Novel CRISPR technologies for generating complex synthetic gene circuits

KFSP Technical journal club
20.09.2016

Audrey Fahrny

Synthetic biology

Motivation:

- In eukaryotic cells a single genome can give rise to many distinct functional and cellular states due to the implementation of complex gene regulatory programs.
  - Ex: Developmental biology, response to environmental/endogenous signals, mitosis, etc.

- Methods to artificially control gene expression in eukaryotic cells provide valuable tools:
  - to probe the mechanisms of cell differentiation and functions.
  - to construct designer cells for therapeutic or biotechnological applications (e.g. production of biofuels and commodity chemicals).

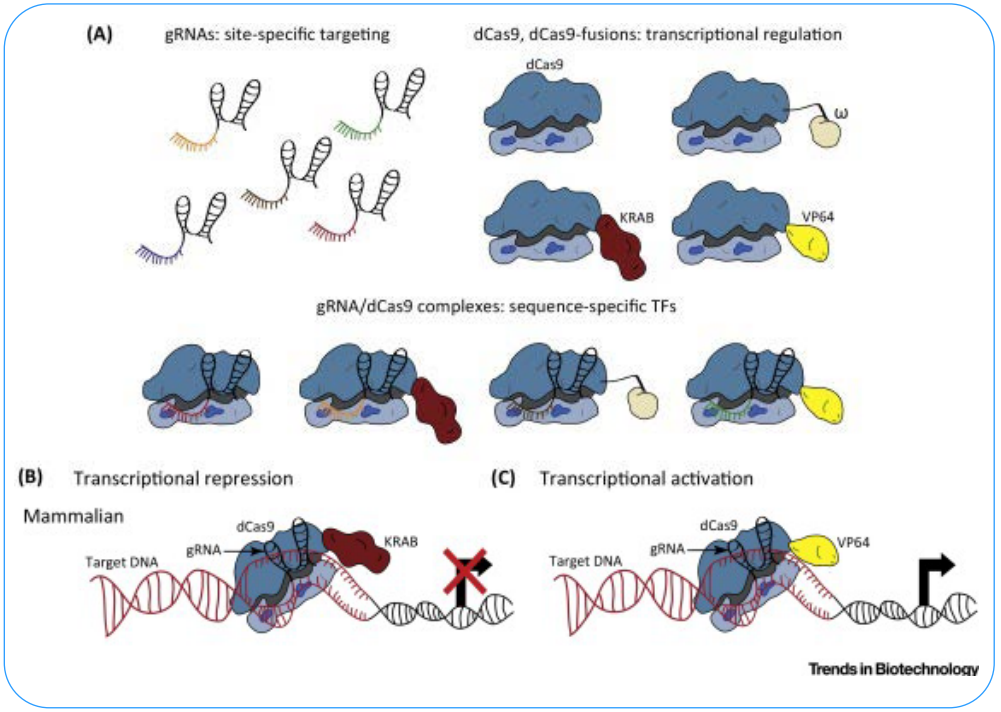
Synthetic biology is the implementation of engineering principles to the fundamental components of biology

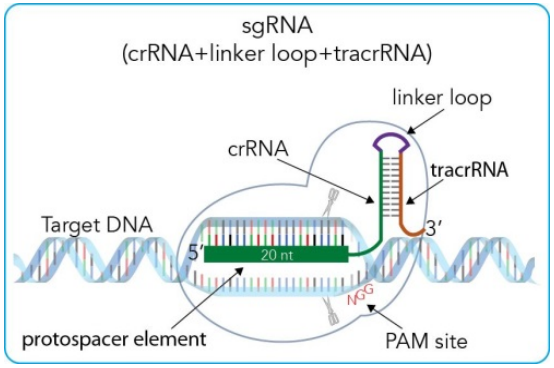
- Fundamental challenge in building gene circuits is the shortage of functional parts that can operate together in **one cell without crosstalk**.
- The ability to build transcriptional regulatory devices depends on **the availability of programmable, sequence-specific, and effective synthetic transcription factors** (TFs).

CRISPR-based synthetic transcription factors

- A CRISPR-based transcriptional regulator comprises two parts: a **single guide RNA** (sgRNA) and a **deactivated Cas9** (dCas9) protein that may be fused to a transcription regulatory domain.
- sgRNA consists of modular domains:
  - crRNA: contains both sequence complementary to the genomic DNA target and additional nucleotides which are complementary to the tracrRNA.
  - tracrRNA: hybridizes to the crRNA and binds to the dCAS9 protein.

→ **The resulting dCas9–sgRNA complex can repress or activate a gene of interest.**





https://eu.idtdna.com/

\+ **Simplicity of design**: can design TFs against any locus of interest as dCas9 targeting depends entirely on the associated sgRNA.

\+ Modified dCas9 variants can achieve **strong transcriptional activation/repression** of mammalian genes with a single gRNA per gene

→ **CRISPR system offers an unprecedented ease in designing synthetic TFs that could regulate multiple genes simultaneously**

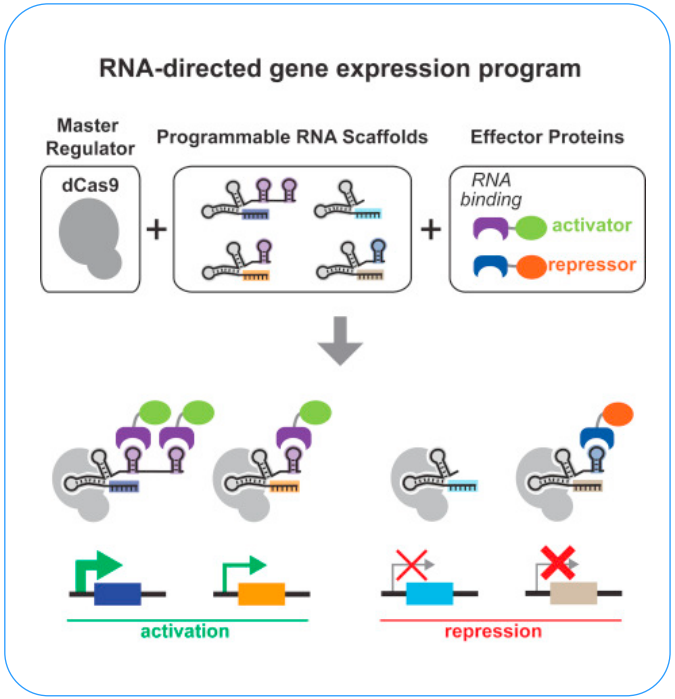
B. Jusiak et al., Trends in Biotechnology (2016)

Resource

Cell

Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds

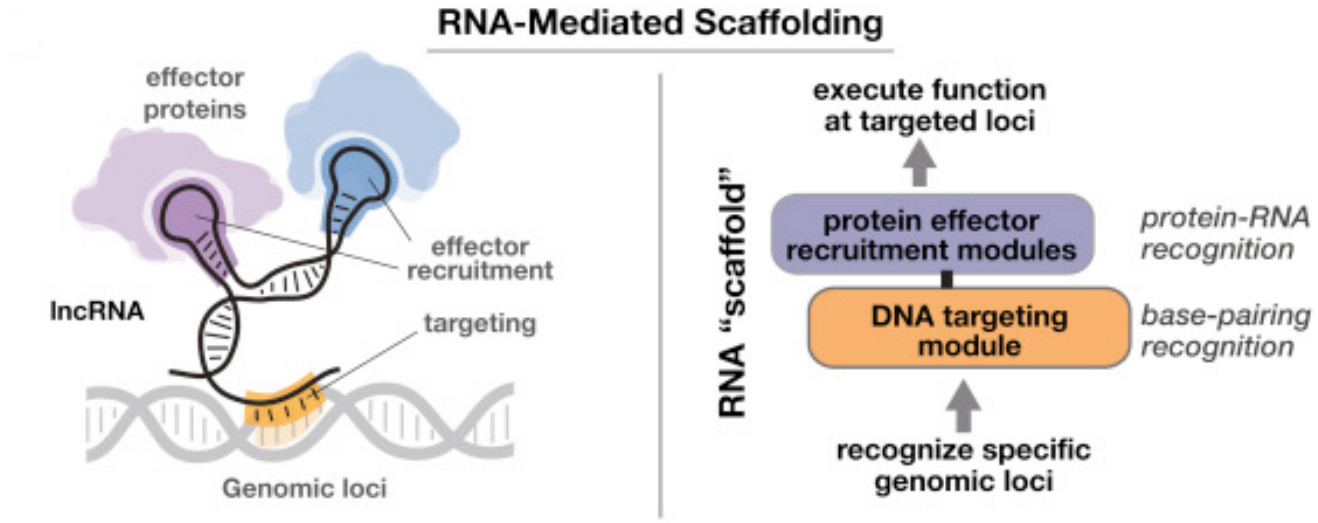
Jesse G. Zalatan,[1,2,9] Michael E. Lee,[5,6] Ricardo Almeida,[1,2] Luke A. Gilbert,[1,2,7] Evan H. Whitehead,[1,3,10] Marie La Russa,[1,3,4,10] Jordan C. Tsai,[1,2] Jonathan S. Weissman,[1,2,7,8] John E. Dueber,[5,6] Lei S. Qi,[1,3,8,10,*] and Wendell A. Lim[1,2,3,8,*]



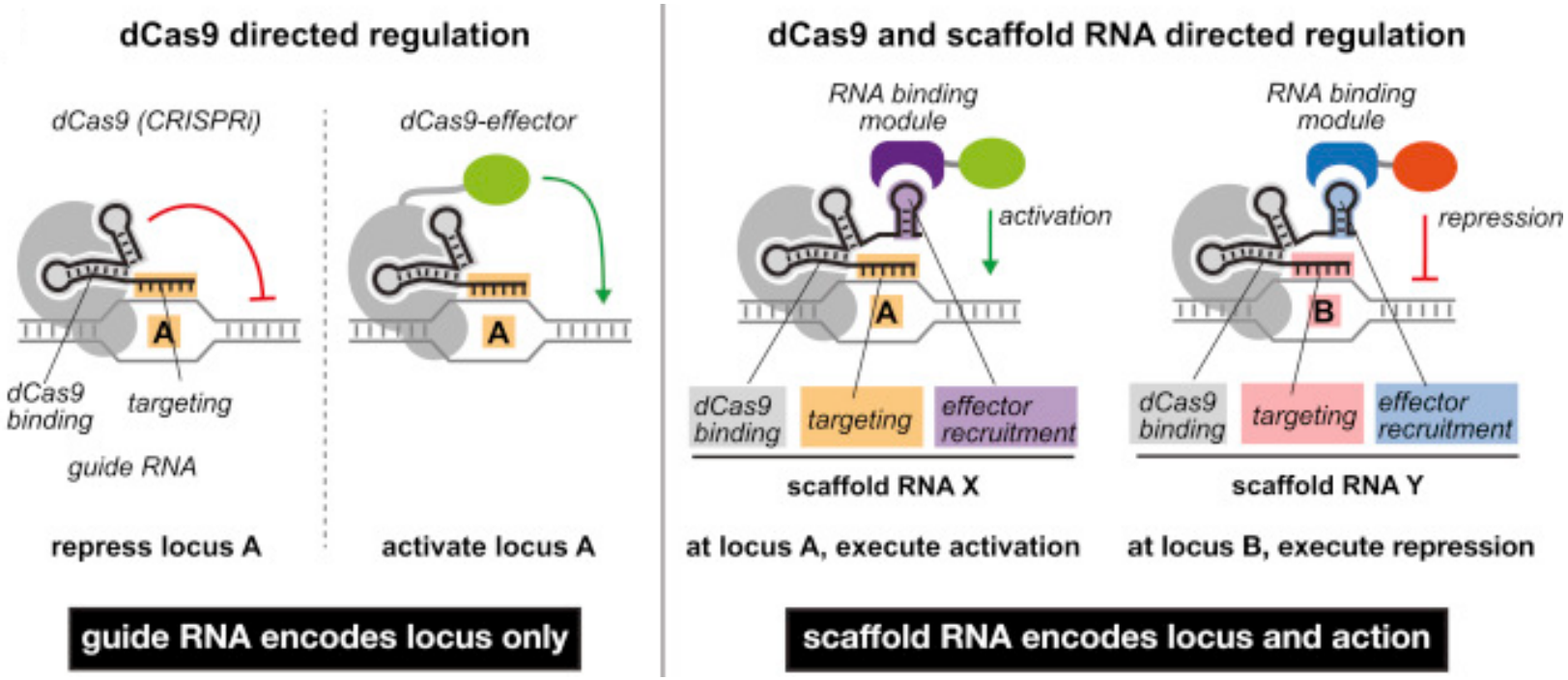
- Transform sgRNA into a **scaffold RNA** by extending sgRNA sequences to include effector protein recruitment sites. → Scaffold RNAs encode both target locus and regulatory action.
- By combing in sets of scaffold RNAs they generated synthetic multigene transcriptional programs in yeast and human cells.

Aim & approach

- Aim: Engineer novel gene regulatory programs in which many loci are targeted simultaneously but with distinct types of regulation at each locus.
- Took inspiration from natural regulatory systems that encode both target specificity and regulatory function in the same molecule
  - Ex.: lncRNA molecules proposed to act as assembly scaffolds that recruit key epigenetic modifiers to specific genomic loci.



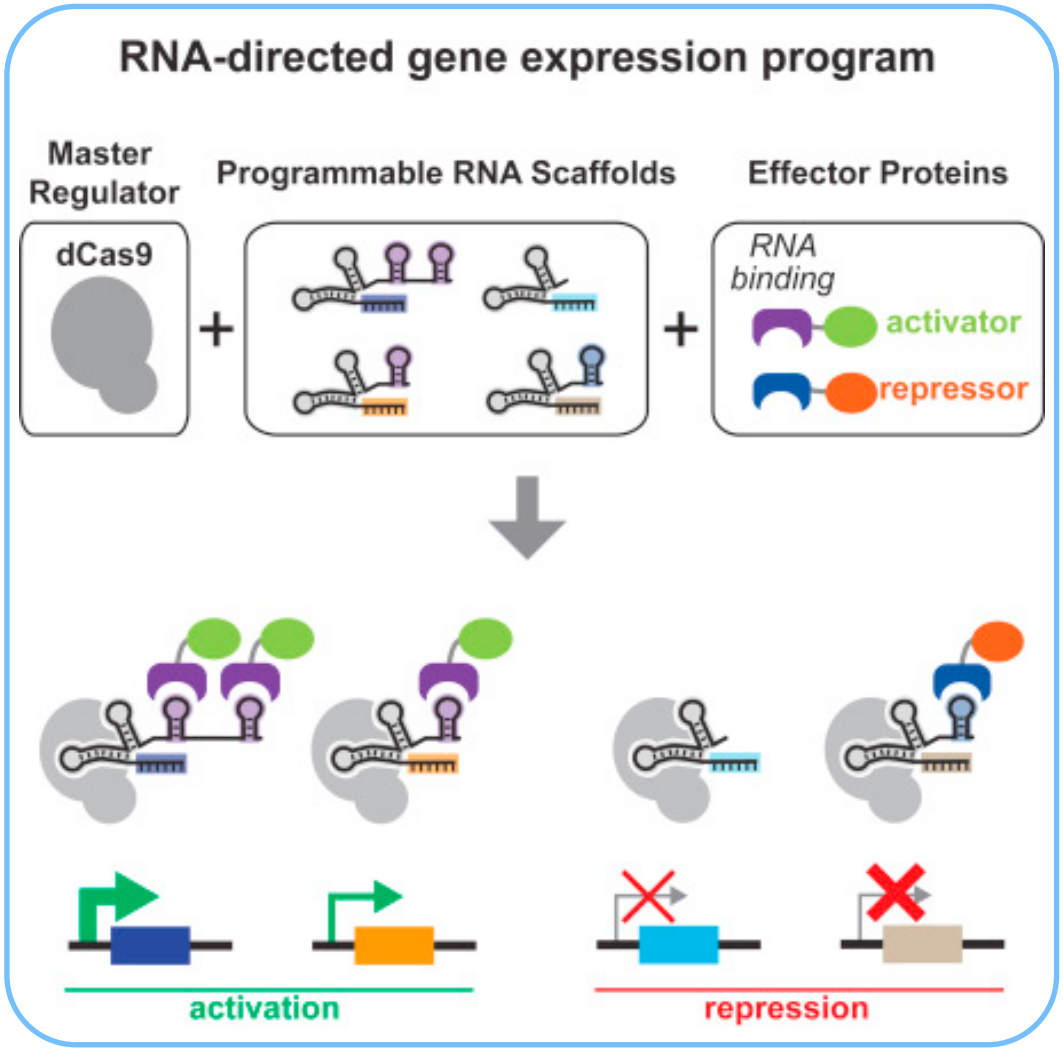
Approach: sgRNA encodes target & function



- When function is encoded in dCas9 or dCas9-effector fusion proteins, the sgRNA recruits the same function to every target site.

- **Convert CRISPR sgRNA into a scaffold RNA (scRNA)** by extending the sequence with modular RNA domains that recruit RNA-binding proteins.
- **scRNA approach allows distinct types of regulation to be executed at individual target loci, thus allowing simultaneous activation and repression**.

Modification of CRISPR/dCas9 technology for locus-specific transcriptional programming

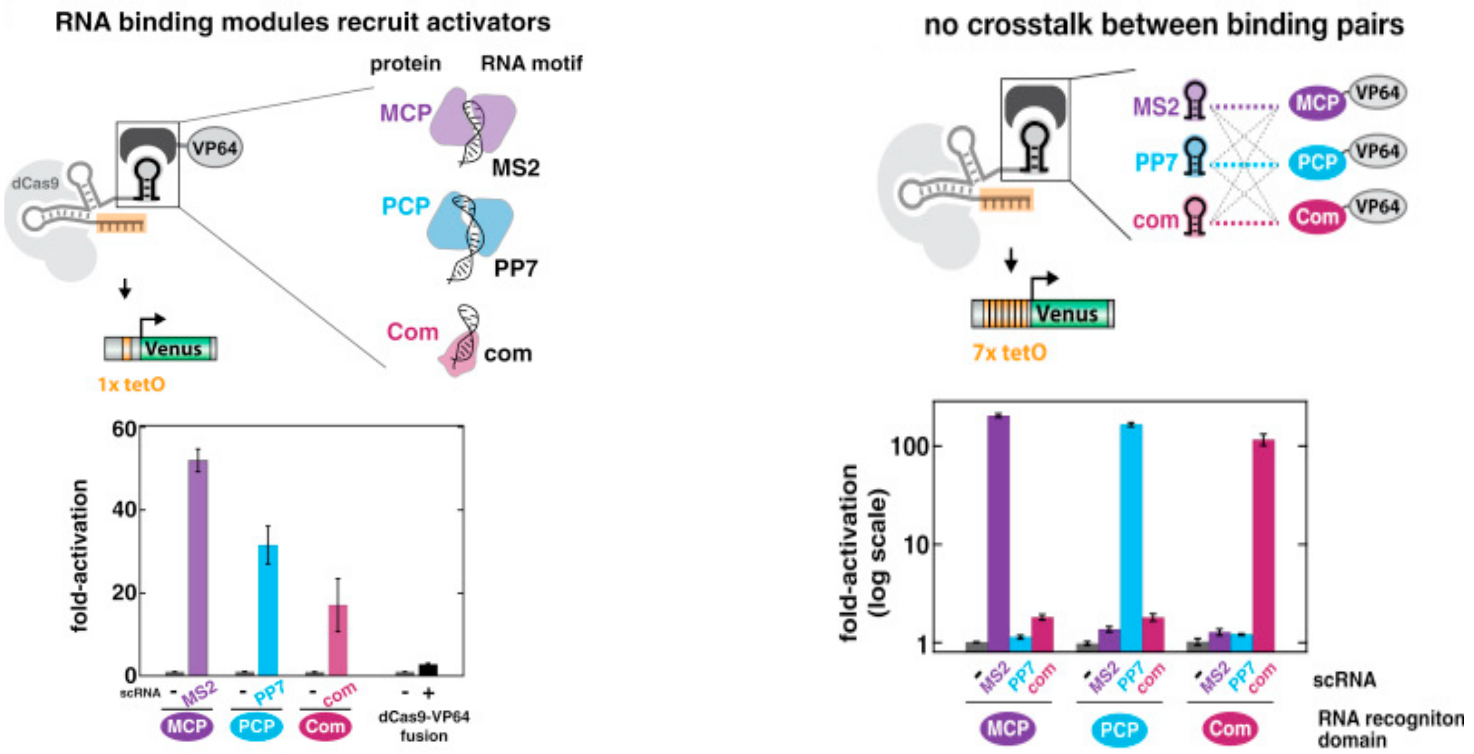


1. **Modify sgRNA**:
   - Introduced single or mutliple RNA hairpin domains to the 3' end of the sgRNA, connected by a two-base linker, which encode specific RNA recruitment modules.
   - RNA recruitment modules: used the well-characterized viral RNA sequences MS2, PP7, and com, which are recognized by the MCP, PCP, and Com RNA-binding proteins, respectively.

2. **Fuse functional transcriptional domains (activators or repressors) to RNA-binding proteins**.

CRISPR scRNAs specifically and efficiently activate gene expression **in yeast**

- Yeast strain with tet promoter driving a reporter gene.
- Transformed to express dCas9, scRNAs targeting the tet operator and corresponding VP64-fusion protein(s).

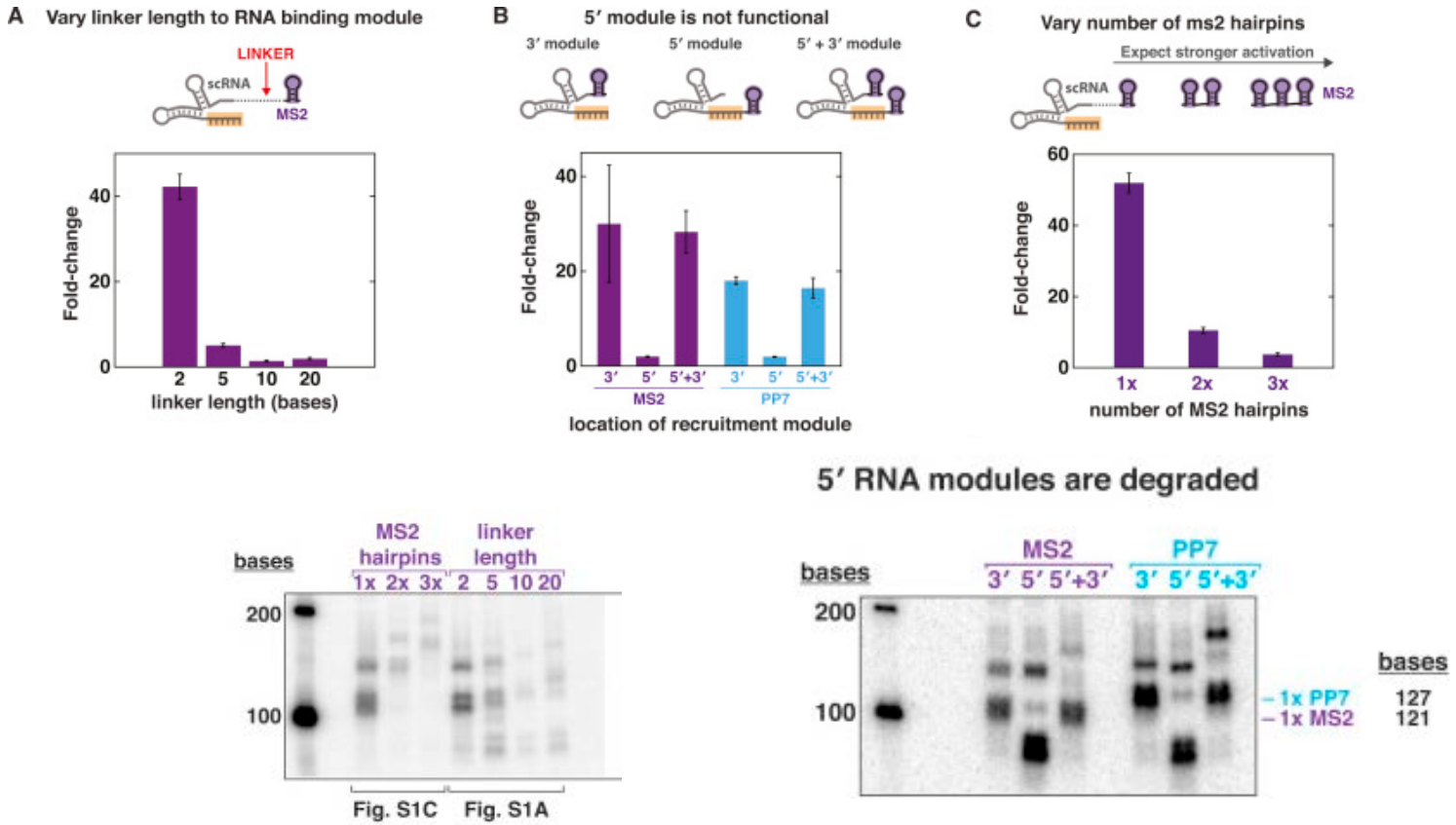


→ Gene activation mediated by scRNA recruitment of VP64 was substantially greater than that for the direct dCas9-VP64 fusion protein.

→ No significant crosstalk detected between mismatched pairs of scRNA sequences and fusion-binding proteins.

→ Potential for simultaneous, independent regulation of multiple target genes.

Optimizations of CRISPR scRNAs

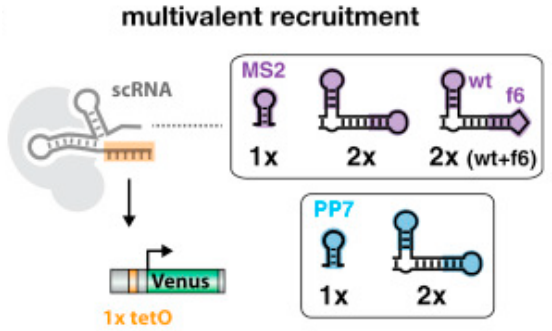


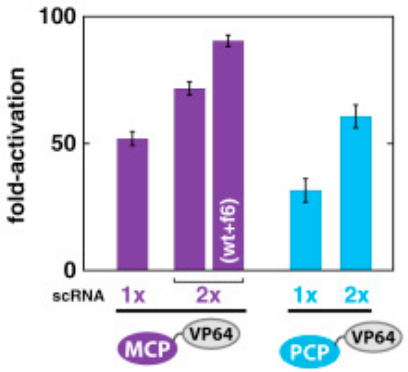
- Increasing length of linker resulted in weaker reporter gene expression.
- scRNAs with recruitment sequences attached to 5’ end of sgRNA gave no significant activation.
- Multivalent recruitment to scRNAs (i.e. increasing the number of RNA hairpins in scRNA) decreased reporter gene expression
- Northern blot analyses:
  - → **RNA levels correlate with functional activity.**
  - → Increasing linker length or number of MS2 hairpins decreases steady-state RNA levels

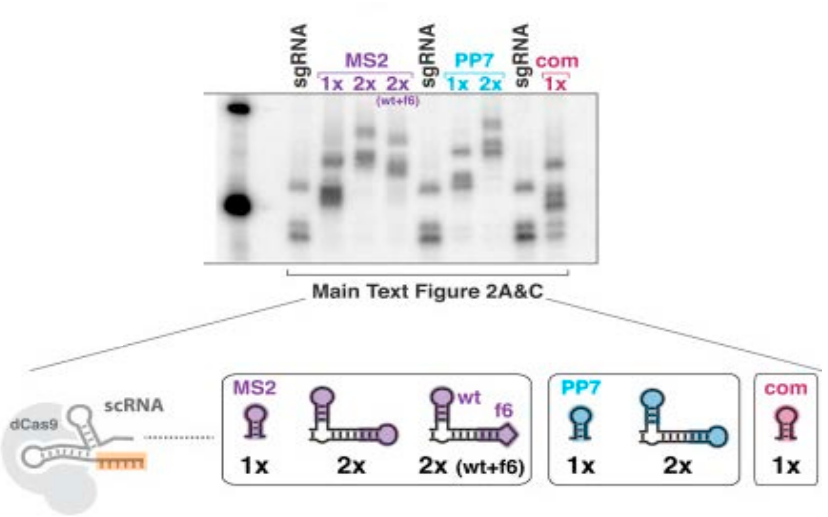
→ RNA stability is limiting for these multi-hairpin scRNAs

Multivalent recruitment to scRNAs

- Constructed alternative scRNAs using double-stranded linkers between the two repeats of the recruitment hairpins to **enforce stable, local hairpin formation**.
- Also constructed a scRNA construct with a mixed hairpin construct containing two different recruitment motifs (*MS2 & f6*) specific for the MCP-VP64 effector protein.



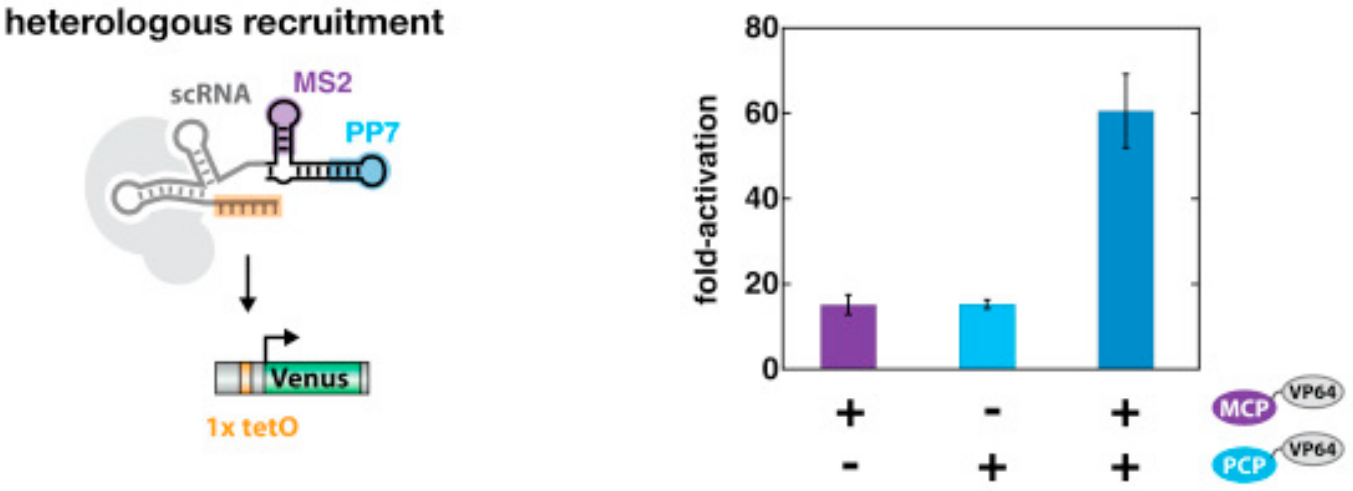




→ Alternative designs produced stronger reporter gene activation for both MS2 and PP7 modules relative to the analogous single-hairpin scRNAs.

→ Northern blot analysis: 2x constructs with double-stranded linkers have steady-state RNA levels comparable to single-hairpin scRNA and unmodified sgRNA constructs.

Heterologous scRNA



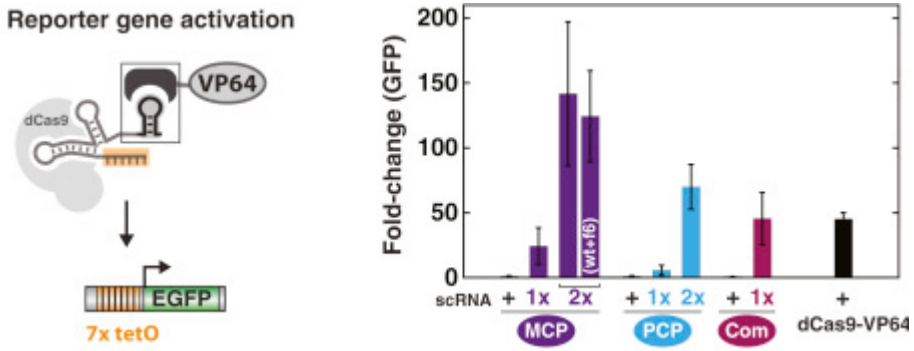
- A mixed MS2-PP7 scRNA using the 2x double-stranded linker architecture recruits both MCP-VP64 and PCP-VP64 fusion proteins → increased reporter gene expression.

→ Distinct RNA-binding proteins can be recruited to the same target site

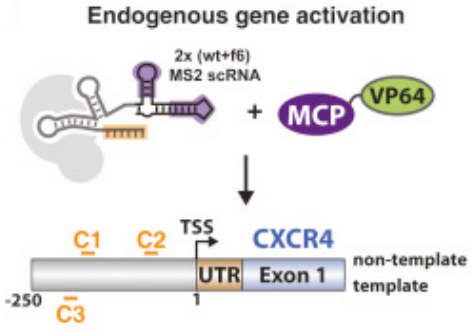
→ Provides an effective approach to combinatorially recruit multiple effectors for the logical control of target genes.

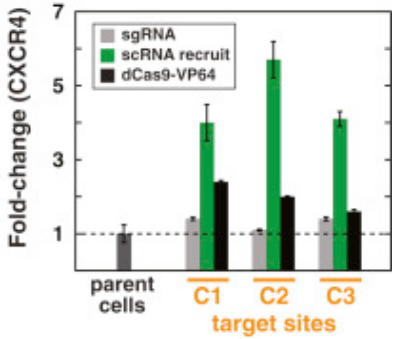
CRISPR scRNAs can mediate activation and repression of reporter and endogenous genes **in human cells**

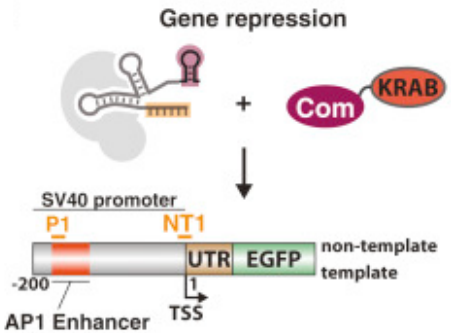
CRISPR scRNAs tested in HEK293T cell lines encoding a tet-driven reporter and stably expressing dCas9:

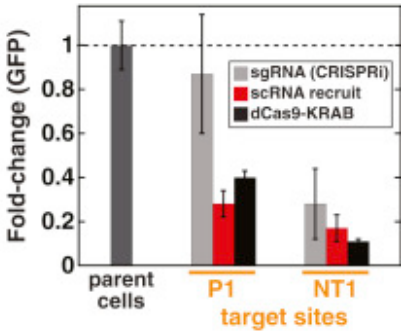


→ Expression of an scRNA (targeting tet) with the corresponding VP64 fusion protein effector produced substantial activation of the GFP-reporter gene for all three RNA-binding modules.







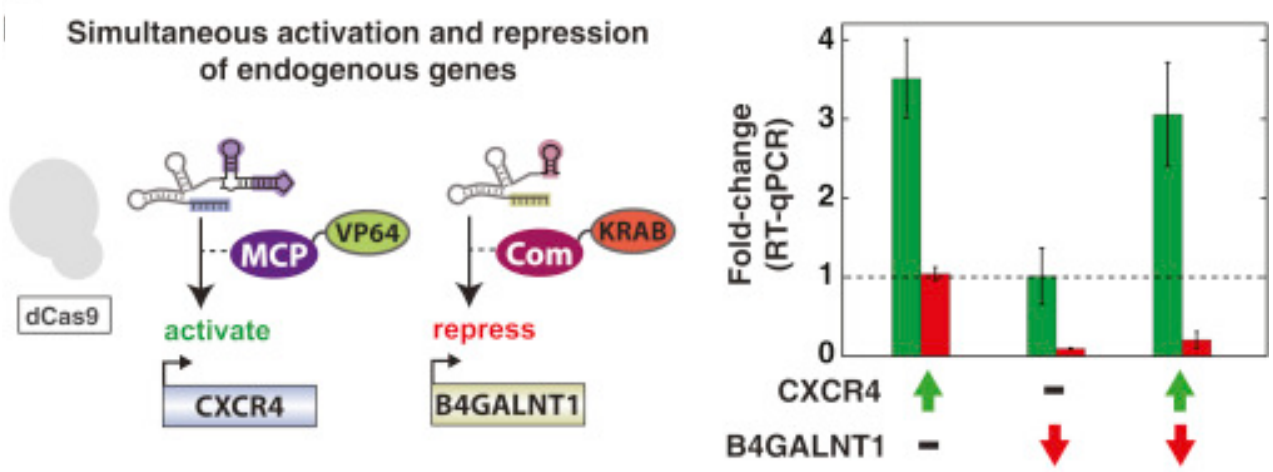


→ scRNAs recruit MCP-VP64 and Com-KRAB to activate and repress expression of endogenous CXCR4, respectively.

→ Comparatively weaker activation/repression observed in cells with dCas9-VP64/KRAB and unmodified sgRNA.

Simultaneous ON/OFF Gene Regulation in Human Cells

- Simultaneous ON/OFF gene regulatory switches mediated by orthogonal RNA-binding proteins fused to transcriptional activators (VP64) or repressors (KRAB).
- Targeted endogenous CXCR4 for activation with MCP-VP64 while simultaneously targeting an additional endogenous B4GALNT1 gene for repression with COM-KRAB in HEK293T cells.



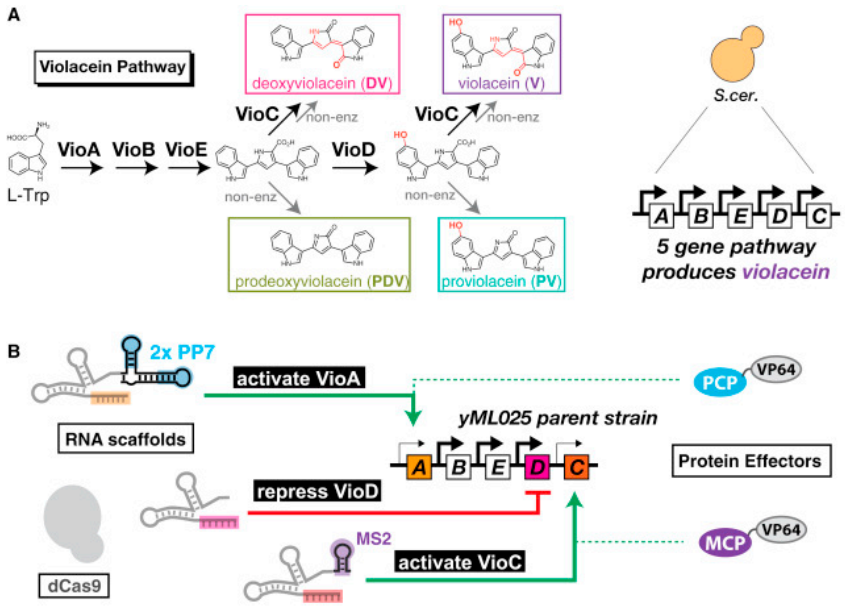
→ Observe **simultaneous activation of CXCR4 and repression of B4GALNT1** measured by RT-qPCR, at similar levels to when single genes were targeted.

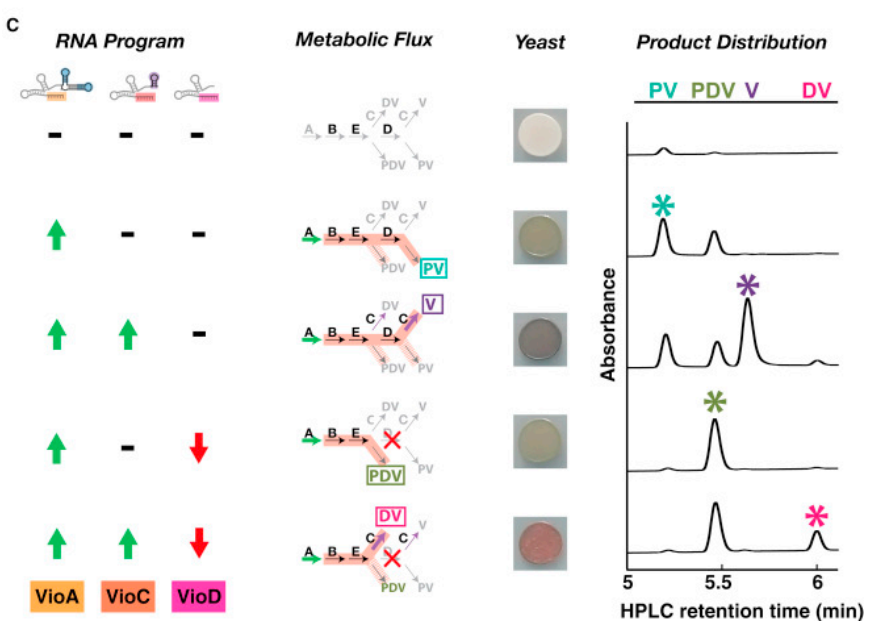
→ Activation and repression are mediated by a single scRNA for each target gene and requires a single pool of dCas9.

→ **Platform could, in principle, be used for largescale screening of pairwise combinations of genes that yield a target phenotype when one gene is activated and the other is repressed.**

scRNA multigene ON/OFF transcriptional programs to redirect metabolic pathway output in yeast

- Violacein pathway 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.
- 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.
  - 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 → **5 distinct output states**





→ The scRNA/dCas9 platform flexibly and efficiently generates each of the multigene transcriptional states necessary to yield all possible metabolic outputs of the violacein pathway.

Summary

- Expanded the CRISPR toolkit to link functional effector domains to specific genomic targets.
  - Modular scaffold RNA: Extended sgRNA to encode both target sequence and the particular regulatory function to be executed at that site, thus creating a modular scaffold RNA.

- Expression of multiple RNA scaffolds simultaneously permits independent, programmable control of multiple genes in parallel.
- Provides a straightforward method to implement simultaneous multigene ON/OFF regulatory switching programs.

NATURE METHODS | ARTICLE

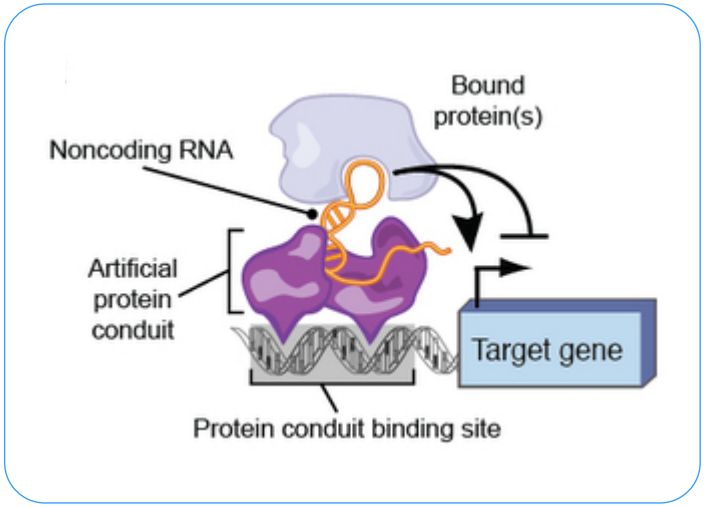
Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display

David M Shechner, Ezgi Hacisuleyman, Scott T Younger & John L Rinn

Affiliations | **Contributions** | **Corresponding author**

Nature Methods **12**, 664–670 (2015) | doi:10.1038/nmeth.3433

Received 03 November 2014 | Accepted 08 May 2015 | Published online 01 June 2015

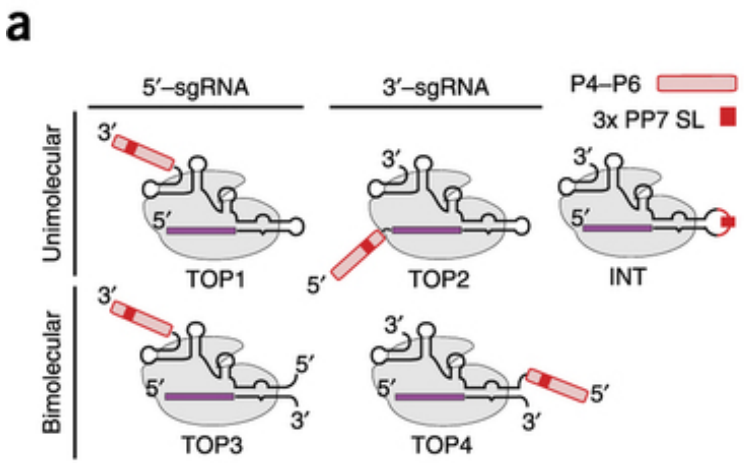


- Similar approach to Zalatan et al. (i.e. also extended sgRNA to encode function), but taken further.
- CRISPR-Display: a targeted localization method that uses dCas9 to **deploy large RNA cargos to DNA loci**.

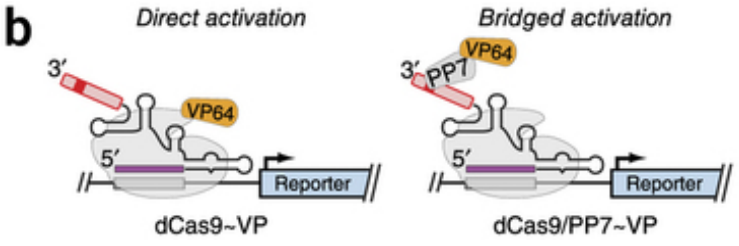
Adapting CRISPR-dCas9 as an RNA display device (CRISP-Disp)

Where can insertions within the dCas9-bound RNAs be tolerated and how large they can be?

- Created 5 model cofactor RNAs (*TOP1–TOP4 and INT*) in which structured, 81- to 250-nt **accessory domains were either inserted within the sgRNA or appended to the termini**.
- The largest of these constructs (357nt) adds **3-fold more sequence than the longest modified sgRNA previously reported**.
- Each accessory domain carried a cassette of stem-loops recognized by the PP7 phage coat.
- GLuc reporter HEK293FT cells transiently transfected or transduced with the new sgRNA constructs, dCas9 and PP7-VP or dCas9-VP plasmids



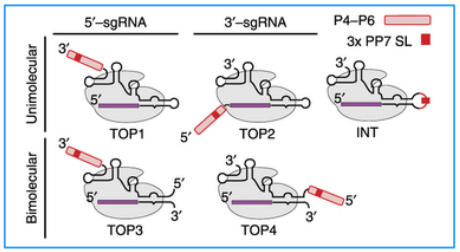
CRISPR transcription activation assay:

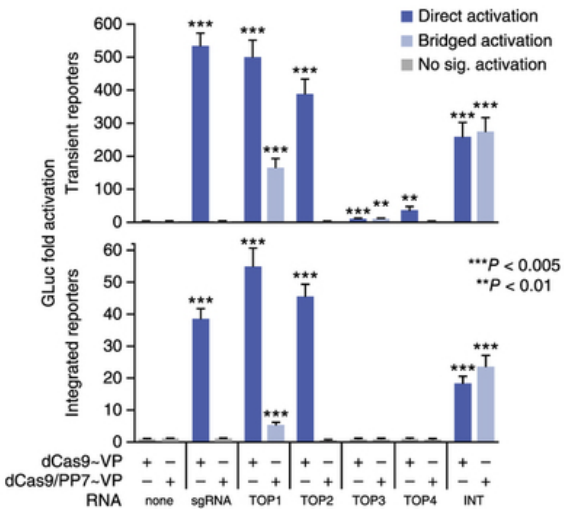


*Reporter gene activation indicates that the **sgRNA variant competently binds and targets dCas9**.*

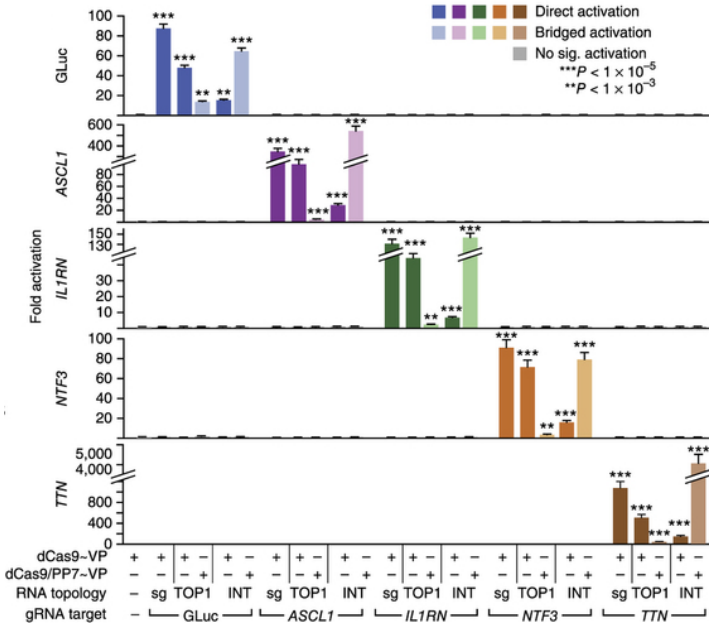
*Activation indicates **accessory domain remains intact** in the mature dCas9 complexes.*

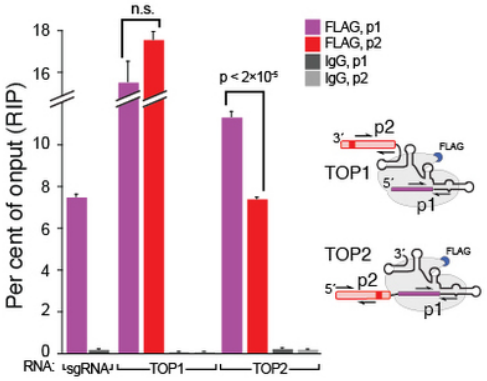
CRISP-Disp enables deployment of large RNA domains to genomic loci





Targeting CRISP-Disp to endogenous loci in HEK293FT cells:



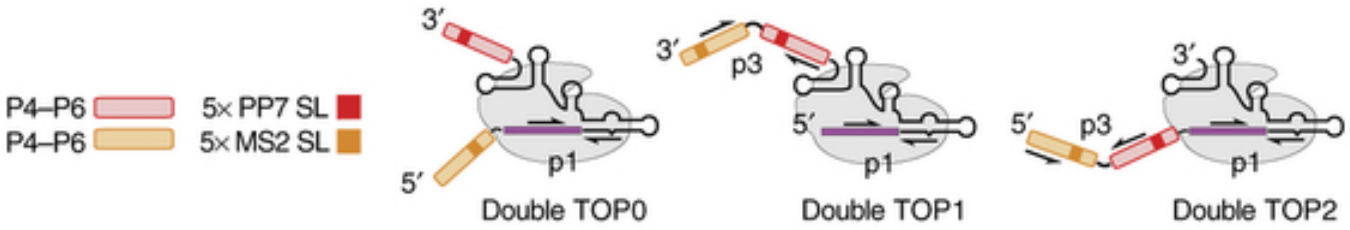


→ Observed varying degrees of direct activation with all five topologies using transient reporters (modified sgRNAs remained functional)

→ Bridged activation only observed with TOP1, TOP3 and INT → other constructs do not retain functional accessory domains in mature dCas9 complexes (*RIP-qRT-PCR: TOP2 accessory domain partially degraded*)

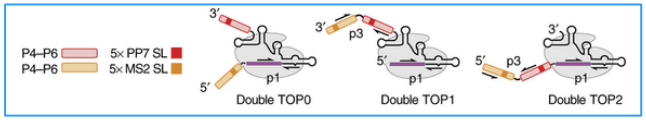
CRISPR-Display supports artificial and natural lncRNAs

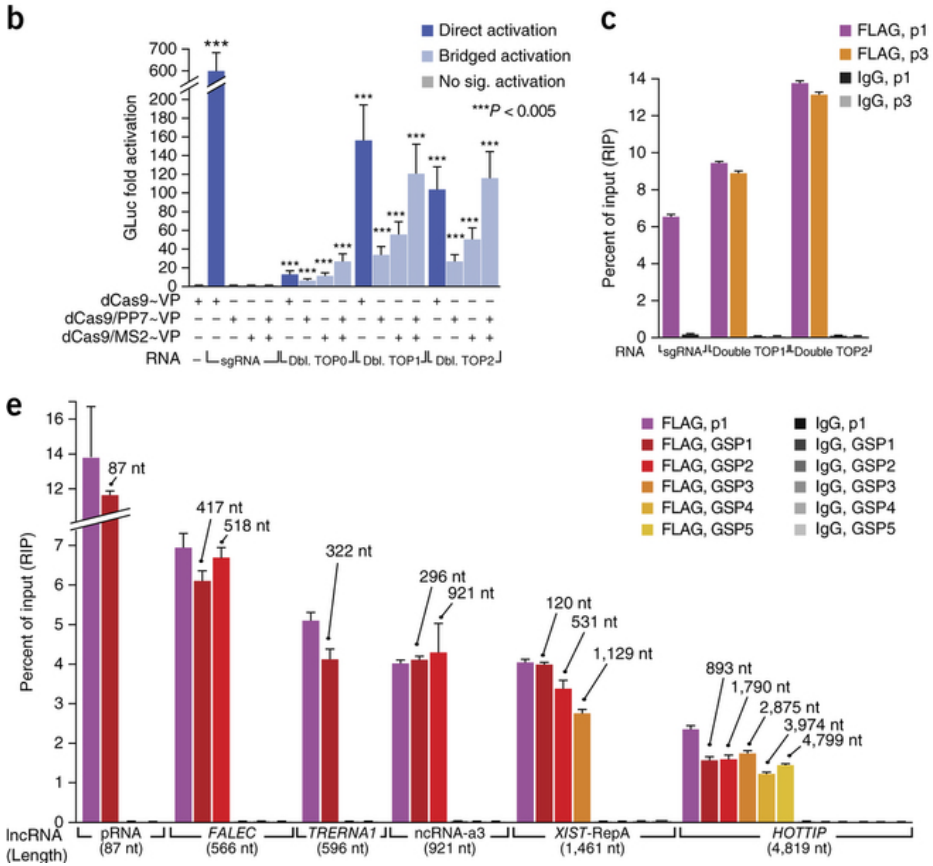
- Engineer CRISP-Disp for **use with even longer ncRNAs** → required replacing the conventional RNA polymerase III promoter used previously with an RNA polymerase II promoter and terminator.
- Generated a series of '**artificial lncRNA**' scaffolds with ~650-nt accessory domains:
  - *Expanded TOP1 and TOP2 constructs with a second P4–P6 domain, bearing stem-loops recognized by the MS2 phage coat protein*



- Also generated a series of sgRNAs modified with different **natural human lncRNA** domains appended

CRISPR-Display supports artificial and natural lncRNAs





B)

- All 3 constructs induced measurable direct & bridged activation.
- Activity increased upon co-expression with PP7~VP, MS2~VP or both activator proteins, indicating that the RNA constructs **retained both P4–P6 domains in mature dCas9 complexes**.

C)

- Integrity of double TOP1&TOP2: essentially stoichiometric yields of sgRNA core and double accessory domain.

E)

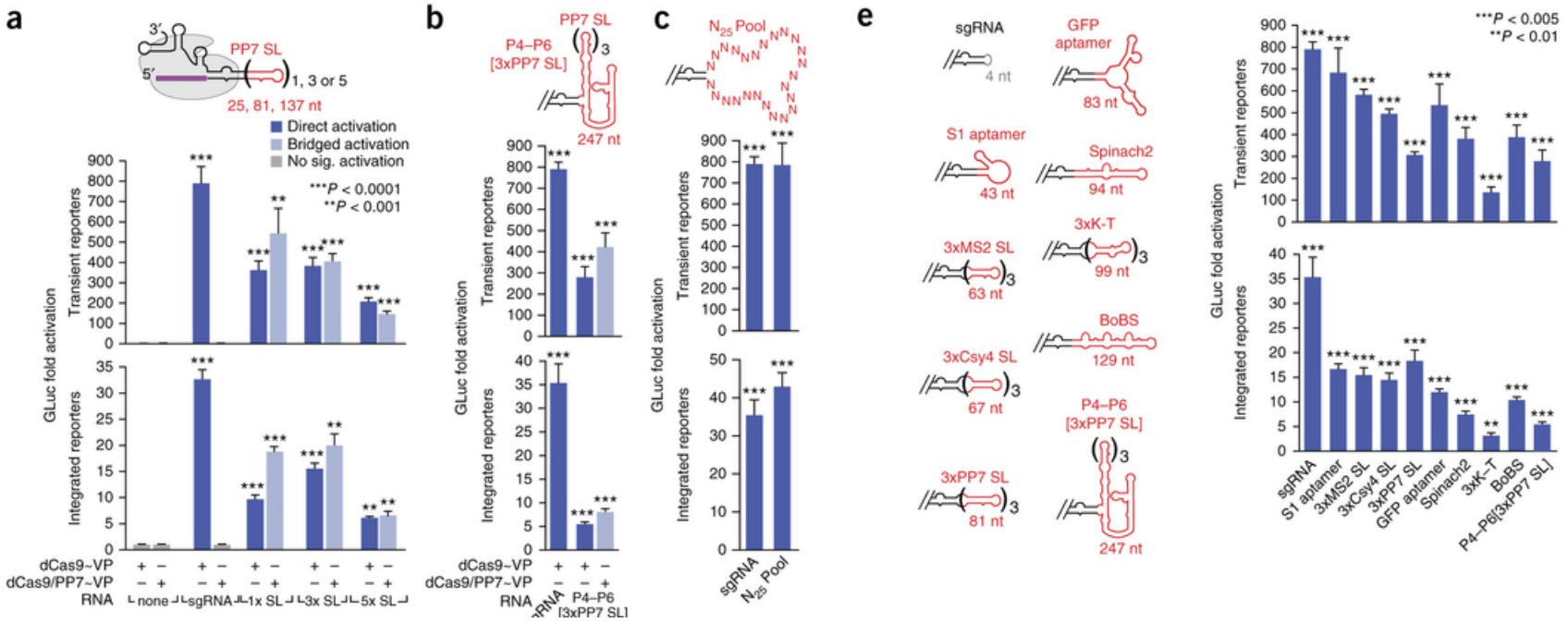
- **LncRNA accessory domains remain intact in CRISP-Disp complexes**.
- Overall dCas9 complexation appeared to decline with increasing sgRNA-lncRNA length.
- Observed nearly quantitative yields of intact lncRNA domains relative to the corresponding sgRNA core.

F)

- Most of the natural lncRNA constructs repressed or activated GLuc expression as expected.

→ Plausibility of large-scale lncRNA functional studies with CRISP-Disp

CRISP-Disp with sgRNA-INT designs



- Very good bridged activation with sgRNA-INT constructs → **sgRNA-engineered loop might provide a universal insertion point for exogenous RNA domains.**
- Generated INT-like constructs bearing:
  - (a) 25–137 nt cassettes of PP7 stem-loops
  - (b) ~250-nt domain equivalent to those of TOP1–4
  - (c) Pool of 1.2x 10^6 unique sgRNAs displaying internal cassette of 25 random nucleotides (INT-N_{25}Pool)
  - (e) functional RNA domains comprising natural protein-binding motifs or artificial aptamers that bind proteins or small molecules

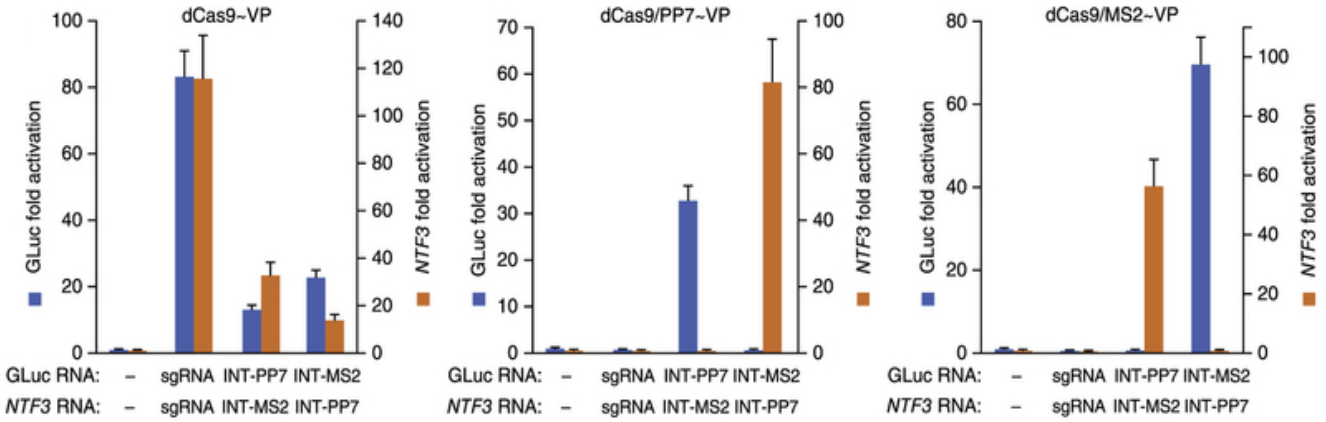
→ Each construct induced activation in all assay formats (*also worked when targeting endogenous loci, not shown*)

→ **dCas9 can easily accommodate even large, structurally discontinuous inserts within the sgRNA core.**

→ **CRISP-Disp is not intrinsically limited by the sequence of an internal insert.**

CRISP-Disp enables multiple concurrent functions

Direct activation *Bridged activation*



→ In presence of RNA-binding proteins fused to effector proteins (PP7-VP64 or MS2-VP64) only the genes targeted by sgRNAs bearing the cognate RNA-binding motif were activated.

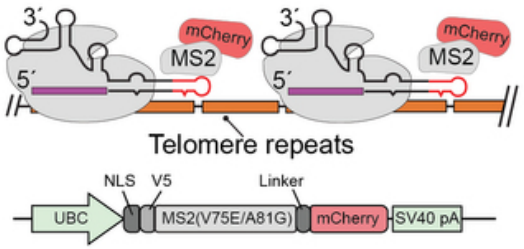
- Co-expressed a set of orthogonal effector-RNA-binding fusion proteins and sgRNAs

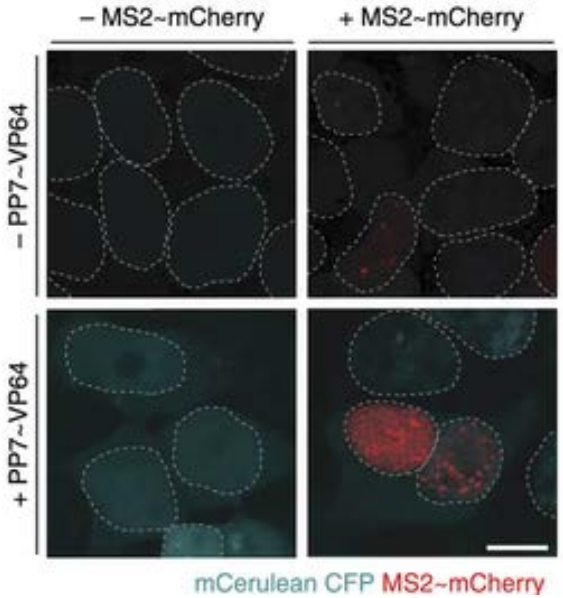
→ No cross-talk

→ **CRISP-Disp enables modular control of gene expression.**

CRISP-Disp enables multiple concurrent functions

- Simultaneously activate one locus and image another, using two CRISP-Disp constructs in GLuc reporter HEK293FT cells:
  - Bridged imaging of targeted loci (*telomeres*) using dCas9, MS2~mCherry and INT-like [3xMS2] constructs
  - Bridged activation at GLuc using a dCas9, MS2~mCherry and INT-like [3xMS2] constructs





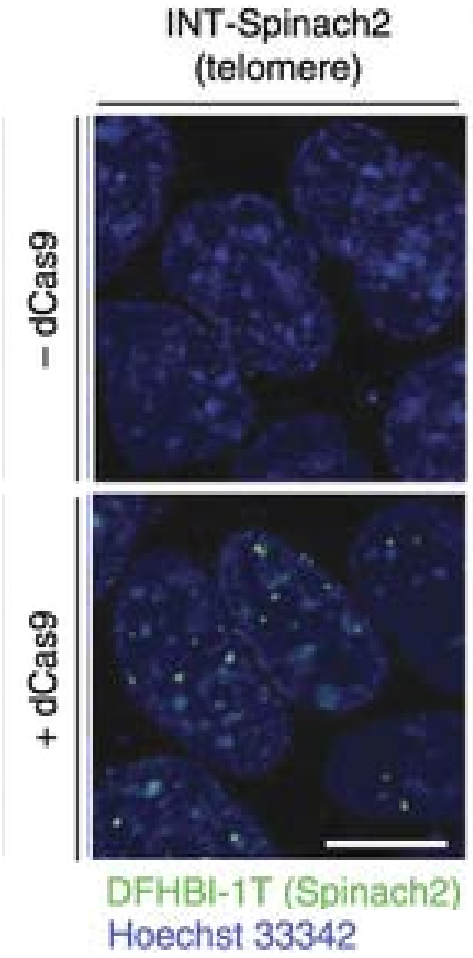
→ **Each function could be individually or concomitantly controlled by expression with PP7~VP64, MS2~mCherry or both**.

CRISP-Disp with autonomous RNA domains

- Use CRISP-Disp to target autonomously functional RNA domains (i.e. ribozymes, aptamers and regulatory devices) to individual loci.
- Targeted the Spinach2 aptamer to telomeres.
  - *Spinach2 fluoresces upon binding the cell-permeable dye DFHBI-1T13*.
- Co-express a Spinach2-appended telomere-targeting sgRNA with dCas9 and treated cells with DFHBI-1T

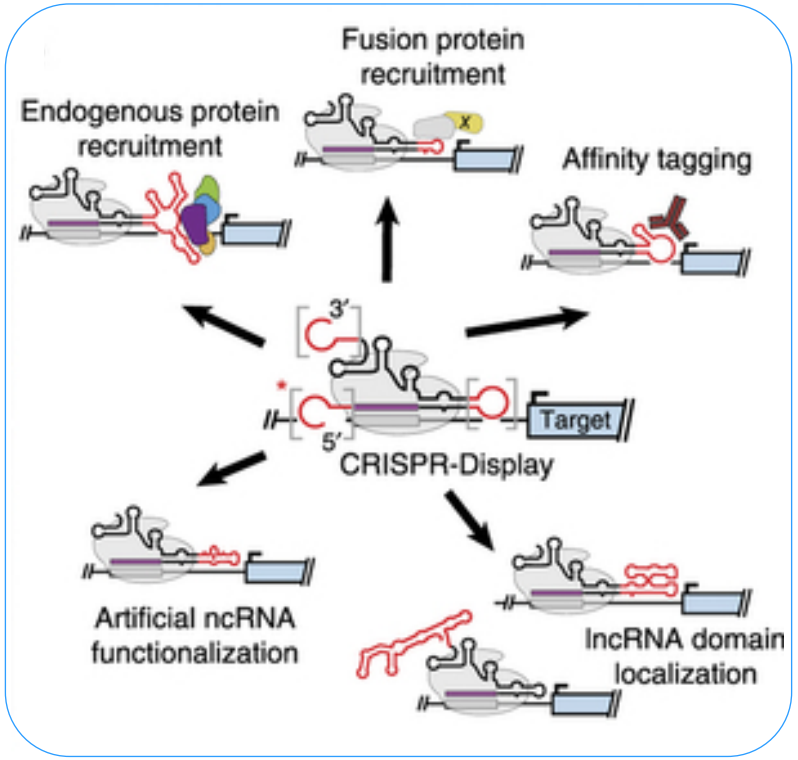
→ Observed numerous nuclear fluorescent foci, only when dCas9 present

→ **Proof-of-principle that artificial RNA domains can provide dCas9 with novel properties.**



Summary CRISPR-Display

- CRISP-Disp limits the function of dCas9 to DNA targeting and ‘outsources’ all other roles to RNA domains → basis for an even wider array of methods
- CRISP-Disp allows independent cargo-specific functions to manifest at several target sites simultaneously using a common toolkit
  - Not limited by RNA cargo length or sequence
- Breadth of distinct functions accessible in a single experiment is limited only by the number of orthogonal RNA domains or RNA–binding-protein pairs available



NATURE METHODS | ARTICLE

Directing cellular information flow via CRISPR signal conductors

Yuchen Liu, Yonghao Zhan, Zhicong Chen, Anbang He, Jianfa Li, Hanwei Wu, Li Liu, Chengle Zhuang, Junhao Lin, Xiaoqiang Guo, Qiaoxia Zhang, Weiren Huang & Zhiming Cai

Affiliations | **Contributions** | **Corresponding authors**

Nature Methods (2016) | doi:10.1038/nmeth.3994

Received 21 March 2016 | Accepted 08 August 2016 | Published online 05 September 2016

- Extending sgRNAs to include modified riboswitches that recognize specific signals → Create **CRISPR–Cas9-based 'signal conductors' that regulate transcription of endogenous genes in response to external or internal signals** of interest.

Aim & approach

- Aim: Engineer synthetic devices for efficient sensing, processing and control of biological signals, to create a **synthetic link allowing the control of endogenous proteins and RNAs**.

- Approach: repurpose CRISPR-dCas9 system for signal control and assignment, by **combining ligand-responsive riboswitches and sgRNAs** thereby creating a new class of engineered **signal conductors.**

Construction of CRISPR–Cas9-based signal conductors

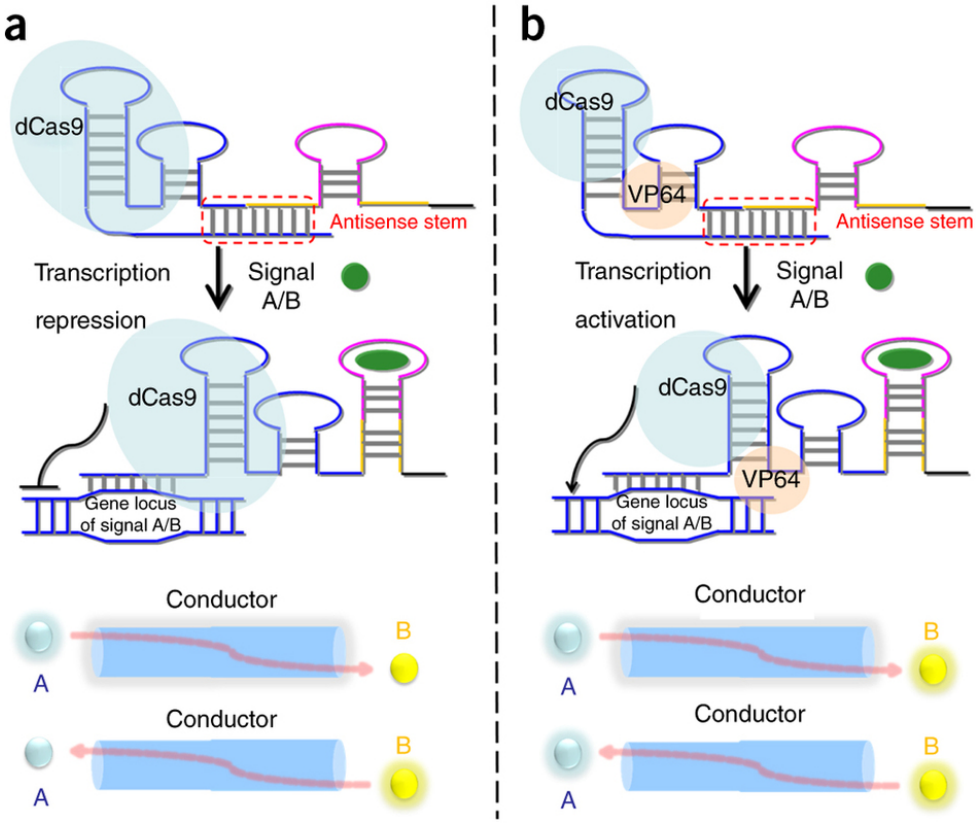
Reengineered the sgRNA :

- Traditional sgRNA modules for DNA substrate targeting and to recruit dCas9 or dCas9-VP64
- Addition of a riboswitch module (signal-responsive aptamer) to recognize a specific signal
- Aptamer stem modified to base pair with guide region of sgRNA

Guide region of sgRNA is paired within the designed antisense stem:
*→ sgRNA is in the **'off' state**.*

Specific signal A (B) binds aptamer → strand displacement

Guide region of the sgRNA can bind to its target DNA:
*→ Conformation of sgRNA is switched to the **'on' state***
→ Modulation of target gene expression through dCas9 or dCas9–VP64 fusion protein

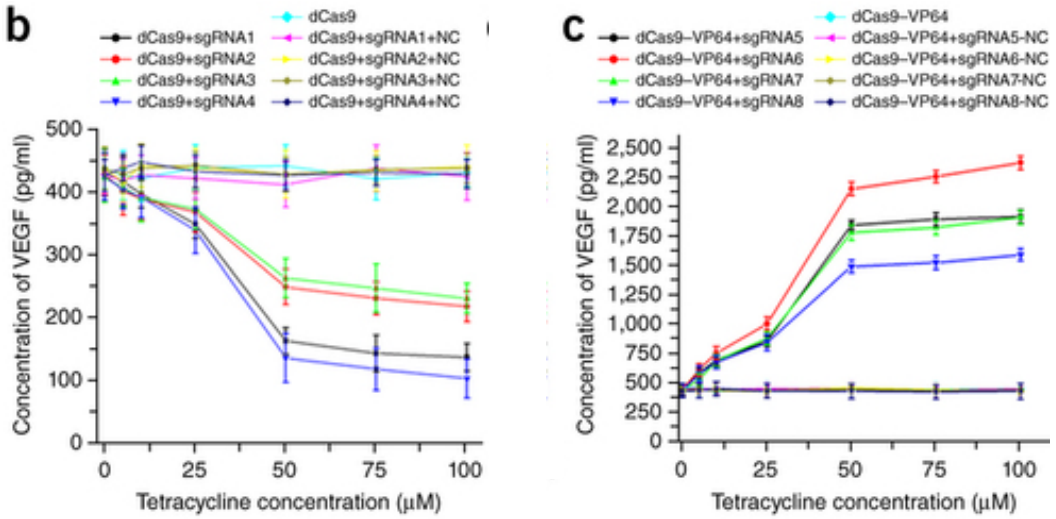


→ **Converts the traditional sgRNA into a scaffold RNA that couples the detection of one input signal with the production of another output signal.**

→ **Thus linking signal concentrations to gene-expression events.**

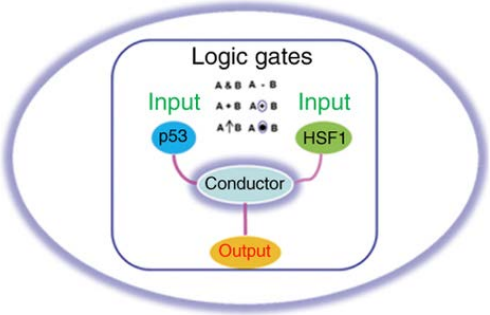
Control of endogenous gene expression

- Generated 4 sgRNAs targeting VEGF & incorporating a tetracycline aptamer at the 3′ end
- Transfected HEK293T cells encoding VEGF protein with sgRNAs and dCas9 or dCas9-VP64:

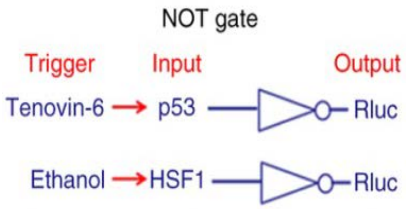


→ Dose-dependent repression (b) or activation (c) of VEGF expression upon treatment with tetracycline

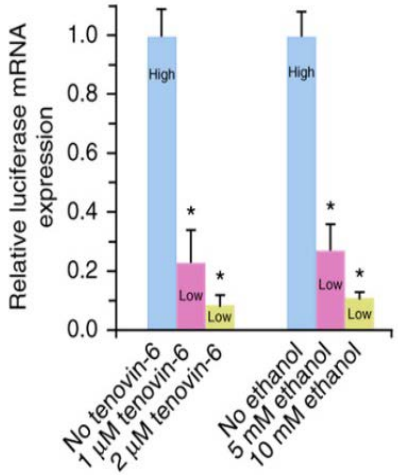
Construction of Boolean logic gates



- Integrated modified aptamers recognizing endogenous p53 or HSF-1 signals into the 3′ end of each sgRNA targeting RLuc
  - Gates use p53 and HSF1 as the input signals and output is a luciferase signal.
  - Tenovin-6 and ethanol can be used to elevate p53 and HSF1 protein levels, respectively.
- 2 NOT logic gates were constructed using signal conductors in HEK293T cells stably expressing Rluc:



Input p53	Output Rluc	Input HSF1	Output Rluc
0	1	0	1
1	0	1	0



→ *Only in the absence of corresponding trigger did each NOT gate produce high luciferase output, at increasing concentrations of input signal dramatically decreased output.*

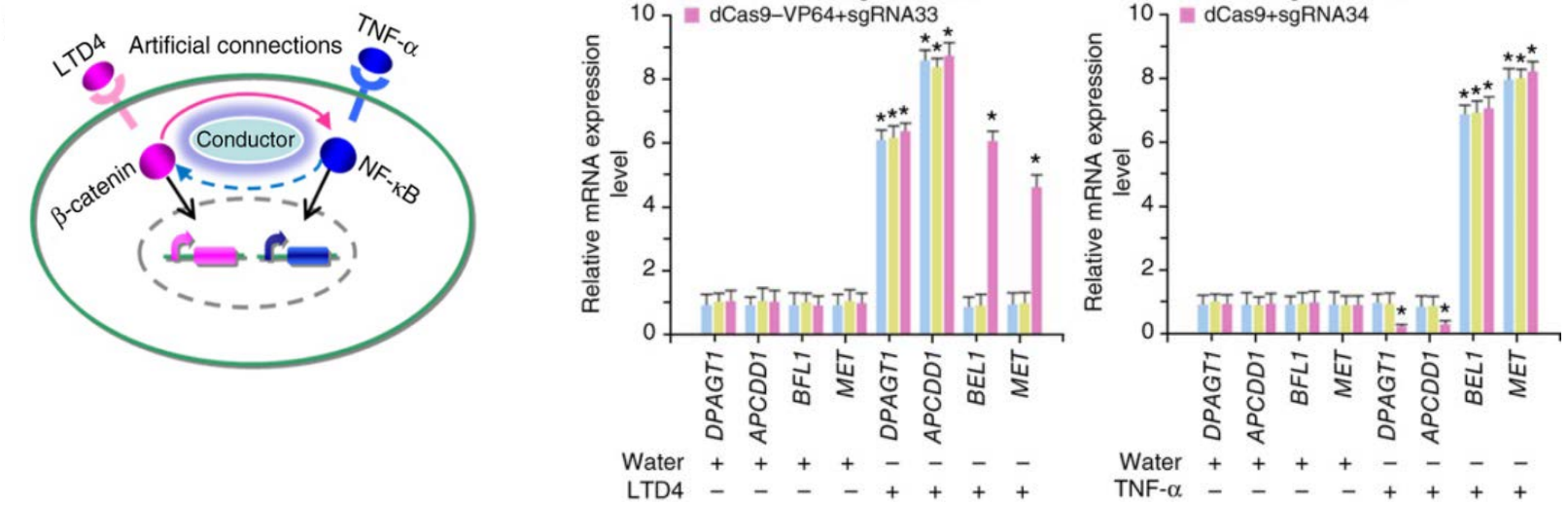
→ **By integrating different combinations of reprogrammed sgRNAs (signal conductors), they could construct various logic gates to combine signals to make a decision in a single cell.**

Note: Also constructed AND, NAND, OR, NOR, XOR, XNOR logic gates (not shown)

Artificial networks connecting two signaling pathways

Can sgRNA based signal conductors redirect the output responses of native signaling pathways by establishing artificial networks?

- Designed two artificial connections between β-catenin and NF-κB, using two signal conductors:
  - *sgRNA targeting NF-κB encoding β-catenin aptamer*)
  - *sgRNA targeting β-catenin encoding NF-κB aptamer*
- Expressed signal conductors and dCas9 or dCas9–VP64 in human bladder cancer T24 cells (*express NF-κB and β-catenin proteins highly*):



LTD4 treatment:

→ mRNA expression levels of two genes downstream of β-catenin (*DPAGT1* and *APCDD1*) were elevated.

→ Compared with dCas9–VP64 or a control device, the **signal conductor increased the mRNA levels of genes downstream of NF-κB** (*BEL1* and *MET*)

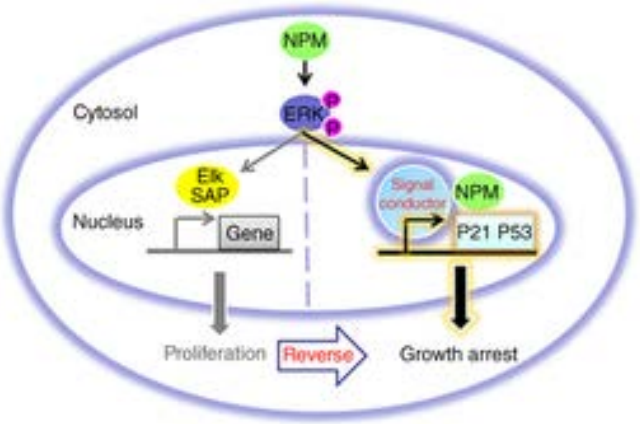
TNF-α treatment:

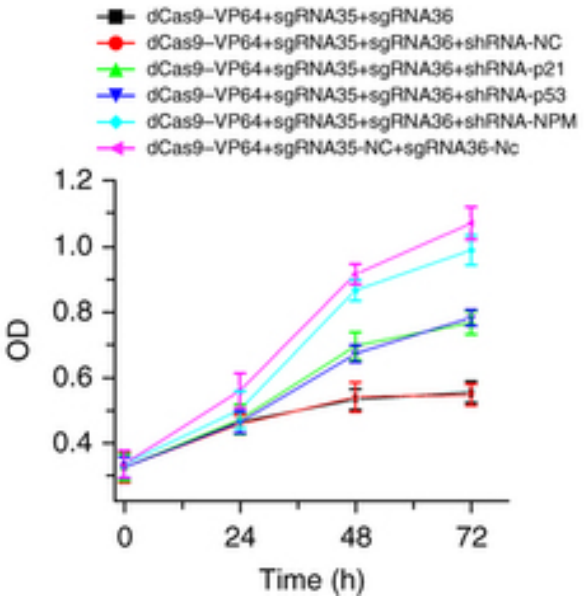
→ Activated expression of *BEL1* and *MET* in cells expressing the signal conductor or the controls.

→ **Only cells expressing the signal conductor displayed decreased expressions of *DPAGT1* and *APCDD1***

Redirecting oncogenic signaling to an anti-oncogenic pathway

- Nucleophosmin (NPM) is an oncogenic signal that activates the MAPK pathway and promotes cell proliferation in progression of bladder cancer.
- Constructed a signal conductor that activates transcription of 2 tumor suppressors (*P53* and *P21*) in response to cellular NPM proteins→ **rewired NPM signaling from the proliferation pathway to quiescence in bladder cancer**.

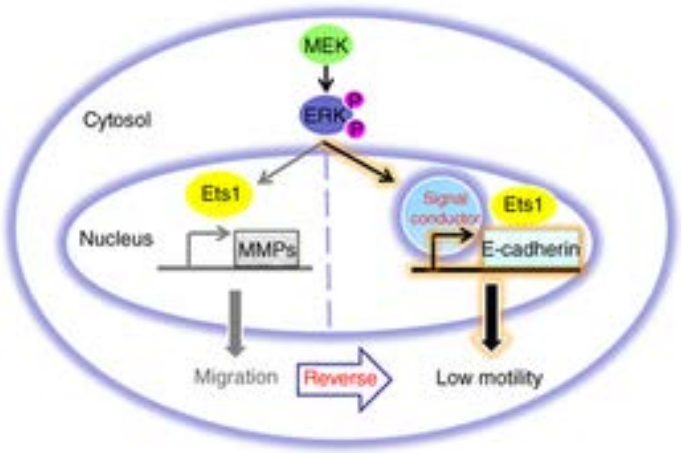


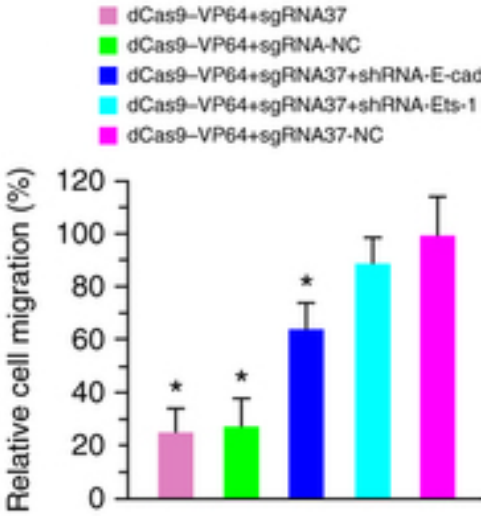


→ Inhibition of cell growth was observed in the group transfected with signal conductor (*knockdown of NPM ablated this effect*).

Redirecting oncogenic signaling to an anti-oncogenic pathway

- Ets-1 is an oncogenic signal that activates the transcription of matrix metalloproteases (MMPs) and promotes cell migration in bladder cancer cells.
- Constructed a signal conductor that activates the transcription of a tumor migration suppressor (E-cadherin) in response to Ets-1 → **rewired Ets-1 signaling from the migration pathway to quiescence in bladder cancer**



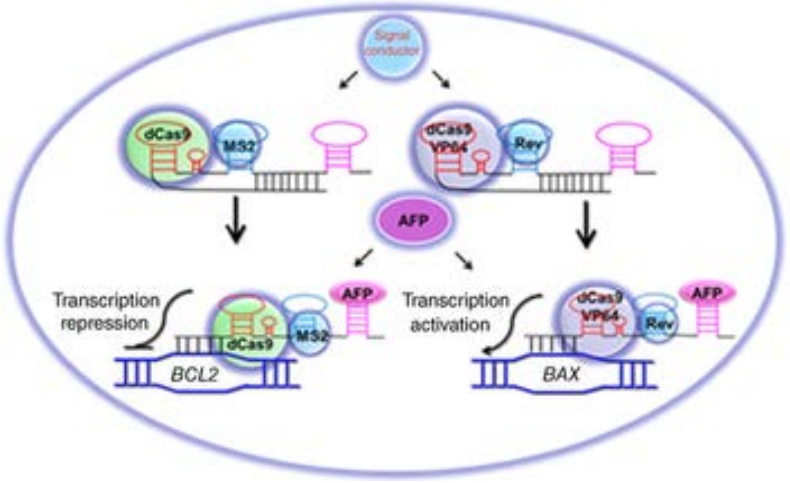


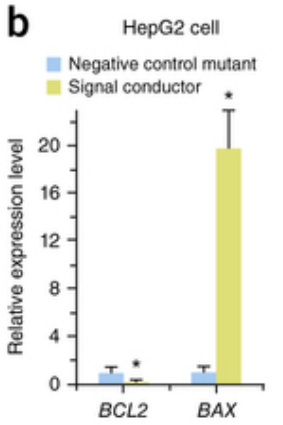
→ Inhibition of cell migration was observed in the group transfected with signal conductor (*knockdown of Ets-1 ablated this effect*).

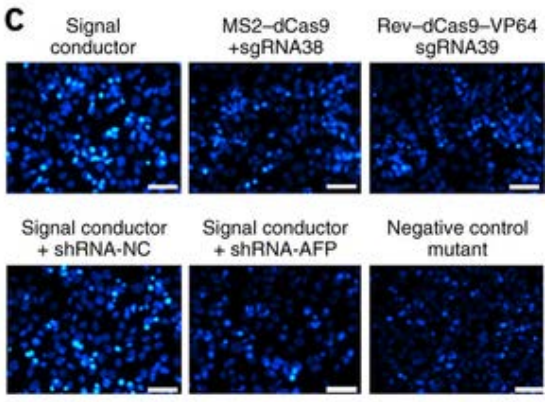
Simultaneous activation and repression of genes

Redirect oncogenic signal by simultaneous activation and repression of genes:

- Bcl2/Bax protein ratio is closely associated with the sensitivity of cells to apoptosis.
- Alpha-fetoprotein (AFP) can be used as an oncogenic signal.
- Constructed reprogrammed sgRNAs to repress *BCL2* with MS2–dCas9 and to activate *BAX* with Rev–dCas9–VP64 in human liver cancer HepG2 cells (*possess a high level of AFP)*:







<u>AFP-treated cells expressing signal conductor</u>:

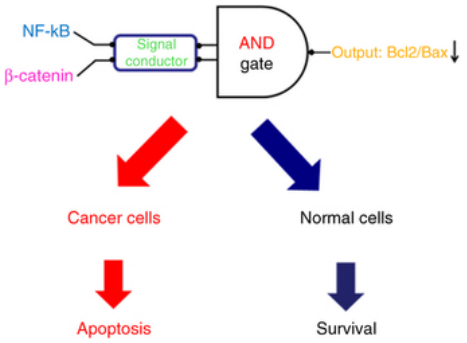
→ (b) Showed a decrease in BCL2 expression and increase in BAX expression

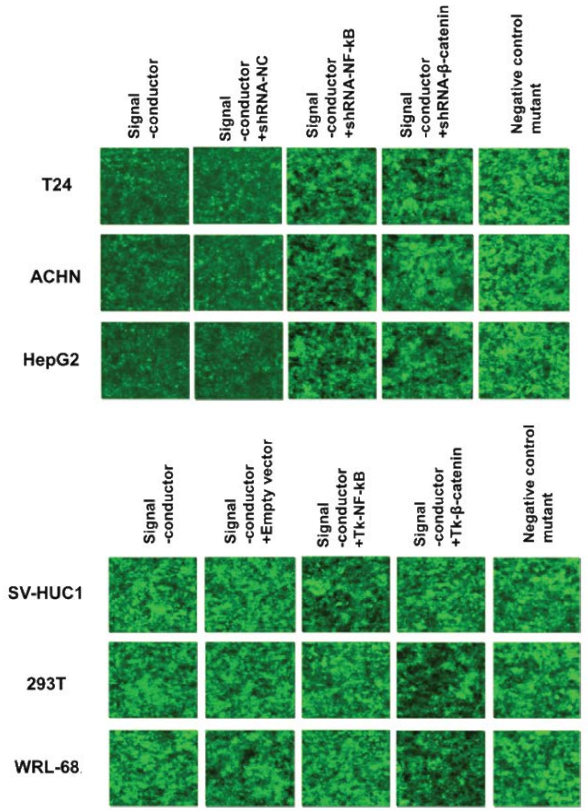
→ (c) Exhibited stronger blue fluorescence, revealing typical apoptotic characteristics (*knockdown of AFP by shRNA ablated these effects*).

→ The control devices (MS2–dCas9+sgRNA38 or Rev–dCas9–VP64+sgRNA39), which activated BAX expression or repressed BCL2 expression alone, didn't increase the apoptotic rate.

Synthetic AND gate circuit selectively & efficiently influences cancer cells

- Constructed a two-input AND gate that used two oncogenic signals as the inputs and had a high output only if both inputs were high, to **test if synthetic AND gate circuit could selectively & efficiently influence cancer cells**.
- Programmed sgRNA targeting NF-κB was used to suppress *BCL2* & programmed sgRNA targeting β-catenin was used to activate *BAX* expression.
- Signal conductor expressed in different cells lines engineered to stably express GFP:
  - Cancer cell lines (T24, HepG2 and ACHN) express high levels of NF-κB and β-catenin
  - Normal cell lines (SV-HUC1, HEK-293T and WRL-68), don't express NF-κB and β-catenin



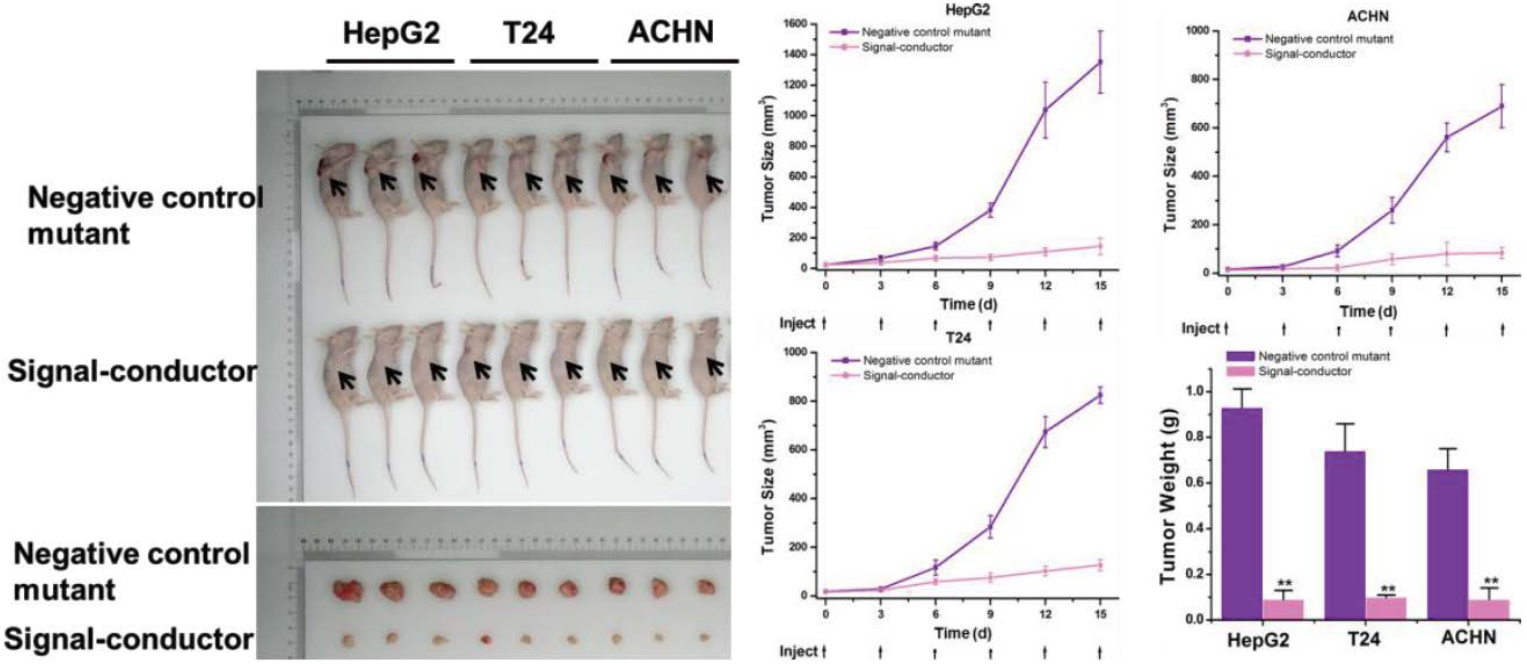


→ Signal conductor was able to reduce GFP intensity (triggered apoptosis) only in the cancer cells (*knockdown of NF-κB or Bcl2 partly reversed effect*)

Synthetic AND gate circuit reduces *in vivo* tumor development

Anticancer effect of the synthetic NF-κB/β-catenin AND gate circuit on *in vivo* tumor development:

- Cancer cells were subcutaneously injected into nude mice and tumors developed several days after cell inoculation.
- Direct intra-tumoral injection of circuit-expressing plasmids:



→ The sizes of tumors treated with the AND gate circuit were dramatically smaller than those treated with the negative control mutant.

Summary

- Combined truncated sgRNAs with various aptamers to create signal conductors.
- Signal conductors allow highly responsive, dose-dependent and dynamic control of gene expression.
- First description of the use of riboswitch-based CRISPR to integrate information about different biomolecules and used in a preclinical tumor study.

Conclusion

- CRISPR-based technologies can be extended and modulated to generate complex gene regulatory circuits.
- General approach: use a deactivated Cas9 and modify sgRNA to recruit effector domains to target loci.

- Applications include basic research, therapeutic purposes and biomolecule production processes.

Photo by SCMP/YouTube/McGovern Institute for Brain Research