Title

Tunable, self-contained gene dosage control via proteolytic cleavage of CRISPR-Cas systems

Author list and affiliations

Noa Katz[1#], Connie An[1#], Yu-Ju Lee[2,3], Josh Tycko[4], Meng Zhang[1], Jeewoo Kang[5], Lacramioara Bintu[6], Michael C Bassik[4], Wei-Hsiang Huang[2,3], Xiaojing J Gao[1,7,8,9*]

[1] Department of Chemical Engineering, Stanford University, Stanford, CA, USA
[2] Department of Neurology & Neurosurgery, Centre for Research in Neuroscience, McGill University, Montréal, QC H3G 1A3, Canada
[3] Brain Repair and Integrative Neuroscience Program, The Research Institute of the McGill University Health Centre, Montréal, QC H3G 1A3, Canada
[4] Department of Genetics, Stanford University, Stanford, CA 94305, USA
[5] Neurosciences Interdepartmental Program, Stanford University, Stanford, CA, 94305, USA
[6] Department of Bioengineering, Stanford University, Stanford, CA, USA
[7] Stanford Bio-X, Stanford University; Stanford, 94305, USA
[8] ChEM-H Chemistry/Biology Interface Training Program, Stanford University, Stanford, CA, USA
[9] Lead contact
*Correspondence: xjgao@stanford.edu
# These authors contributed equally to this work

Summary

Gene therapy holds great therapeutic potential. Yet, controlling cargo expression in single cells is limited due to the variability of delivery methods. We implement an incoherent feedforward loop based on proteolytic cleavage of CRISPR-Cas activation or inhibition systems to reduce gene expression variability against the variability of vector delivery. We demonstrate dosage control for activation and inhibition, post-delivery tuning, and RNA-based delivery, for a genome-integrated marker. We then target the *RAI1* gene, the haploinsufficiency and triplosensitivity of which cause two autism-related syndromes, Smith-Magenis-Syndrome (SMS) and Potocki-Lupski-Syndrome, respectively. We demonstrate dosage control for RAI1 activation in HEK293s, Neuro-2As, and mouse cortical neurons via AAVs and lentiviruses. Finally, we activate the intact *RAI1* copy in SMS patient-derived cells to an estimated two-copy healthy range, avoiding the harmful three-copy regime. Our circuit paves the way for viable therapy in dosage-sensitive disorders, creating precise and tunable gene regulation systems for basic and translational research.

Introduction

For basic research and potential therapies, it is now routine to deliver genetic materials into cells to change the expression level of target genes using vectors such as lentiviral vectors, AAV (adeno-associated virus), and messenger RNA (mRNA)

encapsulated in lipid nanoparticles. However, cargo expression stemming from such delivery methods have remained inherently variable, both between individual cells within an experiment and across different batches. There remains a scarcity of methods to quantitatively maintain the target gene at predefined levels against delivery variability. Such methods would empower large swaths of biomedical applications sensitive to gene dosage. For example, there are many genetic disorders where increased or decreased levels of a gene are both pathological such as in autism spectrum disorders.[1] *RAI1* (retinoic acid-induced 1) is a notable example: its heterozygous loss of function causes Smith–Magenis syndrome (SMS) and its duplication causes Potocki–Lupski syndrome (PTLS), with several opposing phenotypes.[2–7] Another prominent example is the methyl-CpG-binding protein 2 (*MECP2)* gene. Its loss-of-function causes Rett syndrome, while its duplication causes *MECP2* duplication syndrome.[8] Similarly, in basic research, although it is now straightforward to overexpress or knock down target genes and examine the phenotype, it is much more difficult to establish quantitative genotype-phenotype relations, especially when such relations are non-monotonic.[9] We would meet all these challenges if we could bridge the critical gap between the variability of vector delivery and the need for controlling the dosage of the target gene (**Fig. 1A**).

The incoherent feedforward loop (IFFL) is a circuit motif broadly found in natural biological pathways.[10,11] In an IFFL, one common input passes through an activating arm that increases the level/activity of a target gene and a parallel inhibitory arm that decreases it (**Fig. 1B**). These antagonistic arms maintain the output level against input variability. Inspired by nature, synthetic biologists have been engineering IFFLs to buffer against delivery variability. In this context, the level of the vector serves as the input. An ideal engineered IFFL would exhibit the following features. (1) For robust functionality, it would depend minimally on the host cell. (2) It would be broadly compatible with different delivery methods. (3) It would enable both the increase and the decrease of diverse target genes. (4) We should be able to tune it to maintain the target gene at different levels (referred to as “set points” hereafter).

Previous IFFLs relied on an inhibitory factor directly or indirectly targeting a delivered target gene. Over the years various inhibition factors were implemented, such as transcriptional repressors, antagonists, and RNAi.[12–20] While all have achieved impressive dosage control, the most prominent mechanism of repression remains RNAi, which has been proven effective in multiple scenarios, including targeting a neurogenetic disease model *in vivo* while demonstrating some of the features mentioned above.[20] Inspired by RNAi-based IFFLs, we also recognized several limitations. First, because the inhibition factor requires endogenous RNAi machinery while target gene expression requires transcription and translation, final expression levels will depend on cellular context. Consequently, with RNAi machinery saturated at high vector levels, the dosage control behavior is compromised. Second, whereas the set point is tunable through the choice of miRNA and its binding site, it cannot be tuned dynamically post-delivery. Since the optimal dosage and the delivery efficiency likely differ between patients, post-delivery tuning using small molecules might prove crucial for some applications. It is currently nontrivial to build that tuning capability into RNAi-based IFFLs. Third, for such IFFLs to operate as intended, the components constituting both arms must be constantly produced from the vector. This precludes the use of

mRNA vectors. Due to the transient, non-mutagenic nature of mRNA, they are the safer option in certain applications compared to DNA viral vectors,[21] and it would be useful to prepare IFFLs for this emerging delivery modality.

To create IFFLs that meet the above criteria, we reason that proteolytic cleavage is a compelling candidate for the inhibitory arm. We choose viral proteases and their highly specific cleavage sites, where the self-contained cleavage event will not rely on additional host cell machinery other than transcription and translation, which both arms share. Such systems have been extensively engineered and demonstrated for diverse controls,[22,23] offering variants suitable for both *a priori* and post-delivery tuning of set points. Furthermore, such post-translational proteolytic cleavage is compatible with mRNA delivery. As for the activation arm, one could potentially engineer any target protein to be controlled by proteolytic cleavage, but that would require *ad hoc* engineering for each new target to ensure that its function is not impaired by the inclusion of the protease cut site, and that its function is inhibited by protease presence. Furthermore, almost 50% of autosomal human genes predicted to be both haploinsufficient and triplosensitive are too large to fit into an AAV (**Fig. 1C**), arguably the most promising long-term delivery vector for treating such diseases.[24] Therefore, we set out to indirectly regulate the target gene through Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-based mechanisms. This both allows us to overcome the length limitation while targeting various too-large-to-fit genes, and grants us flexibility in regulation strength. Additionally, we can now create dosage-insensitive inhibition via CRISPR-based RNA-interference systems, such as the RNA-targeting Cas13.[25] This allows us to target triplosensitivity disorders, such as PTLS, where dosage control is impossible to achieve with current IFFL-based methods that rely on the delivery of the target gene.

Here, we combined experimental and computational methods for the novel implementation of IFFLs based on proteolytic cleavage and CRISPR activation/repression. Using transient, genome-integrated, and endogenous target genes, we demonstrated the IFFLs' feasibility, tunability, mRNA compatibility, and capability of both increasing and decreasing target gene expression. As a therapeutically pertinent example, we used proteolytic IFFLs to control *RAI1* dosage via various delivery methods and in multiple contexts. We demonstrate that our IFFL decreases the upper-bound gene expression variability when delivered via AAVs to murine cells. Lastly, we adapted our IFFL to target human *RAI1* and demonstrated dosage control in both HEK293 and SMS patient-derived cells when delivered via lentiviruses.

Results

Establishing proteolytic IFFLs using a synthetic reporter

For the IFFL building blocks, we prioritize those that have been well-characterized and extensively engineered: catalytically inactive SpCas9 (dCas9) fused to a transcription activation domain (VP64) and tobacco etch virus (TEV) protease (**Fig. 1D**). In subsequent experiments, we replace VP64 with a more potent activation domain, NZF,[26] whenever we expect the strength of transcription activation to be a limiting factor, and we replace SpCas9 with sadCas9, when we prefer compactness. For

proteolytic inhibition, we inserted a TEV cleavage site between dCas9 and the transcription activation domain.

To qualitatively explore possible outcomes, we developed a computational simulation model. We used ordinary differential equations to describe the reactions, where cleavage between dCas9 and VP64 simply abolishes transcription activation. We used parameter values in experimentally grounded regimes and indeed observed that the presence of the inhibitory TEV protease reduces the range of output levels given the same range of vector levels (**Fig. 1E**). As expected, higher amounts of TEV protease relative to CRISPRa leads to stronger dosage control effects.

To expedite the design-build-test cycles, we adopted a synthetic system for the initial experiments. We created a human embryonic kidney (HEK) 293 cell line with a genomically integrated GFP reporter under the control of seven repetitive tetR-response elements (TRE), which we paired with a TRE-targeting guide RNA. At first, we transiently transfected multiple plasmids encoding the IFFL components, taking advantage of the known high correlation between the uptake of distinct plasmids in lipofection.[27] This multi-plasmid paradigm facilitates component swapping and stoichiometry titration for initial optimizations. To test our system, we conducted two parallel experiments that differ in how we consider the variability in dosage: in the first, we transfected different amounts of our IFFL plasmids (25ng - 400ng) to mimic variable DNA dosage delivery and examined the average output levels (**Fig. 1F**), which we termed “circuit titration”. In the second, we included a co-transfection marker to indicate overall plasmid uptake in individual cells, which we termed “co-transfection method”. This is a widely used technique, and it captures a wider range of vector level variability than circuit titration. Since it is also easier to implement, we included several conditions where we change the amount of TEV plasmid, thus observing its effect on dosage control (**Fig. 1G**). Consistent with the trend predicted by the simulation, both experiments indicated dosage control, with the co-transfection marker method also showing that the presence of TEV protease reduces the range of output levels while creating a distinct dosage-insensitive region (**Fig. 1G**).

The co-transfection marker method revealed a trend of negative correlation between the output and the vector amount towards the high end of co-transfection marker expression (**Fig. 1G**). This was also observed when analyzing the running averages of the circuit titration (Fig. S1A), suggesting that taking the average of all transfected cells (Fig. S1B) would skew the average GFP fluorescence. Thus, to extract the quantities for **Fig. 1F**, we only considered cells that are within the monotonic regime. We hypothesized that the non-monotonic effect may be due to the creation of inactive dCas9 species that can bind to target sites and outcompete functional activators (**Fig. 1H**). To test this hypothesis, we added such a competitive inhibition feature to our simulation, and indeed observed the non-monotonic behavior qualitatively consistent with the experiments (**Fig. 1H**). Because in therapeutic contexts it is often difficult to achieve high vector uptake, this high-vector regime is likely irrelevant there. Nevertheless, it allowed us to demonstrate the utility of computational simulation in testing the intuitive interpretation of experimental outcomes.

Demonstrating unique features of proteolytic IFFLs

Although we achieved set point modulation by varying the amount of the protease (**Fig. 1G**), in the final use cases it would be difficult to keep the stoichiometry constant at the single-cell, and a possible solution would be to vary protease activity instead. Therefore, we tested truncated variants of the TEV protease, and indeed observed set point modulation using transient transfection in HEK293 cells (**Fig. 2A**). While truncated variants "hard code" set point values, we also need a paradigm to externally fine-tune the set point post-delivery. We achieved this by using a split TEV protease with the two halves fused to engineered versions of FRB and FKBP, respectively. The FRB and FKBP domains dimerize in response to a small molecule, rapalog, thus reconstituting protease activity. Since we hypothesized that splitting TEV would weaken the repression arm of the IFFL, we included a second cut site between dCas9 and VP64 to improve TEV sensitivity (Fig. S1C). We demonstrated fine-tuning of the set point by administering different amounts of rapalog (**Fig. 2B**).

To demonstrate the compatibility of proteolytic IFFLs with mRNA delivery, we encode TEV protease and the CRISPR components on separate *in vitro* transcribed mRNAs. Since expression from mRNA is lower than in DNA transfections, we replaced VP64 activation domain with the more potent NZF. We then transfected them into the HEK293 cell line along with the mRNA of a co-transfection marker. Using the same genomic TRE reporter as those used above, we observed mRNA-delivered, IFFL-mediated dosage control (**Fig. 2C)**.

To flip the sign of target gene regulation, we chose the RNA-guided RNase Cas13. Computational simulation suggests that such an IFFL would achieve dosage control in the down regulation of the target gene (**Fig. 2D**). We chose RfxCas13d (referred to as CasRx), due to its small size and interest for therapeutics use.[28] Since it has recently been determined that CasRx has strong collateral activity that only occurs when targeting abundant mRNA transcripts,[29] we chose a regime where collateral activity is minimized (Fig. S2A). To establish proteolytic inhibition of Cas13, we screened for cleavage site positions where the site minimally disrupts Cas13 activity, yet makes the enzyme susceptible to protease cleavage (Fig. S2B). After displaying limited inhibition effect, we added a second cleavage site at its N terminus based on the "N-end rule"[30] such that, once exposed upon cleavage, it will serve as a degradation signal and further reduce Cas13 activity (Fig. 2D and Fig. S2C). We then performed a circuit titration with this construct and observed dosage control for repression (**Fig. 2E** and Fig. S2D and).

IFFL-mediated activation of *Rai1* using a dual AAV system in primary mouse neurons

While synthetic reporters facilitate initial optimization, they do not fully represent the challenges one might encounter in actual applications. Using *Rai1* as an example, we set out to examine whether proteolytic IFFLs can achieve dosage control for a therapeutically relevant endogenous gene in a more realistic context. *RAI1's* full length is ~20kb, including introns, with the cDNA spanning ~7.6kb; such a large gene would be difficult to deliver via AAVs, given the 4.8kb size limitation. Since our current system is also too large to fit on a single AAV, we opted to use a dual AAV system: in the first AAV we replaced spCas9 with its shorter version, sadCas9, and placed it with half of TEV, in the second we placed the gRNA[31] and the other half of TEV, either an active

one (IFFL) or a catalytically dead one (CRISPRa-only) (**Fig. 3A**). The split halves of TEV are fused to heterodimerization leucine zippers that will lead to the reconstitution of TEV when both halves are present.
This method ensures that the strength of the inhibitory arm is coupled to the levels of both AAVs in a cell.

To test whether this dual AAV system can control *Rai1* expression, we transiently transfected it to N2A cells, a mouse neuroblastoma cell line. (**Fig. 3B**). We could see 2.3 fold change for the CRISPRa-only case, while the IFFL produced 1.8 fold change, compared with the negative control, indicating successful modulation of gene activation. To go beyond ensemble measurements, we adopted the hybridization chain reaction (HCR) protocol to quantify mRNA transcripts in single cells (**Fig. 3C**).[32] We then performed HCR on N2A cells transduced with our dual AAV system, this time with a BFP gene on the second AAV, to be used as a transduction marker. We use HCR to visualize mRNA transcripts for the housekeeping gene *GAPDH*, *Rai1*, and exogenous BFP using flow cytometry. We observe that our IFFL system has a lower normalized Rai1 activation than CRISPRa alone (**Fig. 3D, left** and Fig. S3A). Most notably, the CRISPRa alone condition exhibits a long tail of highly activated cells, curtailed in the IFFL condition, as is quantified using right-hand kurtosis (**Fig. 3D, right**).

Lastly, we transduced primary mouse neurons with the dual AAV system and used HCR to visualize mRNA transcripts using confocal microscopy (**Fig. 3E**). As can be seen, *Rai1* transcripts appear as puncta and are visually more abundant in the CRISPRa and IFFL conditions compared to non-transduced cortical neurons, indicating that transduction was successful and there is indeed *Rai1* activation (**Fig. 3F** and Fig. S3B). To quantify the effect and study the variability at the single cell level, we calculated the percent area of the cell that is positive for *Rai1* and plotted the three histograms (**Fig. 3G**), as described in Star Methods and Fig. S3F. While overall activation was relatively low, both IFFL and CRISPRa histograms have peaks higher than the non-transduced median (>0.86%), indicating activation, and peaks in the lower region (<0.86%), likely from cells that have not been or are inadequately transduced. Notably, the IFFL histogram has a drop in density on the right side, while the right-hand tail of the CRISPRa condition is longer and shows cells with normalized *Rai1* levels near the 3x non-transduced mark (2.6%). Taken together, our data indicates that the IFFL reduces upper-bound variability of *Rai1* activation in both N2A cells and primary cortical neurons when transduced with AAVs.

IFFL-mediated activation of *RAI1* in human SMS patient cell lines

Finally, we investigate whether proteolytic IFFLs can activate *RAI1* to an intermediate level without overshooting into the potentially detrimental regime in a human patient context. First, we screened for human *RAI1*-targeting guide RNAs in HEK293 cells (Fig. **S**3A-D), and identified one that enabled CRISPR-based activation of endogenous *RAI1,* which we used for all subsequent experiments. When targeting *RAI1* in HEK cells, the activation strength can be countered by 5ng of the TEV protease, which constitutes 1% of the total amount of the CRISPRa system delivered (**Fig. 4A**). Interestingly, the variability between independent replicates was also minimized with TEV expression, although more replicates would need to be tested to validate this effect.

While our AAV results above hinted at the feasibility of proteolytic IFFLs, the additional variability and reduced efficiency caused by dual vector delivery were likely the reason for the suboptimal activation. Therefore, to achieve adequate circuit expression and to "lock" the stoichiometry between all components, we switched to encoding them on a single lentivirus vector. As can be seen in **Fig. 4B**, while the range of activation and tunability is significantly smaller than demonstrated using transient transfection, we were nonetheless able to observe IFFL-mediated dosage control of *RAI1* overexpression in HEK293s, where the set point is tunable using truncated variants of the TEV protease. We note that, because lentiviral delivery is much less variable than plasmid transfection, the ranges of vector and output variability here are much smaller than in transient transfection. We consider the relatively modest variability here to be more pertinent to the intended use cases, where there would be a critical difference between two vs. three copies of a causal gene underlying a genetic disorder. Nevertheless, it is promising that the dosage control behavior and tunability hold true when examined over orders of magnitude of vector levels as well as merely several folds.

Finally, we tested the lenti-encoded IFFL in a B-lymphoblastoid cell line (B-LCL) derived from an SMS patient (**Fig. 4C**). This cell line enables us to quantitatively evaluate circuit performance on the only intact copy of *RAI1* in the haploinsufficient scenario. Such B-LCLs have been used to assess the potential immune phenotype of *RAI1* haploinsufficiency compared to B-LCLs derived from healthy individuals, and are therefore of potential clinical relevance.[33] As can be observed in **Fig. 4D**, CRISPR activation alone increases *RAI1* levels by almost threefold, whereas the IFFL brings it back down to around twofold, indicating the functionality of the proteolytic inhibition. Most notably, and similarly to the previous figure, the CRISPRa condition exhibits a long tail of highly activated cells; this tail is less populated in one replicate of the IFFL, and completely gone in the other. Overall, our data indicates that the IFFL reduces upper-bound variability of *RAI1* activation in SMS patient-derived cell lines when delivered via lentivirus.

Discussion

Scientists have recognized the importance of dosage control via IFFLs for their potential in biomedical applications sensitive to gene dosage, as well as to offer better control over gene-manipulation techniques for scientific research in general. While most successful approaches rely on a delivered gene inhibited by a self-targeting microRNA, we adopted a different approach. By implementing an IFFL via proteolytic cleavage of CRISPR-Cas systems, we tackled a few fundamental limitations of current systems. First, our use of the protease as the inhibitory arm allowed us to create a closed system that does not rely on an endogenous pathway, which can vary between cells, break down as the system gets saturated, and interfere with cell functions. Second, we demonstrated post-delivery tuning by utilizing a small molecule-activated split-TEV, a difficult feat with an miRNA-based system. Third, delivery via mRNA was possible due to the constant production of both TEV and CRISPR-Cas from their respective mRNAs, enabling first-ever dosage control for RNA delivery, a safer and less costly avenue for transient gene therapy. Finally, by utilizing the CRISPR-Cas systems as the activating arm, we both removed the length limitation of the target gene and opened the door for

dosage control for repression. This also grants us compatibility with high-throughput CRISPR-based screens. Such screens are highly affected by dosage variance between constructs which is hard to account for in downstream analysis, and therefore could benefit from a more uniform expression. Taken together, our study represents an important milestone in the design and characterization of proteolytic IFFLs. We demonstrate, for the first time to our knowledge, post-delivery dynamic tuning, RNA-based delivery, and dosage control for target gene repression. We also demonstrate dosage-control in a therapeutic context, for a potential avenue for dosage-sensitive neurogenic genetic disorders.

Limitations of the Study

Using our protein-based IFFL, we demonstrated dosage control over several orders of magnitude, in various cell types and via different delivery methods. Yet, a few aspects require additional work. First, in our demonstration for dosage control via the dual AAV system, the activation efficiency was relatively low (**Fig. 3D & 3F**), most probably because of low transduction efficiency following the requirement for a double-AAV infection. While the dual AAV approach is a common solution for large cargos,[34,35] and Cas13 is small enough as is, we can miniaturize our activation system to fit on a single AAV by using shorter promoters,[36,37] or utilizing smaller CRISPRa systems.[38] In addition, in a few of the synthetic system cases, we see decreased activation towards the high end of the transfection marker, due to competitive inhibition (**Fig. 1H**). While this behavior is not observed for AAV and lentivirus delivery methods, we could place the TEV cut site location so that the CRISPR-cas system binding ability is abolished. Lastly, the therapeutic potential of our IFFL for SMS would be best characterized in SMS mouse models, with the viral vector directly delivered to the relevant brain region. Nonetheless, our work demonstrates robust and tunable dosage control for creating precise gene regulation systems for basic and translational research.

Acknowledgments

This work was funded by the National Institutes of Health (4R00EB027723-02; to X.J.G.), Simons Foundation Autism Research Initiative - 2021 Genomics of ASD: Pathways to Genetic Therapies (to W.H.H & X.J.G.), Seed Grant from Brain Research Foundation (to X.J.G.), NARSAD Young Investigator Grant from the Brain and Behavior Research Foundation (to X.J.G.), Longevity Impetus Grant (to X.J.G.), Stanford Science Fellows (to N.K.), Fulbright Foundation (to N.K.), N.K. is an Awardee of the Weizmann Institute of Science—Israel National Postdoctoral Award Program for Advancing Women in Science. We thank the Gao lab members for their feedback. We thank C. Liou for technical advice on qPCR and A. Vlahos for technical advice on microscopy analysis.

Author Contributions

N.K., C.A., and X.J.G. designed the study. N.K., C.A., Y.J.L performed and analyzed all the experiments, with support from W.H.H for the Rai1 human gRNA and general Rai1 work, and from J.T, L.B., and M.C.B for the NZF activation domain. M.Z. and J.K. helped with mRNA synthesis and primary mouse neuron work, respectively. N.K., C.A., and X.J.G. wrote the manuscript with input from all authors.

Declaration of Interests

X.J.G. is a co-founder and serves on the scientific advisory board of Radar Tx. All other authors declare no conflict of interest.

Resource Availability

Lead Contact

Correspondence: xjgao@stanford.edu

Materials Availability

Plasmids and plasmid maps will be deposited to Addgene. Raw flow cytometry data is available upon request from the corresponding author.

Data and Code Availability

Code for the Simulations can be found here https://data.mendeley.com/drafts/7n2j6ytrt5. Any additional information regarding data or code used in this paper is available from the lead contact upon request.

Figure 1

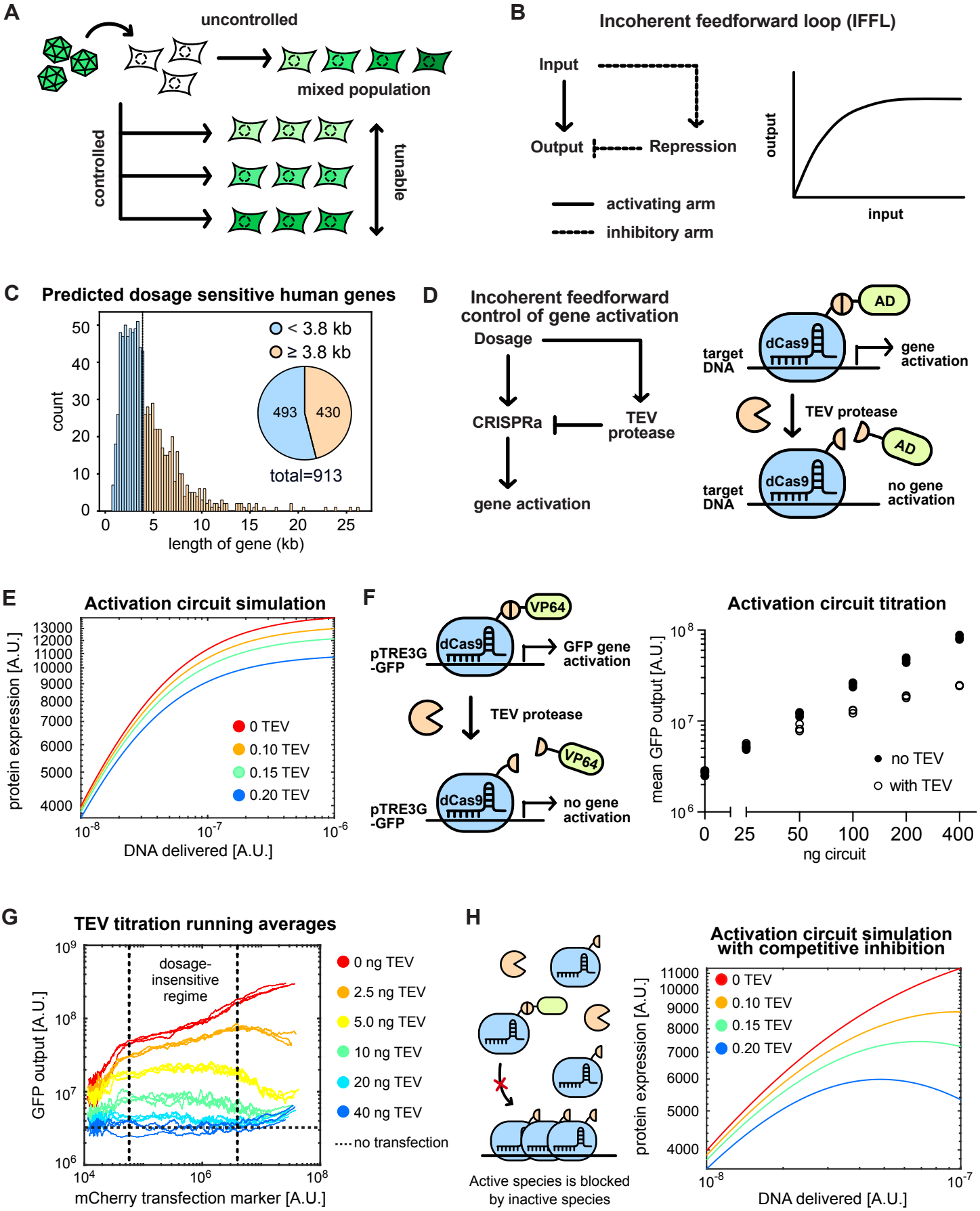

A
uncontrolled
mixed population
controlled
tunable
B
Incoherent feedforward loop (IFFL)
Input
Output
Repression
activating arm
inhibitory arm
output
input
C
Predicted dosage sensitive human genes
< 3.8 kb
≥ 3.8 kb
493
430
total=913
count
length of gene (kb)
D
Incoherent feedforward control of gene activation
Dosage
CRISPRa
TEV protease
gene activation
AD
dCas9
target DNA
gene activation
TEV protease
no gene activation
E
Activation circuit simulation
protein expression [A.U.]
DNA delivered [A.U.]
0 TEV
0.10 TEV
0.15 TEV
0.20 TEV
F
VP64
pTRE3G -GFP
GFP gene activation
TEV protease
no gene activation
Activation circuit titration
mean GFP output [A.U.]
ng circuit
no TEV
with TEV
G
TEV titration running averages
dosage-insensitive regime
GFP output [A.U.]
mCherry transfection marker [A.U.]
0 ng TEV
2.5 ng TEV
5.0 ng TEV
10 ng TEV
20 ng TEV
40 ng TEV
no transfection
H
Active species is blocked by inactive species
Activation circuit simulation with competitive inhibition
protein expression [A.U.]
DNA delivered [A.U.]
0 TEV
0.10 TEV
0.15 TEV
0.20 TEV

Figure 2

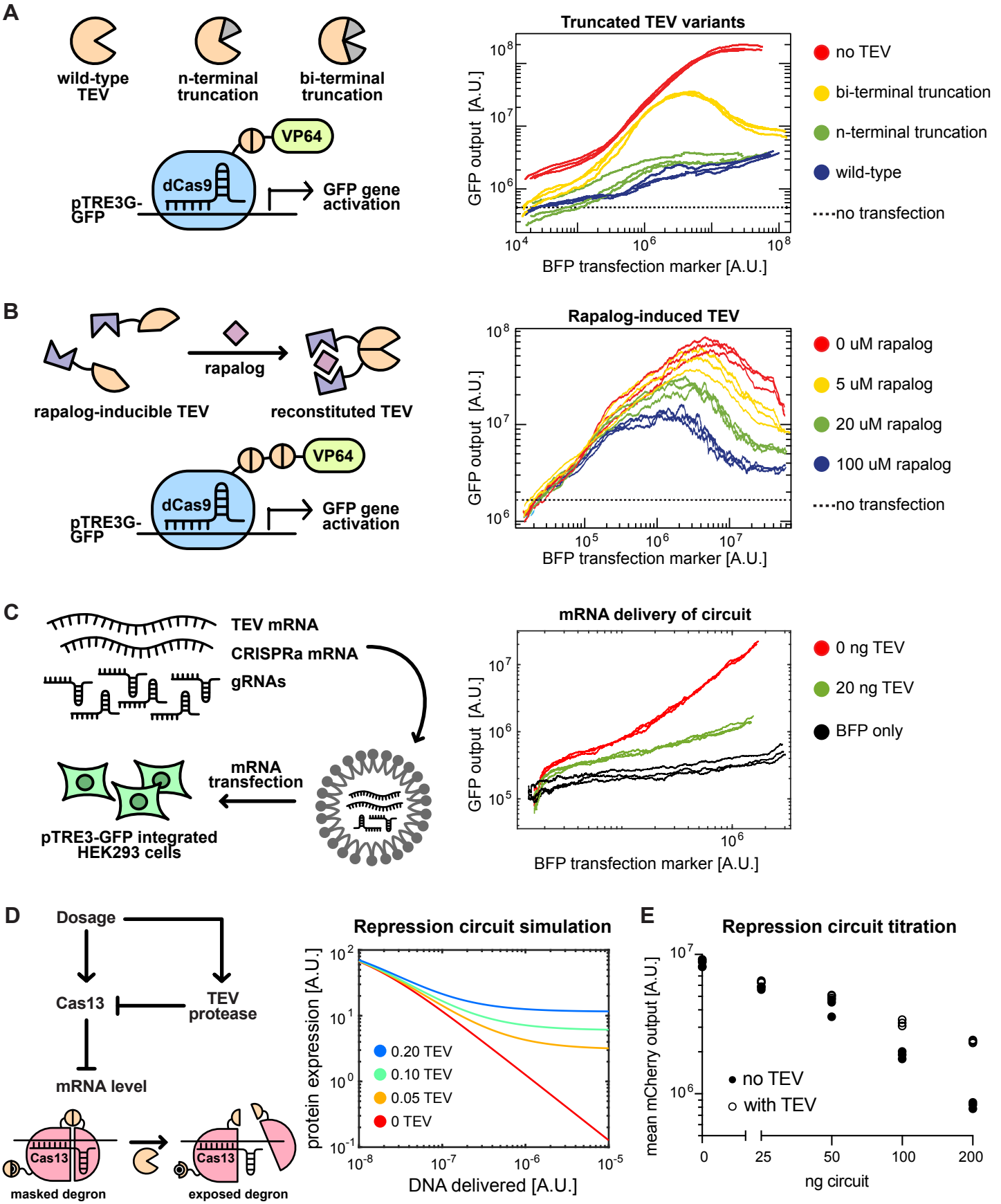

A
wild-type TEV
n-terminal truncation
bi-terminal truncation
VP64
dCas9
pTRE3G-GFP
GFP gene activation
Truncated TEV variants
GFP output [A.U.]
BFP transfection marker [A.U.]
no TEV
bi-terminal truncation
n-terminal truncation
wild-type
no transfection
B
rapalog
rapalog-inducible TEV
reconstituted TEV
VP64
dCas9
pTRE3G-GFP
GFP gene activation
Rapalog-induced TEV
GFP output [A.U.]
BFP transfection marker [A.U.]
0 uM rapalog
5 uM rapalog
20 uM rapalog
100 uM rapalog
no transfection
C
TEV mRNA
CRISPRa mRNA
gRNAs
mRNA transfection
pTRE3-GFP integrated HEK293 cells
mRNA delivery of circuit
GFP output [A.U.]
BFP transfection marker [A.U.]
0 ng TEV
20 ng TEV
BFP only
D
Dosage
Cas13
TEV protease
mRNA level
Cas13
masked degron
exposed degron
Repression circuit simulation
protein expression [A.U.]
DNA delivered [A.U.]
0.20 TEV
0.10 TEV
0.05 TEV
0 TEV
E
Repression circuit titration
mean mCherry output [A.U.]
ng circuit
no TEV
with TEV

Figure 3

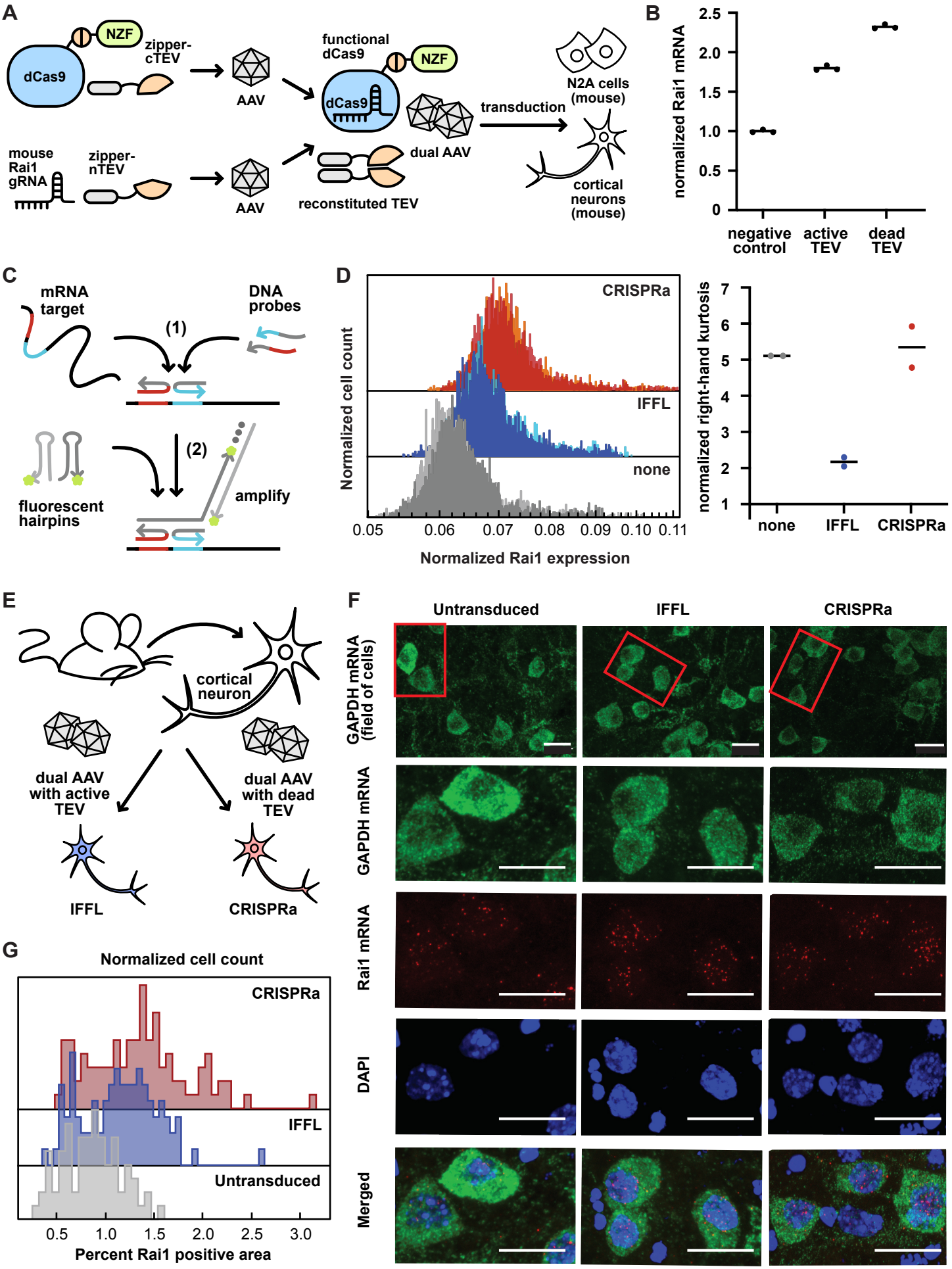
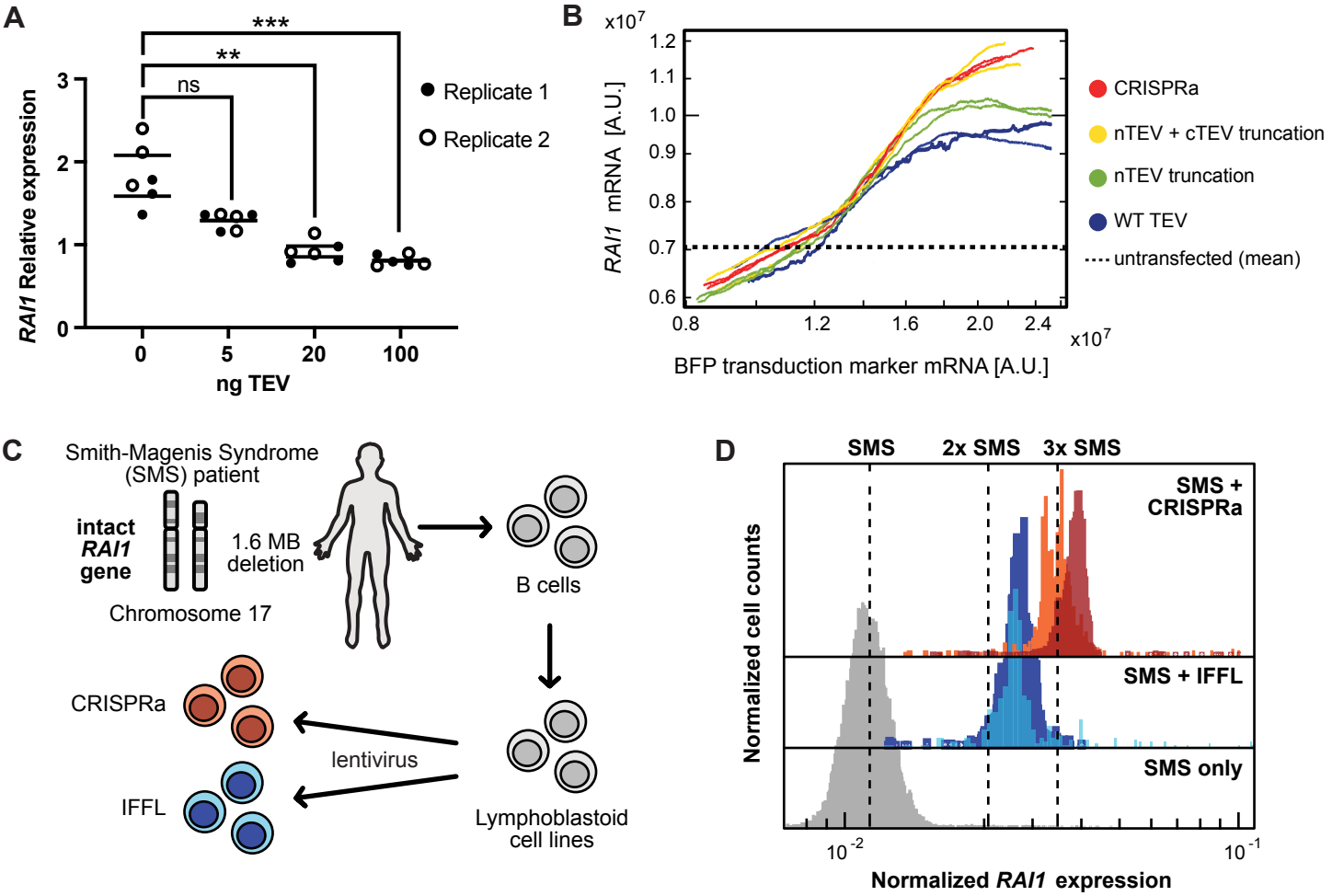

A
NZF
dCas9
zipper-cTEV
AAV
functional dCas9
NZF
dCas9
dual AAV
transduction
N2A cells (mouse)
mouse Rai1 gRNA
zipper-nTEV
AAV
reconstituted TEV
cortical neurons (mouse)
B
normalized Rai1 mRNA
2.5
2.0
1.5
1.0
0.5
0
negative control
active TEV
dead TEV
C
mRNA target
DNA probes
(1)
(2)
fluorescent hairpins
amplify
D
CRISPRa
IFFL
none
Normalized cell count
0.05
0.06
0.07
0.08
0.09
0.10
0.11
Normalized Rai1 expression
normalized right-hand kurtosis
7
6
5
4
3
2
1
none
IFFL
CRISPRa
E
cortical neuron
dual AAV with active TEV
dual AAV with dead TEV
IFFL
CRISPRa
F
Untransduced
IFFL
CRISPRa
GAPDH mRNA (field of cells)
GAPDH mRNA
Rai1 mRNA
DAPI
Merged
G
Normalized cell count
CRISPRa
IFFL
Untransduced
0.5
1.0
1.5
2.0
2.5
3.0
Percent Rai1 positive area

Figure 4


A
***
**
ns
Replicate 1
Replicate 2
RAI1 Relative expression
3
2
1
0
0
5
20
100
ng TEV
B
x10^7
1.2
1.1
1.0
0.9
0.8
0.7
0.6
RAI1 mRNA [A.U.]
0.8
1.2
1.6
2.0
2.4
x10^7
BFP transduction marker mRNA [A.U.]
CRISPRa
nTEV + cTEV truncation
nTEV truncation
WT TEV
untransfected (mean)
C
Smith-Magenis Syndrome (SMS) patient
intact RAI1 gene
1.6 MB deletion
Chromosome 17
B cells
CRISPRa
IFFL
lentivirus
Lymphoblastoid cell lines
D
SMS
2x SMS
3x SMS
SMS + CRISPRa
SMS + IFFL
SMS only
Normalized cell counts
10^-2
10^-1
Normalized RAI1 expression

Main figure titles and legends

Figure 1: Establishing proteolytic IFFLs using a synthetic reporter

1A: Schematic describing dosage-independent gene expression. Current gene delivery methods lead to a stochastic distribution of gene expression levels due to inherent differences in single-cell dosage. An ideal gene delivery system would eliminate this heterogeneity while allowing for tuning of the gene expression level.
1B: Schematic of an incoherent feedforward loop. The output is able to remain invariable to the input due to the input driving both an activating and inhibitory arm.
1C: Histogram of sizes of human genes predicted to be dosage sensitive. Of the 913 genes predicted to be dosage-sensitive, 430 are longer than 3.8 kb and are likely too large to be packaged into an AAV.
1D: Schematic of engineered activation circuit. TEV protease is used to inhibit CRISPRa, creating an IFFL with dosage as the input and CRISPRa activity as the output. CRISPRa activates gene expression.
1E: Schematic of activation circuit components and simulation of how the components interact using ordinary differential equations. Without TEV, as DNA delivered increases, protein expression continues to increase. With TEV, as DNA delivered increases, protein expression increases and eventually plateaus at a 'set point'. Increasing amounts of TEV decreases the 'set point'.
1F: Dot plot of average GFP (output) as a function of circuit transfected in piggyBac-ed pTRE3G-GFP HEK293T cells.
1G: Running averages of GFP (output) as a function of mCherry transfection marker with increasing amounts of TEV plasmid transfected. In highly transfected cells, there is a non-monotonic effect where GFP decreases. The running window is based on mean values, and its length is relative to the number of cells (typically 1%-5% of total events).
1H: Schematic and simulation of the circuit with competitive inhibition. The cleaved dCas9 species can also bind to the target DNA, blocking uncleaved species from binding and activating the target gene, leading to a non-monotonic effect at high dosages.
All experiments shown are based on three independent replicates.

Figure 2: Demonstrating unique features of proteolytic IFFLs

2A: Schematic of truncated TEV variants and running average of GFP (output) as a function of BFP transfection marker using the different TEV variants while keeping TEV amount constant.
2B: Schematic of rapalog-inducible TEV and running average of GFP (output) as a function of BFP transfection marker using rapalog-inducible TEV. Rapalog-inducible TEV consists of split TEV with each half fused to either FRB or FKBP. Upon introduction of the small-molecule drug rapalog, the two halves are brought together, forming a functional TEV protease. There are two TEV cut sites between dCas9 and the transcriptional activator.
2C: Schematic of mRNA delivery and running average of GFP (output) as a function of BFP transfection marker when the circuit is delivered using mRNA.
2D: Schematic of engineered repression circuit based on Cas13 and its simulation using ordinary differential equations. Cas13 is an mRNA cleaving enzyme.TEV protease is used to inhibit Cas13, creating an IFFL with dosage as the input and Cas13 activity as

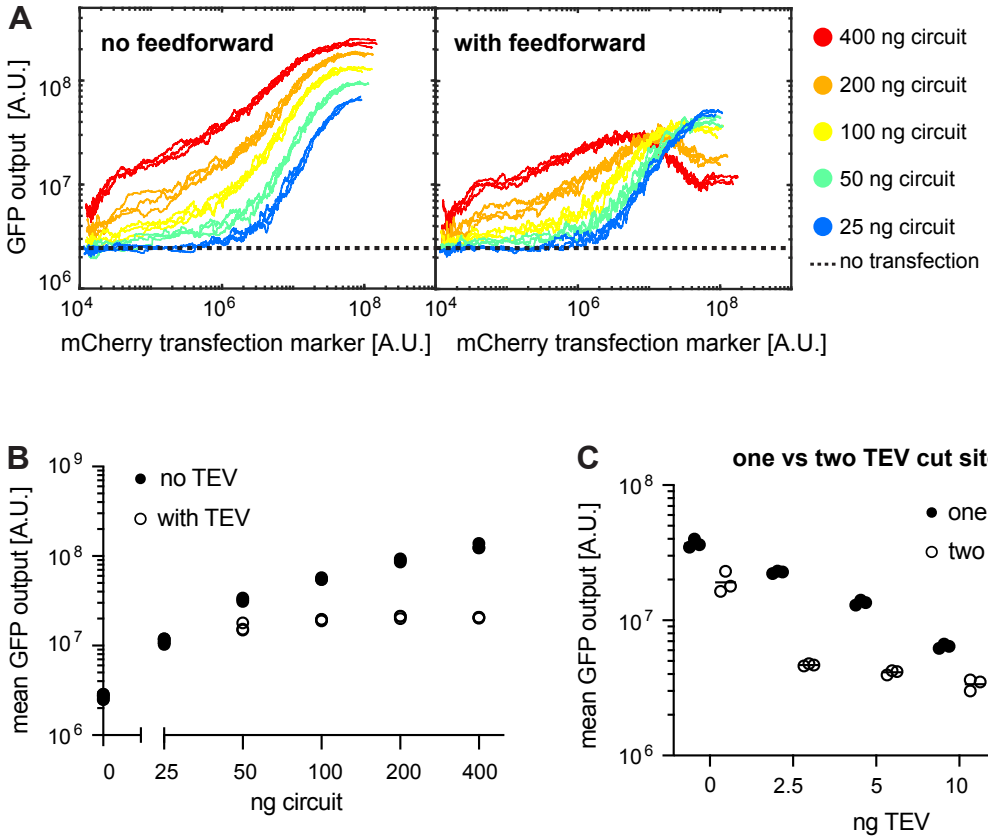
the output. The simulation indicates that without TEV, as DNA delivered increases, protein expression continues to decrease. With TEV, as the delivered DNA amount increases, protein expression decreases and eventually plateaus at a 'set point'. Increasing amounts of TEV increases the 'set point'.
2E: Schematic of TEV-cleavable Cas13 construct with an N-end degron and dot plot of average mCherry (output) as a function of circuit transfected in HEK293T cells gated on highly transfected cells.
All experiments shown are based on three independent replicates.

Figure 3: IFFL-mediated activation of *Rai1* using a dual AAV system
3A: Schematic of dual AAV system.
3B: Dot plot of qPCR results of N2A cells transiently transfected with two plasmids encoding the dual AAV system.
3C: Schematic of HCR steps. (1) DNA probes bind to mRNA targets. (2) Fluorescent hairpins bind to the DNA probes and amplify, leading to fluorescence which can be measured using flow cytometry or confocal microscopy.
3D: Histograms and corresponding right hand kurtosis of running medians for normalized Rai1 expression in N2A cells transduced with the dual AAV system containing either CRISPRa or IFFL conditions. Cells were gated on positive-stained cells, FSC-SSC live cells, with high GAPDH expression, and positive BFP signal. BFP histograms are shown in Fig. S3A. All conditions had two technical replicates with at least 25,000 cells in each histogram.
3E: Schematic of primary mouse cortical neuron dual AAV experiment.
3F: Confocal microscopy images of primary mouse cortical neurons transduced with the dual AAV system. GAPDH (AlexaFluor-488), Rai1 (AlexaFluor-647), BFP (AlexaFluor-546) and Dapi expressions were measured using HCR. Red boxes indicate the field of view of zoomed-in images. The range for all red channel images was uniformly adjusted to better display Rai1 mRNA puncta. Scale bars represent 20um.
3G: Histograms depicting the percentage of *RAI1* positive pixels out of total soma region of individual neurons. Data for all three experiments was combined from two independently transduced wells. Untransduced histogram contains ~80 cells, and both CRISPRa and IFFL ~100 cells. Median percentages were increased by 35% (IFFL) and 58% (CRISPRa) relative to non transduced.

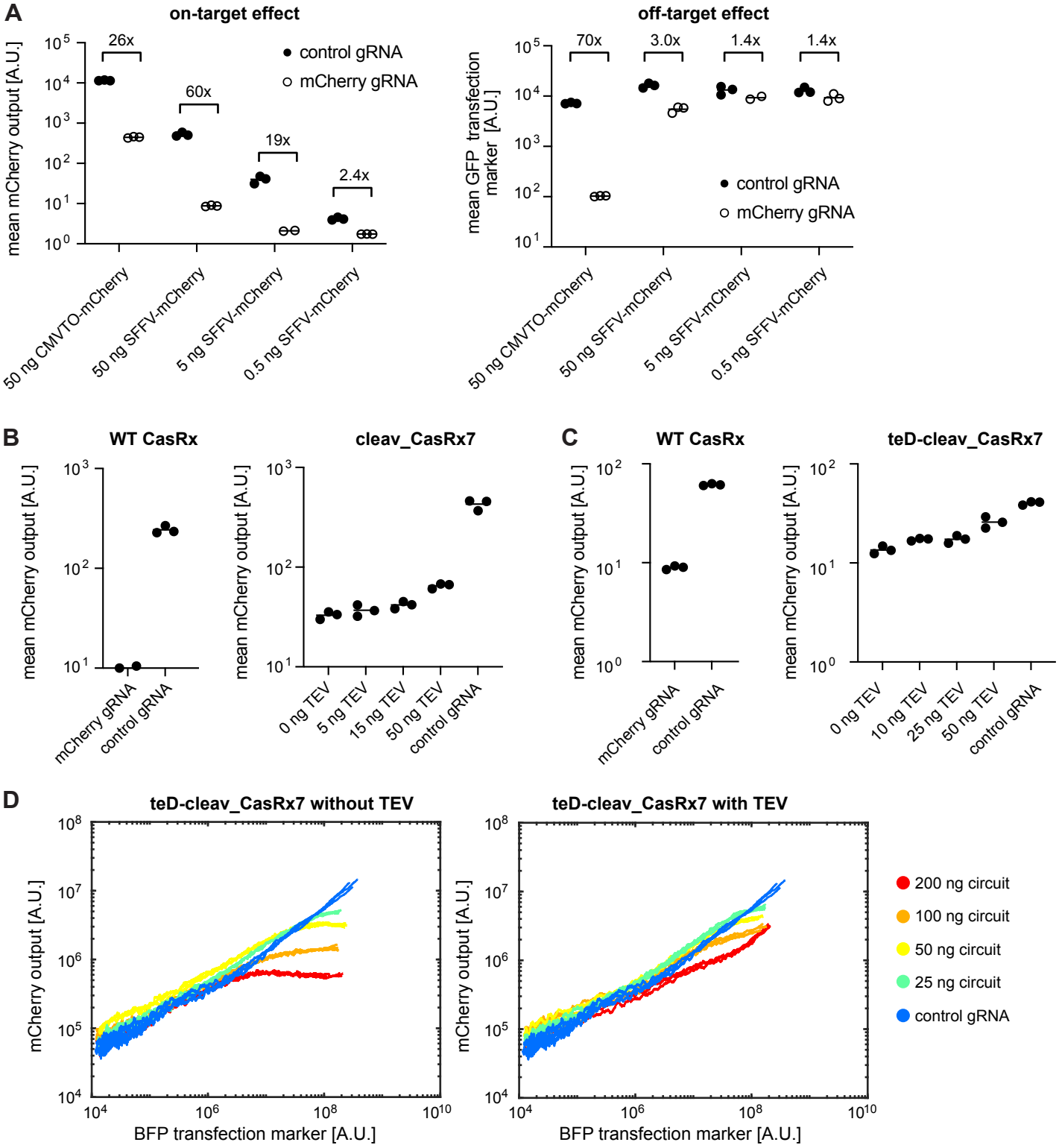
Figure 4: IFFL-mediated activation of RAI1 in patient cell lines
4A: Dot plot of qPCR results using gRNA targeting *RAI1* in transiently transfected HEK293 cells. Results are based on two biological replicates (full and empty circles) and three technical replicates.
4B: Running medians of *RAI1* mRNA as a function of BFP transduction marker mRNA for lentivirus transduction of circuit in HEK293 cells.
4C: Diagram of creation of SMS-patient lymphoblastoid cell lines and lentivirus transduction of CRISPRa circuit and IFFL targeting *RAI1*.
4D. Histograms of normalized *RAI1* expression for each SMS patient lymphoblastoid cell line transduced with the CRISPRa circuit, IFFL, or no transduction. Two similar shades represent two independent replicates. Cells were gated on positive-stained cells, live SSC-FSC plots, high GAPDH expression, and positive BFP signal. BFP

histograms are shown in Fig. S4E. Median levels for CRISPRa, IFFL and SMS-only histograms are 0.365, 0.276, and 0.117, respectively, averaged over the two replicates. CRISPRa and IFFL conditions have two independent replicates with a minimum of 1000 cells in each histogram. SMS-only condition has 90,000 cells.

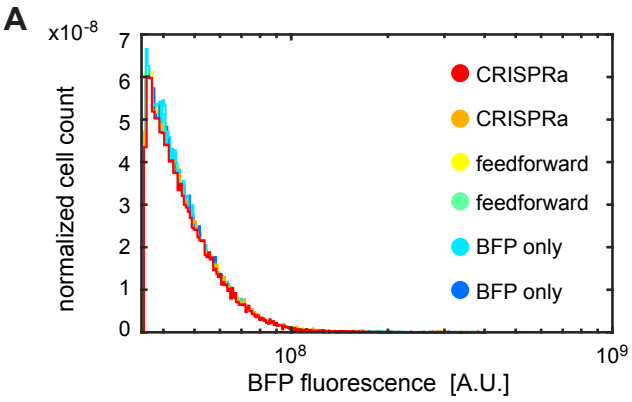
Supplementary Figure 1

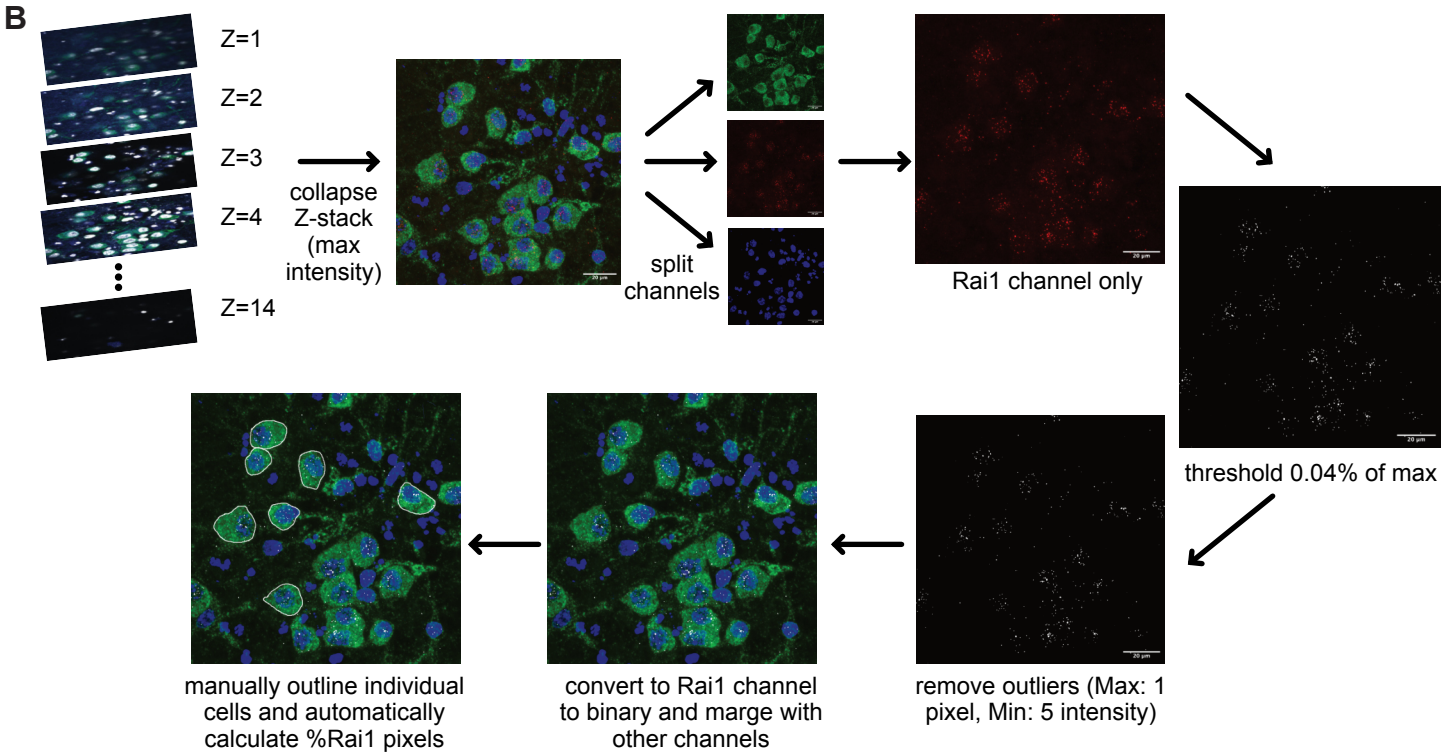


Supplementary Figure 2

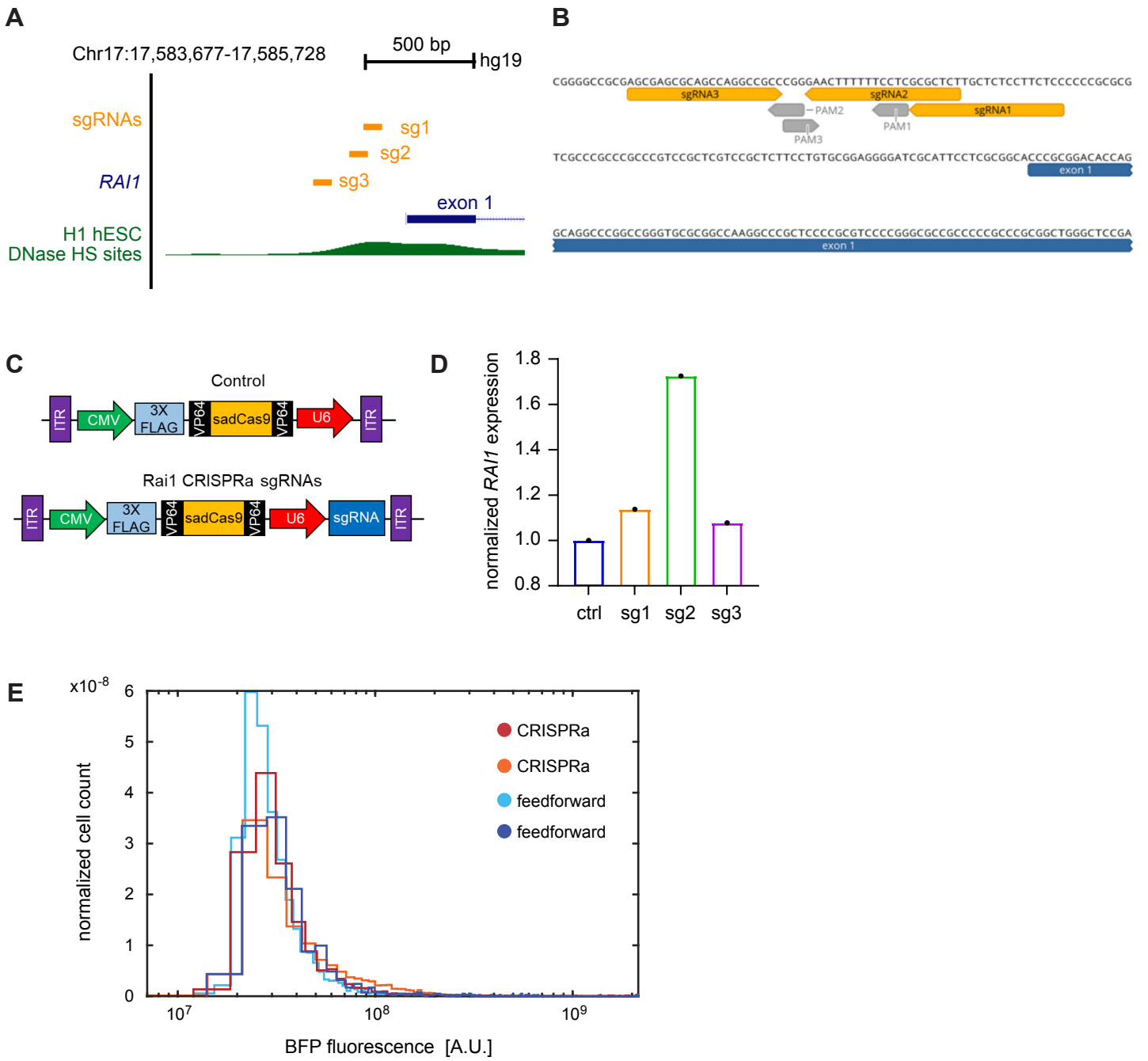


Supplementary Figure 3





Supplementary Figure 4



Supplemental figure titles and legends

Supplementary Figure 1: Activation circuit analysis and optimization

S1A: Dot plot of mean GFP (output) of all transfected cells as a function of circuit transfected in piggyBac-ed pTRE3G-GFP HEK293T cells.

S1B: Running means of GFP (output) as a function of mCherry transfection marker for activation circuit titration. In highly transfected cells, there is a non-monotonic effect where GFP decreases.

S1C: Dot plot comparing TEV sensitivity of activation circuit where dCas9 has either one or two TEV cut sites linking it to the activation domain.

Supplementary Figure 2: Engineering a TEV cleavable Cas13

S2A: Dot plots for the on-target and off-target effects of WT CasRx using different promoters and amounts of mCherry (target) plasmid transfected.

S2B: Dot plots comparing cleav_CasRx7 activity as TEV plasmid amount is titrated compared to WT CasRx.

S2C: Dot plots comparing teD-cleav_CasRx7 activity as TEV plasmid amount is titrated compared to WT CasRx.

S2D: Running means for teD-cleav_CasRx7 whole circuit titration.

Supplementary Figure 3: Targeting mouse *RAI1*

S3A: Histograms of BFP expression for N2A cells transduced with dual AAV system.

S3B: Schematic of confocal microscopy image processing steps for transduced and non transduced primary mouse cortical neurons.

Supplementary Figure 4: Targeting human *RAI1*

S4A: Identification of a human RAI1 sgRNA that upregulates RAI1 expression in 293A cell line. schematic representation of the positions of sgRNAs (orange) in the human RAI1 promoter. DNase hypersensitive sites (HS) in H1 human embryonic stem cells with open chromatin configurations are shown in green.

S4B: positions and sequences of three sgRNAs (orange) followed by Sa protospacer adjacent motif (PAM) sequences (gray) within human RAI1's proximal promoter.

S4C: schematic representation of the sadCas9-2 × VP64 constructs used.

S4D: qRT-PCR showing that sadCas9-2 × VP64-sg2 vector increases RAI1 mRNA expression in 293A cell line.

S4E: Histograms of BFP fluorescence for SMS patient B-cells transduced with either CRISPRa or the feedforward circuit.

Materials and Methods

Construction of plasmids

Plasmids were generated using standard molecular cloning practices, including the following methods: InFusion, ligations, PCR, and annealing. Briefly, all transgene sequences were ordered from Twist Biosciences, or PCR amplified via plasmids from Addgene, by primers ordered from IDT and Thermo Fisher Scientific. Backbone plasmids were restricted using restriction enzymes from Fisher or NEB, and purified. Then transgenes and backbones were put together via in-fusion (Takara) or ligation (NEB), transformed via the heat-shock protocol into competent E.coli cells (Turbo- NEB) and plated on Carb resistance containing agar plates. Plasmids were extracted using Qiagen miniprep kits and verified by either Sanger sequencing or whole-plasmid sequencing (Primordium) . Plasmid sources for dCas9 and TRE-targeting gRNA were generous gifts from Prof. Stanley Qi at Stanford.

Tissue culture work

All cells were cultured in a humidity-controlled incubator under standard culture conditions (37 °C with 5% CO2).

HEK293 and N2A cells

HEK293 cells (catalog no. CRL-1573) and N2A cells (VWR, catalog n. MSPP-CCL131) were purchased from ATCC, and cultured using Dulbecco's modified Eagle's medium (Fisher, 501015428) , supplemented with 10% fetal bovine serum (FBS) (Fisher Scientific catalog no. FB12999102), 1 mM sodium pyruvate (EMD Millipore catalog no. TMS-005-C), 1× penicillin–streptomycin (Genesee catalog no. 25-512), and non essential amino acids (Genesee catalog no. 25-536). The HEK-pTR3G-GFP cell line was cultured in similar conditions with the addition of Hygromycin.

Lymphoblastoid cell lines (LCLs)

LCLs were purchased from the Coriell Institute (catalog no. GM23786) and cultured according to their protocol, using RPMI medium supplement with 10% FBS and 1% Penicillin/Streptomycin. They were grown in vessels made for floating cells with a breathable film cover. Cells were split once every three days, and cultured at a 1:5 ratio in fresh media.

Mouse cortical neurons

Primary mouse neurons were purchased from Thermo Fisher Scientific (catalog no. A15585), thawed, and cultured according to their guidelines. Medium was prepared as follows: Complete Neurobasal™ Medium supplemented with GlutaMAX™ -I Supplement (Cat. no. 35050) and B-27™ Supplement (Cat. no. 17504) to final concentrations of 0.5mM and 2%, respectively. Plates were prepared as follows: 24-well plates were incubated with Poly-D-Lysine in PBS at 4.5 µg/cm2 for 2 hours in a biological hood, followed by 3x washes with water. Plates were then left open to dry for 1-2 hours and kept at 4C until use. Cells were thawed and cultured as follows: cells went through rapid thawing (< 2 minutes) of the frozen vial in a 37℃ water bath, then transferred to the cell culture hood. Pipette tip was rinsed with complete medium and

very gently transferred the cells to the pre-rinsed 50-mL tube 3 mL of complete medium (pre-warmed to 37℃) was added to the cells in the 50-mL tube extremely slowly at the rate of one drop per second. The suspension was then mixed and seeded in a pre-coated glass bottom 24-well plate (ibidi, 82426) with 1ml complete medium pre-warmed to 37℃ at 200k cells per well. 50% of medium was changed 24 hours post-seeding, and repeated every 2-3 days.

Transient DNA transfections

HEK293 cells were cultured in 24-well tissue culture-treated plates under standard culture conditions. When cells were 60–90% confluent, they were transiently transfected with plasmids via the jetOPTIMUS DNA transfection reagent (Polyplus catalog no. 117-15), as per manufacturer's instructions using 0.4 µl of reagent per 50 µl of jetOPTIMUS buffer for 500 ng total DNA transfections in the 24-well plate. To create HEK-pTR3G-GFP cell line, we transfected HEK293 cells with NK90 piggybac plasmid with a matching transposase, let them grow for 72 hours and then added Hygromycin for screening. After 7 days, surviving cells were frozen and used in all activation synthetic-based experiments.

mRNA synthesis and transfection

The DNA templates for in vitro transcription contained dCas9 and TEV coding sequences flanked by optimal 5' and 3' UTRs as well as a T7 promoter and PolyA tail (NK409, NK411). The plasmid template with optimal UTRs was a gift from Prof. Michael Elowitz. In vitro transcribed mRNA was produced and purified by GeneScript Inc. based on the DNA template. gRNA were produced based on Precision gRNA Synthesis Kit (Thermo Fisher Scientific, A29377). The purified mRNAs and gRNA were then transfected to HEK293-pTRE3G-EGFP at 90% confluency via the Mirus TransIT®-mRNA Transfection Kit following the manufacturer's instructions. Briefly, 3 µl of TransIT-mRNA reagent as well as mRNA boost were mixed with mRNA and gRNA per well in 24 well plate. Expression of fluorescent proteins on the mRNA was assayed via flow cytometry 48 hours post transfection.

Flow cytometry and data analysis

HEK293 were harvested approximately 48 h post transfection by trypsinization and resuspended in flow buffer (HBSS + 2.5 mg ml−1 bovine serum albumin). Cells were analyzed by flow cytometry (Biorad ZE5 Cell Analyzer), and data was processed with Matlab software. Cells were gated for live and single cells based on FSC/SSC and FSC-A/FSC-W plots. They were then sorted according to the transfection marker signal (BFP or mCherry, as mentioned) and their EGFP signal averaged over a running window of 1%-5% of total count. EGFP running averages were then plotted on as a function of transfection marker signal. For HCR experiments, the fluorescent signal detected was altered to AlexaFluor-488, AlexaFluor-546, and AlexaFluor-647, for detection of GAPDH, BFP and Rai1, respectively.

RNA extractions and quantitative PCR

Cells were grown as detailed above, and were spun down at 500g for 5 minutes. RNA was then extracted using RNAasy mini kit (Qiagen), RNase-Free DNase Set (Qiagen), and QIAshredder (Qiagen). After extraction RNA concentration was measured via nanodrop. 5000 ng of purified RNA was then reverse transcribed using SuperScript™ IV First-Strand Synthesis System (Invitrogen). cDNA were then diluted and added to a qPCR compatible 96-well plate (Fisher, 4346907) together with qPCR master mix: SENSIMIX SYBR HI-ROX (VWR, 490017-866 / 868) and primers targeting human or mouse Rai1 gene. qPCR was carried out on a QuantStudio3 (Applied Biosystems) using SYBR-Green. RNA estimation was calculated on the basis of the calibration curve of purified plasmid and normalized by the *Ct* threshold. The following primer pair sequences for mouse Rai1, human Rai1 and normalizing gene (β-actin) were used: Rai1-mouse-F, XX; Rai1-mouse-r, XX; Rai1-human-F, CCCAGGAGCACTGGGTGCATGA; Rai1-human-r, GCAGCTGGAACACATCATGTCCACG; β-actin-F, CGTCCACCGCAAATGCTT; β-actin-R, GTTTTCTGCGCAAGTTAGGTTTTGT.

gRNA screening

The sgRNA oligonucleotides targeting human RAI1 promoter regions were designed using the Benchling gRNA Design Tool.[39] The sadCas9-2 × VP64 vector (Addgene #135338)[40] was used as a backbone vector. It carries mutations in the endonuclease catalytic residues (D10A, N580A) of a FLAG-tagged saCas9, which was fused on both N and C termini with transcriptional activators VP64 (four copies of VP16). sgRNAs were then cloned into the backbone using the ligation cloning method previously described. After validating the sequences of constructs, they were transfected into 293A cells using polyethylenimine, and cells were harvested in TRIzol reagent 72 hours after transfection. Total RNA was extracted using phenol-chloroform extraction method, reverse-transcribed using the SuperScript III First-Strand Synthesis System (Thermo Fisher), and qPCR reactions were conducted using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) in StepOnePlus real-time PCR system (Applied Biosystems) with GAPDH as a housekeeping control.

The primers used for quantitative PCR were as follows: human RAI1-F 5'-CCTCAGCATTCCCAGTCCTTC-3', human RAI1-R 5'-CTGTGCAACTCTTATAGGAGTGG-3'; GAPDH-F 5'-AAGGTGAAGGTCGGAGTCAA-3', GAPDH-R 5'-AATGAAGGGGTCATTGATGG-3'.

CasRx repression work

<u>Determining the minimal collateral activity regime</u>

To facilitate screening of cleavable CasRx designs, we use a synthetic system where we use the fluorescent protein mCherry as the output. Since it has been reported that CasRx can cleave non-target RNA, we first determine the regime in which this collateral activity is minimal. Since collateral activity for CasRx depends on target mRNA level,[29] we titrate mCherry expression level (the target mRNA) by changing the promoter of the

mCherry plasmid (CMVTO or SFFV) and the amount of plasmid transfected. We also measure GFP fluorescence which is expressed on the same transcript as CasRx (EF1a-GFP-T2A-CasRx) to determine the amount of off-target effect (SFig. 2A). Transfecting 50 ng of SFFV-mCherry plasmid resulted in the greatest fold change in mean mCherry fluorescence between control gRNA and mCherry-targeting gRNA and minimal difference between mean GFP fluorescence.

Engineering and screening cleavable CasRx variants

Based on the results from the previous section, we transfect 50 ng of SFFV-mCherry, 200 ng of the EF1a-GFP-T2A-CasRx_variant, and 200 ng of the mCherry-targeting gRNA for the following cleavable Cas13 experiments.

To engineer a cleavable CasRx, we first screen for locations within CasRx to place the TEV cut site. Ideally, we would like the cleavable CasRx to (1) have similar mRNA cleaving activity as wild-type CasRx and (2) be inhibited by the presence of TEV. Since there is no crystal structure for CasRx, we use AlphaFold to predict its structure and find flexible loops that we can replace with the TEV cut site or where we can insert the TEV cut site. We screened 8 constructs, and our best candidate was cleav_CasRx7 which had similar mRNA cleaving activity as WT CasRx7 (Fig. S2B). However, there was not appreciable TEV repression of mRNA cleavage activity. To improve the ability of TEV to repress activity of cleav_CasRx7, we added a TEV-cleavable N-end degron to the construct. Degrons are used to regulate degradation rates in cells. In the absence of TEV, the degron is masked. After TEV cleaves the N-end degron, the degron will be exposed, marking the protein for degradation, and decreasing its half-life. Since the N-end degron is appended to the N-terminal of the protein, it should not interfere much with the activity of cleav_CasRx7. Indeed, this construct (teD-cleav_CasRx7) was more TEV sensitive than cleav_CasRx7 and retained its efficiency (Fig. S2C).

Repression circuit titration

Lastly, we used teD-cleav_CasRx7 in a whole circuit titration. We transfect 50 ng of SFFV-mCherry and varying amounts of the teD-cleav_CasRx7 circuit. We use BFP as a transfection marker and average the mCherry fluorescence for cells that are highly transfected (BFP fluorescence > 10^8 A.U.) to produce Fig. 2E. Note that in the running average plots (Fig. S2D), the output mCherry fluorescence increases as dosage increases (BFP transfection marker) which is expected. This is because the mCherry gene is not integrated into the cell's genome and is instead transcribed from a transfected plasmid. Thus, the relationship between mCherry fluorescence and dosage is not the same as in Fig. 2D which is derived from the simulation which assumes that the target gRNA is produced from an integrated gene.

AAV production and transductions

AAVs were ordered and manufactured with the help of the Stanford Gene Vector and Virus Core (GVVC). Plasmids were amplified and extracted via the Qiagen plasmid midiprep kit (Qiagen), their concentration measured and sent to GVVC for production. The three AAVs were measured for their ITR: NK355: 1.12e13vg/ml, NK404: 2.57e13vg/ml, NK405: 2.05e13vg/ml. For a 24-well of primary neurons, 4ul of total virus

was with NK355 always in access, since dCas is the limiting factor and has a lower ITR. Cells were monitored via the EVOS M7000 Cell Imaging System every day, and were fixed 2.5 weeks after AAV transduction, once a BFP signal was detected.

Lentivirus production and transductions

LentiX cells were seeded in 6-well tissue culture plates one day prior to transfections. When cells were 80%-90% confluent, the cells were transiently transfected with plasmid constructs (600 ng PAX2, pMD2g, and 1,100 ng transgene plasmid) as detailed earlier. Cells were incubated for 24 h under standard culture conditions and were supplemented with 3ml of complete DMEM media to each well. Lentivirus was concentrated 24 hrs afterward using viral precipitation: for each lentiviral prep, media was filtered using a syringe and 0.45 µm filter into 15 mL conical tubes. 5x Lentivirus Precipitation Solution (Alstem 480 catalog no. VC100) was mixed with each prep and incubated at 4 °C for 48-72 hours. Viruses were then spun down at 1500xg for 30 min at 4 °C. Supernatant was aspirated, and virus was resuspended using 200 µL complete media (either DMEM for HEK293 and N2A, or RPMI for LCLs). Virus was then added to cells according to the following protocols, and leftover was frozen in -80 °C for further use.

<u>HEK293 and N2A cells</u>

Virus was added dropwise onto 24-well tissue culture plates containing HEK293 or N2A cells seeded at 200k cells/well. Media was changed 24 horse post transduction, and cells were grown for an additional 5 days, at which time positive BFP signal was detected, and then fixed and stained according to the following protocols.

<u>Lymphoblastoid cell lines (LCLs)</u>

Patient LCLs were split a day prior to transduction at a ⅓ ratio for optimal growth. On the day of transfection, cells were counted and 50k cells were transferred to multiple wells of a 96-well plate. Cells were then spun down at 500xg for 5 min and the media was gently taken out. Virus solutions were then supplemented by Polybrene (Merc) to a final concentration of 10ug/ml, and the mixed solutions were gently added to the cells without resuspension. Cells were then placed in the incubator overnight, and the virus solution removed by centrifugation (500xg for 5min). Cells were then merged and expanded, transferring from 96-well plates to a 48-well plate and finally a 24-well plate. Once cells reached confluency in 24-well plates, they were fixed and stained according to the following protocols.

Cell fixation and Hybridization Chain Reaction (HCR)

DNA probes for GAPDH, RAI1 and BFP were designed and ordered from Molecular Instruments based on the full-length genes with maximal available probe set size of ~7500bp (Fig. 3D+3F and Fig. 4B+4D). Gating was done based on positive probe signals, FSC-SSC for live cells, high GAPDH for active cells, and positive BFP mRNA signal for HCR-flow. All protocols were conducted according to Molecular Instruments HCR protocols, brief versions are enclosed below.

<u>Cells in suspension: HEK293 and N2A cells</u>

Day 1: Growth media was aspirated from culture plates and cells were washed with DPBS. Cells were then trypsinized and quenched by the addition of growth media. Cells were then transferred to 15 mL conical tubes and centrifuged for 5 min at 180xg. Cells were then resuspended in 4% formaldehyde to reach approximately 10 million cells/mL. Cells were fixed for 1 hr at room temperature, and then centrifuged and washed with PBST for 4 times. Cells were then re-suspended in cold 70% ethanol overnight.
Day 2: Cells were then centrifuged for 5 min and washed twice with 500 µL of PBST. PBST was then removed and the pellets were resuspended the with 400 µL of probe hybridization buffer and pre-hybridized for 30 min at 37 ◦C. In the meantime, probe solution was prepared by adding a fixed amount of each probe set to 100 µL of probe hybridization buffer pre-heated to 37◦C. Amounts used: 4 pmol for GAPDH, 16 pmol for RAI1, and 4 pmol for BFP. This solution was added to cells, and then they were incubated overnight (>12 h) at 37◦C.
Day 3: Probe solution was then removed, and the cells resuspended with 500 µL of probe wash buffer, incubated for 10 min at 37 ◦C and remove the wash solution by centrifugation for 5 min. This step was repeated for three additional times. Next, the pellet was resuspended with 500 µL of 5× SSCT, incubated for 5 min at room temperature, then centrifuged and re-suspended with 150 µL of amplification buffer and let sit for 30 minutes at room temperature. Then the amplification hairpins were separately prepared: 15 pmol of hairpin h1 and 15 pmol of hairpin h2 for each primary hairpin (6 total) by heating to 95 ◦C for 90 seconds and then cooled to room temperature in a dark drawer for 30 min 5. THe hairpins were then added to the sample and incubated overnight (>12 h) in the dark at room temperature.
Day 4: Cells were then centrifuged for 5 min and hairpin solution removed. The cells were then washed with 5× SSCT for six total times. Finally, cells were resuspended with flow buffer, filtered, and analyzed via flow cytometry as detailed previously.

<u>Cell on a slide: mouse cortical neurons</u>
Day 1: Growth media was aspirated from culture plates and cells were washed with PBST twice. 4% formaldehyde was then added to the cells and fixed for 1.5 hr at room temperature. Next, cells were washed 2x times with PBST and incubated 10min with Triton-X-100. Samples were then washed with 2 × 300 µL of 2× SSC and pre-hybridized samples in 300 µL of probe hybridization buffer for 30 min at 37 ◦C. Probe solution was then prepared similar to the previous protocol, except amounts of probes were increased to: 8 pmol for GAPDH, 32 pmol for RAI1, and 10 pmol for BFP. They were then added to probe hybridization buffer, placed on the cells, and left overnight (>12 hrs).
Day 2: Excess probes were removed by washing 4 × 5 min with 300 µL of probe wash buffer at 37 ◦C. Then, samples were washed with 2 × 5 min with 5× SSCT at room temperature, and then pre-amplified in 300 µL of amplification buffer for 30 min at room temperature. Hairpin amplification solution was prepared similar to the previous protocol, and then added to the cells. Cells were then left in room temperature for 2 hours, and then washed 5 × 5 min with 300 µL of 5× SSCT at room temperature. Finally, 5× SSCT was aspirated and 300ul of ProLong® Gold Antifade Reagent with DAPI was added (Cell Signaling). Samples were then stored at 4 ◦C protected from light prior to imaging. Imaging took place over multiple days starting at one day post storage.

Microscopy experiments and analysis
All images were taken with the Inverted Zeiss LSM 780 multiphoton laser scanning confocal microscope. The following channels were included in the images: AlexaFluor-488, AlexaFluor-546, AlexaFluor-647, and DAPI. Z-stack captures were implemented, with 10-15 slices of each field of view, taking into account the entire range of cells in focus. Analysis was conducted via the FIJI software, and included the following steps: all Z-stacks were combined together using the MAX pixel intensity. GAPDH and DAPI channels remained untouched, while RAI1 channels went through thresholding of 0.4% of intensity to MAX intensity, and outlier pixels were removed (1 pixel size and 5 minimum intensity). Images were then turned into masked images (binary). Finally, the three channels were then combined together, and the outline of each cell was manually drawn, based on the combined GAPDH + DAPI signals. The percentage of positive RAI1 pixels in each cell was then automatically calculated and documented. Data was collected from two individual wells and combined together.

Analysis of predicted HI/TS genes
Since copy number variations are rare mutations, there is insufficient clinical data to understand dosage sensitivity for almost all genes. To address this problem, Collins et al. analyzed copy number variation data from 950,278 individuals and used machine learning to predict haploinsufficiency and triplosensitivity for all autosomal protein-coding genes within the human genome.[24] In this paper, we use data from Table S7: Haploinsufficiency and triplosensitivity predictions for all autosomal protein-coding genes. First, we screen for genes which have both a haploinsufficiency score (pHaplo) >= 0.86 and a triplosensitivity score (pTriplo) >= 0.94, indicating that the genes are predicted to be both haploinsufficient and triplosensitive (as defined by the authors). This resulted in a total of 915 genes after screening. Next, we used the Proteins REST API which provides data from UniProt to determine the protein sequence for each gene.[41] There were two genes with no return (AP000783.1, AC008443.1) which reduces the total number of dosage sensitive genes to 913. Lastly, we took the length of each protein sequence and multiplied by three to calculate the shortest possible gene length for each gene, which is an underestimate of the final sequence which would be necessary for proper expression that would need to fit within an AAV.

Mathematical modeling for activation circuit simulations
To predict the behavior of the activation feedforward circuit, we construct an ordinary differential equation model assuming Michaelis-Menten kinetics, steady-state, and quasi-equilibrium. We incorporate interactions between dCas9 and the DNA, mRNA and protein production, first-order degradation of mRNA and proteins, and dCas9 cleavage by TEV.
We define the probability of active dCas9 binding to the target DNA (e.g. the *RAI1* gene) as:

$$Pb_{dCas9} = \frac{[dCas9]}{K_D + [dCas9]} \quad (1)$$

where K_D is the dCas9 dissociation constant and [*dCas9*] is the concentration of the active dCas9 protein. When we assume that cleaved dCas9 species competes with

active dCas9 species to bind to the DNA, we add another term in the denominator to account for this competitive inhibition:

$$Pb_{dCas9} = \frac{[dCas9]}{K_D + [dCas9] + [dCas9_{cleaved}]} \tag{2}$$

Where [$dCas9_{cleaved}$] is the concentration of the cleaved dCas9 species.
Therefore, the change in corresponding target mRNA concentration is described by:

$$\frac{d[target\ mRNA]}{dt} = k_{on\ target} \cdot Pb_{dCas9} + k_{bac\ target} - k_{deg\ target\ mRNA} \cdot [target\ mRNA] \tag{3}$$

where $k_{bac\ target}$ is the background transcription rate of the target gene, $k_{on\ target}$ is the transcription rate when dCas9 is bound, and $k_{deg\ target\ mRNA}$ is the degradation rate of the target mRNA. The change in corresponding target protein concentration is described by:

$$\frac{d[target]}{dt} = k_{trans\ target} \cdot [target\ mRNA] - k_{deg\ target} \cdot [target] \tag{4}$$

where $k_{trans\ target}$ is the translation rate of the target protein and $k_{deg\ target}$ is the degradation rate of the target protein. The steady-state solutions for equation (3) and (4) are:

$$[target\ mRNA] = \frac{\left(k_{on\ target} \cdot Pb_{dCas9} + k_{bac\ target}\right)}{k_{deg\ target\ mRNA}} \tag{5}$$

$$[target] = \frac{k_{trans\ target} \cdot [target\ mRNA]}{k_{deg\ target}} \tag{6}$$

The dCas9 and TEV components are delivered to the cell such as in plasmid transfection and are therefore a function of dosage (DNA delivered). The TEV mRNA transcription rate can be represented by:

$$k_{on\ TEV\ mRNA} = a \cdot dosage \cdot DNA_{TEV} \tag{7}$$

where *a* is a constant representing the TEV transcription rate in units of concentration/(time*plasmid amount) (e.g. nM/h-ng plasmid), *dosage* is the amount of plasmid that is delivered to the cell, and DNA_{TEV} is the fraction of plasmid that encodes for TEV. The change in TEV mRNA concentration can therefore be described by:

$$\frac{d[TEV\ mRNA]}{dt} = k_{on\ TEV\ mRNA} - k_{deg\ TEV\ mRNA} \cdot [TEV\ mRNA] \tag{8}$$

Similarly, the dCas9 mRNA transcription rate is represented by:

$$k_{on\ dCas9\ mRNA} = a \cdot dosage \cdot DNA_{dCas9} \tag{9}$$

and the change in dCas9 mRNA concentration is described by:

$$\frac{d[dCas9\ mRNA]}{dt} = k_{on\ dCas9} - k_{deg\ dCas9\ mRNA} \cdot [dCas9\ mRNA] \tag{10}$$

The change in TEV protein concentration is described by:

$$\frac{d[TEV]}{dt} = k_{trans\ TEV} - k_{deg\ TEV} \cdot [TEV] \tag{11}$$

Lastly, the change in active dCas9 protein concentration is described by:

$$\frac{d[dCas9]}{dt} = k_{trans\ dCas9} \cdot [dCas9\ mRNA] - k_{deg\ dCas9} \cdot [dCas9] - k_{cut} \cdot [TEV] \cdot [dCas9] \tag{12}$$

where k_{cut} is the TEV k_{cat}/K_M, while the change in cleaved dCas9 protein concentration is described by:

$$\frac{d[Cas9_{cleaved}]}{dt} = k_{cut} \cdot [TEV] \cdot [dCas9] - k_{deg\ dCas9} \cdot [dCas9_{cleaved}] \tag{13}$$

The steady-state solutions for equations (8), (10), (11), (12), and (13) are respectively:

$$[TEV\ mRNA] = \frac{k_{on\ TEV}}{k_{deg\ TEV\ mRNA}} \tag{14}$$

$$[TEV] = \frac{k_{trans\ TEV} \cdot [TEV\ mRNA]}{k_{deg\ TEV}} \tag{15}$$

$$[dCas9\ mRNA] = \frac{k_{on\ dCas9}}{k_{deg\ dCas9\ mRNA}} \tag{16}$$

$$[dCas9] = \frac{k_{trans\ dCas9} \cdot [dCas9\ mRNA]}{k_{deg\ dCas9} + k_{cut} \cdot [TEV]} \tag{17}$$

$$[dCas9_{cleaved}] = \frac{k_{cut} \cdot [dCas9] \cdot [TEV]}{k_{deg\ dCas9}} \tag{18}$$

All numerical values, units, and references used for the parameters are found in Table S1.

Gating strategy for activation circuit titration

We observe a non-monotonic trend at high transfection levels for the running means of the activation circuit titration (SFig. 1A). Measuring the average GFP fluorescence of all transfected cells masks this non-monotonic effect (SFig. 1B). Since this non-monotonic effect biases the average GFP fluorescence of all transfected cells, we present the activation circuit titration in **Fig. 1F** as the average of transfected cells with an mCherry fluorescence below 10^7.

Increasing TEV sensitivity of the activation circuit

Another method of tuning the circuit is to change the efficiency of the TEV cut site between dCas9 and the transcriptional activator. An easy way to increase the efficiency is to use tandem TEV cut sites. Thus, we created a construct that had two TEV cut sites and compared it to our original construct which contains one (SFig. 1C). The double TEV cut site construct was much more sensitive to TEV, requiring much less to repress gene activation. We note that the double TEV cut site construct also has lower gene activation ability, possibly due to the transcriptional activator being further away from the DNA due to the longer linker.

Engineering a TEV cleavable Cas13

Mathematical modeling for repression circuit simulations

To predict the behavior of the repression feedforward circuit, we construct an ordinary differential equation model assuming Michaelis-Menten kinetics, steady-state, and quasi-equilibrium. We incorporate interactions between Cas13 and the target mRNA, mRNA and protein production, first-order degradation of mRNA and proteins, and Cas13 cleavage by TEV. Unfortunately, there is little kinetic data for CasRx in contrast to TEV. Furthermore, much of the kinetic data available measures off-target mRNA cleavage rather than on-target mRNA cleavage. Therefore, we also assume that k_{cat}/K_M for Cas13 is the same as TEV, though this is likely not the case. The change in corresponding target mRNA concentration is described by:

$$\frac{d[target\ mRNA]}{dt} = k_{bac\ target} - k_{cut\ Cas13} \cdot [Cas13] \cdot [target\ mRNA] - k_{deg\ target\ mRNA} \cdot [target\ mRNA] \tag{19}$$

where $k_{bac\ target}$ is the background transcription rate of the target gene, $k_{cut\ Cas13}$ is the k_{cut}/K_M for Cas13, and $k_{deg\ target\ mRNA}$ is the degradation rate of the target mRNA. The change in corresponding target protein concentration is described by:

$$\frac{d[target]}{dt} = k_{trans\ target} \cdot [target\ mRNA] - k_{deg\ target} \cdot [target] \tag{20}$$

where $k_{trans\ target}$ is the translation rate of the target protein and $k_{deg\ target}$ is the degradation rate of the target protein. The steady-state solutions for equation (19) and (20) are:

$$[target\ mRNA] = \frac{k_{bac\ target}}{k_{cut\ Cas13} \cdot [Cas13] + k_{deg\ target\ mRNA}} \tag{21}$$

$$[target] = \frac{k_{trans\ target} \cdot [target\ mRNA]}{k_{deg\ target}} \tag{22}$$

The Cas13 and TEV components are delivered to the cell such as in plasmid transfection and are therefore a function of dosage (DNA delivered). The TEV mRNA transcription rate can be represented by:

$$k_{on\ TEV\ mRNA} = a \cdot dosage \cdot DNA_{TEV} \tag{23}$$

where *a* is a constant representing the TEV transcription rate in units of concentration/(time*plasmid amount) (e.g. nM/h-ng plasmid), *dosage* is the amount of plasmid that is delivered to the cell, and DNA_{TEV} is the fraction of plasmid that encodes for TEV. The change in TEV mRNA concentration can therefore be described by:

$$\frac{d[TEV\ mRNA]}{dt} = k_{on\ TEV\ mRNA} - k_{deg\ TEV\ mRNA} \cdot [TEV\ mRNA] \tag{24}$$

Similarly, the Cas13 mRNA transcription rate is represented by:

$$k_{on\ Cas13\ mRNA} = a \cdot dosage \cdot DNA_{Cas13} \tag{25}$$

and the change in dCas9 mRNA concentration is described by:

$$\frac{d[Cas13\ mRNA]}{dt} = k_{on\ Cas13} - k_{deg\ Cas13\ mRNA} \cdot [Cas13\ mRNA] \tag{26}$$

The change in TEV protein concentration is described by:

$$\frac{d[TEV]}{dt} = k_{trans\ TEV} - k_{deg\ TEV} \cdot [TEV] \tag{27}$$

Lastly, the change in active dCas9 protein concentration is described by:

$$\frac{d[Cas13]}{dt} = k_{trans\ Cas13} \cdot [Cas13\ mRNA] - k_{deg\ Cas13} \cdot [dCas9] - k_{cut\ TEV} \cdot [TEV] \cdot [Cas13] \tag{28}$$

where $k_{cut\ TEV}$ is the TEV k_{cat}/K_M.

The steady-state solutions for equations (24), (26), (27), and (28) are respectively:

$$[TEV\ mRNA] = \frac{k_{on\ TEV}}{k_{deg\ TEV\ mRNA}} \tag{29}$$

$$[TEV] = \frac{k_{trans\ TEV} \cdot [TEV\ mRNA]}{k_{deg\ TEV}} \tag{30}$$

$$[Cas13\ mRNA] = \frac{k_{on\ Cas13}}{k_{deg\ Cas13\ mRNA}} \tag{31}$$

$$[Cas13] = \frac{k_{trans\ Cas13} \cdot [Cas13\ mRNA]}{k_{cut\ TEV} \cdot [TEV] + k_{deg\ Cas13}} \tag{32}$$

All numerical values, units, and references used for the parameters are found in Table S1.

Calculating the plasmid transcription rate for the simulations

To estimate the value of the plasmid transcription ate *a*, as mentioned in equation (7), (9), (23), and (25), we use experimental data and data reported for GFP fluorescence in the supplementary information of a paper previously published by the Gao lab.[42] We conducted an experiment where 10 ng of CMVTO-GFP was transfected for a single 24-well of HEK293 cells and measured an average GFP fluorescence of 1.3*10^6 A.U. From Fig. S2A and S2C from Kaseniit *et al*., we estimate that 300 copies of GFP mRNA corresponds to a measured fluorescence of 7*10^5 A.U. (we multiply the reported value by 100 due to differences in normalization between analysis software). 300 copies of mRNA is 0.5 nM of mRNA using Avogadro's number and an estimated cell volume of 1 pL. Assuming a linear relationship between average GFP fluorescence and ng of DNA transfected, we would estimate that an average GFP fluorescence of $7*10^5$ would correspond to 5.38 ng of plasmid transfected. Since these measurements were taken at steady state, we calculate *a* using the following equation set equal to zero:

$$\frac{d[plasmid\ mRNA]}{dt} = a \cdot (ng\ plasmid) - k_{deg}[plasmid\ mRNA] = 0 \tag{33}$$

$$a = \frac{k_{deg}[plasmid\ mRNA]}{[plasmid]} \tag{34}$$

We estimate the ng plasmid per cell by dividing 5.38 ng by $0.24*10^6$, the number of cells within a confluent well of a 24-well plate,[43] to determine (*ng plasmid*). We previously calculated [*plasmid mRNA*] to be 0.5 nM. Lastly, we use an estimate for the median mRNA degradation rate (Table S1). We then solve for *a* resulting in $1.71*10^3$ nM/(h-ng plasmid).

Supplemental Excel table titles and legends

Supplementary Table 1: Model Parameters

References

1. Rice, A.M., and McLysaght, A. (2017). Dosage-sensitive genes in evolution and disease. BMC Biol. *15*, 78. 10.1186/s12915-017-0418-y.
2. Huang, W.-H., Guenthner, C.J., Xu, J., Nguyen, T., Schwarz, L.A., Wilkinson, A.W., Gozani, O., Chang, H.Y., Shamloo, M., and Luo, L. (2016). Molecular and Neural Functions of Rai1, the Causal Gene for Smith-Magenis Syndrome. Neuron *92*, 392–406. 10.1016/j.neuron.2016.09.019.
3. Javed, S., Lee, Y.-J., Xu, J., and Huang, W.-H. (2021). Temporal dissection of Rai1 function reveals brain-derived neurotrophic factor as a potential therapeutic target for Smith-Magenis syndrome. Hum. Mol. Genet. *31*, 275–288. 10.1093/hmg/ddab245.
4. Potocki, L., Chen, K.S., Park, S.S., Osterholm, D.E., Withers, M.A., Kimonis, V., Summers, A.M., Meschino, W.S., Anyane-Yeboa, K., Kashork, C.D., et al. (2000). Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. Nat. Genet. *24*, 84–87. 10.1038/71743.
5. Walz, K., Paylor, R., Yan, J., Bi, W., and Lupski, J.R. (2006). Rai1 duplication causes physical and behavioral phenotypes in a mouse model of dup(17)(p11.2p11.2). J. Clin. Invest. *116*, 3035–3041. 10.1172/JCI28953.
6. Smith, A.C., McGavran, L., Robinson, J., Waldstein, G., Macfarlane, J., Zonona, J., Reiss, J., Lahr, M., Allen, L., and Magenis, E. (1986). Interstitial deletion of (17)(p11.2p11.2) in nine patients. Am. J. Med. Genet. *24*, 393–414. 10.1002/ajmg.1320240303.
7. Slager, R.E., Newton, T.L., Vlangos, C.N., Finucane, B., and Elsea, S.H. (2003). Mutations in RAI1 associated with Smith-Magenis syndrome. Nat. Genet. *33*, 466–468. 10.1038/ng1126.
8. Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat. Genet. *23*, 185–188. 10.1038/13810.
9. Wan, Y., Cohen, J., Szenk, M., Farquhar, K.S., Coraci, D., Krzysztoń, R., Azukas, J., Van Nest, N., Smashnov, A., Chern, Y.-J., et al. (2023). Nonmonotone invasion landscape by noise-aware control of metastasis activator levels. Nat. Chem. Biol. *19*, 887–899. 10.1038/s41589-023-01344-z.
10. Reeves, G.T. (2019). The engineering principles of combining a transcriptional incoherent feedforward loop with negative feedback. J. Biol. Eng. *13*, 62. 10.1186/s13036-019-0190-3.
11. Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. Cell *138*, 760–773. 10.1016/j.cell.2009.06.013.
12. Strovas, T.J., Rosenberg, A.B., Kuypers, B.E., Muscat, R.A., and Seelig, G. (2014). MicroRNA-based single-gene circuits buffer protein synthesis rates against perturbations. ACS Synth. Biol. *3*, 324–331. 10.1021/sb4001867.
13. Segall-Shapiro, T.H., Sontag, E.D., and Voigt, C.A. (2018). Engineered promoters enable

constant gene expression at any copy number in bacteria. Nat. Biotechnol. *36*, 352–358. 10.1038/nbt.4111.

14. Yang, J., Lee, J., Land, M.A., Lai, S., Igoshin, O.A., and St-Pierre, F. (2021). A synthetic circuit for buffering gene dosage variation between individual mammalian cells. Nat. Commun. *12*, 4132. 10.1038/s41467-021-23889-0.
15. Mallozzi, A., Fusco, V., Ragazzini, F., and di Bernardo, D. (2024). The CRISPRaTOR: a biomolecular circuit for Automatic Gene Regulation in Mammalian Cells with CRISPR technology. BioRxiv. 10.1101/2024.03.30.587417.
16. Love, K.S., Johnstone, C.P., Peterman, E.L., Gaglione, S., and Galloway, K.E. (2024). Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells. BioRxiv. 10.1101/2024.06.25.600629.
17. Shimoga, V., White, J.T., Li, Y., Sontag, E., and Bleris, L. (2013). Synthetic mammalian transgene negative autoregulation. Mol. Syst. Biol. *9*, 670. 10.1038/msb.2013.27.
18. Du, R., Flynn, M.J., Honsa, M., Jungmann, R., and Elowitz, M.B. (2024). miRNA circuit modules for precise, tunable control of gene expression. BioRxiv. 10.1101/2024.03.12.583048.
19. Jones, R.D., Qian, Y., Siciliano, V., DiAndreth, B., Huh, J., Weiss, R., and Del Vecchio, D. (2020). An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. Nat. Commun. *11*, 5690. 10.1038/s41467-020-19126-9.
20. Flynn, M.J., Mayfield, A.M.H., Du, R., Gradinaru, V., and Elowitz, M.B. (2024). Synthetic dosage-compensating miRNA circuits allow precision gene therapy for Rett syndrome. BioRxiv. 10.1101/2024.03.13.584179.
21. Huang, X., Kong, N., Zhang, X., Cao, Y., Langer, R., and Tao, W. (2022). The landscape of mRNA nanomedicine. Nat. Med. *28*, 2273–2287. 10.1038/s41591-022-02061-1.
22. Gao, X.J., Chong, L.S., Kim, M.S., and Elowitz, M.B. (2018). Programmable protein circuits in living cells. Science *361*, 1252–1258. 10.1126/science.aat5062.
23. Fink, T., Lonzarić, J., Praznik, A., Plaper, T., Merljak, E., Leben, K., Jerala, N., Lebar, T., Strmšek, Ž., Lapenta, F., et al. (2019). Design of fast proteolysis-based signaling and logic circuits in mammalian cells. Nat. Chem. Biol. *15*, 115–122. 10.1038/s41589-018-0181-6.
24. Collins, R.L., Glessner, J.T., Porcu, E., Lepamets, M., Brandon, R., Lauricella, C., Han, L., Morley, T., Niestroj, L.-M., Ulirsch, J., et al. (2022). A cross-disorder dosage sensitivity map of the human genome. Cell *185*, 3041-3055.e25. 10.1016/j.cell.2022.06.036.
25. Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., and Hsu, P.D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell *173*, 665-676.e14. 10.1016/j.cell.2018.02.033.
26. Tycko, J., Van, M., Aradhana, DelRosso, N., Yao, D., Xu, X., Ludwig, C., Spees, K., Liu, K., Hess, G.T., et al. (2023). Development of compact transcriptional effectors using high-throughput measurements in diverse contexts. BioRxiv. 10.1101/2023.05.12.540558.

27. Cohen, R.N., van der Aa, M.A.E.M., Macaraeg, N., Lee, A.P., and Szoka, F.C. (2009). Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. J. Control. Release *135*, 166–174. 10.1016/j.jconrel.2008.12.016.

28. Gupta, R., Ghosh, A., Chakravarti, R., Singh, R., Ravichandiran, V., Swarnakar, S., and Ghosh, D. (2022). Cas13d: A new molecular scissor for transcriptome engineering. Front. Cell Dev. Biol. *10*, 866800. 10.3389/fcell.2022.866800.

29. Shi, P., Murphy, M.R., Aparicio, A.O., Kesner, J.S., Fang, Z., Chen, Z., Trehan, A., Guo, Y., and Wu, X. (2023). Collateral activity of the CRISPR/RfxCas13d system in human cells. Commun. Biol. *6*, 334. 10.1038/s42003-023-04708-2.

30. Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci. *20*, 1298–1345. 10.1002/pro.666.

31. Chang, H.-C., Lee, Y.-J., Javed, S., Haque, M., Chang, Y.-T., Lin, Y.C., Oram, C., and Huang, W.-H. (2023). rAAV-CRISPRa therapy corrects Rai1 haploinsufficiency and rescues selective disease features in Smith-Magenis syndrome mice. J. Biol. Chem. *299*, 102728. 10.1016/j.jbc.2022.102728.

32. Choi, H.M.T., Schwarzkopf, M., Fornace, M.E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., and Pierce, N.A. (2018). Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. Development *145*. 10.1242/dev.165753.

33. Carmona-Mora, P., Canales, C.P., Cao, L., Perez, I.C., Srivastava, A.K., Young, J.I., and Walz, K. (2012). RAI1 transcription factor activity is impaired in mutants associated with Smith-Magenis Syndrome. PLoS ONE *7*, e45155. 10.1371/journal.pone.0045155.

34. McClements, M.E., and MacLaren, R.E. (2017). Adeno-associated Virus (AAV) Dual Vector Strategies for Gene Therapy Encoding Large Transgenes. Yale J. Biol. Med. *90*, 611–623.

35. Albini, S., Palmieri, L., Dubois, A., Bourg, N., Lostal, W., and Richard, I. (2023). Assessment of therapeutic potential of a dual AAV approach for duchenne muscular dystrophy. Int. J. Mol. Sci. *24*. 10.3390/ijms241411421.

36. Wang, J., Lin, J., Chen, Y., Liu, J., Zheng, Q., Deng, M., Wang, R., Zhang, Y., Feng, S., Xu, Z., et al. (2023). An ultra-compact promoter drives widespread neuronal expression in mouse and monkey brains. Cell Rep. *42*, 113348. 10.1016/j.celrep.2023.113348.

37. Preece, R., Georgiadis, C., Gkazi, S.A., Etuk, A., Christi, A., and Qasim, W. (2020). “Mini” U6 Pol III promoter exhibits nucleosome redundancy and supports multiplexed coupling of CRISPR/Cas9 effects. Gene Ther. *27*, 451–458. 10.1038/s41434-020-0142-z.

38. Xu, X., Chemparathy, A., Zeng, L., Kempton, H.R., Shang, S., Nakamura, M., and Qi, L.S. (2021). Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. Mol. Cell *81*, 4333-4345.e4. 10.1016/j.molcel.2021.08.008.

39. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832. 10.1038/nbt.2647.

40. Kemaladewi, D.U., Bassi, P.S., Erwood, S., Al-Basha, D., Gawlik, K.I., Lindsay, K., Hyatt, E., Kember, R., Place, K.M., Marks, R.M., et al. (2019). A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. Nature *572*, 125–130. 10.1038/s41586-019-1430-x.

41. Nightingale, A., Antunes, R., Alpi, E., Bursteinas, B., Gonzales, L., Liu, W., Luo, J., Qi, G., Turner, E., and Martin, M. (2017). The Proteins API: accessing key integrated protein and genome information. Nucleic Acids Res. *45*, W539–W544. 10.1093/nar/gkx237.

42. Kaseniit, K.E., Katz, N., Kolber, N.S., Call, C.C., Wengier, D.L., Cody, W.B., Sattely, E.S., and Gao, X.J. (2023). Modular, programmable RNA sensing using ADAR editing in living cells. Nat. Biotechnol. *41*, 482–487. 10.1038/s41587-022-01493-x.

43. Useful Numbers for Cell Culture | Thermo Fisher Scientific - US https://www.thermofisher.com/us/en/home/references/gibco-cell-culture-basics/cell-culture-protocols/cell-culture-useful-numbers.html.

44. Sims, C.E., and Allbritton, N.L. (2007). Analysis of single mammalian cells on-chip. Lab Chip *7*, 423–440. 10.1039/b615235j.

45. Xu, C., Zhou, Y., Xiao, Q., He, B., Geng, G., Wang, Z., Cao, B., Dong, X., Bai, W., Wang, Y., et al. (2021). Programmable RNA editing with compact CRISPR-Cas13 systems from uncultivated microbes. Nat. Methods *18*, 499–506. 10.1038/s41592-021-01124-4.

46. Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342. 10.1038/nature10098.

47. Sun, P., Austin, B.P., Tözsér, J., and Waugh, D.S. (2010). Structural determinants of tobacco vein mottling virus protease substrate specificity. Protein Sci. *19*, 2240–2251. 10.1002/pro.506.

48. Josephs, E.A., Kocak, D.D., Fitzgibbon, C.J., McMenemy, J., Gersbach, C.A., and Marszalek, P.E. (2015). Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage. Nucleic Acids Res. *43*, 8924–8941. 10.1093/nar/gkv892.