRNA construct constructed using the 2x double-stranded linker architecture recruits both MCP and PCP.

Fold-change values in (A)–(D) are fluorescence levels relative to parent yeast strains lacking scRNA. Values are median ± SD for at least three measurements. RNA sequences are reported in Table S1.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

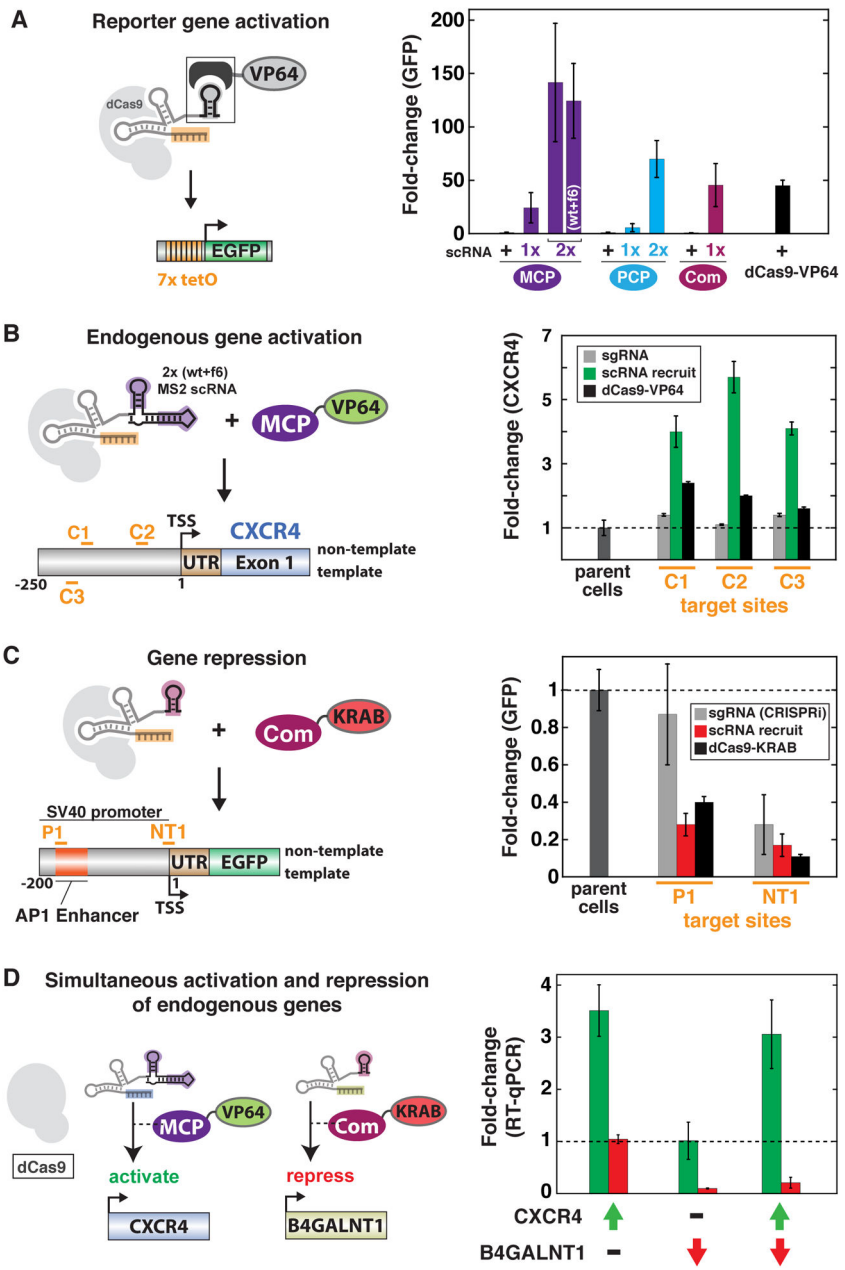

Figure 3. CRISPR RNA Scaffold Recruitment Can Activate or Repress Gene Expression in Human Cells

(A) scRNA constructs with MS2, PP7, or com RNA hairpins recruit corresponding RNA-binding proteins fused to VP64 to activate reporter gene expression in HEK293T cells. scRNA and RNA binding proteins were expressed in a cell line with dCas9 and a TRE3G-EGFP reporter containing a 7x repeat of a tet operator site. For comparison, an unmodified sgRNA targeting the same reporter gene was expressed in a cell line with the dCas9-VP64 fusion protein. Representative flow cytometry data are presented in Figure S3.

(B) The 2x MS2 (wt+f6) scRNA construct recruits MCP-VP64 to activate expression of endogenous CXCR4 in HEK293T cells expressing dCas9. Comparatively weak activation is

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

observed in cells with dCas9-VP64 and unmodified sgRNA. There is no significant activation of CXCR4 in cells with dCas9 and unmodified sgRNA. Similar effects were observed at each of three individual target sites located within ~200 bases of the transcriptional start site (TSS). Cell surface expression of CXCR4 was measured by antibody staining.

(C) The com scRNA construct recruits Com-KRAB to silence a SV40-driven EGFP reporter gene in HEK293T cells expressing dCas9. At the P1 site, upstream of the TSS, recruitment of dCas9 (i.e. CRISPRi) does not silence EGFP, but scRNA-mediated KRAB recruitment does. At the NT1 site, overlapping the TSS, CRISPRi partially silences EGFP, and scRNA-mediated KRAB recruitment further enhances silencing. The P1 and NT1 target sites were selected from a panel of sites examined in a prior study (Gilbert et al., 2013).

Fold-change values in (A)–(C) are fluorescence levels relative to a parent cell line lacking scRNA. Values are median ± SD for at least three measurements.

(D) scRNA constructs mediate simultaneous activation and repression at endogenous human genes in HEK293T cells, measured by RT-qPCR. A 2x MS2 (wt+f6) scRNA construct recruits MCP-VP64 to activate CXCR4, and a 1x com scRNA construct recruits COM-KRAB to silence B4GALNT1. Fold-change values are gene expression levels (mean ± SD) from two RT-qPCR measurements, relative to negative control cell lines. The observed change in CXCR4 mRNA level measured by RT-qPCR corresponds to an increased protein level (Figure S3D).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

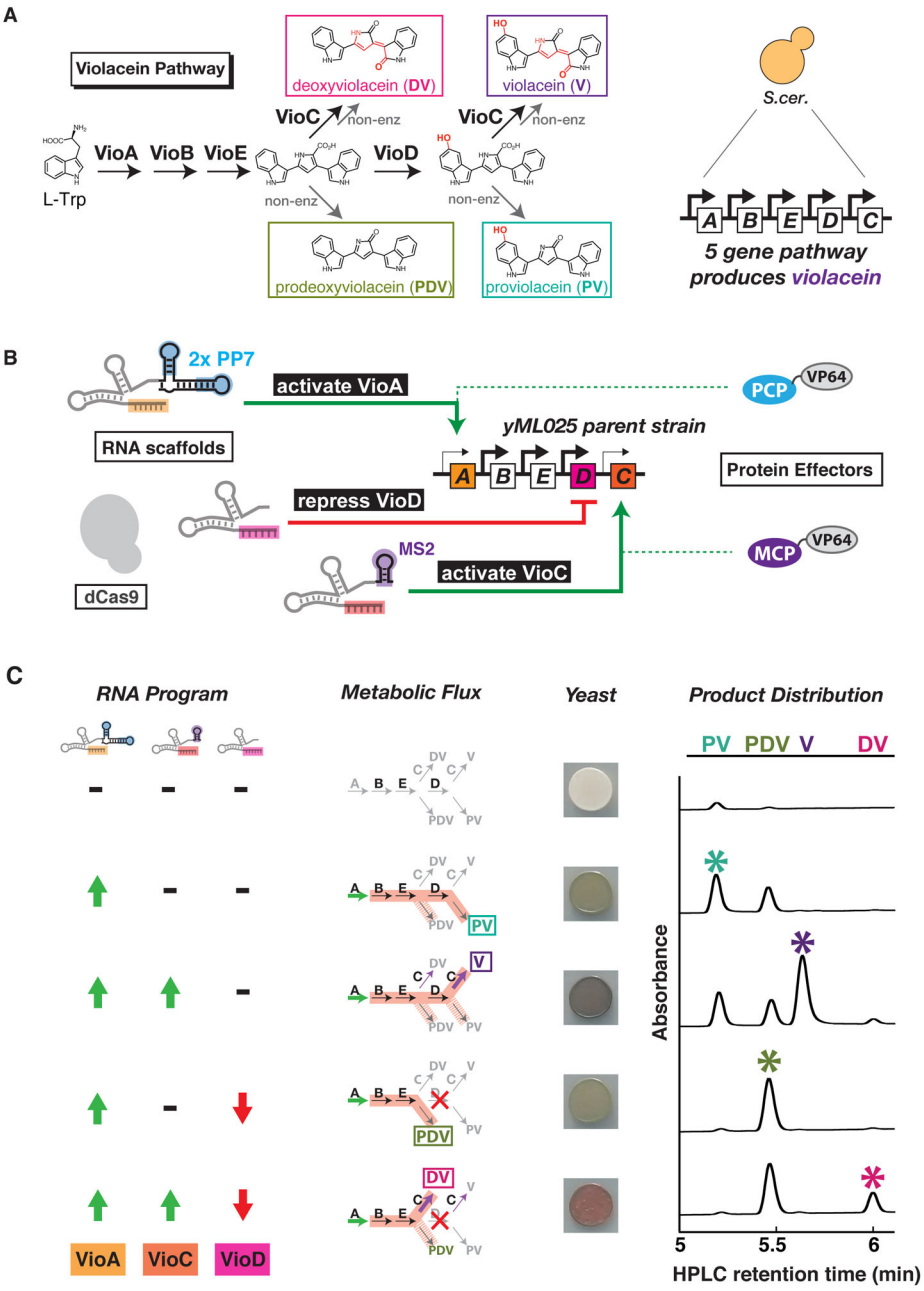

Figure 4. Reprogramming the Output of a Branched Metabolic Pathway with a 3-Gene scRNA CRISPR ON/OFF Switch

(A) Heterologous expression of bacterial violacein biosynthesis pathway in yeast produces violacein from L-Trp with five enzymatic steps and one non-enzymatic step. Branch points at the last two enzymatic transformations catalyzed by *VioD* and *VioC* produce four possible pathway outputs.

(B) An scRNA program regulates three genes simultaneously to control flux into the pathway and to direct the choice of product. The yML025 yeast strain (Table S4) has *VioBED* genes strongly expressed (ON), and *VioAC* genes weakly expressed (OFF). A 2x

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

PP7 scRNA targets *VioA* and a 1x MS2 scRNA targets *VioC* for activation. An unmodified sgRNA targets *VioD* for repression by CRISPRi.

(C) scRNA programs flexibly redirect the output of the violacein pathway. The yML025 yeast strain expressing dCas9, MCP-VP64, and PCP-VP64 was transformed with an empty parent vector (pRS316) or with a plasmid containing one, two, or three scRNA constructs to route the pathway to all four product output states (Table S6). Yeast strains were grown on SD –Ura agar plates. Product distribution was analyzed by HPLC. Stars on the chromatograms indicate the expected product of the engineered pathway. Quantitative values for changes in gene expression (by RT-qPCR) and product distributions are reported in Figure S4B.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

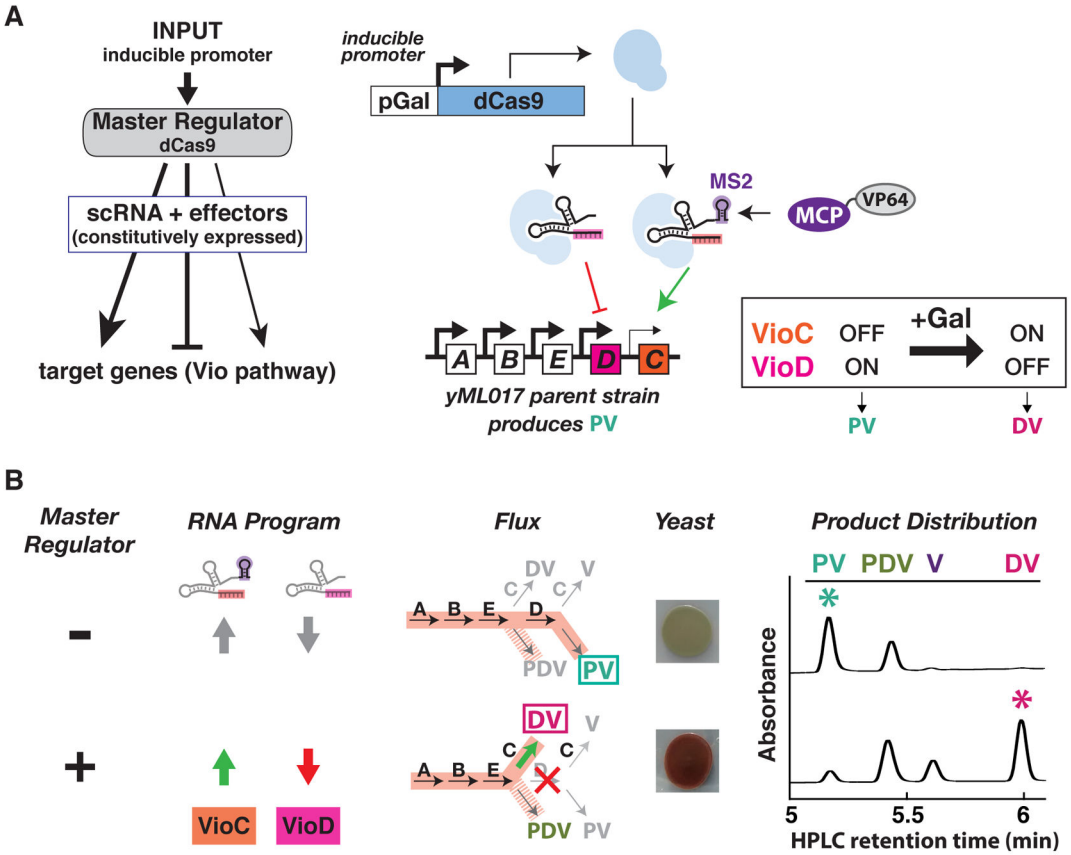

Figure 5. The dCas9 Master Regulator Inducibly Executes scRNA-Encoded Programs

(A) dCas9 can act as a synthetic master regulator of scRNA-encoded circuits. We placed dCas9 under the control of an inducible Gal10 promoter. The yML017 yeast strain (Table S4) has *VioABED* genes strongly expressed (ON), and *VioC* weakly expressed (OFF). A 1x MS2 scRNA targets *VioC* for activation. An unmodified sgRNA targets *VioD* for repression by CRISPRi.

(B) The presence of the master regulator dCas9 controls execution of the scRNA program. Yeast expressing a two-component scRNA program and MCP-VP64 were grown on agar plates in the presence or absence of galactose to control dCas9 expression. When the dCas9 master regulator is not present (−Gal), Vio pathway gene expression remains in the basal state and pathway flux proceeds to the PV product. When dCas9 is present (+Gal), *VioC* switches ON, *VioD* switches OFF, and pathway flux diverts to the DV product.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

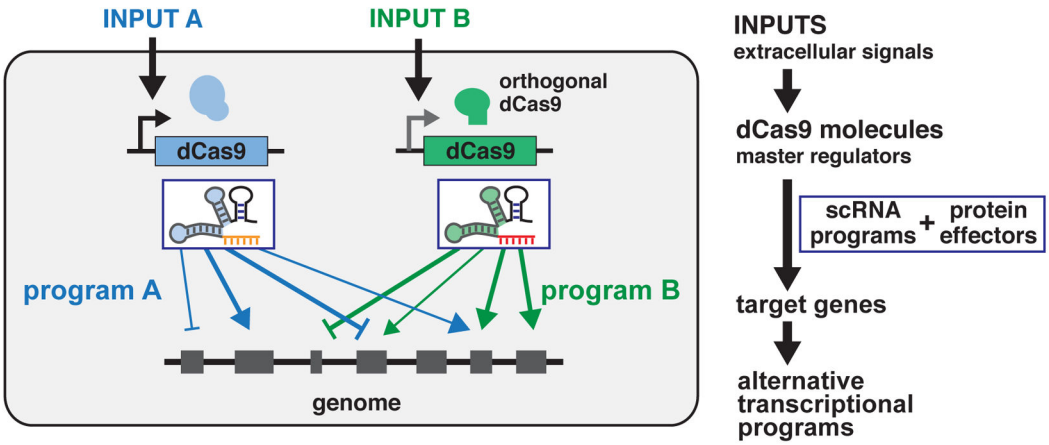

Figure 6. Encoding Complex dCas9/scRNA Regulatory Programs

scRNAs can be combined with dCas9 to construct designer transcriptional programs in which distinct target genes can be simultaneously activated or repressed, or subject to other types of regulation. Temporal control of the synthetic program can be achieved by inducing the dCas9 protein as a master regulator. Alternative scRNA gene expression programs could be achieved in the same cell by harnessing orthogonal dCas9 proteins that recognize their guide RNAs through distinct sequences (Esvelt et al., 2013). Each orthogonal dCas9 protein could control a distinct set of scRNAs, allowing independent control over distinct gene expression programs. Each scRNA, in turn, allows independent control at the level of an individual gene. Distinct dCas9 proteins could be placed under the control of different extracellular signals or inducible promoters.