iScience, Volume *23*

Supplemental Information

Robust Filtering and Noise Suppression in Intragenic miRNA-Mediated Host Regulation

Taek Kang, Tyler Quarton, Chance M. Nowak, Kristina Ehrhardt, Abhyudai Singh, Yi Li, and Leonidas Bleris

Reference	Sequence
sgRNA 1 (5')	**AGCTTGTCGACGATATCTCC**
sgRNA 2 (3')	**CAAACGATATGGGCTGAATAC**
FF3 target	**TTTGCTATACCCGATCTTA**

Figure S1. Guide RNA sequences for removing FF3 miRNA targets. Related to Figure 1**.** In order to remove the FF3 miRNA target in the genomically-integrated synthetic circuit, the nuclease was targeted directly to the target site with 2 different sgRNAs. The sgRNA targets are shown in green brackets.

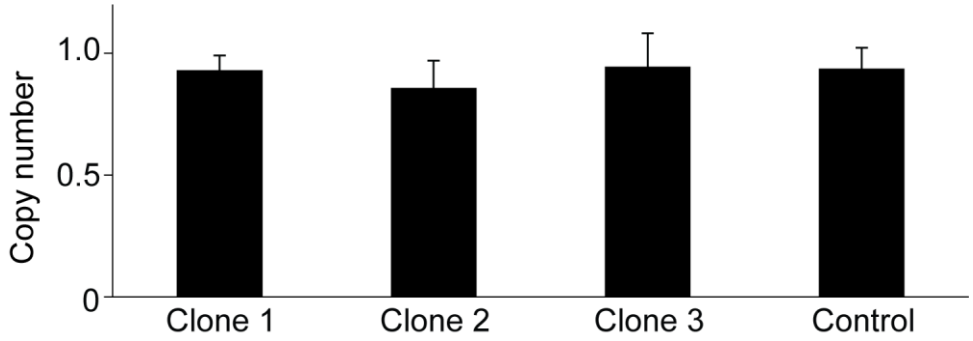


Figure S2. Calculated copy numbers for each transgenic clones and control. Related to Transparent Methods.

For each stable clone, 50 ng of genomic DNA were used performed and the average copy numbers were calculated as the mean ± SD. For statistical analysis, z scores were calculated against estimated integer copy numbers, and $-1.96<z<1.96$ was determined as no statistical difference (corresponding to 95% confidence interval).

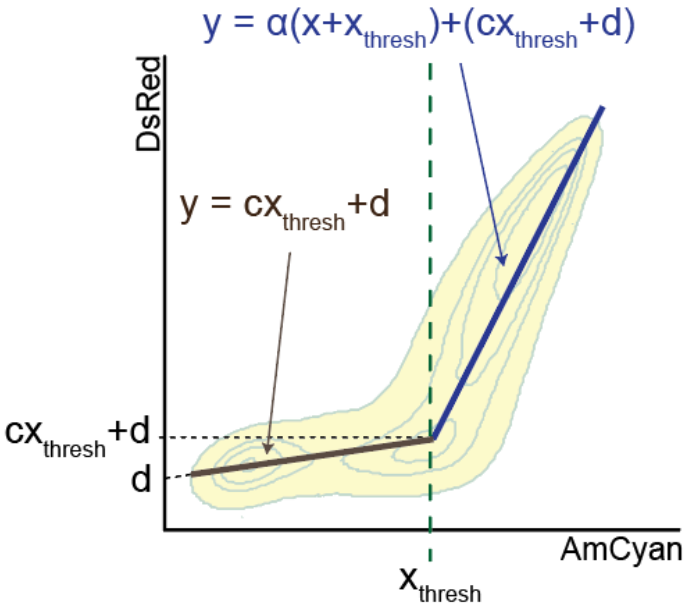


Figure S3. Graphical illustration of the promoter activity threshold calculation. Related to Figure 2. Flow cytometry population of the stable cell lines can be separated into population above and below the DsRed expression threshold (x_{thresh}). Each of these populations can be described by a linear curve after fitting.

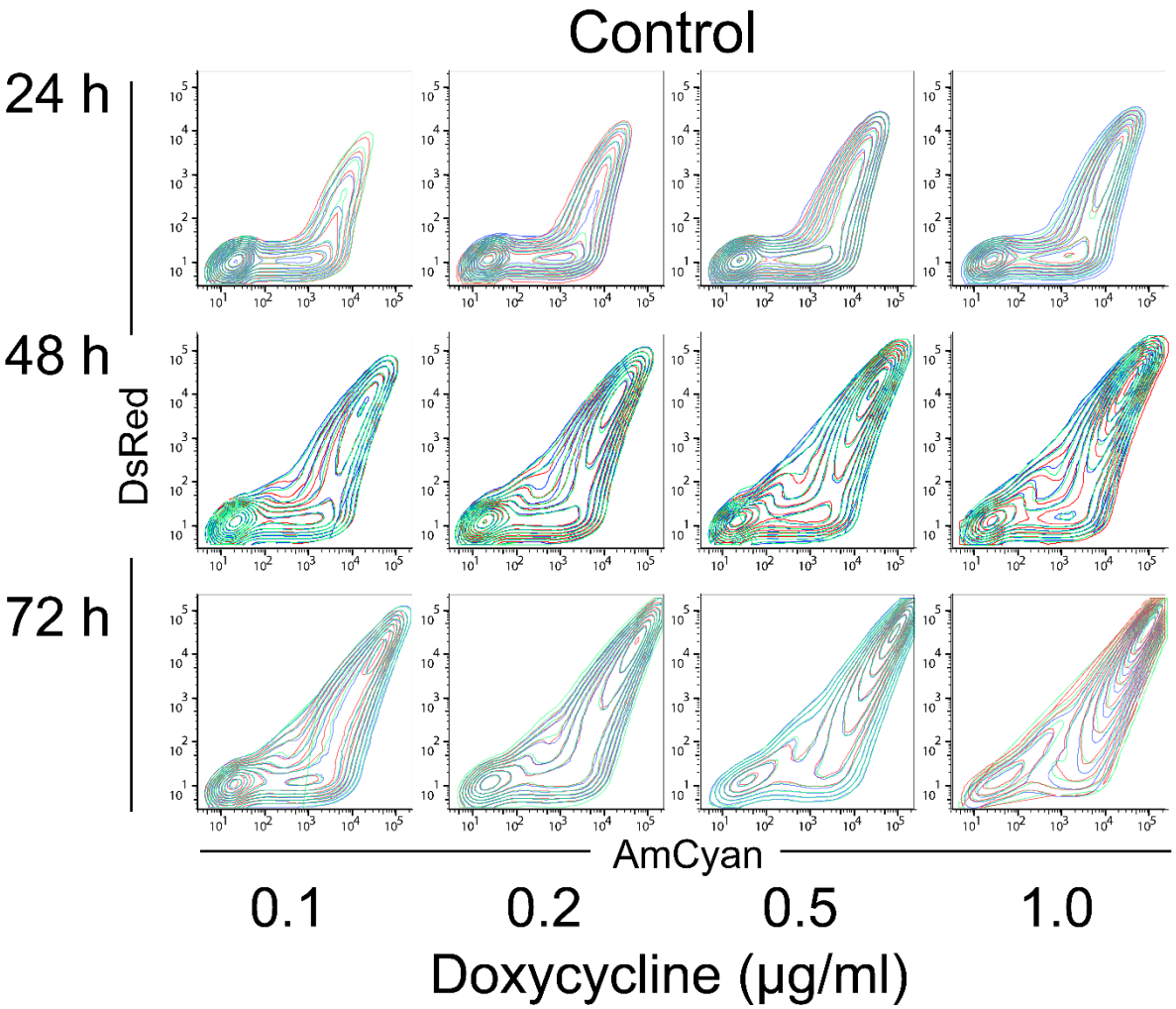


Figure S4. Flow cytometry measurement of populations without intragenic miRNA regulated DsRed. Related to Figure 2. Each panel contains the triplicates indicated by green, red and blue contours.

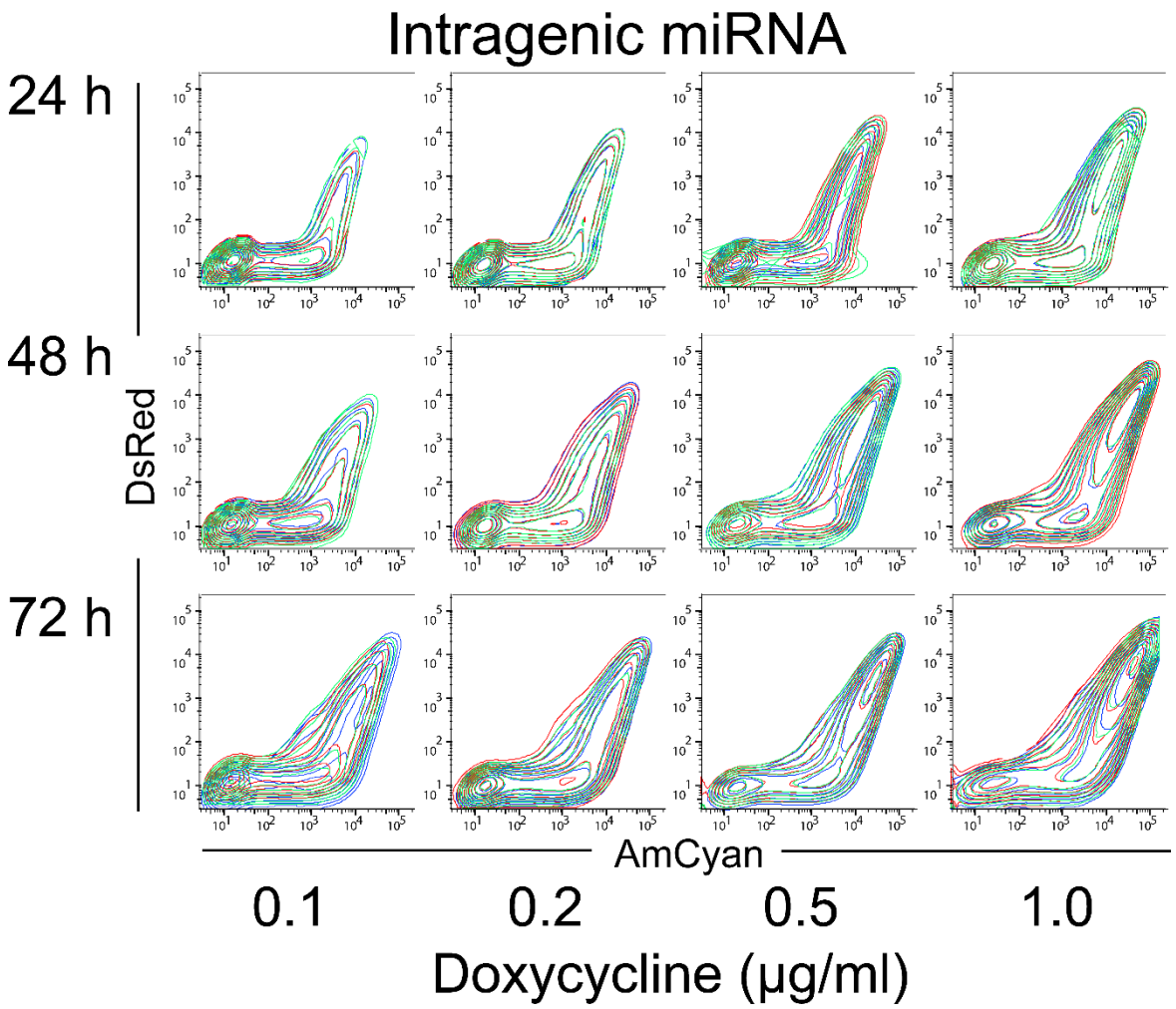


Figure S5. Flow cytometry measurement of populations intragenic miRNA regulated DsRed. Related to Figure 2**.** Each panel contains the triplicates indicated by green, red and blue contours.

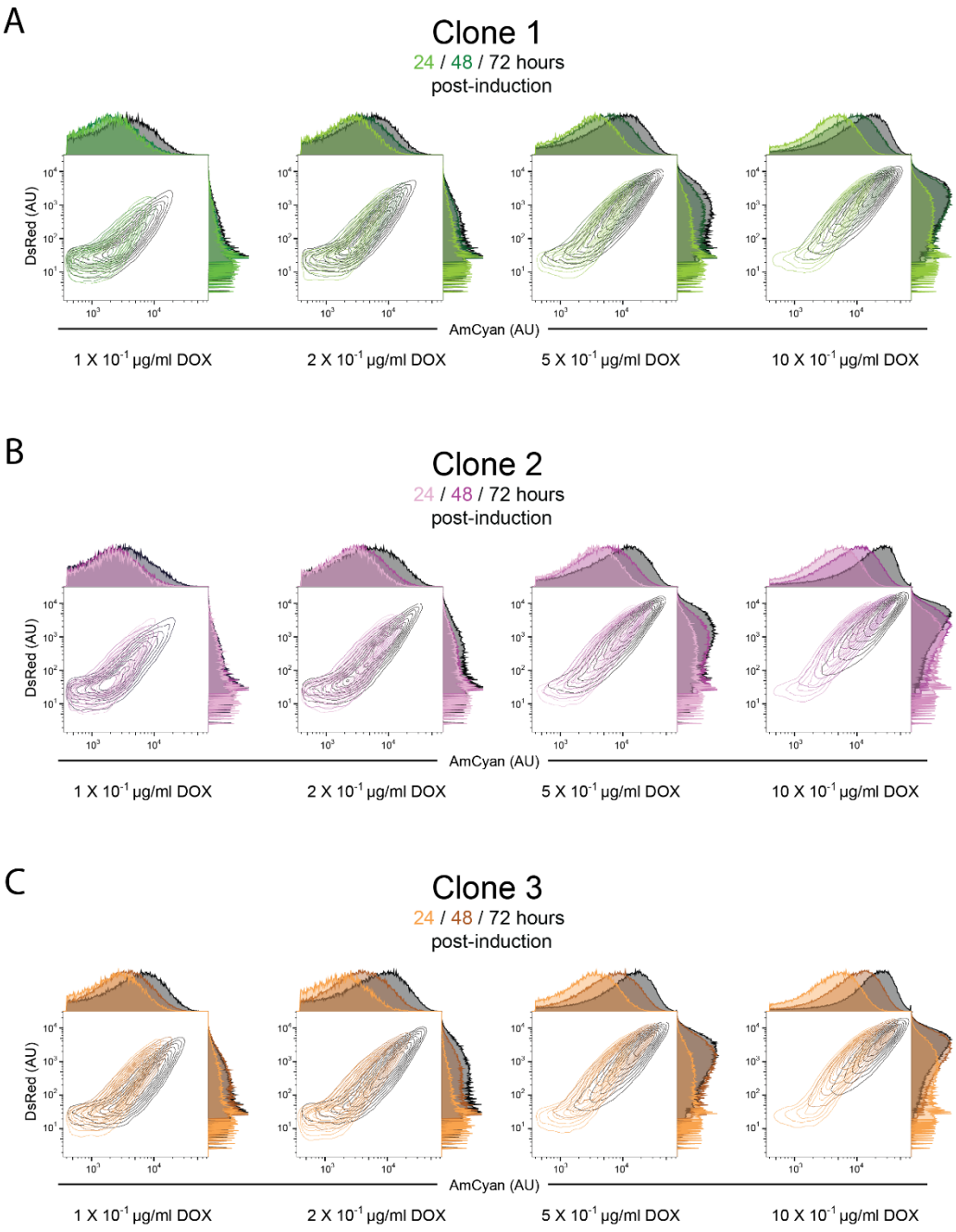


Figure S6. Doxycycline titration of the synthetic gene circuit for three isogenic clones. Related to Figure 3. For each isogenic clones containing the synthetic gene circuit, the fluorescent reporter expression was analyzed at 4 different doxcycyline levels (0.1, 0.2, 0.5, 1.0 µg, left to right, respectively) at 24, 48 and 72 hours.

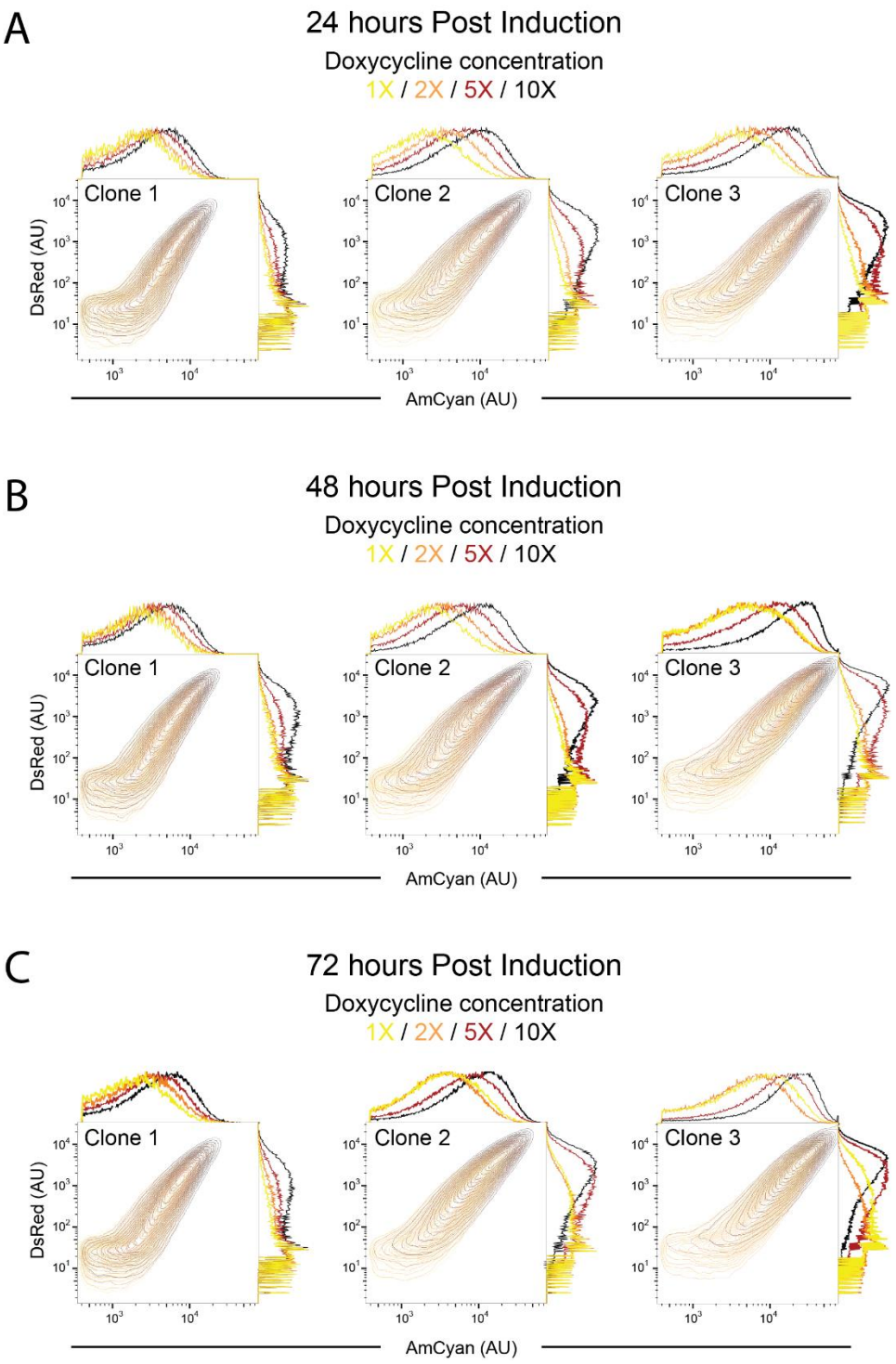


Figure S7. Doxycycline titration of the synthetic gene circuit for three different time points. Related to Figure 3. See Figure S6 for description. Doxycycline concentrations of 1X, 2X, 5X,10X correspond to 0.1, 0.2, 0.5, 1.0 µg, respectively).

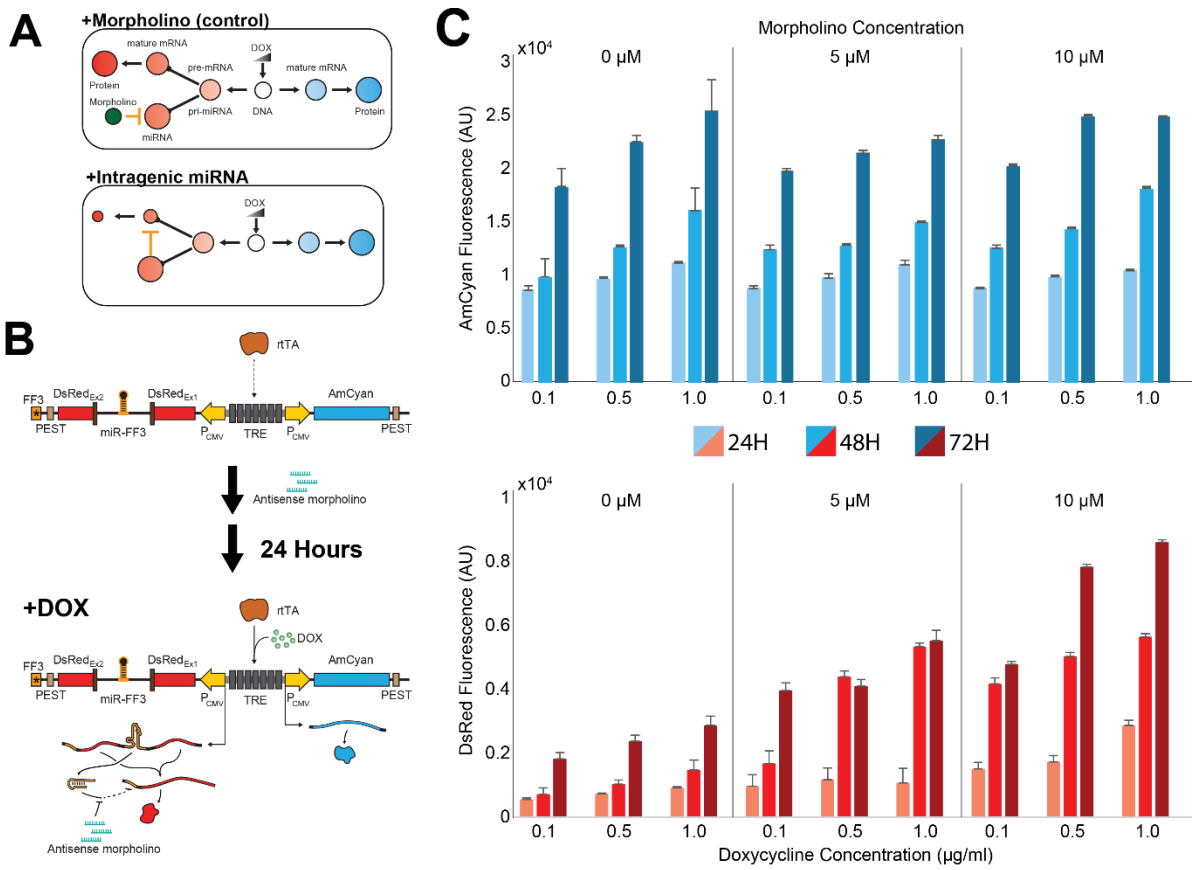


Figure S8. Steric inhibition of intragenic FF3 miRNA action via Morpholino. Related to Figure 4. (A) Morpholino oligomers are nucleic acid analogs that sterically inhibit RNA activity. (B) Customized Morpholino oligomers antisense to mature FF3 miRNA were introduced 24 hours to the cell lines harboring the synthetic construct. (C) Flow cytometry results of the circuits 24, 48, and 72 hours after induction.

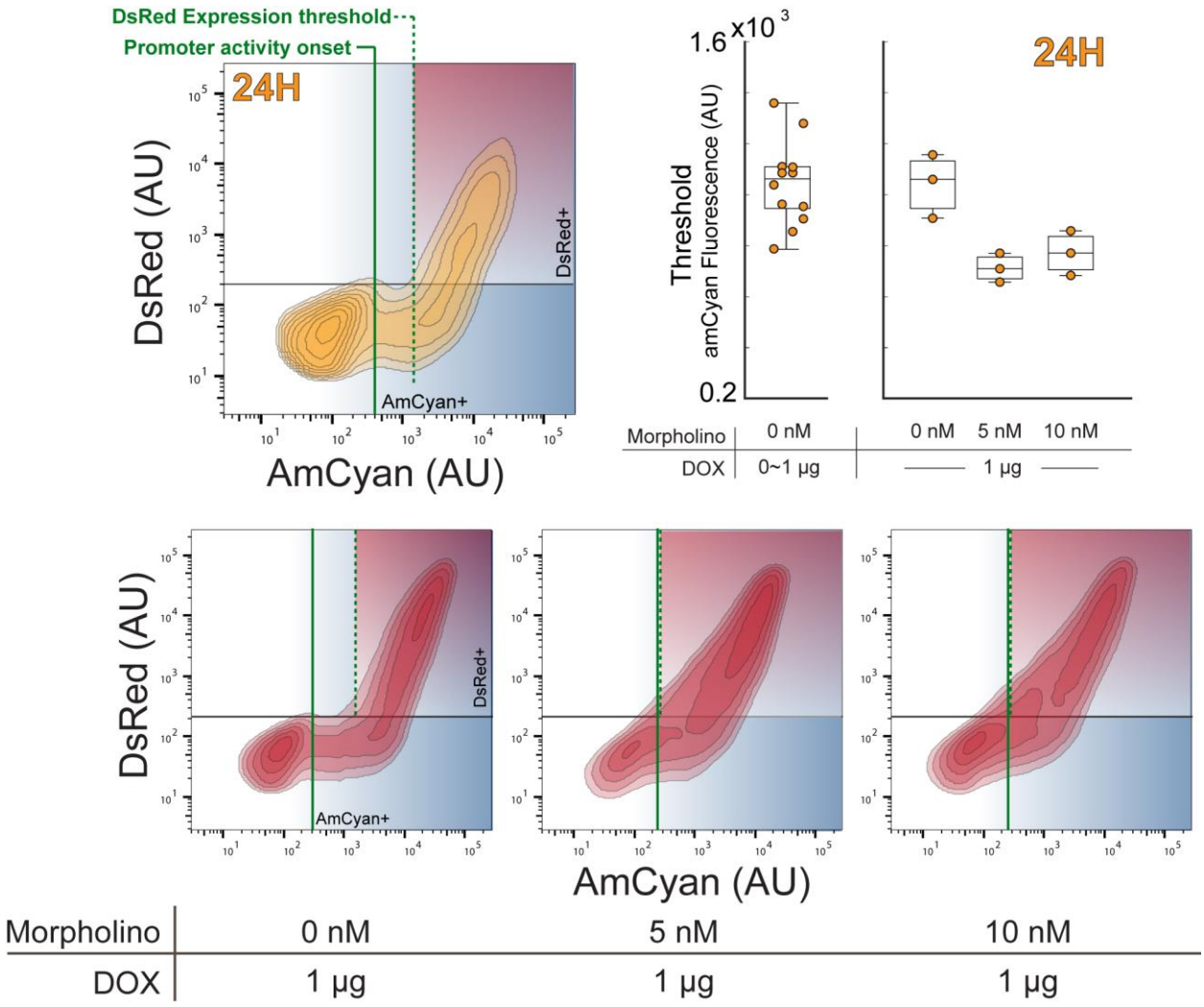


Figure S9. Steric inhibition of intragenic FF3 miRNA action via Morpholino alters DsRed expression threshold. Related to Figure 2. With intragenic miRNA regulation, the threshold promoter activity required for active DsRed expression, as indicated by AmCyan fluorescence, (green dotted line, top left) is reached after the onset of promoter activity (green solid line). This threshold decreases with the addition of Morpholino oligomer 24 hours after induction (top right). The boxplot on the left indicates data presented in the Figure 2C. Eventually, measured at 48h and 72h, the threshold disappears and DsRed expression is observed immediately at the onset of promoter activity (bottom).

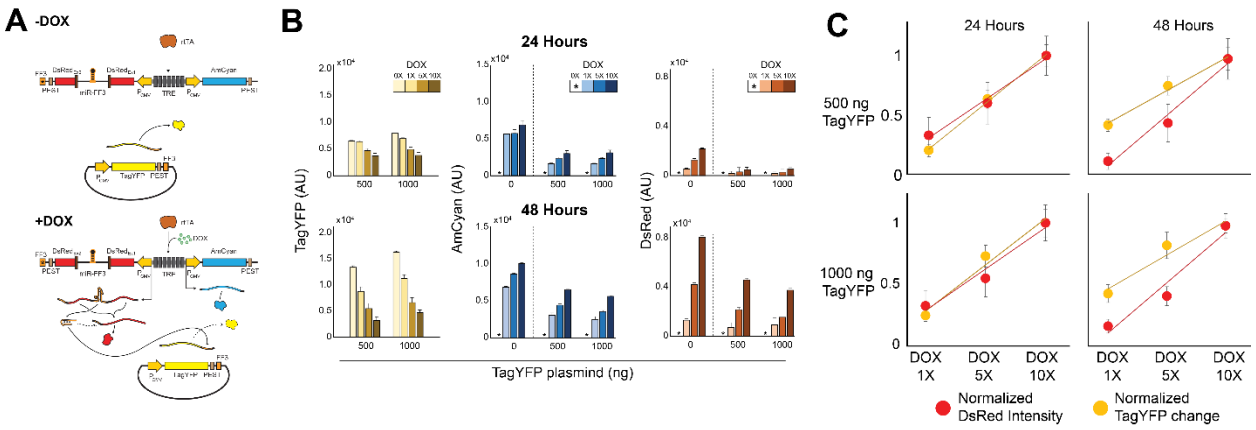


Figure S10. Isolated effect of intragenic FF3 miRNA, as measured by an independent fluorescent reporter. Related to Figure 4. (A) TagYFP fluorescent reporter plasmid with constitutive expression was introduced to the control circuit where DsRed FF3 target has been scrambled (top). Once induced, the mature FF3 miRNA directly effects TagYFP production. (B) Fluorescence of cells expressing all three fluorescent reporters. (C) DsRed fluorescence intensity at each doxycycline concentration plotted alongside the magnitude of decrease in TagYFP at the corresponding doxycycline concentration.

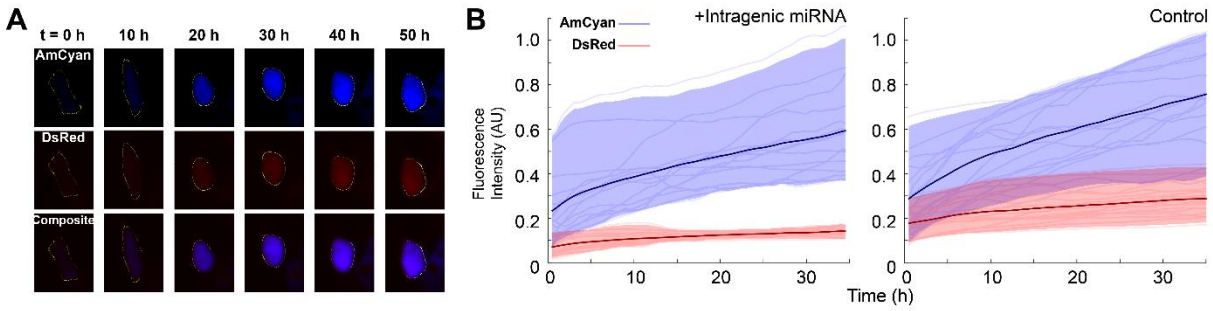


Figure S11. Single-cell time lapse microscopy of the synthetic circuit. Related to Figure 4. Fluorescence activity of cells harboring the synthetic gene circuit were captured every 30 minutes for 72 hours. (A) Snapshots of representative fluorescent microscopy images. Rows represent fluorescence of each cells and columns represent various time points. (B) Plot composed of 20 individual traces single-cells tracked minimum of 35 hours. Dark colored line and the shaded area indicates mean fluorescence intensity and 95% confidence interval for each time point.

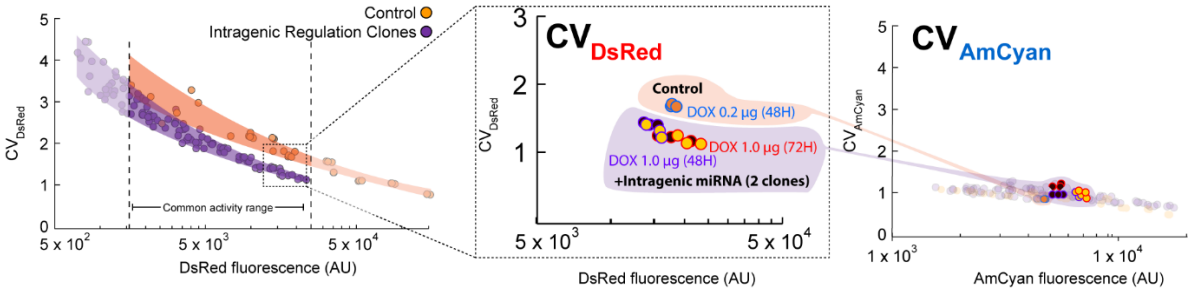


Figure S12. Comparison of fluorescent reporter noises at a given DsRed fluorescence. Related to Figure 5. Noise of expression in DsRed for all doxycycline concentration and time (left). Doxycycline concentration and time period for comparable expression but different noise for DsRed expression (center). Noise of AmCyan expression for corresponding population (right).

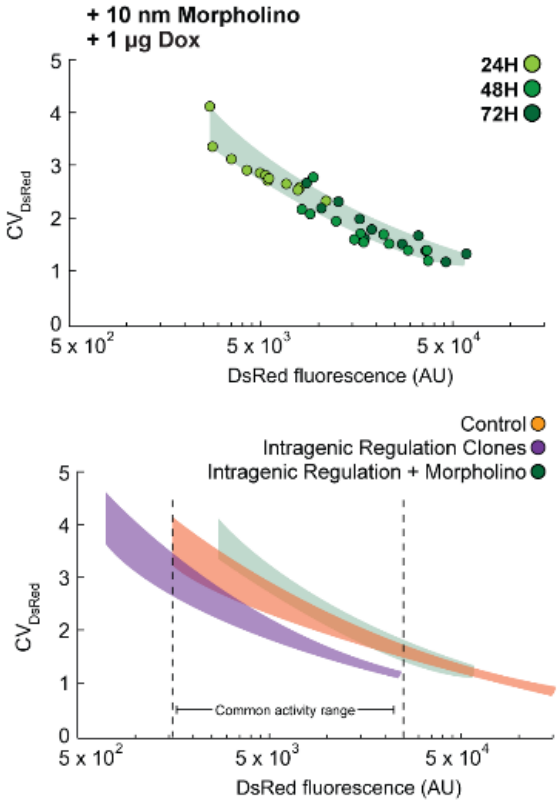


Figure S13. Steric inhibition of intragenic FF3 miRNA action via Morpholino reduces expression noise. Related to Figure 6. Mean expression of the DsRed plotted against the coefficient of variation. Aggregated plot of DsRed noise vs. mean expression for three isogenic clones with intragenic regulation (purple), control population with disrupted miRNA target (orange) and Morpholino (green).

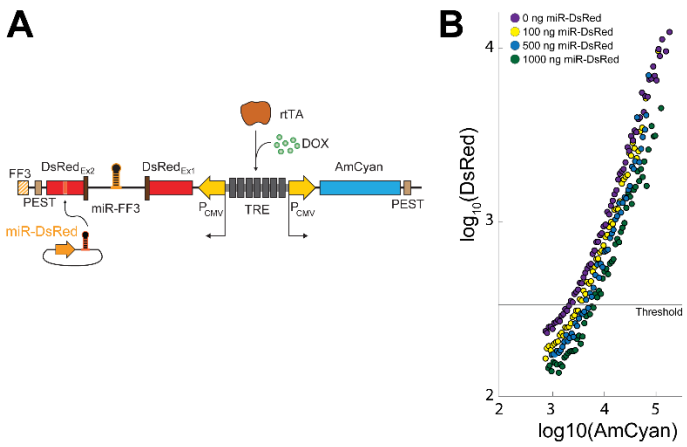
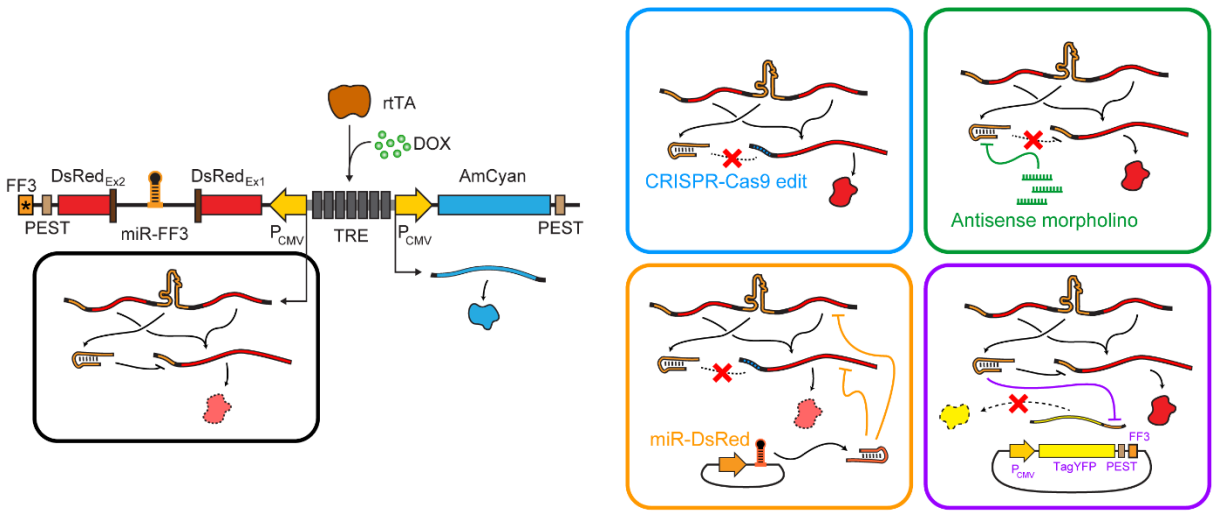


Figure S14. Effect of intergenic miRNA targeting DsRed. Related to Figure 2. (A) The stable cell line with scrambled FF3 miRNA was transfected with varying amounts of constitutively-expressed miRNA expression plasmid. The intergenic miRNA is fully complementary to the exon of DsRed. The promoter was induced with a saturating concentration of doxycycline (1 µg/ml) and the fluorescence was measured 48 hours post induction. (B) Scatter plot representing the flow cytometry results of the experiment. Each data point represents a mean fluorescence of the binned population. Increasing amounts of intergenic miRNA plasmid causes the AmCyan threshold for a given DsRed expression to increase.



Circuit Type Modification	Experimental Active intragenic miR	Control Disrupted miR target	Morpholino Steric miR blocking	DsRed miRNA Intergenic miRNA
Presence of a threshold	Yes	Yes	Yes	No
Threshold robustness	Yes; to time and dox	No	No	N/A
Noise	Low	High	High	N/A

Figure S15. Summary of control experiments performed for intragenic miRNA and subsequent effect on threshold and noise. Related to Figure 1. DsRed expression noise was not calculated for the experimental diagram shown in the purple box. The detailed results of this experiments are shown in Figure S12.

List of intragenic miRNAs		
Known protein-coding genes		20 530
Known miRNA precursors		1871
	Intragenic miRNAs	**1072**
	Host genes	930
	Sense miRNAs in respect to host orientation	902
	Antisense miRNAs in respect to host orientation	170
	Intergenic miRNAs	**799**
Expressed coding genes		18 442
Expressed miRNAs		1111

Table S1. Distribution of intragenic miRNAs in human genome. Related to Figure 1. (Hinske et al., 2014)

Left homology arm from the adeno-associated virus integration site (5' HA)	GAGCACTTCCTTCTCGGCGCTGCACCACGTGATGTCC TCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGG GCATCTCTCCTCCCTCACCCAACCCCATGCCGTCTTC ACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCT GTGCCATCTCTCGTTTCTTAGGATGGCCTTCTCCGAC GGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAG GCCTGCATCATCACCGTTTTTCTGGACAACCCCAAAGT ACCCCGTCTCCCTGGCTTTAGCCACCTCTCCATCCTC TTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGAT TCGGGTCACCTCTCACTCCTTTCATTTGGGCAGCTCC CCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTC CAGCCCCCTGTCATGGCATCTTCCAGGGGTCCGAGA GCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTAT GTCCACTTCAGGACAGCA
Right homology arm from the adeno-associated virus integration site (3' HA)	ACTAGGGACAGGATTGGTGACAGAAAAGCCCCATCCT TAGGCCTCCTCCTTCCTAGTCTCCTGATATTGGGTCTA ACCCCCACCTCCTGTTAGGCAGATTCCTTATCTGGTG ACACACCCCCATTTCCTGGAGCCATCTCTCTCCTTGC CAGAACCTCTAAGGTTTGCTTACGATGGAGCCAGAGA GGATCCTGGGAGGGAGAGCTTGGCAGGGGGTGGGA GGGAAGGGGGGGATGCGTGACCTGCCCGGTTCTCAG TGGCCACCCTGCGCTACCCTCTCCCAGAACCTGAGCT GCTCTGACGCGGCCGTCTGGTGCGTTTCACTGATCCT GGTGCTGCAGCTTCCTTACACTTCCCAAGAGGAGAAG CAGTTTGGAAAAACAAAATCAGAATAAGTTGGTCCTGA GTTCTAACTTTGGCTCTTCACCTTTCTAGTCCCCAATTT ATATTGTTCCTCCGTGCGTCAGTTTTACCTGTGAGATA AGGCCAGTAGCCAGCCCCGTCCTGGCAGGGCTGTGG TGAGGAGGGGGGTGTCCGTGTGGAAAACTCCCTTTGT GAGAATGGTGCGTCCTAGGTGTTCACCAGGTCGTGGC CGCCTCTACTCCCTTTCTCTTTCTCCATCCTTCTTTCCT TAAAGAGTCCCCAGTGCTATCTGGGACATATTCCTCC GCCCAGAGCAGGGTCCCGCTTCCCTAAGGCCCTGCT CTGGGCTTCTGGGTTTGAGTCCTTGGCAAGCCCAGGA GAGGCGCTCAGGCTTCCCTGTCCCCCTTCCTCGTCCA CCATCTCATGCCCCTGGCTCTCCTGCCCCTTCCCTAC AGGGGTTCCTGGCTCTGCTCT
AAVS1-specific sgRNA	GTCCCCTCCACCCCACAGT

Table S2. Integration of the circuit plasmid into AAVS1 locus. Related to Figure 1. Left and right homology arm of the plasmid used to gnomically integrate synthetic construct, and the AAVS1-specific sgRNA used to induce NHEJ.

DsRed forward	5’-ctccaccacggtgtagtcct-3’
DsRed reverse	5’-agaccgtgtacaaggccaag-3’
BRCA1 forward	5’-gagcgtcccctcacaaataa-3’
BRCA1 reverse	5’-tgctccgtttggttagttcc-3’

Table S3. **Primers used for qPCR-based transgene copy number validation.** Related to Figure 1.

Transparent Methods

Cell Culture: Tet-On cells and the derived stable clones were maintained at 37°C, 5% CO_2 and 100% humidity. The cells were grown in Dulbecco's modified Eagle's medium (DMEM, Life Technologies, #11965-118) supplemented with 10% Fetal Bovine Serum (FBS, Atlanta Biologicals, #S11550), 0.1 mM MEM non-essential amino acids (Life Technologies, #11140-050), 0.045 units/mL of Penicillin and 0.045 units/mL of Streptomycin (Penicillin-Streptomycin liquid, Life Technologies, #15140-122). To pass the cells, the adherent culture was washed with DPBS (Dulbecco's Phosphate-Buffered Saline, 1X with calcium and magnesium, Corning, #21-030-CM) and then trypsinized with Trypsin-EDTA (Trypsin-EDTA (0.25%), phenol red, Life Technologies, #25200-114) and finally diluted in fresh medium upon reaching 50–90% confluence. The derived stable cells were maintained in 2ug/mL puromycin.

Generation of stable cell line: To generate the monoclonal stable cell lines, ~10 million of the Tet-On cells were seeded onto a 10 cm petri dish. 16 hours later, the cells were transiently transfected with 5 μg of the donor plasmid, 5 μg of the U6-gRNA-PCMV-Cas9-t2A-mKate plasmid using the JetPRIME reagent (Polyplus Transfection). 48 hours later, puromycin (Life Technologies, catalog number: A1113803) was added at the final concentration of 2 μg/mL. The selection lasted ~2 weeks, after which the surviving clones were pooled to generate the polyclonal stable cells, from which monoclonal stable cells were isolated using the serial dilution method.

Fluorescence Microscopy: Microscopy was performed 24–72 hours post induction with doxycycline. The live cells were grown on 12-well plates (Greiner Bio-One) in the complete medium. Cells were imaged using an Olympus IX81 microscope in a Precision Control environmental chamber. The images were captured using a Hamamatsu ORCA-03 Cooled monochrome digital camera. The filter sets (Chroma) are as follows: ET436/20x (excitation) and ET480/40 m (emission) for AmCyan, ET545/30x (excitation) and ET620/60 m (emission) for DsRed. Data collection and processing was performed in the software package Slidebook 5.0. All images within a given experimental set were collected with the same exposure times and underwent identical processing.

Time-Lapse Fluorescence Microscopy: Microscopy was performed immediately following induction with doxycycline. The live cells were grown on 6-well plates (Greiner Bio-One) in the complete medium. Cells were imaged using Zeiss Axio Observer Z1 microscope in a Incubation control environmental chamber. The images were captured using an Andor iXon X3 camera. The CoolLED pE-2 lasers were used for excitation. The filter sets (Semrock) are as follows: 438/24 (excitation) and 483/32 (emission) for AmCyan, 543/22 (excitation) and 593/40 (emission) for DsRed. Data collection was performed using Micro-Manager 1.4. Image processing was performed in ImageJ. All images were collected with the same exposure times and underwent identical processing.

Flow Cytometry: Flow cytometry was performed 24-72 hours post induction with doxycycline. Each well of the 12-well plates were washed with 1.0 mL of DPBS and then trypsinized with 0.3 mL 0.25% Trypsin-EDTA at 37°C for 5 minutes. Trypsin-EDTA was then neutralized by adding 0.7 mL of complete medium. The cell suspension was centrifuged at 1,000 rpm for 5 minutes and after removal of supernatants, the cell pellets were resuspended in 0.5 mL DPBS. The cells were analyzed on a BD LSRFortessa flow analyzer. AmCyan was measured with a 445-nm laser, a 505

long-pass filter, and 525/50 band-pass filter, and DsRed with a 561-nm laser, and 582/15 band-pass filter. For all experiments ~200,000 events were collected. All wells within a given experimental set had the same number of events collected with the same voltages and underwent identical processing. The flow cytometry data was processed using FlowJo. A FSC (forward scatter)/SSC (side scatter) gate was generated applied to all cell samples. The gated cells were then further gated into AmCyan positive cells by gating above the AmCyan levels of uninduced cells. For some analysis a DsRed positive gate was then applied, again by gating above the DsRed levels of uninduced cells. All data underwent identical processing for each analysis.

Single Cell Serial Dilution: To generate clonal population, serial dilution method used with a 96-well plate. First, a confluent cell culture was detached using trypsin and resuspended using conditioned media. The conditioned media prepared by harvesting at approximately 75% confluency, then clarified by centrifugation at 1000 x g for 10 minutes then passed through a sterile filter. Resuspension was diluted to a final concentration of 2 x 10 4 cells/mL, then added inoculated to a single well in a 96-well plate.

Minimum Promoter Activity Threshold: To model this thresholding relationship between the intragenic miRNA-regulated DsRed and the unregulated AmCyan using flow cytometry data, we devised a three-parameter piecewise function of the form

$$\mu(x) = \begin{cases} c\,x + d & x \leq x_{thresh} \\ a\,(x - x_{thresh}) + (c\,x_{thresh} + d) & x > x_{thresh} \end{cases}, \quad (1)$$

where μ represents the miRNA repressed DsRed expression, x represents the unregulated AmCyan expression, x_{thresh} captures the critical point in which active DsRed expression begins and a represents the proportionality between DsRed and AmCyan above the AmCyan threshold. The parameters c and d each represent slope and y-intercept of the linear curve that describes the cytometry population below the promoter theshold **(Figure S2)**. Operationally, the promoter threshold is an inflection point at which the slope of the curve describing the cytometry population changes. The curve beyond the inflection point ($x > x_{thresh}$) can be now described with the equation $a\,(x - x_{thresh}) + (c\,x_{thresh} + d)$ this curve, where a is the slope of this curve. This can be described as a three-parameter piecewise function as shown below:

$$y = \begin{cases} c\,x + d & x \leq x_{thresh} \\ a\,(x - x_{thresh}) + (c\,x_{thresh} + d) & x > x_{thresh} \end{cases}$$

The x_{thresh} that satisfies the above function corresponds to the promoter threshold value, i.e. the expression of AmCyan at which the observable DsRed production begins. In order to find x_{thresh}, we use Mathematica function NonlinearModelFit with the piecewise function as the desired output form. The process cycles through iterations of nonlinear regression fitting procedure until the result reaches 95% confidence, i.e. there is a 95% probability that the population lies within the confidence interval.

Transgene copy number validation: Real-time quantitative PCR (qPCR) has been recognized as a viable alternative to Southern blot or fluorescence *in situ* hybridization for evaluation of gene copy numbers (Hoebeeck et al., 2007). As such, we performed real-time quantitative PCR to determine the absolute copies of integration for our circuits using $\Delta\Delta C_t$ method **(Table S3)**. The HEK293 cell line that was previously verified to contain a single copy of DsRed transgene was

used as the control cell line (Shimoga et al., 2013). All genomic DNA samples were extracted using DNeasy Blood and Tissue kit (Qiagen).

Briefly, the procedure for determining the copy number is as follows: Harvest genomic DNA for the transgenic clones as well as the control cell line, perform qPCR and obtain the C_t values for each stable clone for the transgene (DsRed) and the endogenous reference gene (BRCA1), calculate the copy number with $2^{-\Delta\Delta Ct} = ((1 + E_{DsRED})^{-\Delta Ct,DsRED}) / ((1 + E_{BRCA1})^{-\Delta Ct,\ BRCA1})$, where E_{DsRED} is the PCR amplification efficiency for DsRed (1.07) and E_{BRCA1} for endogenous reference gene BRCA1 (0.98) (Zheng et al., 2011) (**Figure S2)**.

Supplementary References

Hinske, L.C., Franca, G.S., Torres, H.A.M., Ohara, D.T., Lopes-Ramos, C.M., Heyn, J., Reis, L.F.L., Ohno-Machado, L., Kreth, S., and Galante, P.A.F. (2014). miRIAD--integrating microRNA inter- and intragenic data. Database *2014*, bau099–bau099.

Hoebeeck, J., Speleman, F., and Vandesompele, J. (2007). Real-time quantitative PCR as an alternative to southern blot or fluorescence in situ hybridization for detection of gene copy number changes. Methods Mol. Biol. *353*, 205–226.

Shimoga, V., White, J.T., Li, Y., Sontag, E., and Bleris, L. (2013). Synthetic mammalian transgene negative autoregulation. Mol. Syst. Biol. *9*, 670.

Zheng, S., Houseman, E.A., Morrison, Z., Wrensch, M.R., Patoka, J.S., Ramos, C., Haas-Kogan, D.A., McBride, S., Marsit, C.J., Christensen, B.C., et al. (2011). DNA hypermethylation profiles associated with glioma subtypes and EZH2 and IGFBP2 mRNA expression. Neuro. Oncol. *13*, 280–289.