miRNA circuit modules for precise, tunable control of gene expression

Rongrong Du*[1], Michael J. Flynn*[1], Monique Honsa[2], Ralf Jungmann[2], Michael B. Elowitz[1†]

*These authors contributed equally to this work
[†]correspondence: melowitz@caltech.edu

[1]Howard Hughes Medical Institute and Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[2]Max Planck Institute of Biochemistry, Martinsried, Germany; Faculty of Physics, Ludwig Maximilian University, Munich, Germany

Abstract

The ability to express transgenes at specified levels is critical for understanding cellular behaviors, and for applications in gene and cell therapy. Transfection, viral vectors, and other gene delivery methods produce varying protein expression levels, with limited quantitative control, while targeted knock-in and stable selection are inefficient and slow. Active compensation mechanisms can improve precision, but the need for additional proteins or lack of tunability have prevented their widespread use. Here, we introduce a toolkit of compact, synthetic miRNA-based circuit modules that provide precise, tunable control of transgenes across diverse cell types. These circuits, termed DIMMERs (Dosage-Invariant miRNA-Mediated Expression Regulators) use multivalent miRNA regulatory interactions within an incoherent feed-forward loop architecture to achieve nearly uniform protein expression over more than two orders of magnitude variation in underlying gene dosages or transcription rates. They also allow coarse and fine control of expression, and are portable, functioning across diverse cell types. In addition, a heuristic miRNA design algorithm enables the creation of orthogonal circuit variants that independently control multiple genes in the same cell. These circuits allowed dramatically improved CRISPR imaging, and super-resolution imaging of EGFR receptors with transient transfections. The toolbox provided here should allow precise, tunable, dosage-invariant expression for research, gene therapy, and other biotechnology applications.

One sentence description: Compact synthetic miRNA-based regulatory circuits enable tunable, orthogonal, and generalizable dosage-invariant gene expression control for research and biotechnology.

Introduction

Biotechnology and biomedical research rely heavily on ectopic expression of transgenes in living cells. Popular expression systems produce a broad range of expression levels in individual cells. This is true for non-integrating approaches such as DNA transfection and AAV vectors, as well as integrating systems such as lentivirus[1] or piggyBac transposons[2]. This heterogeneity reflects unavoidable variation in the number of gene copies taken up, integrated, and expressed by each cell, as well as gene expression noise[3,4]. Variability or noise may be tolerable or even useful[5] in some situations, but more often presents an obstacle to accurate analysis and precise control of cell behaviors. Selecting for stable clones can reduce expression variation but is time-consuming, and can also be susceptible to stochastic silencing[6]. An ideal gene regulation system would compensate for this variation, allowing more precise control of expression level, reduced toxicity in gene and cell therapies, and lower backgrounds with reporters, among other applications (**Figure 1A**).

The incoherent feed-forward loop (IFFL) circuit motif provides an ideal mechanism to provide these capabilities[7]. In an IFFL, a target gene and its negative regulator are co-regulated by the same input (**Figure 1B**). Gene dosage can be considered such an input, and proportionately affects expression of both the regulator and its target. In some parameter regimes, these two effects cancel out, and target expression approaches a fixed level at high dosage (**Figure 1C, Supplementary Modeling**). An ideal IFFL system should further allow tuning of this expression set point, the creation of multiple orthogonal regulation systems for simultaneous control of multiple genes, and the ability to operate in multiple cell types (**Figure 1C**).

Earlier work demonstrated IFFL circuits could generate dosage-invariant expression over a 50- or 100-fold range in bacteria and mammalian cells, respectively[8,9]. However, these systems required expression of additional proteins, complicating their routine use. miRNA could be an ideal regulator for an IFFL dosage compensation system, as it can be expressed from within an intron, or from compact transcripts. In pioneering work, Bleris et al. demonstrated that a miRNA-based IFFL could achieve dosage compensation[10]. Strovas et al. improved on this by incorporating a natural miRNA and multiple repeats of its target sequence within the gene[11]. This reduced expression variation in single-copy integrations, and achieved dosage compensation over a ~20-fold range, at the cost of potential crosstalk with endogenous genes. Most recently, Yang et al introduced an “equalizer” architecture that combined transcriptional negative feedback through the TetR protein with miRNA[12]. This extended the range of effective dosage compensation, but required expression of the bacterial TetR protein, adding complexity and potential immunogenicity. Despite much work, it has remained unclear what sequence features are sufficient to enable orthogonal, tunable, dosage compensating miRNA IFFLs, and as a result a broadly useful toolkit of such circuits does not yet exist.

Here, we engineered a set of miRNA-based dosage compensation systems termed DIMMERs (Dosage Invariant miRNA-Mediated Expression Regulators) that fulfill this need. These circuits use specific configurations of the miRNA expression cassette and its target sequences, and take advantage of the ability to achieve multivalent miRNA regulation through the natural TNRC6 scaffold system. They allow systematic tuning of expression levels by modulating the number of miRNA cassettes, numbers of target binding sites, and

miRNA-target site complementarity. Further, they can be used to orthogonally regulate multiple genes in the same cell, and operate similarly across different cell types. We constructed a toolkit of ten mutually orthogonal ready-to-use expression systems that can be incorporated into diverse systems. Finally, we demonstrated their utility for CRISPR imaging and super-resolution protein imaging modalities. DIMMERs should allow routine research and biotechnology applications to operate with greater precision, control, and predictability.

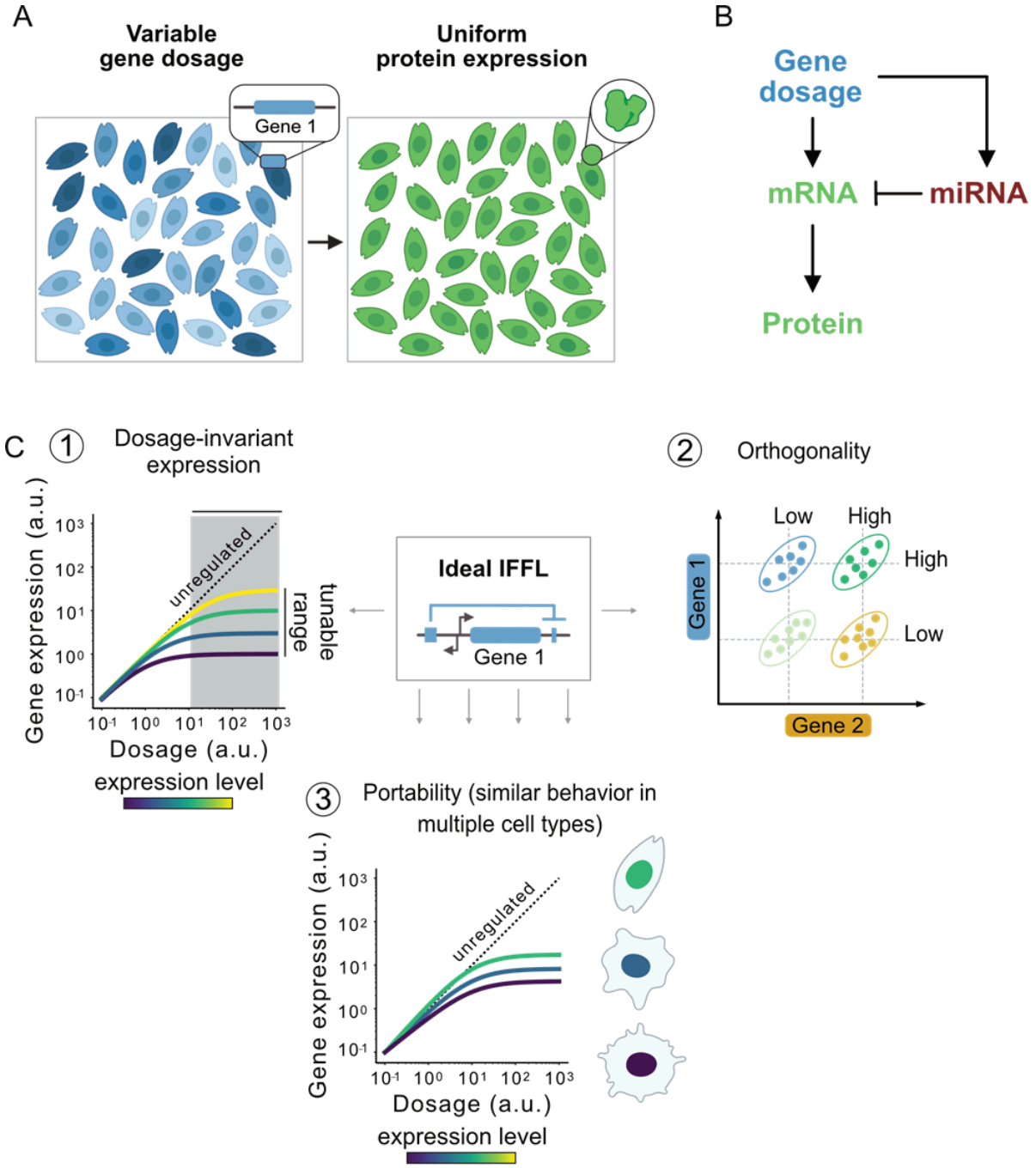

Figure 1. miRNA incoherent feedforward circuits enable dosage-invariant gene expression.
(A) An ideal gene expression system generates uniform protein expression levels despite variable gene dosage delivered. The blue gradient shown in the nucleus indicates the copy numbers of gene delivered, in which darker blue represents higher dosage, and lighter blue represents lower dosage. Green indicates the desired uniform output protein expression levels.
(B) The architecture of the incoherent feedforward loop (IFFL). The input is gene dosage in arbitrary units, which activates the expression of both the mRNA and the microRNA. microRNA inhibits mRNA translation. Output is the resulting protein expression.
(C) IFFLs enable tunable **(①)**, orthogonal **(②)** control of the target and also operate in multiple cell contexts **(③)**. **(①)** shows the modeling of the dosage compensation system which permits tuning of the setpoint levels. **(②)** depicts schematic flow cytometry data of two genes of interest regulated by orthogonal, tunable, miRNA-controlled dosage compensation circuits. Cells are poly-transfected with two independently-regulated constructs. Each color represents one set of the designs used in the

poly-transfection. Each ellipse indicates where most cells are located. Dashed lines indicate the centroids of expression. ③ schematically depicts relatively similar expression behavior across diverse cell types. The central diagram introduces the circuit architecture of the microRNA(miRNA)-based IFFL, in which the two arrows indicate the divergent promoter, the small rectangle on the left indicates the miRNA, and the long rectangle on the right indicates the regulated gene, the short rectangle on the right indicates the miRNA binding sites.

Results

Mathematical modeling identifies parameter regimes for dosage compensation

We first set out to identify the general conditions under which miRNA-based IFFL circuits can achieve dosage compensated target expression (**Figure 1B**, **Supplementary Modeling**). We developed a minimal model based on several assumptions: First, we assumed that pri-miRNA and the target mRNA are each transcribed constitutively at a fixed ratio, and in direct proportion to gene dosage. Second, we assumed a constant total rate, per gene copy, of RISC complex production. This rate incorporates the combined process of pri-miRNA transcription, post-transcriptional processing, and binding to Argonaute proteins[13]. Third, we assumed that the RISC and its target mRNA bind reversibly to form a RISC-mRNA complex. Finally, we assumed that this complex can cause degradation of the associated mRNA.

This model exhibited dosage-invariant expression profiles (**Figure 2A**). More specifically, target protein expression levels initially increased linearly at low gene dosage but approached a dosage independent limiting expression level, at a value of K/k_c (**Figure 2A**). Altering the values of the binding or unbinding rates of the mRNA and the miRNA, or the inhibiting strength of the miRNA could be used to tune the steady state level of the mRNA while maintaining compensation (**Figure S1A and B**). We note that ultrasensitive or sub-linear regulation of mRNA by RISC can lead to non dosage-invariant profiles (**Supplemental Modeling, Figure S1C**). Nevertheless, these results suggest that under general conditions, miRNA-based IFFLs could achieve gene dosage invariant protein expression.

miRNA-based IFFLs generate dosage-independent expression

Based on these modeling results, we designed and built a set of miRNA-based IFFLs. After exploring multiple circuit configurations, we focused on a design in which the miRNA is expressed on a separate adjacent transcript from the same DNA region[10] (**Figure 2B**) in part because it provides flexibility in the relative miRNA and target gene expression levels (**Figure S2**).

To implement the circuit, we constructed a platform comprising a divergent promoter pair (**Figure 2B**). One promoter (oriented left in **Figure 2B**) constitutively expressed a fluorescent protein dosage indicator, mRuby3, along with a synthetic miRNA. To create the miRNA expression cassette, we inserted the miR-E backbone for pri-miRNA expression[14] into a synthetic intron[15]. The second promoter (oriented right in **Figure 2B**) constitutively expressed the target gene, EGFP. To allow regulation by the synthetic miRNA, we

incorporated one or more target miRNA sites of varying complementarity in the 3'UTR of the target gene. This compact two-transcript construct allowed systematic analysis of different miRNAs and target site configurations.

To implement miRNA regulation orthogonal to natural miRNAs, we first used a previously described miRNA sequence targeting Renilla luciferase (miR-L), together with a single copy of its fully complementary 21bp target site. We transiently transfected U2OS cells with the resulting construct, analyzed cells by flow cytometry after 48h, and plotted target EGFP expression versus gene dosage, as indicated by mRuby3 (**Figure 2C**). Compared to an unregulated control with no miRNA target site, the IFFL strongly reduced target EGFP expression, as expected (**Figure 2C**). However, the circuit failed to achieve the dosage compensation behavior anticipated from mathematical modeling (**Figure 2A**).

We next asked whether the lack of dosage compensation could relate to the strongly repressing regime produced by full miRNA-target complementarity. We designed a set of IFFL variants which progressively reduced complementarity of the single target site from 21 to 17 bp (**Table S1**). Designs with reduced 3' complementarity showed weak or no repression of target expression, particularly below 19bp (**Figure 2D**), while those that did efficiently repress retained strong repression comparable to that of the full length 21bp construct. Nevertheless, within this set, no construct achieved full dosage invariance. Thus, modulation of complementarity alone was not sufficient to provide dosage compensation.

The loss of regulation at reduced complementarity contrasted with the well known regulatory capacity of miRNAs with much shorter complementary regions of only ~8bp[13]. One mechanism to enable specific regulation with short sequences involves multivalent recognition of multiple target binding sites on the same mRNA[16]. TNRC6 is a scaffold protein that enables multivalent recognition by simultaneously binding to multiple miRNA-loaded Argonaute (Ago) complexes (**Figure 2G**)[17,18,19]. Consistent with a role for multivalent regulation, tandem repeats of two to four copies of the 17bp target site progressively increased regulation, and strongly reduced dosage sensitivity at higher expression levels (**Figure 2E**). For example, with 4 tandem binding sites, target expression increased by only 4-fold over a 200-fold range of dosages (**Figure 2F**). A "tail" of elevated expression at the highest dosages may reflect saturation of miRNA-associated machinery, as observed in other studies[12]. Together, these results show that a miRNA-based IFFL based on 17bp of miRNA-target complementarity and 4 tandem binding sites can achieve nearly complete compensation over more than two orders of magnitude of dosage variation.

To find out whether this compensation behavior was dependent on TNRC6, we took advantage of a previously identified fragment of the natural TNRC6B protein, the T6B peptide that competitively inhibits TNRC6 activity (**Figure 2G**)[20]. When co-transfected with the 4×17bp IFFL, the T6B inhibitory peptide abolished regulation, producing dosage-dependent expression nearly identical to that produced by an unregulated construct (**Figure 2H**). By contrast, the T6B inhibitory peptide had little effect on the single fully complementary 21bp construct, suggesting that it regulates in a TNRC6-independent manner (**Figure 2I**). Finally, negative control mutant variants of T6B lacking the Ago2-binding domain failed to abolish regulation, as expected (**Figure S3A**). Together, these results suggest that the 4×17bp and 1×21bp designs respectively operate through TNRC6-dependent and TNRC6-independent mechanisms. Because of their ability to limit

expression, we term these circuits DIMMERs. More generally, these results indicate that multivalent regulation through multiple, individually weak, miRNA binding sites can achieve strong regulation and dosage compensation within the context of the IFFL circuit.

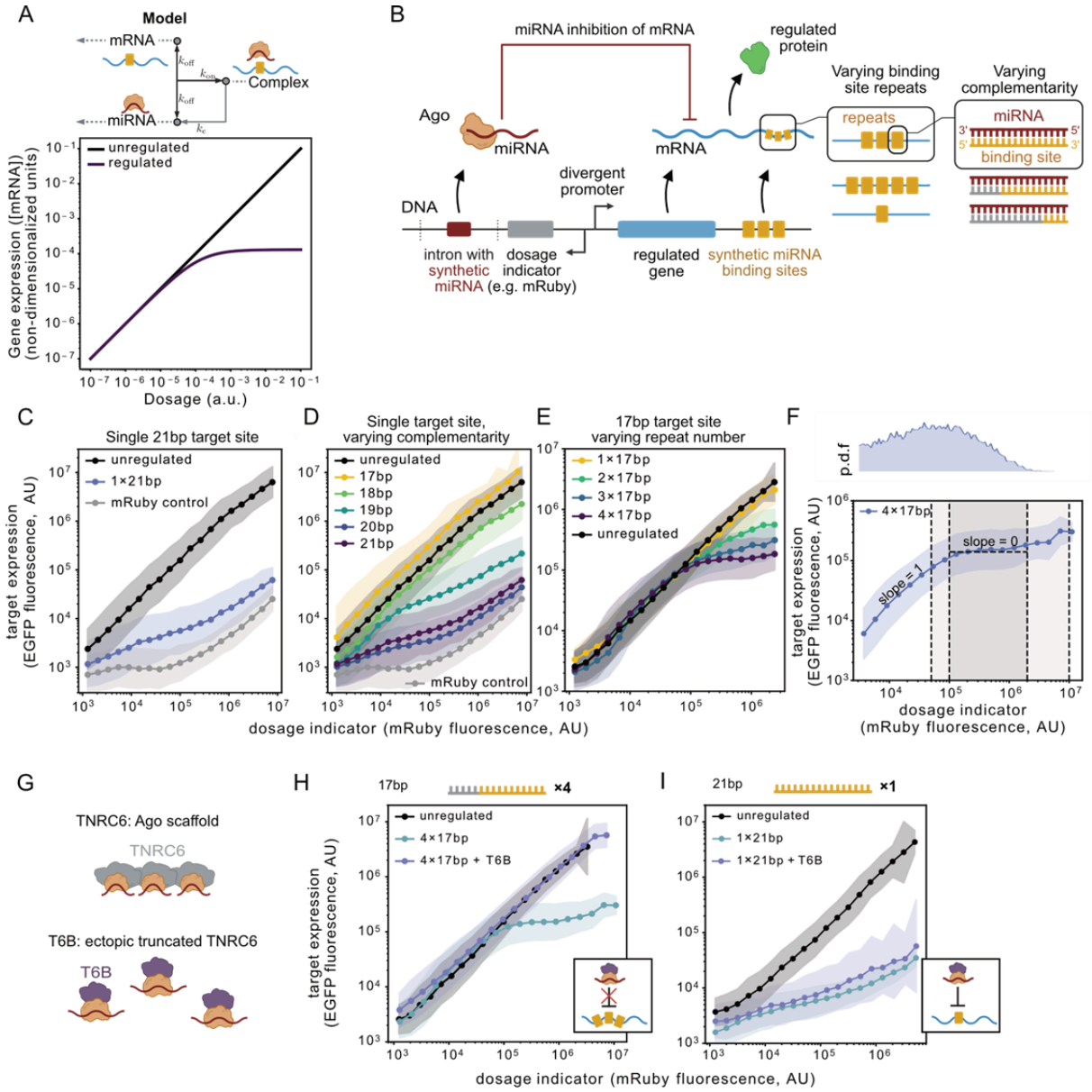

Figure 2. miRNA-based IFFLs achieve linear regulation and dosage independence through TNRC6-mediated repression.

(A) To identify the parameter regimes allowing dosage compensation, we built a simplified model of miRNA inhibition. Upper panel, gray dashed arrows in the diagram represent the natural dilution and degradation of the mRNA or miRNA. Black, double-directed arrows represent the association and dissociation of the mRNA-miRNA complex. Gray arrows represent the decay of the mRNA. The rates of all the reactions are labeled on the side of the arrows. Lower panel, modeling results of the miRNA-mRNA interaction. The unregulated curve suggests the expression when there's no miRNA.

(B) To implement the miRNA-based dosage compensation system, we designed the divergent promoter circuit. We varied the target-miRNA complementarity and target site numbers to explore the parameter space that gives rise to the dosage independence.

(C) We performed flow cytometry to verify the behavior of the circuit shown in **(A)** with a single, fully complementary miR-L target site versus an unregulated control, which has neither the miR-L nor the miR-L target site. We used mRuby3 fluorescent protein as the dosage indicator, and EGFP fluorescent protein as the target. Cells are gated and binned by mRuby3 intensities. Each dot corresponds to the geometric mean fluorescence intensity of the mRuby3 bin breaks and median fluorescence intensity of EGFP in the bin. Shaded regions denote the range from

$\frac{\text{EGFP median}}{\text{geometric standard deviation}}$ to $\text{EGFP median} \times \text{geometric standard deviation}$. The gray curve shows an mRuby3-only transfection control, which suggests the bleedthrough from the mRuby3 channel to the EGFP channel.
(D) We used flow cytometry to understand how the complementarity of a single target site affects the regulation. Single miR-L target sites with the complementarity ranging from 17 bp to 21 bp, starting from the seed region, were used in this experiment. mRuby3 serves as the dosage indicator, and EGFP is the target. Cells are gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **(C)**. The gray curve shows an mRuby3-only transfection control.
(E) We performed flow cytometry to dissect the effect of increasing target repeats while keeping the complementarity constant. A miR-L 17 bp complementary site starting from the seed region was repeated from 1 to 4 times. Cells are gated and binned by mRuby3 intensities, a constitutively expressed iRFP plasmid was co-transfected as a cotransfection marker. Dots and shaded regions were calculated as described in **(C)**.
(F) A zoom-in of the 4×17 bp construct used in **(E)**. p.d.f, probability distribution function, which measured the distribution of mRuby3 fluorescence intensity. Cells are gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **(C)**. The gray rectangle in the scope from 10^5 to 2×10^6 (shown by the vertical dashed lines) on the mRuby3 axis indicates the range where the fluorescence intensity of EGFP does not change. The light gray rectangle in the scope from 5×10^4 to 10^7 (shown by the vertical dashed lines) on the mRuby3 axis indicates the range where the fluorescence intensity of EGFP has minor change (i.e., less than 5 fold). The dashed lines illustrate the curves with a slope of either 1 or 0, revealing either the linear dependence of the dosage or independence of the dosage.
(G) A diagram of the T6B peptide inhibition mechanism. T6B is a short, dominant-negative version of the TNRC6B. T6B competes with TNRC6 by binding to the Ago2-miRNA complex.
(H) To unravel the mechanism of the inhibition that gives rise to the dosage compensation behavior, we co-transfected cells with or without T6B peptide and the 4×17 bp construct described in **(F)** and performed flow cytometry. The T6B peptide is fused to the iRFP fluorescence protein, and the control group was co-transfected with the constitutively expressed iRFP. Cells are gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **(C)**. T6B transfected cells lost the inhibition, as illustrated in the small graph located in the plot.
(I) To compare the mechanism of the strong inhibition with the TNRC6-dependent inhibition, we co-transfected cells with or without T6B peptide and the single site, fully complementary construct described in **(C)** and performed flow cytometry. The experimental setting is the same as described in **(H)**. Cells are gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **(C)**. T6B brought minor effects to the inhibition, and therefore suggested that the inhibition is TNRC6-independent, as illustrated in the small graph located in the plot.

DIMMER circuits allow tuning of dosage-independent expression levels

An ideal control system would allow not only dosage invariance but also control of saturating expression level. In principle, this could be achieved by modulating the complementarity of repeated target sites, the number of miRNA cassettes within the synthetic intron, or the strength of the promoter controlling expression of the miRNA cassette. We systematically varied each of these parameters and analyzed its effects on expression.

First, we varied complementarity of the miR-L target sites from 8 to 21bp, in each case incorporating 4 tandem target site copies (**Figure 3A, Figure S4**). At each length, we quantified the expression of the IFFL-regulated target (**Figure 3A**, left panel) at high gene

dosage (**Figure 3A**, right panel, shaded region). Repression was modest at 8bp, diminished with increased complementarity in the central region, and then strengthened again as more complementarity was added after the central region (**Figure 3A,** left panel). These results are broadly consistent with previous observations showing that miRNA inhibition does not increase monotonically with complementarity[21,22,23]. For the miR-L target site, repression was most sensitive at 16-20 base pairs of complementarity (**Figure 3A**). In fact, three designs—4×17, 4×18, and 4×19—achieved evenly spaced, dosage invariant expressions spanning more than an order of magnitude in saturating expression level (**Figure 3B**).

We next varied the number of copies of the miRNA expression cassettes in the synthetic intron, effectively modulating the stoichiometric ratio of miRNA to mRNA (**Figure 3C**). Compared to a single copy, two or three copies of the miRNA reduced expression by 2-fold and 3-fold, respectively, while preserving dosage compensation, providing a means of fine-tuning expression control.

Finally, we varied the promoter controlling expression of the miRNA cassette. Compared to EF1α, the weaker PGK promoter allowed ~3.5-fold more target gene expression at a given dosage level (**Figure 3D**). Nevertheless, dosage compensation was preserved here too. Further, we questioned whether our circuit works in an inducible system setting. We used the 4-epi-tetracycline tuned promoter to express the target and tested the circuit's performance in TRex cells (**Figure 3E**). By modulating the concentration of 4-epi-tetracycline, we were able to vary the setpoint of the circuit by ~2 orders of magnitude while preserving dosage compensation (**Figure 3E, Figure S5**). Taken together, these results demonstrate that four distinct mechanisms can be used to tune expression level in a coarse or fine manner.

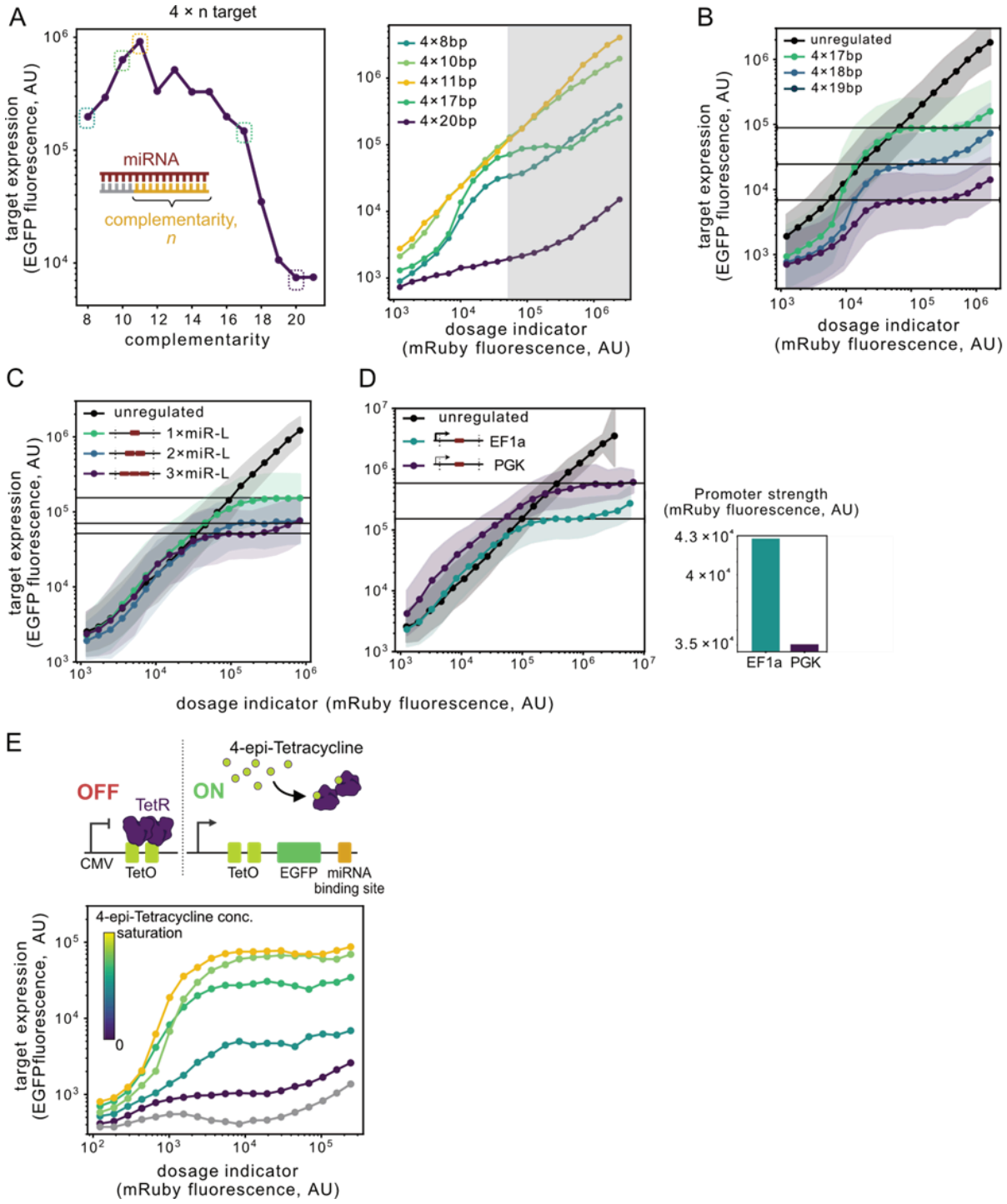

Figure 3. Sequence complementarity, miRNA cassette copy number, and promoter strength tune expression level.

(A) We designed and constructed circuits shown in **Figure 2B** harboring miR-L targeting sites with the complementarity ranging from 8 bp to 21 bp starting from the seed region, with each targeting site repeated four times. In the left panel, we performed flow cytometry on the cells and each dot represents one well. Cells were gated based on the mRuby3 fluorescence intensity. Median fluorescence intensities of the EGFP in the mRuby3 positive cells were plotted. In the right panel, the dosage response curves of the five selected constructs (dashed boxes) from the left panel are displayed. Gray rectangle indicates the gated region of the left panel. Cells were binned by mRuby3 intensities. Dots were calculated as described in **Figure 2C**.

(B) We performed flow cytometry on the cells transfected with the circuits shown in **Figure 2B** harboring miR-L targeting sites with the complementarity ranging from 17 bp to 19 bp starting from the seed region, with each targeting site repeated four times. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. Black solid horizontal lines denote the setpoint levels of the constructs.
(C) We performed flow cytometry on the cells transfected with the circuits shown in **Figure 2B** harboring miR-L cassettes with 1-3 repeats incorporated in the same intron, the target sites of which were miR-L targeting sites with 17 bp complementarity starting from the seed region, repeated four times. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. Black solid horizontal lines denote the setpoint levels of the constructs.
(D) We performed flow cytometry on the cells transfected with the circuits shown in **Figure 2B** harboring miR-L cassettes expressed from the EF1α or PGK promoter, the target sites of which were miR-L targeting sites with 17 bp complementarity starting from the seed region, repeated four times. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. Black solid horizontal lines denote the setpoint levels of the constructs. The barplot at the bottom right corner shows the mRuby3 median fluorescence intensities of the transfected cells.
(E) Upper panel, design of the 4-epi-Tetracycline inducible DIMMER. The promoter that drives the EGFP is a CMV promoter harboring two TetO sites. When expressing the constructs in the TRex cell line, the TetR transcriptional repressor binds to TetO and shuts off the transcription. 4-epi-tetracycline removes the TetR from TetO in a gradient-dependent manner, and therefore restores the transcription. Lower panel, we performed flow cytometry on the TRex cell line which was transfected with the 4×19 bp DIMMER construct, and the EGFP was driven by the CMV-2×TetO promoter. The concentrations of the 4-epi-Tetracycline, from purple to yellow, were 0, 10, 33.3, 100, 333.3 ng/mL. The gray curve denotes the mRuby-only transfection control, and the 4-epi-Tetracycline concentration is 333.3 ng/mL.

Orthogonal dosage compensation circuits allow independent control of target genes

Engineered genetic systems increasingly require multiple genes and transcripts, many or all of which may benefit from dosage-independent control. We therefore set out to construct an expanded set of orthogonal synthetic miRNA-target site pairs, which we term *synmiRs*. To design synmiRs, we started by generating five random miRNA sequences, each with an initial A in the miRNA, a cognate U at the 3' end of a single target site and 25% GC content, similar to the structure of miR-L. We then measured the dose-response curve for each candidate miRNA. To do so, we designed an "open loop" system, which allows independent control of miRNA expression and measurement of its effect on a target miRNA reporter gene (**Figure 4A**). For a single fully complementary miR-L site, we see the inhibition increases with the dosage of the miRNA (**Figure 4B**), consistent with the earlier closed loop results (**Figure 2C**). Of the five sequences initially considered, synmiRs 1,4, and 5 repressed by at least an order of magnitude relative to a control lacking the miRNA (**Table S1, Figure 4D**). By contrast, synmiRs 2 and 3 achieved less substantial repression (**Figure S6A**). Inspection of these designs revealed subsequences with two or more A/T pairs in the extensive region. We hypothesized that these A/T dinucleotides could potentially trigger the miRNA degradation machinery to recognize the miRNA tailing and trimming signal, resulting in miRNA destabilization[24,25]. Consistent with this hypothesis, an A to G substitution at position 20 in synmiR-2 or at position 19 in synmiR-3 restored miRNA inhibition of target gene expression (**Figure 4D**).

Based on these results, we formulated an empirical two-step sequence design algorithm (**Figure 4C**). In the first step, we generate random 21bp candidate miRNA sequences constrained to have a 5'U in the mature miRNA. In the second step, based on known requirements for miRNA loading[13,26], we constrain sequences to a total 5-8 GC nucleotides, 1-4 in the seed region, and 1-2 in the extensive region (**Figure 4C**). We generated and tested five candidate sequences (synmiRs 6-10) based on this simple algorithm and analyzed their open loop behavior (**Figure 4D**, **Figure S6B**). When used as a single fully complementary site, all five sequences generated strong repression, comparable to that of miR-L. However, we discarded one sequence, synmiR-6, due to its similarity to the human endogenous miRNA hsa-mir-5697 (**Methods**). Altogether, these results produced a set of ten miRNA sequences that were capable of strong repression in their fully complementary form (**Figure 4D**). Pairing each of these ten miRNAs with all ten of the target sequences in the open loop system revealed strong orthogonality in regulation, as desired (**Figure 4E**).

Using these synmiR sequences, we developed a set of ten orthogonal dosage compensation systems, using the framework in **Figure 2B**. Based on our earlier analysis of miR-L (**Figures 2, 3**), we varied complementarity and the number of tandem target sites for each sequence design. Different sequences required alternate configurations to produce dosage compensation. For example, with synmiR-4 and synmiR-5, the 4×17bp configuration produced a strong inhibition profile akining more with the fully complementary target (**Figure S7**). We reasoned that this could reflect higher GC content in the seed and supplementary regions of these two miRNAs compared to miR-L. This increase in GC content allowed shorter 8 or 9 bp target sites, present in 4-8 repeats, to successfully produce dosage compensating designs for these miRNAs (**Figure 4F**). We repeated a similar process for the other sequences and eventually identified nine additional closed-loop systems that exhibit substantial levels of dosage compensation (**Figure 4F**). In general, regulatory behavior was sensitive to both the number and complementarity of target sites (**Figure S8**). Critically, however, each of these nine synmiR sequences, as well as the original sequence, was able to generate at least partial dosage compensation in one or more configurations (**Figure 4F**), and many could be tuned through these features to different expression setpoints (**Figure S8**). These results provide a toolkit of dosage compensating systems and more generally suggest that it should be possible to engineer additional systems with varying expression setpoints.

By using multiple dosage compensation systems together, one should be able to simultaneously and independently specify the expression of multiple target genes (**Figure 1C, panel 2**). To test this possibility, we constructed a second set of dosage compensation expression systems using distinct fluorescent reporters (**Figure 4G**). We transfected cells with pairs of systems that had different regulatory setpoints, and analyzed the resulting expression profiles of the two regulated target genes (**Figure 4H**). We analyzed four pairs of systems. Each produced a distinct two-dimensional expression distribution based on the set-points for the two reporters. This demonstrates that the engineered dosage compensation systems make it possible to specify two-dimensional expression distributions, and suggests that control of higher dimensional distributions of more genes should also be accessible.

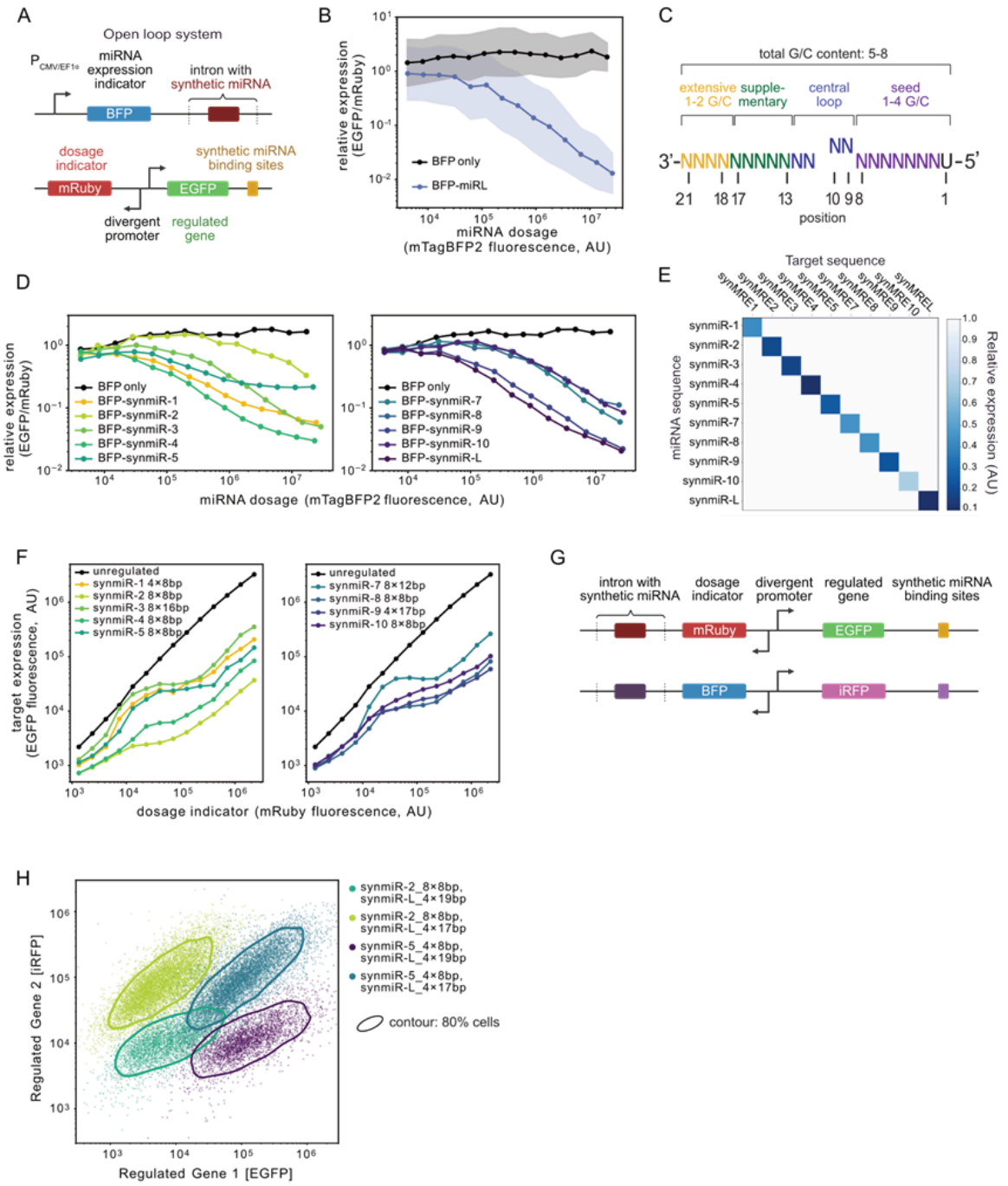

Figure 4. Algorithm-based miRNA designs are orthogonal to each other, and enable independent regulation of multiple genes.

(A) Diagram of the open loop system. miRNA is expressed from the intron of the 3'UTR of the mTagBFP2. mRuby3 serves as the dosage indicator, while EGFP is the target.

(B) We performed flow cytometry on the cells co-transfected with the circuit shown in **(A)** with or without the miR-L (the BFP only control does not have the 3'UTR miRNA), the target site of which is a single, fully complementary miR-L site. Cells were gated by mRuby3 and mTagBFP2 intensities and binned by mTagBFP2 intensities. We defined the relative expression as the median of the ratio of EGFP fluorescence intensity to mRuby3 fluorescence intensity in the mTagBFP2 bin. Dots and shaded regions were calculated as described in **Figure 2C**.

(C) Diagram of a mature synthetic miRNA and the sequence design constraints used to design them.

(D) We performed flow cytometry on the cells co-transfected with the circuit shown in **(A)** with the synthetic miRNAs, and the corresponding target site of which is a single, fully complementary

synthetic miRNA site. The BFP only control does not have the 3'UTR miRNA and uses a target without any target sites. Cells were gated by mRuby3 and mTagBFP2 intensities and binned by mTagBFP2 intensities. Relative expression levels were quantified as described in **(B)**. Dots were calculated as described in **Figure 2C**.
(E) To test if the synthetic miRNAs and their targets are orthogonal to each other, we poly-transfected the cells with all the combinations of the 10 miRNAs and the 10 single, fully complementary sites (10×10 combinations) and performed flow cytometry. Each grid represents one well. Cells were gated by mRuby3 and mTagBFP2 intensities. Relative expression levels were first quantified in the transfected cells as described in **(B)**, and then normalized by dividing the minimal values that were larger than 1.
(F) A gallery of the circuits with the architecture shown in **Figure 2B** established based on different miRNA and targets. The repeat number and the complementarity between the target and the miRNA were denoted by $m \times n$ bp, where m represents the repeat number and n represents the complementarity. Cells were measured by flow cytometry, gated and binned by mRuby3 intensities. Dots were calculated as described in **Figure 2C**.
(G) Diagram of a double IFFL reporter system. In addition to the construct illustrated in **Figure 2B**, which uses mRuby3 as the dosage indicator and EGFP as the target, we introduced a distinct construct which uses mTagBFP2 as the dosage indicator and emiRFP670 as the target. These two reporters were designed to be established on orthogonal miRNA and targets to implement independent regulations.
(H) We performed flow cytometry on cells that were poly-transfected with the double reporter system shown in **(G)**. Cells were gated by mRuby3 and mTagBFP2 intensities. The miRNA and target designs of the double reporter system are listed on the top right corner of the plot, the nomenclature of which is described in **(F)**, where the former part denotes the cassette regulating EGFP, and the latter part denotes the cassette regulating iRFP. The solid contour lines denote where 80% of the cell masses were located centering the expression centroids.

Dosage compensation systems are portable and minimally perturbative

An ideal dosage compensation system would be portable, able to operate similarly across different cell types, function in both transient transfection and genomic integration, and minimally perturb the host cell. To examine these features, we transiently transfected several circuit variants, including the 4×17 miR-L system (**Figure 2E**), in four mammalian cell lines: U2OS[27], CHO-K1[28], HEK293[29], and N2A[30]. In each cell line, we observed strong and qualitatively similar dosage compensation (**Figure 5A**). Cell lines varied in the threshold dosage at which expression saturated (**Figure 5A**, gray vertical line), and in the saturating expression level (**Figure 5A**, gray horizontal line), as measured in arbitrary fluorescence units. However, the ratio of these values was conserved (**Figure 5B**). We obtained similar results for other circuits as well, including synmiR-4, with 8 repeats of a 9 bp target site, as well as both synmiR-L and synmiR-5, each with 8 repeats of an 8bp target site (**Figure S9A**). Again, the ratio of the saturating expression level to the threshold dosage was similar, for each construct, across cell lines (**Figure S9B**). This suggests a model in which the miRNA circuit functions equivalently in different cell types, but protein expression strengths vary, possibly due to differences in translational capacity or basal protein degradation rates[31–34]. Together, these results indicate that the dosage compensation circuits can function across different cell contexts.

Stable cell lines are important in research as well as applications like cell therapy. To find out whether dosage compensation circuits could also function in a stable integration context, we

used PiggyBac transposition together with the iON system that allows expression only from constructs that have successfully integrated in the genome and undergone site-specific recombination[35]. We then selected mono-clones, and analyzed reporter expression by flow cytometry (**Figure 5C**). Integration copy numbers varied among clones by over two orders of magnitude, as indicated by mRuby3 fluorescence intensity (**Figure 5D**, x-axis). Nevertheless, the cargo EGFP expression remained nearly constant (**Figure 5D,** y-axis). Thus, dosage compensation circuits function in stable integration settings.

The expression of synthetic miRNAs could in principle perturb endogenous gene expression. To identify such effects, we performed bulk RNAseq on cells transfected with miR-L and each of the 9 orthogonal synmiRs, and compared them to a negative control transfection of a BFP expression vector. Only a few genes were significantly up- or down-regulated by the miRNA (**Figure 5E**). These were enriched for heat shock proteins such as HSPA6. Critically, the gene sets up-regulated by different miRNAs exhibited strong overlap (**Figure 5F**, **Table S2, Figure S10**). Thus, for the synmiRs described here, off-target regulation appears to only reflect non-specific effects of miRNA expression, rather than sequence-specific perturbations.

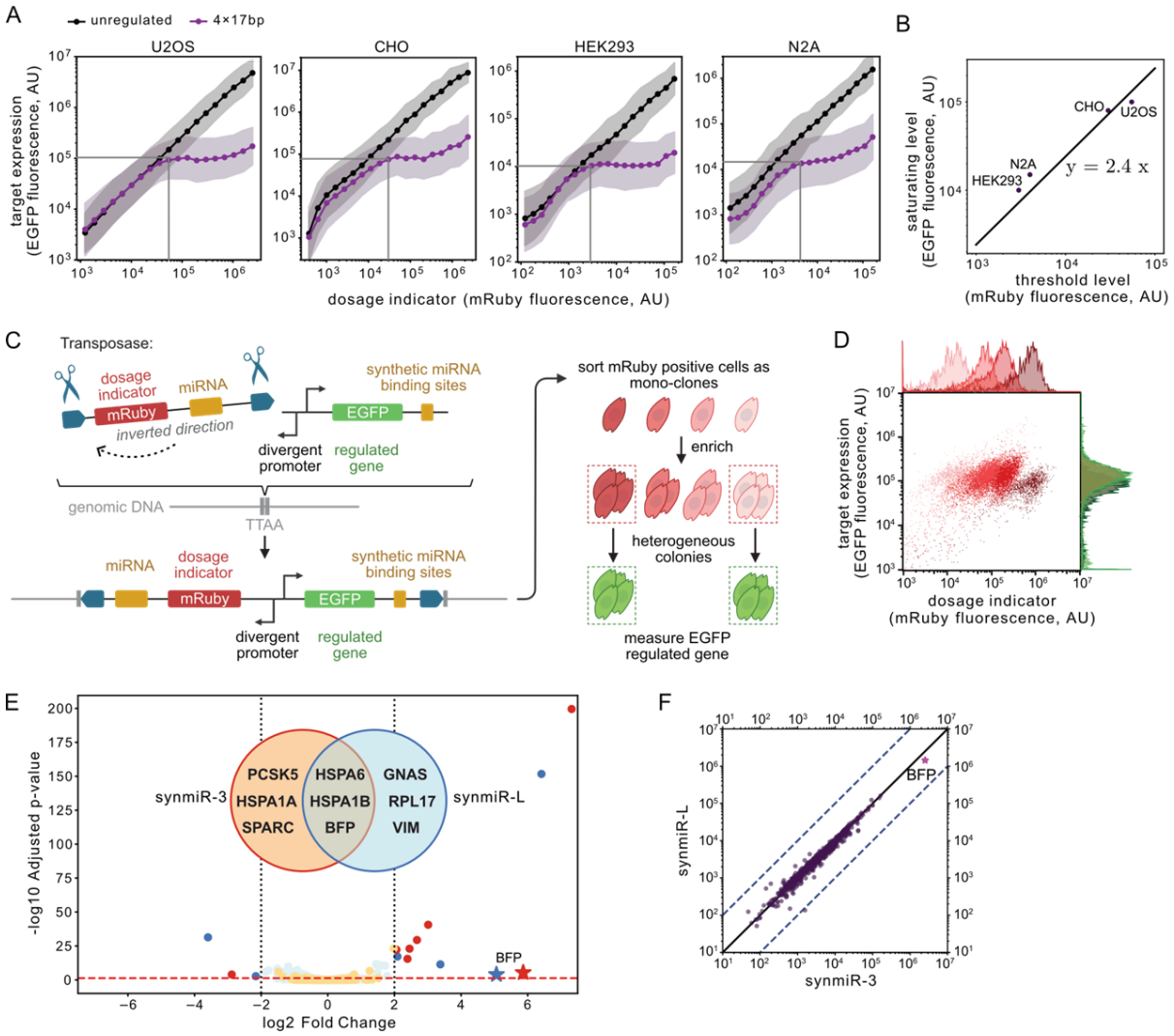

Figure 5. DIMMER circuits work across different cell lines and don't perturb the transcriptome.
(A) We transfected the circuit described in **Figure 2B** regulated by the miR-L with the 4×17 bp target into different cell lines and performed flow cytometry. Cells were gated and binned by mRuby intensities. Dots and shaded regions were calculated as described in **Figure 2C**. Gray vertical lines intersecting the x axis mark the mRuby3 fluorescence intensities where the circuit started to show dosage independence behavior. Gray horizontal lines intersecting the y axis suggest the EGFP fluorescence intensities of the setpoints.
(B) The values of the intersection points of the gray lines with the axis in **(A)** were plotted as scatters. The black solid line indicates a linearity between the threshold level and the saturating level, which was described by $y = k\,x$, where k was fitted by applying a linear regression model.
(C) Workflow of the DIMMER mono-clone sorting and measurement.
(D) We performed flow cytometry on the DIMMER mono-clones that showed varied mRuby distributions. The histograms of the mRuby3 and the EGFP intensities of each mono-clone are displayed on the top and the right of the scatter plot, respectively.
(E) We performed bulk RNAseq on the cells that were transfected with the BFP-miRNA construct. The orange and red dots depict the transcripts of the synmiR-3 expressing cells, while the light blue and dark blue dots show the transcripts of the synmiR-L expressing cells. The space out of the dashed black lines indicate where the absolute value of log2 fold change was larger than 2. The space above the red dashed line indicates where the adjusted p-value is less than 0.05. The Venn diagram above denotes the shared and nonshared significantly differentially expressed genes between synmiR-3 and synmiR-L.

(F) The normalized transcripts per million (TPM) of the miR-L expressing cells were plotted against those of the synmiR-3 expressing cells. Solid line indicates where the TPM of the two samples were equal to each other. Dashed lines indicate where the TPM of the synmiR-3 expressing cells is 10 fold of synmiR-L expressing cells, or where the TPM of the latter is 10 fold of the former.

Dosage compensation systems can suppress background to enhance biological imaging

Dosage compensation could have diverse applications across biology. In quantitative imaging, limiting the expression of fluorescent proteins fused to other proteins of interest could reduce background fluorescence and allow more precise spatial localization of fusion protein distributions, and make the exogenous protein expressions matching the endogenous expressions, facilitating single-molecule imaging, even with transient transfection[36,37]. To test this, we transfected an EGFR-mEGFP membrane marker fusion protein expression construct, either unregulated or controlled by the 4×17 or 4×18 circuit (**Figure 6A**). Both circuits limited expression, with 4×18 providing the lowest expression levels (**Figure S11A**). Further, the unregulated control showed heterogeneous expression, with ~20% of the cells in the field overexpressing EGFR, making its membrane localization difficult to perceive (**Figure 6B,** left). By contrast, in the 4×18bp DIMMER system, most cells exhibited a more clearly defined pattern, with strong plasma membrane localization (**Figure 6B,** right).

The advantage of dosage compensated expression can be observed directly and quantitatively with super-resolution imaging using DNA-PAINT (Point Accumulation for Imaging in Nanoscale Topography), which allows analysis of single protein molecules on the surface of cells[38]. We transfected CHO-K1 cells with unregulated and regulated EGFR-mEGFP plasmids. After transient transfection for 48 hours, cells were fixed and prepared for DNA-PAINT image acquisition using DNA-conjugated anti-GFP nanobodies targeting the intracellular mEGFP tag of the EGFR receptor (**Figure 6C**). Raw DNA-PAINT localizations (**Figure 6D**) were subjected to recently developed clustering procedures to yield the positions of single EGF receptor proteins in the cell membrane (**Figure 6E**)[38,39]. From these datasets, we determined the overall receptor density for different plasmid constructs.

As expected, the receptor densities from unregulated constructs were significantly higher than those of the regulated plasmids (**Figure 6F**). Specifically, the receptor density for the unregulated plasmid was 29 ± 40 μm^{-2} (mean ± s.d.), while the densities for the receptors in the cells transfected with the regulated plasmids were ~15 times lower (4×17bp: 1.9 ± 1.2 μm^{-2}, 4×18bp: 1.0 ± 0.8 μm^{-2}, 4×19bp: 0.6 ± 0.4 μm^{-2}). Excitingly, not only did the mean densities follow the expected trend, but distributions were considerably narrower for the regulated plasmids as compared to the unregulated ones (**Figure 6F**).

Dosage compensation circuits also improved CRISPR imaging. For example, it is possible to image telomeres by expressing dCas9-EGFP, with or without dosage compensation, along with a gRNA targeted to repetitive telomeric sequences[37,40,41] (**Figure 6G**). Flow cytometry showed that circuits reduce dCas9-EGFP expression, and its dosage sensitivity, although not to the extent seen with other genes (**Figure S11B**). With the unregulated system,

dCas9-EGFP seldom labeled the telomeres but rather formed bright aggregations in the nucleolus, consistent with previous observations[42] (**Figure 6H, left panel**). By contrast, the 4×17bp circuit restricted most fluorescence to dots, consistent with telomeric labeling, with reduced labeling of the nucleolus. Further, the stronger 4×19bp circuit removed nearly all labeling in the nucleolus, while maintaining apparent telomere labeling. The circuits thus improved signal-to-background ratio, as evident in line scans of the images (**Figure 6H, middle and right panel**) as well as contrast in each dot. Taken together, these results demonstrate how the dosage compensation circuits developed here improve imaging of proteins and cellular structures.

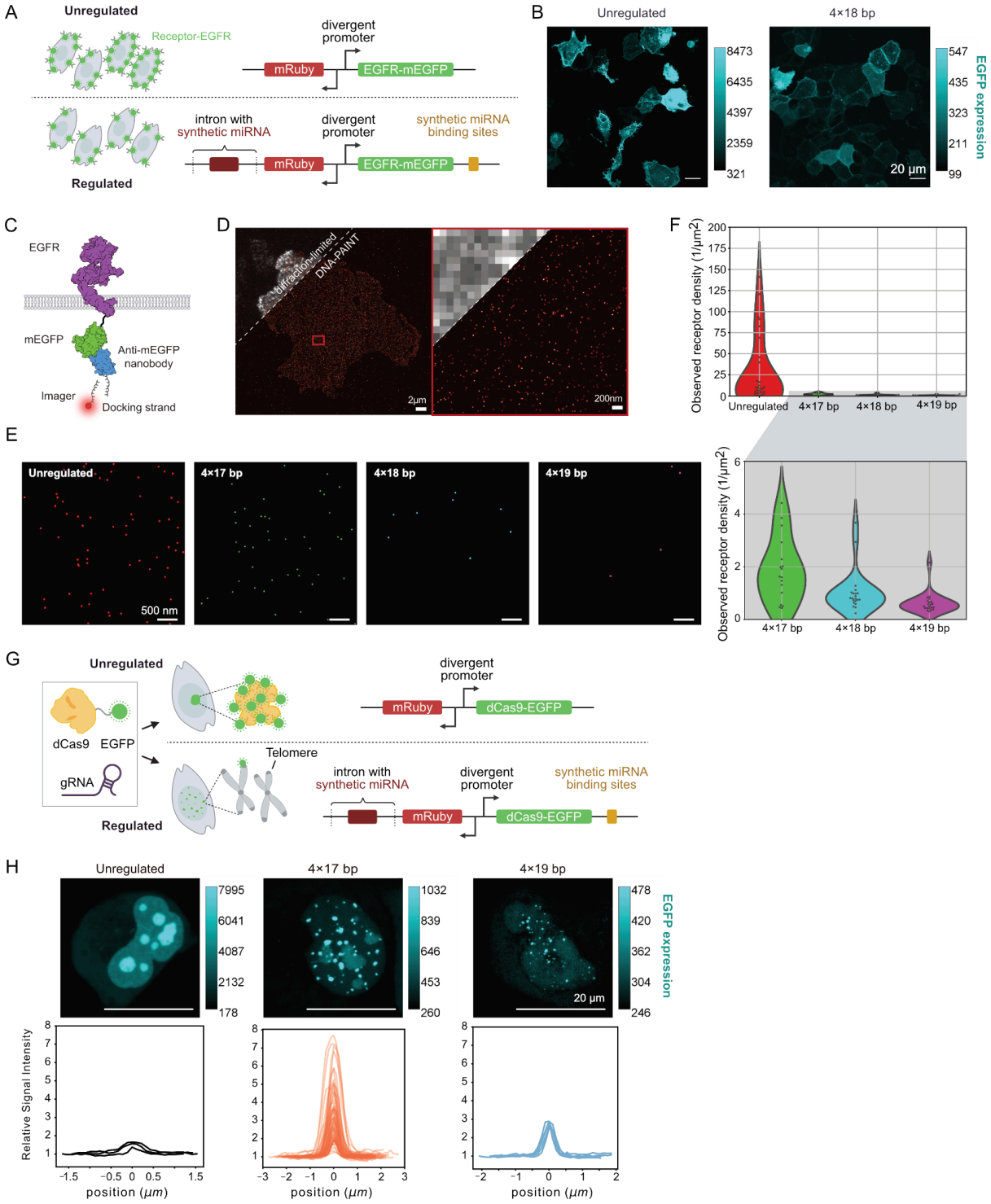

Figure 6. DIMMER circuits improve fluorescence imaging.
(A) Diagram of the EGFR membrane receptor imaging experiment. The regulated EGFR-mEGFP is expected to show a more homogenous distribution of fluorescence signals.
(B) Representative fluorescence images of the EGFR-mEGFP with (right) or without (left) the DIMMER regulation module. miR-L cassette was chosen to build the circuit for cloning convenience. Target site nomenclature was described in **Figure 4F**. Numbers on the colorbars indicate the fluorescence intensities measured by imageJ.
(C) Illustration of EGFR in the cell membrane (adapted from PDB 7SYE). EGFR is linked to an mEGFP which is labeled with a DNA-conjugated anti-GFP nanobody (PDB 6LR7).

(D) Diffraction-limited and DNA-PAINT image of CHO-K1 cells expressing EGFR-mEGFP. The mEGFP is labeled with DNA-conjugated anti-mEGFP nanobodies. The right panel shows the zoom-in of the boxed area.
(E) Representative single-protein-resolved DNA-PAINT images of cells transfected with EGFR-mEGFP with or without the DIMMER module. miR-L cassette was used, and the target site nomenclature was described in **Figure 4F**.
(F) Quantitation of the receptor density for the cells transfected with EGFR-mEGFP with or without the DIMMER module. Insert shows zoom-in of the cells transfected with the DIMMER circuit-regulated EGFR-mEGFP.
(G) Diagram of the dCas9 telomere imaging experiment. The protein controlled by the circuit was switched to dCas9-EGFP. The gRNA is targeted to the telomeres and expressed from the same vector as dCas9-EGFP, but is driven by the U6 promoter.
(H) Upper panel, Representative fluorescence images of the dCas9-EGFP with or without the DIMMER module. miR-L cassette was chosen to build the circuit for cloning convenience. Target site nomenclature was described in **Figure 4F**. Numbers on the color bars indicate the fluorescence intensities measured by imageJ. Lower panel, quantitation of the relative signal intensity of the dots in the cells transfected with the dCas9-EGFP with or without the DIMMER module. Each line represents one dot inside a cell. It should be noted that it's extremely hard to find dots inside the cells transfected with the unregulated dCas9-EGFP, therefore, only three dots in one cell were found and analyzed among all the cells in the scope.

Discussion

Ectopic gene expression is a cornerstone of modern biology, as well as gene and cell therapy, but most approaches produce variable and dosage-sensitive expression distributions. Here, we engineered a set of miRNA-based circuits that allow more precise, and gene dosage-invariant, control of protein expression. These circuits also allow tuning of expression level through sequence (**Figure 2 and 3**), orthogonal control of multiple target genes (**Figure 4**), and portability across cell types and modes of delivery (**Figure 5**). We therefore anticipate that they could become standard systems for controlled gene expression in diverse areas of biomedical science and biotechnology, including imaging (**Figure 6**).

The process of engineering robust dosage invariant circuits required identifying design principles for miRNA regulation. The fully complementary single target site produced an extremely low setpoint of expression (**Figure 2C**). Reducing complementarity is believed to tune the binding affinity between the miRNA and the target, and also the inhibition strength, therefore could have raised the expression level of the circuit. However, systematic variation of single target site complementarity showed that reducing complementarity removed regulation altogether (**Figure 2D**). Remarkably, however, TNRC6-dependent cooperativity among multiple target sites can restore regulation to sites that were individually weak, and implement dosage invariance (**Figure 2E**). Thus, while cooperativity in other regulatory systems is associated with ultrasensitivity, here it allows linear dependence of target expression on miRNA levels, which is required for dosage invariance (**Supplementary Modeling**). These results suggest that weak, multivalent miRNA regulation, possibly by avoiding direct cleavage, can provide the linear responses required for effective dosage compensation. We also note that the behavior observed here is broadly consistent with evidence for an important role for weak cooperative interactions in natural miRNA regulatory systems[43,44].

We found that varying miRNA regulation strength by complementarity was straightforward for one case of miR-L, but was more complex for other synmiRs generated, indicating that additional factors besides base pairing are important for determining the activity of a miRNA on its target. It will be interesting to determine whether the molecular regulatory mechanism in the dosage compensated regime is catalytic, and possibly slicer-independent[19,45].

A second design feature of these circuits is the separation of miRNA and target gene into separate, divergently transcribed genes. Early designs incorporating the miRNA and target into a single transcript failed to achieve full dosage invariance (**Figure S2**). The divergent design allows strong expression of the miRNA and avoids the miRNA effectively inhibiting its own production. Further engineering may allow even more compact designs.

The system may be extended or improved in different ways. First, we identified empirical rules that are effective for designing orthogonal miRNA sequences. However, better miRNA sequence-function models could potentially allow predictive control of set points. Second, current expression distributions exhibit significant variability. This expression noise could reflect transcriptional bursting of the target gene as well as the miRNA encoding gene. Differences in the time delays for miRNA production and processing compared to RNA splicing and nuclear export could also disrupt ideal regulation. Going forward, it will be important to quantify these and other contributions to overall variability[3,4,46,47] and to identify the limits of precision possible with miRNA dosage compensation systems. Third, while circuits achieve dosage compensation across multiple cell lines (**Figure 5A, Figure S9**), total protein levels vary. Regulation of translation or protein degradation could help to achieve cell type independent control of protein concentration. Finally, the ability to further modulate expression with inducers could make these systems even more useful.

The DIMMER circuits developed here represent a powerful toolbox for researchers and biomedical engineers. They can reduce background in imaging (**Figure 6**), allow probing of concentration dependent effects in numerous studies, could ensure a fixed expression level for receptors or other components in cell therapies[48], and could also help make ectopic transcription factor expression more precise in regenerative medicine[49,50]. One major application category is gene therapy. Many diseases that could be targeted by gene therapy exhibit toxicity at high levels of the therapeutic gene, making it critical to suppress overexpression[51–55]. Thus, we anticipate these systems becoming useful components in a wide range of engineered research and therapeutic contexts.

Acknowledgement

We thank Phillip Zamore (UMass), Viviana Gradinaru, Acacia Mayfield, James Linton, Dongyang Li, Bo Gu, Kaiwen Luo, Martin Tran, Duncan Chadly, Felix Horns, Lucy Chong, Leah Santat, Sheng Wang, Shiyu Xia, Yodai Takei for discussion and technical support; Inna-Marie Strazhnik (Caltech) for graphical design; Fei Chen (Broad Institute), Jacob Parres-Gold, Evan Mun, Judy Shon, and other members of the Elowitz lab for critical feedback, and administrative support. This work is supported by the National Institute Of Biomedical Imaging And Bioengineering of the National Institutes of Health under Award Number R01EB030015. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. M.B.E. is a

Howard Hughes Medical Institute Investigator. We apologize for incomplete citations due to space limits.

Author contributions

R.D., M.F., and M.B.E. conceived and designed the study. M.B.E. directed and supervised the study. R.D., M.F., and M.H. performed or assisted with experiments and analyzed data. R.D. and M.B.E. wrote the manuscript with input from all authors.

Declaration of interests

Patent applications related to this work have been filed by the California Institute of Technology (US application numbers CIT-9070-P/CIT-9070-P2). M.B.E is a scientific advisory board member or consultant at TeraCyte, Primordium, and Spatial Genomics.

Data and materials availability

Data needed to evaluate the conclusions are available in the paper or the Supplementary Materials. Raw data and computer codes used for data analysis are available from the corresponding author.

Methods

Supplementary Modeling

Plasmids construction

Constructs used in this study are listed in **Table S1**. Some constructs were generated using standard cloning procedures. The inserts were generated using PCR or gBlock synthesis (IDT) and were ligated either by T4 ligase (NEB #M0202M) or In-Fusion (Takara #102518) assembly with backbones that are linearized using restriction digestion. The rest of the constructs were designed by the authors and synthesized by GenScript. Selected constructs will be deposited at Addgene and the maps are available.

miRNA alignment to the database

Each synthetic miRNA sequence (mature miRNA, 22 nt) is aligned to the known miRNA sequence database[56–63] (https://mirbase.org/) to identify if there are any similarities existing between the synthetic sequences and the natural sequences.

Tissue Culture

U2OS cells, T-Rex cells, CHO cells, and N2A cells were cultured at 37°C in a humidity-controlled chamber with 5% CO2. The growth media consisted of DMEM (Dulbecco's Modified Eagle Medium, ThermoFisher #11960-069) supplemented with 10% FBS, 1 U/ml penicillin, 1 µg/ml streptomycin, 1 mM sodium pyruvate, 1X NEAA (ThermoFisher #11140-050), 1 mM L-glutamine, and 0.1 mg/mL Normocin (InvivoGen #ant-nr).

Transient Transfection

Cells were seeded at a density of 50,000 cells in each well of a 24 well plate, either standard for flow cytometry or glass-bottom for imaging experiments, and cultured under standard conditions overnight. The following day, the cells were transiently transfected using Fugene HD (Promega #E2311), according to the manufacturer's protocol.

Flow Cytometry
Cells were incubated 2 days after transient transfection, and the culture media was replaced 24 hours post-transfection. Cells were trypsinized with 75 µL of 0.25% trypsin for 5 minutes at 37 C. After digestion, cells were resuspended with 125 µL of HBSS containing 2.5 mg/ml BSA and 1 mM EDTA. Cells were then filtered through a 40 µm cell strainer and analyzed using a CytoFLEX S instrument (Beckman Coulter). We used the self-build python code to analyze the flow data.

Cell sorting
To prepare the mono-clones that expressed the genomic-integrated DIMMER circuit, cells were harvested and resuspended in sorting buffer (BD FACS Pre-Sort Buffer) supplemented with 1 U/ml DNAse I by the cell sorter (Sony MA900) as mono-clones. Cells were sorted into 96 well plates in the normal U2OS culture media. Cells were expanded in the 24 well plate before measurement using the flow cytometer.

In-vitro image analysis
The transiently transfected U2OS cells were imaged using a Nikon confocal microscope at 60× magnification, such that each image was spaced by 0.5 microns in the z-direction. Images were processed by the Fiji software[64].
To analyze the relative signal intensity in the dCas9 imaging experiment, maximum intensity projection of 11 slices of the z-stacks were applied. To determine the signal intensity of the dots, freehand lines were drawn to select the dot regions. To determine the signal intensity in the background, ~ 5 micron-long straight lines centering the dots were drawn. The signal intensities were generated by the Fiji ROI mean intensity function. The relative signal intensity is calculated by normalizing the background intensity to be 1.
To analyze the signal to noise ratio (SNR), freehand lines were drawn to select the dot regions and the nucleus regions. The noise intensity was calculated by the intensity in the nucleus excluding the dots area. The SNR was calculated dividing the mean intensities of each dot by the noise.

Bulk RNA sequencing
Sample preparation and sequencing
To verify the off-target effect of all the synmiRs, U2OS cells were plated on 6-well plates with 300,000 cells per well. Cells were transfected the following day with 1,000 ng of either the control plasmid or the BFP-miRNA plasmid using Fugene HD (Promega #E2311) according to the manufacturer's instructions. Media was replaced with 2 mL of fresh media 24 hours post-transfection. Cells were harvested 48 hours post-transfection by digestion with 0.25% Trypsin-EDTA, centrifugation at 300g for 5 minutes, and removal of the supernatant by aspiration. The cell's pellet was stored in -80 C prior to the purification.
RNA was extracted using the RNeasy kit (Qiagen #74106) according to the manufacturer's instructions. RNA was treated with Turbo DNase (ThermoFisher #AM2238) and purified using the RNeasy kit RNA cleanup protocol. mRNA sequencing libraries were prepared by Novogene.

Preprocessing of sequencing data
Reads from the RNA sequencing were aligned to a custom reference genome using kallisto (0.48.0)[65]. This reference consisted of the human genome GRCh38 cDNA (https://github.com/pachterlab/kallisto-transcriptome-indices/releases) and mTagBFP2

coding sequences. Weakly expressed genes were filtered out if they exhibited fewer than 3 samples expressing at least 10 transcripts per million (TPM), or if the maximum TPM among all samples was less than 105. Then, filtered counts were input to DEseq to eliminate the impact of size factors. As the BFP-only cells were used as a reference to evaluate the off-target effect of the miRNAs, genes that showed fluctuating expressions among the three biological replicates of BFP-only cells should be removed from analysis. To achieve this, we computed log(1+x), where x denotes the normalized TPM among the three biological replicates of BFP-only cells. The Fano factors[66,67] of these logarithmic expressions were determined and ranked. Transcripts that ranked as the largest 2.5% in logarithmic Fano factors were eliminated from further analysis. Finally, we computed log(1+x), where x denotes the normalized TPM among all the samples. A difference function was defined to compute the absolute value of the log(1+x) difference between each sample and the non-transfected sample. The medians of the difference function of the BFP-only groups and the experimental groups were calculated and used for comparison. The difference between those two difference functions were ranked and similarly, transcripts that ranked as the largest 3% were removed from further analysis.

The Fano factor is defined as:

$$\text{Fano factor} = \frac{\text{variance}(\log(1+x))}{\text{mean}(\log(1+x))}$$

The equation of the difference function is defined as:

$$\Delta(\text{BFP, untransfected}) = |\log(1+x_{\text{BFP}}) - \log(1+x_{\text{untransfected}})|$$

$$\Delta(\text{experimental, untransfected}) = |\log(1+x_{\text{experimental}}) - \log(1+x_{\text{untransfected}})|$$

$$\Delta_{\text{ranked}} = \text{median}(\Delta(\text{BFP}, \text{untransfected})) - \text{median}(\Delta(\text{experimental}, \text{untransfected}))$$

Differential gene expression analysis

To characterize the perturbations that synthetic miRNA brought to the endogenous transcriptome, differential expression analysis was performed using DESeq2 (1.40.1)[68] in R (4.3.1) comparing transcript counts in miRNA transfected cells and BFP-only cells.

DNA-PAINT

Buffers

Blocking buffer: 1×PBS, 1 mM EDTA (Thermo Fisher, no. AM9260G), 0.02% Tween-20 (Life Science, no. P7949), 2% BSA (Sigma-Aldrich, no. A9647-100G)

Washing buffer: 1×PBS, 1 mM EDTA, 0.01% Tween-20

Imaging buffer: 1×PBS, 1 mM EDTA, 500 mM NaCl (Thermo Fisher, no. AM9760G)

Cloning

An mEGFP Gblock (obtained from Integrated DNA Technologies) was inserted into a pcDNA3.1(+) backbone (ThermoFisher, no. V79020) via Gibson assembly. Two codon-optimized fragments of human EGFR (obtained from Integrated DNA Technologies) were fused to the mEGFP-pcDNA3.1(+) backbone via Gibson assembly.

Cell culture

CHO-K1 cells (ATCC: CCL-61) were cultured in Gibco Ham's F-12K (Kaighn's) medium supplemented with 10% FBS (Gibco, no. 11573397) at 37°C and 4% CO2. Cells were split via trypsinization using trypsin-EDTA (Gibco, no. 25300096) every 2–3 days.

Nanobody-DNA conjugation

First, the GFP nanobody (clone 1H1, Nanotag Biotechnologies) was conjugated to a DBCO-PEG4-Maleimide linker (no. CLK-A108P, Jena Bioscience). After removing unreacted linker with Amicon centrifugal filters (10,000 MWCO), the DBCO-nanobody was conjugated via DBCO-azide click chemistry to the docking strand 5'- azide - CTC TCT CTC TCT CTC TCT C -3' (Metabion). A detailed description of the conjugation can be found in our previous protocol[69].

DNA-PAINT imaging

5,000 cm-2 CHO-K1 cells were seeded on an ibidi eight-well high glass-bottom chambers (no. 80807) one day before transfection. The cells were transfected with EGFR-mEGFP plasmids with a ThermoFisher Lipofectamine 3,000 reagent (no. L3000008) with a lower Lipofectamine concentration as indicated by the manufacturer and 250 ng plasmid per well (200µL solution per well and 25µL transfection solution). After 48h of transfection, the cells were fixed with 250 µL of pre-warmed methanol-free 4% PFA (ThermoFisher, no. 043368.9M) in 1×PBS for 15 min. The cells were washed 3 times with 1×PBS and then permeabilized with 0.125% TritonX-100 (Sigma Aldrich, no. 93443) in 1×PBS for 2 min. After washing 3 times with 1×PBS, the cells were blocked with the blocking buffer overnight at 4 °C.

The cells were then washed three times with the washing buffer. 25 nM anti-GFP nanobodies were incubated in the blocking buffer for 1h at RT. After washing 3 times with the washing buffer, the sample was incubated with the imaging buffer for 5 min. The nanobodies were then post-fixed with 4% PFA in 1×PBS for 5 min. The cells were then washed 3 times with the washing buffer and once with 1×PBS. 90 nm gold-nanoparticles (Absource, no. G-90-100) in 1:1 in 1×PBS were incubated for 5 min at RT. After washing three times with 1×PBS, the cells were washed once with the imaging buffer. The samples were imaged in the imaging buffer with 100 pM imager strand (5'- GAG AGA G - Cy3B 3', obtained from Metabion) for 40k frames with 100 ms exposure time per frame and a readout rate of 200 MHz.

Microscope setup

The samples were measured on inverted total internal reflection fluorescence (TIRF) microscopes (Nikon Instruments, Eclipse Ti2) which are equipped with an oil-immersion objective (Nikon Instruments, Apo SR TIRF ×100/numerical aperture 1.49, oil) and a perfect focusing system. The mRuby3 signal was bleached by the 560 nm laser (MPB Communications, 1 W) by using Highly inclined and laminated optical sheet (HILO) illumination. Afterwards, the TIRF mode was established. The Cy3B imagers were excited with the 560 nm laser. The laser beam was cleaned with a filter (Chroma Technology, no. ZET561/10) and coupled into the microscope with a beam splitter (Chroma Technology, no. ZT561rdc). The fluorescent signal was filtered with an emission filter (Chroma Technology, nos. ET600/50m and ET575lp) and projected onto a sCMOS camera (Hamamatsu, ORCA-Fusion BT) without further magnification.

The acquired region of interest has a size of 576 × 576 pixels. The resulting effective pixel size is 130 nm. The raw microscopy data was acquired via µManager.

DNA-PAINT analysis

Obtained fluorescent data was reconstructed with Picasso software[38]. The data was first drift-corrected with redundant cross-correlation, after that with picked gold particles as fiducials.
In order to determine the receptor density, a homogeneous area of the cells was picked and the DNA-PAINT data was clustered with the SMLMS clustering algorithm of Picasso[38,39].
The determined cluster centers were used to calculate the receptor density per μm^2 (number of cluster centers per area).

Supplementary Figures

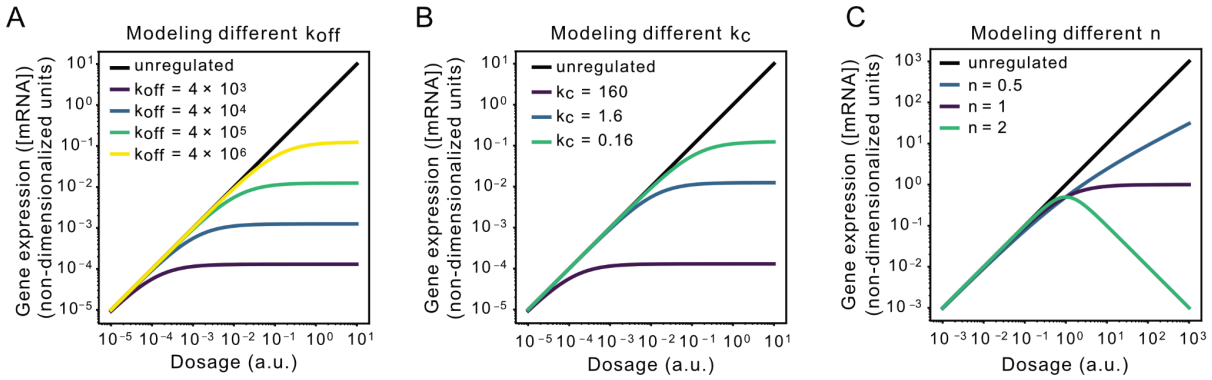

Figure S1. Modeling of the IFFL.

A, modeling the steady-state, non-dimensionalized mRNA concentration under different k_{off} while maintaining $k_{\mathrm{on}} = 2 \times 10^5$, $k_{\mathrm{c}} = 160$. B, modeling the steady-state, non-dimensionalized mRNA concentration under different k_{c} while maintaining $k_{\mathrm{on}} = 2 \times 10^5$, $k_{\mathrm{off}} = 4 \times 10^3$. C, modeling the steady-state, non-dimensionalized mRNA concentration under different Hill coefficient n while maintaining $\mathrm{K} = 1$.

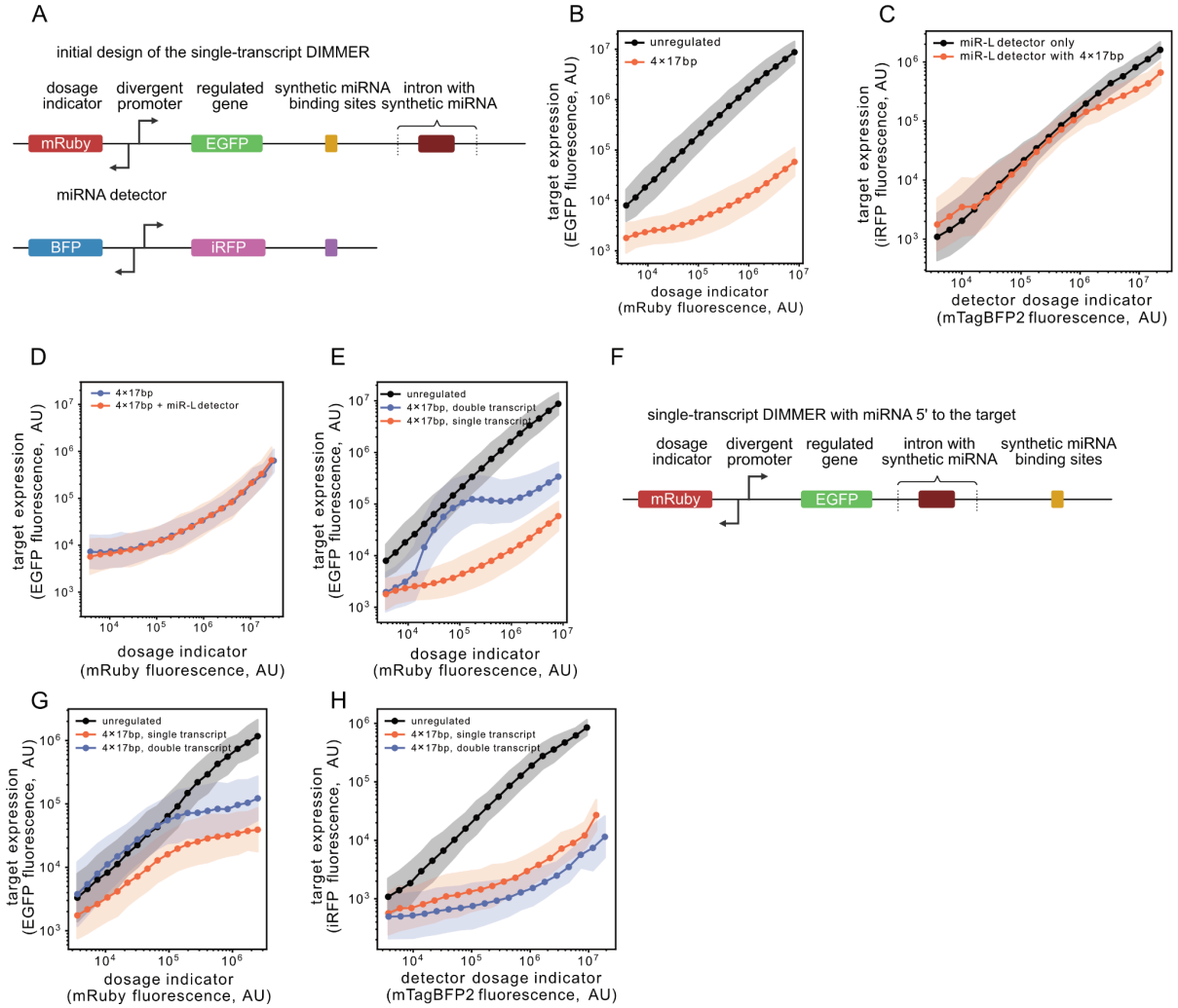

Figure S2. Comparisons between circuit configurations in the single transcript and the double transcript design.

(A) The diagram of the initial design of the miR-L 4×17bp circuit (upper) and the miRNA detector circuit (lower). The miRNA detector has the corresponding miRNA target site on the 3'UTR of the iRFP.

(B) We performed flow cytometry on the cells transfected with the initial design of the miR-L 4×17 bp circuit displayed in the top panel of **(A)**. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. The initial design shows a very low EGFP expression level.

(C) We performed flow cytometry on the cells transfected with the 4×17bp miRNA detector alone (**A**, bottom) or with the initial miR-L 4×17bp circuit (**A**, top). Cells were gated and binned by mTagBFP2 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. The detector of the initial circuit just shows very weak inhibition at very high dosage, and this inconsistency in inhibition suggests that the low expression EGFP level of the initial circuit doesn't come from miRNA inhibition.

(D) We performed flow cytometry on the cells transfected with the initial miR-L 4×17bp circuit, or the initial miR-L 4×17bp circuit along with the 4×17bp miRNA detector. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. The addition of the detector doesn't affect the response of the initial miR-L 4×17bp circuit.

(E) We performed flow cytometry on the cells transfected with either the initial miR-L 4×17bp circuit or the double transcript design illustrated in **Figure 2A**. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**.

(F) The diagram of the new single transcript design of the miR-L 4×17bp circuit. The intronic miR-L cassette was shifted to the 5' to the target site.
(G-H) We performed flow cytometry on the cells co-transfected with the double transcript miR-L 4×17bp circuit described in **Figure 2A** and the 10×21bp miRNA detector, versus the new single transcript design of the miR-L 4×17bp shown in **(F)** and the 10×21bp miRNA detector. Compared to the 4×17bp miRNA detector, the 10×21bp miRNA detector is able to detect less miRNA. Cells were gated and binned by mRuby3 **(G)** or mTagBFP2 **(H)**, respectively. Dots and shaded regions were calculated as described in **Figure 2C**. Unregulated represents either a mRuby3-divergent promoter-EGFP construct **(G)** or a mTagBFP2-divergent promoter-iRFP670 construct **(H)** without the regulation cassettes and the target sites. The new single transcript design shows some dosage-independence ranges, but still presents the miRNA-independent inhibition to some extent.

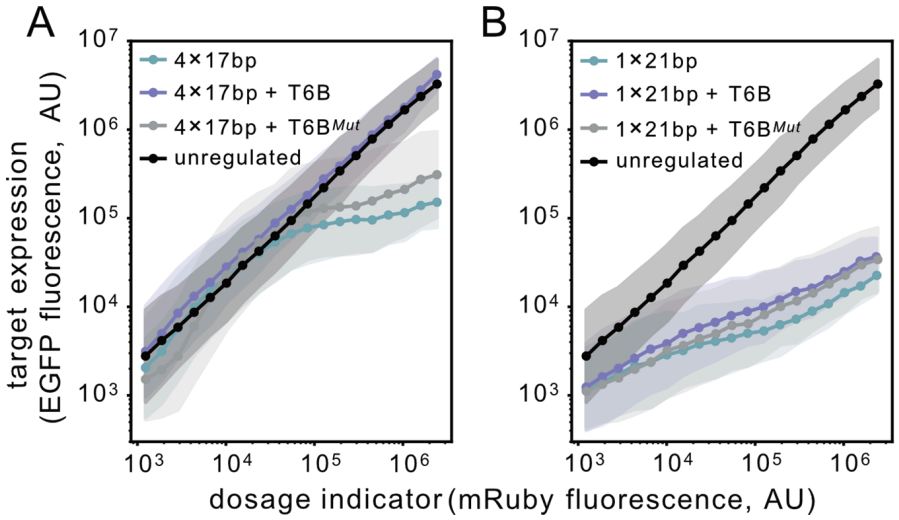

Figure S3. Biologically dead T6B fails to remove TNRC6-dependent inhibition of the miR-L 4×17bp circuit.

(A-B) We performed flow cytometry on the cells transfected with the miR-L 4×17bp circuit **(A)** or miR-L 1×21bp circuit **(B)** described in **Figure 2B** along with the fluorescent protein-only negative control, the T6B peptide, or the catalytically dead T6B peptide (denoted as $T6B^{Mut}$), respectively. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. The design of the catalytically dead T6B peptide was consistent with the previous literature[20].

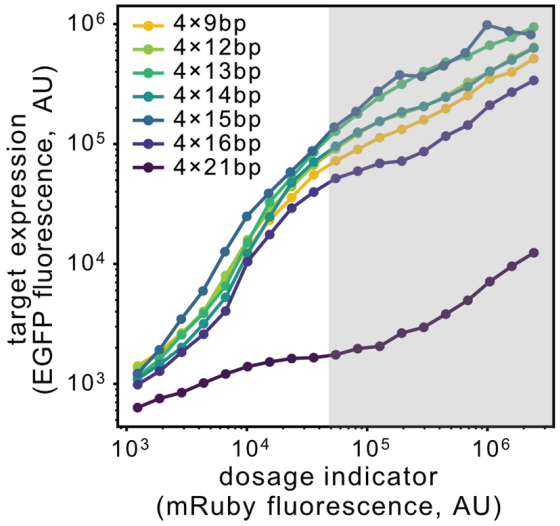

Figure S4. The dosage response curves of multiple miR-L 4×n circuits.
We performed flow cytometry on the cells transfected with the miR-L regulating 4×n circuits described in **Figure 3A**. Gray rectangle indicates the gated region of the **Figure 3A** left panel. Cells were gated and binned by mRuby3 intensities. Dots were calculated as described in **Figure 2C**.

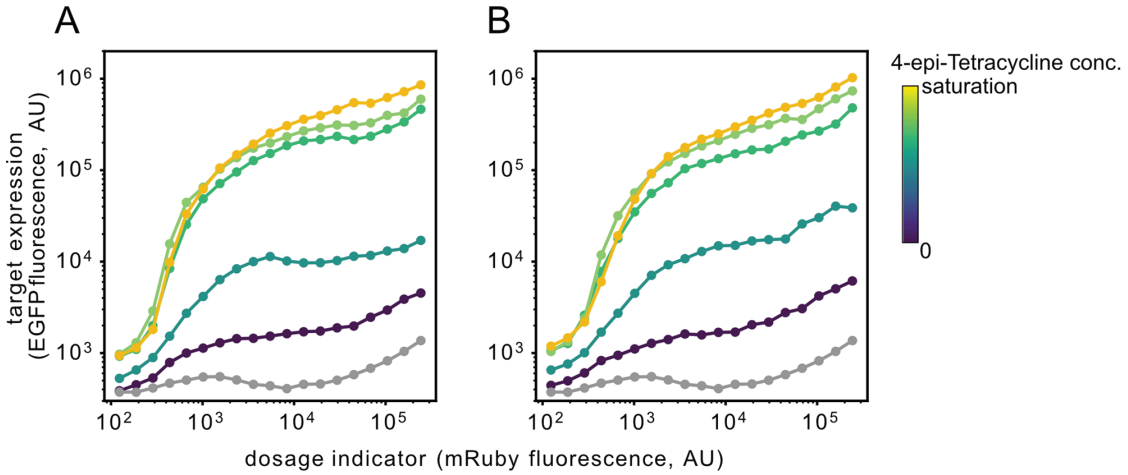

Figure S5. The dosage response curves of the inducible DIMMER circuits in TRex cells. **(A-B)** We performed flow cytometry on the TRex cell line which was transfected with the 4×17 bp **(A)** or 4×18bp **(B)** DIMMER construct, and the EGFP was driven by the CMV-2×TetO promoter. The concentrations of the 4-epi-Tetracycline, from purple to yellow, were 0, 10, 33.3, 100, 333.3 ng/mL. The gray curve denotes the mRuby-only transfection control, and the 4-epi-Tetracycline concentration is 333.3 ng/mL.

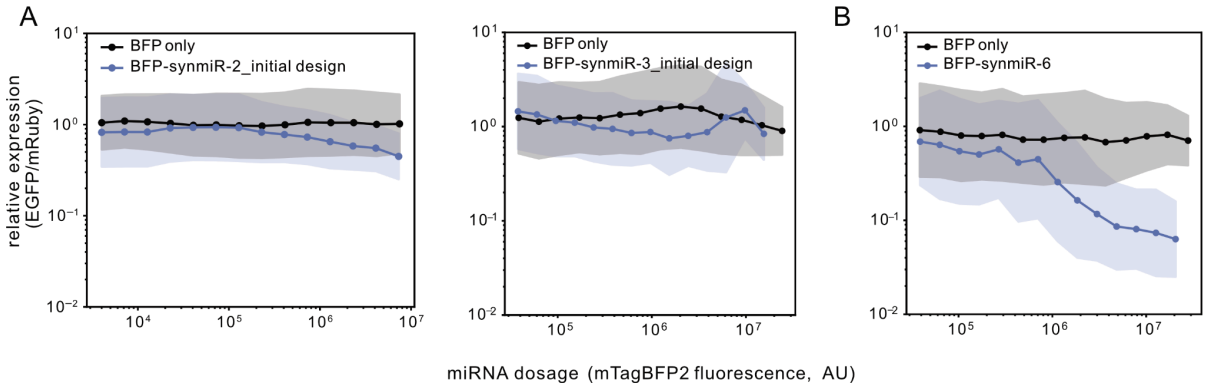

Figure S6. Initial designs of synmiR-2 and synmiR-3 didn't work. synmiR-6 worked but showed a sequence similarity to the endogenous miRNA.

We performed flow cytometry on the cells co-transfected with the circuit shown in **Figure 4A** using the initial sequence of synmiR-2 (**A, left panel**), synmiR-3 (**A, right panel**), synmiR-6 (**B**), and their corresponding targets with a single, fully complementary target site, respectively. The BFP only control does not have the 3'UTR miRNA. Cells were gated by mRuby3 and mTagBFP2 intensities and binned by mTagBFP2 intensities. Relative expression levels were quantified as described in **Figure 4B**. Dots and shaded regions were calculated as described in **Figure 2C**.

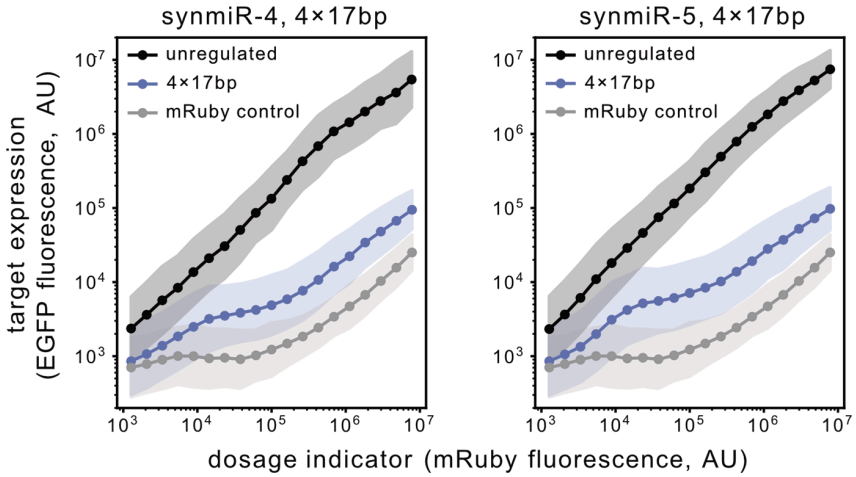

Figure S7. 4×17 bp target of synmiR-4 and synmiR-5 showed strong regulation.
We transfected the cells with the circuit architecture in **Figure 2A** harboring either the synmiR-4 cassette and the corresponding 4×17 bp target, or synmiR-5 cassette and the corresponding 4×17 bp target and performed flow cytometry. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**.

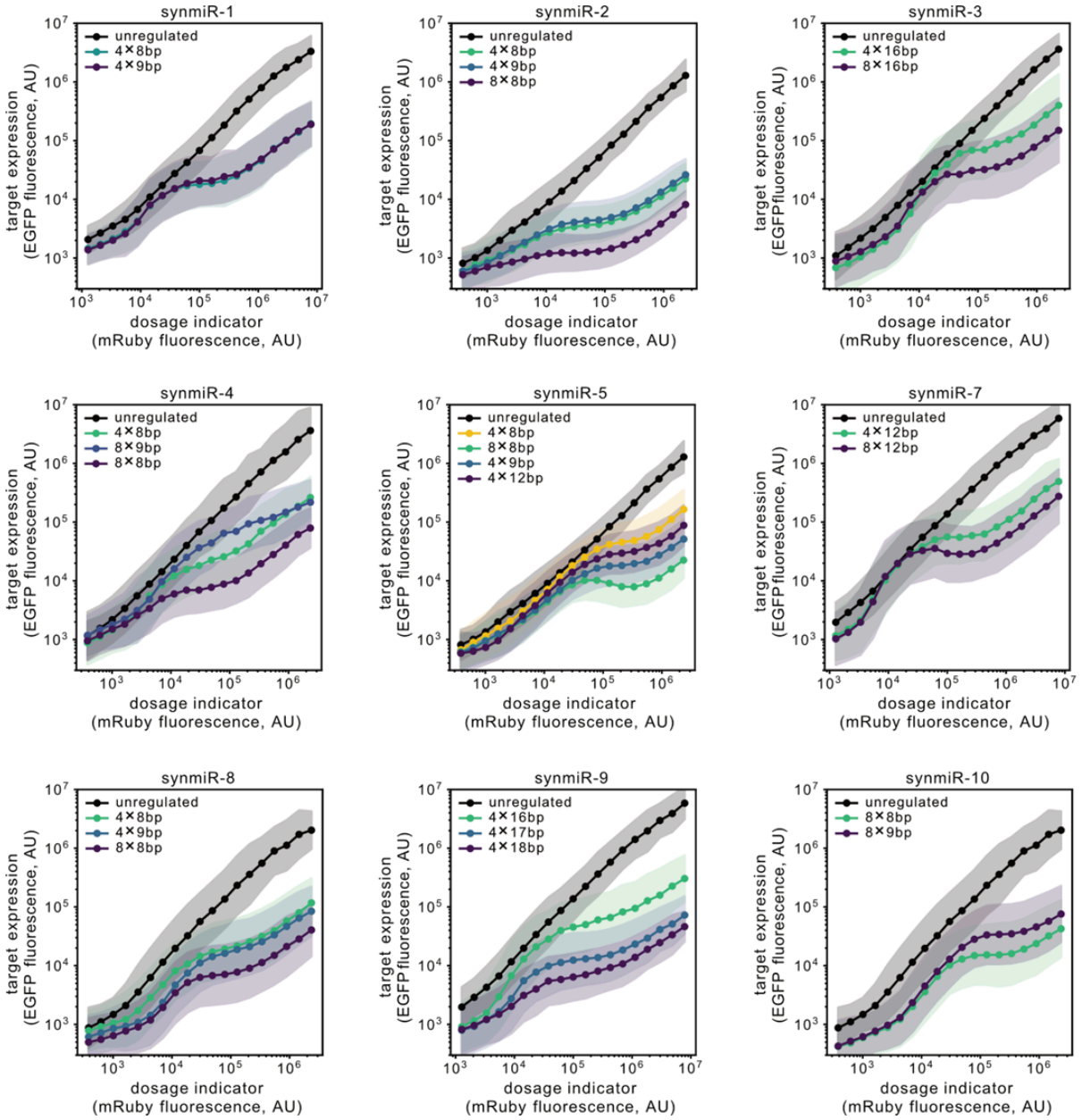

Figure S8. A gallery of all the IFFL designs based on different miRNAs and targets.
We transfected the cells with the circuit architecture in **Figure 2B** harboring the miRNA cassette denoted on the top of the plot and the corresponding target. Dots and shaded regions were calculated as described in **Figure 2C**. Target site nomenclature was described in **Figure 4F**.

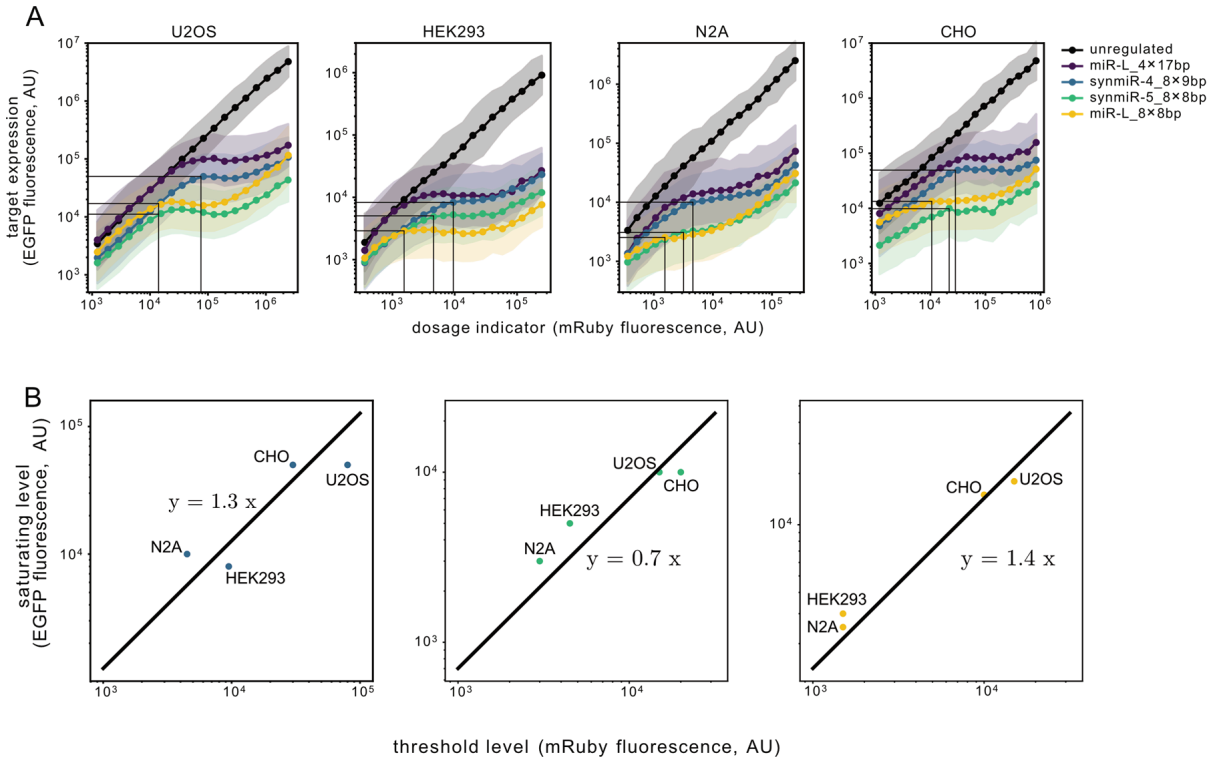

Figure S9. IFFL works across different cell lines.
(A) We transfected the circuit described in **Figure 2B** regulated by the miRNA and target site denoted into different cell lines and performed flow cytometry. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. Black vertical lines intersecting the x axis suggest the threshold levels of mRuby3 fluorescence intensities. Black horizontal lines intersecting the y axis suggest the saturating levels of EGFP fluorescence intensities. **(B)** The values of the intersection points of the gray lines with the axes in **(A)** were plotted as scatters. The black solid lines were calculated as described in **Figure 5B**.

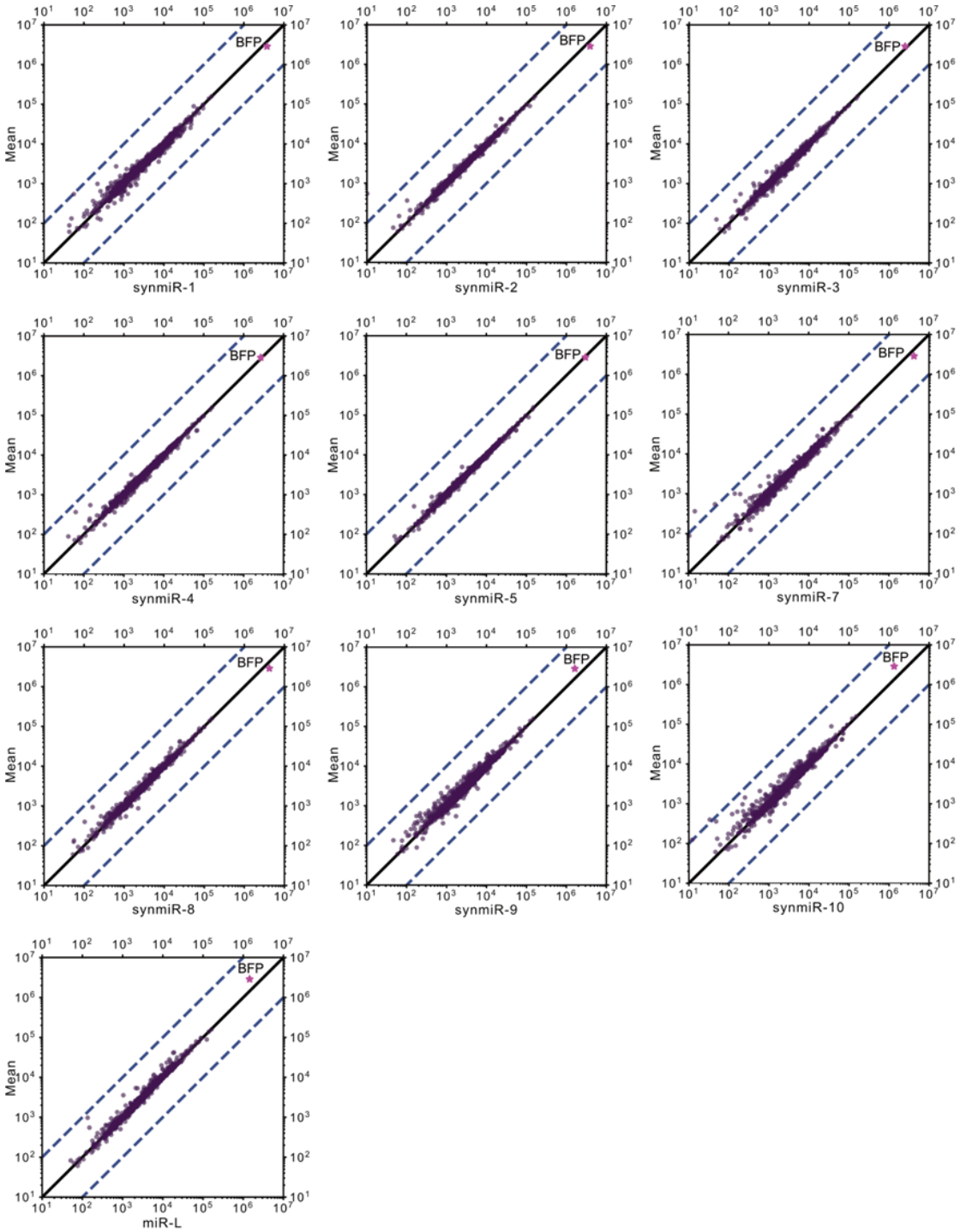

Figure S10. Plotting of the normalized transcripts per million (TPM) of the synthetic miRNA expressing cells versus the mean TPM.

The mean TPM was calculated by averaging all the TPM of all the microRNA-expressing cell samples. Solid line indicates where the TPM of the two samples were equal to each other. Dashed lines indicate where the TPM of the synmiR expressing cells is 10 fold of the mean, or where the TPM of the latter is 10 fold of the former.

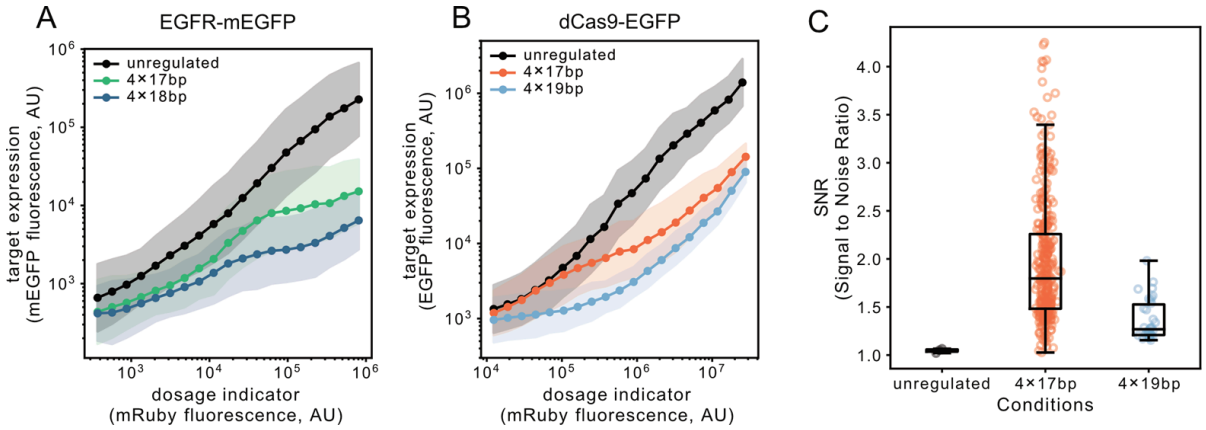

Figure S11. IFFL improves fluorescence imaging.

We performed flow cytometry on cells that were transfected with the EGFR-mEGFP with or without the DIMMER module **(A)** and the dCas9-EGFP with or without the DIMMER module **(B)**, respectively. Cells were gated and binned by mRuby3 intensities. Dots and shaded regions were calculated as described in **Figure 2C**. **(C)** shows the quantification of the signal-to-noise ratio (SNR) of the dots in the cells transfected with the dCas9-EGFP with or without the DIMMER module. Each scatter represents one dot inside a cell.

Supplementary Tables

Table S1 List of miRNA and targets used in this study and their sequences, uses in the figures

miRNA/target name	Sequence	Usage in the figures
synmiR-1	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCCAACATAAGCATAAACTA CGATAGTGAAGCCACAGATG TATCGTAGTTTATGCTTATGTT GATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	4D, 4E, 4F, S8, S10
synmiR-2_initial version	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCTAATCTACAAAGTGACAT AAATAGTGAAGCCACAGATG TATTTATGTCACTTTGTAGATT AATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	S6A
synmiR-2_20G	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCTGATCTACAAAGTGACAT AAATAGTGAAGCCACAGATG TATTTATGTCACTTTGTAGATC	4D, 4E, 4F, 4H, S8, S10

	AATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	
synmiR-3_initial version	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCGAAAAGTTTATAATATCTT GATAGTGAAGCCACAGATGT ATCAAGATATTATAAACTTTTC ATGCCTACTGCCTCGGACTT CAAGGGGCTAGAATTCGAGC AATTATCTTGTTTACTAAAACT GAATACCTTGCTATCTCTTTG ATACATTTTTACAAAGCTGAAT TAAAATGGTATAAATTAAATCA CTTTTCCTGACCATTCATCCT CTTTCTTTTTCCT	S6A
synmiR-3_19G	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCGAGAAGTTTATAATATCT TGATAGTGAAGCCACAGATG TATCAAGATATTATAAACTTCT CATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	4D, 4E, 4F, 5E, 5F, S8, S10
synmiR-4	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCCGTAAAGATCATGAAATT GAATAGTGAAGCCACAGATG TATTCAATTTCATGATCTTTAC	4D, 4E, 4F, S7-10

	GATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	
synmiR-5	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCAGCATCTATCTAACGGTT TGATAGTGAAGCCACAGATG TATCAAACCGTTAGATAGATG CTATGCCTACTGCCTCGGAC TTCAAGGGGCTAGAATTCGA GCAATTATCTTGTTTACTAAAA CTGAATACCTTGCTATCTCTT TGATACATTTTTACAAAGCTG AATTAAAATGGTATAAATTAAA TCACTTTTCCTGACCATTCAT CCTCTTTCTTTTTCCT	4D, 4E, 4F, 4H, S7-10
synmiR-6	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCCTCTTATCAAGCAGTTTC ATATAGTGAAGCCACAGATGT ATATGAAACTGCTTGATAAGA GATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	S6B
synmiR-7	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCAGAAATGGTGTAATTTAG CAATAGTGAAGCCACAGATG TATTGCTAAATTACACCATTTC	4D, 4E, 4F, S8, S10

	TATGCCTACTGCCTCGGACTT CAAGGGGCTAGAATTCGAGC AATTATCTTGTTTACTAAAACT GAATACCTTGCTATCTCTTTG ATACATTTTTACAAAGCTGAAT TAAAATGGTATAAATTAAATCA CTTTTCCTGACCATTCATCCT CTTTCTTTTTCCT	
synmiR-8	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCCGTTGAGATTTAAAGATC GAATAGTGAAGCCACAGATG TATTCGATCTTTAAATCTCAAC GATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	4D, 4E, 4F, S8, S10
synmiR-9	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCCAACACCTTTCTTAAAAC TTATAGTGAAGCCACAGATGT ATAAGTTTTAAGAAAGGTGTT GATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	4D, 4E, 4F, S8, S10
synmiR-10	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCGGTAATCTTAAGCATAGA TGATAGTGAAGCCACAGATG TATCATCTATGCTTAAGATTAC	4D, 4E, 4F, S8, S10

	CATGCCTACTGCCTCGGACT TCAAGGGGCTAGAATTCGAG CAATTATCTTGTTTACTAAAAC TGAATACCTTGCTATCTCTTT GATACATTTTTACAAAGCTGA ATTAAAATGGTATAAATTAAAT CACTTTTCCTGACCATTCATC CTCTTTCTTTTTCCT	
miR-L	TTGAATGAGGCTTCAGTACTT TACAGAATCGTTGCCTGCAC ATCTTGGAAACACTTGCTGG GATTACTTCGACTTCTTAACC CAACAGAAGGCTCGAGAAGG TATATTGCTGTTGACAGTGAG CGCAGGAATTATAATGCTTAT CTATAGTGAAGCCACAGATGT ATAGATAAGCATTATAATTCCT ATGCCTACTGCCTCGGACTT CAAGGGGCTAGAATTCGAGC AATTATCTTGTTTACTAAAACT GAATACCTTGCTATCTCTTTG ATACATTTTTACAAAGCTGAAT TAAAATGGTATAAATTAAATCA CTTTTCCTGACCATTCATCCT CTTTCTTTTTCCT	2C-F, 2H-I, 3A-E, 4B, 4D, 4E, 4H, 5A, 5B, 5D-F, 6B, 6D-F, 6H, S2-5, S9-11
Note: the mature miRNA sequence (including the sense chain and the antisense chain) is shaded with light blue.		
miR-L_1×21bp	AGGAATTATAATGCTTATCTA	2C, 2D, 2I, 4B, 4D, 4E, S3B
miR-L_1×17bp	TACTATTATAATGCTTATCTA	2D
miR-L_1×18bp	CTTAATTATAATGCTTATCTA	2D
miR-L_1×19bp	TTGAATTATAATGCTTATCTA	2D
miR-L_1×20bp	TGGAATTATAATGCTTATCTA	2D
miR-L_2×17bp	TACTATTATAATGCTTATCTATA CTATTATAATGCTTATCTA	2E
miR-L_3×17bp	TACTATTATAATGCTTATCTATA CTATTATAATGCTTATCTATAC TATTATAATGCTTATCTA	2E
miR-L_4×17bp	TACTATTATAATGCTTATCTATA CTATTATAATGCTTATCTATAC TATTATAATGCTTATCTATACT ATTATAATGCTTATCTA	2E, 2F, 2H, 3A-D, 4H, 5A, 5B, 5D, 6E, 6F, 6H, S2, S3A, S5A, S9-S11
miR-L_4×8bp	TACTTGGCATTATCTTATCTAT ACAACAATTTACTTGGCATTAT CTTATCTATAATCACTTGTACT	3A

	TGGCATTATCTTATCTATACAA CTGGATACTTGGCATTATCTT ATCTA	
miR-L_4×9bp	TACTTGGCATTAGCTTATCTAT ACAACAATTTACTTGGCATTA GCTTATCTATAATCACTTGTAC TTGGCATTAGCTTATCTATACA ACTGGATACTTGGCATTAGCT TATCTA	3A, S4
miR-L_4×10bp	TACTTGGCATTTGCTTATCTAT ACAACAATTTACTTGGCATTT GCTTATCTATAATCACTTGTAC TTGGCATTTGCTTATCTATACA ACTGGATACTTGGCATTTGCT TATCTA	3A
miR-L_4×11bp	TACTTGGCATATGCTTATCTAT ACAACAATTTACTTGGCATAT GCTTATCTATAATCACTTGTAC TTGGCATATGCTTATCTATACA ACTGGATACTTGGCATATGCT TATCTA	3A
miR-L_4×12bp	TACTTGGCAAATGCTTATCTA TACAACAATTTACTTGGCAAA TGCTTATCTATAATCACTTGTA CTTGGCAAATGCTTATCTATA CAACTGGATACTTGGCAAATG CTTATCTA	3A, S4
miR-L_4×13bp	TACTTGGCTAATGCTTATCTAT ACAACAATTTACTTGGCTAAT GCTTATCTATAATCACTTGTAC TTGGCTAATGCTTATCTATACA ACTGGATACTTGGCTAATGCT TATCTA	3A, S4
miR-L_4×14bp	TACTTGGATAATGCTTATCTAT ACTTGGATAATGCTTATCTATA CTTGGATAATGCTTATCTATAC TTGGATAATGCTTATCTA	3A, S4
miR-L_4×15bp	TACTTGTATAATGCTTATCTAT ACAACAATTTACTTGTATAATG CTTATCTATAATCACTTGTACT TGTATAATGCTTATCTATACAA CTGGATACTTGTATAATGCTTA TCTA	3A, S4
miR-L_4×16bp	TTATAATGCTTATCTAAATATTA TAATGCTTATCTAAATATTATAA TGCTTATCTAAATATTATAATG CTTATCTA	3A, S4

miR-L_4×18bp	AATTATAATGCTTATCTAAATAAATTATAATGCTTATCTAAATAAATTATAATGCTTATCTAAATAAATTATAATGCTTATCTA	3A, 3B, 6B, 6E, 6F, S5B, S11A
miR-L_4×19bp	GAATTATAATGCTTATCTAAATAGAATTATAATGCTTATCTAAATAGAATTATAATGCTTATCTAAATAGAATTATAATGCTTATCTA	3A, 3B, 3E, 4H, 6E, 6F, 6H, S11B-C
miR-L_4×20bp	TGGAATTATAATGCTTATCTATACAACAATTTGGAATTATAATGCTTATCTATAATCACTTGTGGAATTATAATGCTTATCTATACAACTGGATGGAATTATAATGCTTATCTA	3A
miR-L_4×21bp	AGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTA	3A, S4
miR-L_8×8bp	TACTTGGCATTATCTTATCTATACAACAATTTACTTGGCATTATCTTATCTATAATCACTTGTACTTGGCATTATCTTATCTATACAACTGGATACTTGGCATTATCTTATCTATAACTAGTTCTTTGATACTTGGCATTATCTTATCTATACAACAATTTACTTGGCATTATCTTATCTATAATCACTTGTACTTGGCATTATCTTATCTATACAACTGGATACTTGGCATTATCTTATCTA	S9
miR-L_10×21bp	AGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTAAGGAATTATAATGCTTATCTA	S2H
synmiR-1_1×21bp	CAACATAAGCATAAACTACGA	4D, 4E
synmiR-2_1×21bp_initial	TAATCTACAAAGTGACATAAA	S6A
synmiR-2_1×21bp_20G	TGATCTACAAAGTGACATAAA	4D, 4E
synmiR-3_1×21bp_initial	GAAAAGTTTATAATATCTTGA	S6A
synmiR-3_1×21bp_19G	GAGAAGTTTATAATATCTTGA	4D, 4E

synmiR-4_1×21bp	CGTAAAGATCATGAAATTGAA	4D, 4E
synmiR-5_1×21bp	AGCATCTATCTAACGGTTTGA	4D, 4E
synmiR-6_1×21bp	CTCTTATCAAGCAGTTTCATA	S6B
synmiR-7_1×21bp	AGAAATGGTGTAATTTAGCAA	4D, 4E
synmiR-8_1×21bp	CGTTGAGATTTAAAGATCGAA	4D, 4E
synmiR-9_1×21bp	CAACACCTTTCTTAAAACTTA	4D, 4E
synmiR-10_1×21bp	GGTAATCTTAAGCATAGATGA	4D, 4E
synmiR-1_4×8bp	ACAATATTCGTATAACTACGAT ACAACAATTACAATATTCGTAT AACTACGATAATCACTTGAAA ATATTCGTATAACTACGATACA ACTGGAACAATATTCGTATAA CTACGA	4F, S8
synmiR-1_4×9bp	ACAATATTCGTAAAACTACGA TACAACAATTACAATATTCGTA AAACTACGATAATCACTTGAA AATATTCGTAAAACTACGATA CAACTGGAACAATATTCGTAA AACTACGA	S8
synmiR-2_4×8bp	CATGAAGTTTTCAGACATAAA TACAACAATTCATGAAGTTTT CAGACATAAATAATCACTTGC ATGAAGTTTTCAGACATAAAT ACAACTGGACATGAAGTTTTC AGACATAAA	S8
synmiR-2_4×9bp	CATGAAGTTTTCTGACATAAA TACAACAATTCATGAAGTTTT CTGACATAAATAATCACTTGC ATGAAGTTTTCTGACATAAAT ACAACTGGACATGAAGTTTTC TGACATAAA	S8
synmiR-2_8×8bp	CATGAAGTTTTCAGACATAAA TACAACAATTCATGAAGTTTT CAGACATAAATAATCACTTGC ATGAAGTTTTCAGACATAAAT ACAACTGGACATGAAGTTTTC AGACATAAATAACTAGTTCTTT GACATGAAGTTTTCAGACATA AATACAACAATTCATGAAGTT TTCAGACATAAATAATCACTT GCATGAAGTTTTCAGACATAA ATACAACTGGACATGAAGTTT TCAGACATAAA	4F, 4H, S8

synmiR-3_4×16bp	TCTTTGTTTATAATATCTTGATACAACAATTTCTTTGTTTATAATATCTTGATAATCACTTGTCTTTGTTTATAATATCTTGATACAACTGGATCTTTGTTTATAATATCTTGA	S8
synmiR-3_8×16bp	TCTTTGTTTATAATATCTTGATACAACAATTTCTTTGTTTATAATATCTTGATAATCACTTGTCTTTGTTTATAATATCTTGATACAACTGGATCTTTGTTTATAATATCTTGATAACTAGTTCTTTGATCTTTGTTTATAATATCTTGATACAACAATTTCTTTGTTTATAATATCTTGATAATCACTTGTCTTTGTTTATAATATCTTGATACAACTGGATCTTTGTTTATAATATCTTGA	4F, S8
synmiR-4_4×8bp	GTACTCTGATCATAAATTGAATACAACAATTGTACTCTGATCATAAATTGAATAATCACTTGGTACTCTGATCATAAATTGAATACAACTGGAGTACTCTGATCATAAATTGAA	S8
synmiR-4_8×8bp	GTACTCTGATCATAAATTGAATACAACAATTGTACTCTGATCATAAATTGAATAATCACTTGGTACTCTGATCATAAATTGAATACAACTGGAGTACTCTGATCATAAATTGAATAACTAGTTCTTTGAGTACTCTGATCATAAATTGAATACAACAATTGTACTCTGATCATAAATTGAATAATCACTTGGTACTCTGATCATAAATTGAATACAACTGGAGTACTCTGATCATAAATTGAA	4F, S8
synmiR-4_8×9bp	GTACTCTGATCAGAAATTGAATACAACAATTGTACTCTGATCAGAAATTGAATAATCACTTGGTACTCTGATCAGAAATTGAATACAACTGGAGTACTCTGATCAGAAATTGAATAACTAGTTCTTTGAGTACTCTGATCAGAAATTGAATACAACAATTGTACTCTGATCAGAAATTGAATAATCACTTGGTACTCTGATCAGAAATTGAATACAACTGGAGTACTCTGATCAGAAATTGAA	S8, S9
synmiR-4_4×17bp	GTACAAGATCATGAAATTGAA	S7

	TACAACAATTGTACAAGATCATGAAATTGAATAATCACTTGGTACAAGATCATGAAATTGAATACAACTGGAGTACAAGATCATGAAATTGAA	
synmiR-5_4×8bp	CCATGAGTAATCTCGGTTTGATACAACAATTCCATGAGTAATCTCGGTTTGATAATCACTTGCCATGAGTAATCTCGGTTTGATACAACTGGACCATGAGTAATCTCGGTTTGA	4H, S8
synmiR-5_4×9bp	CCATGAGTAATCACGGTTTGATACAACAATTCCATGAGTAATCACGGTTTGATAATCACTTGCCATGAGTAATCACGGTTTGATACAACTGGACCATGAGTAATCACGGTTTGA	S8
synmiR-5_4×12bp	CCATGAGTACTAACGGTTTGATACAACAATTCCATGAGTACTAACGGTTTGATAATCACTTGCCATGAGTACTAACGGTTTGATACAACTGGACCATGAGTACTAACGGTTTGA	S8
synmiR-5_8×8bp	CCATGAGTAATCTCGGTTTGATACAACAATTCCATGAGTAATCTCGGTTTGATAATCACTTGCCATGAGTAATCTCGGTTTGATACAACTGGACCATGAGTAATCTCGGTTTGATAACTAGTTCTTTGACCATGAGTAATCTCGGTTTGATACAACAATTCCATGAGTAATCTCGGTTTGATAATCACTTGCCATGAGTAATCTCGGTTTGATACAACTGGACCATGAGTAATCTCGGTTTGA	S8, S9
synmiR-5_4×17bp	CCATTCTATCTAACGGTTTGATACAACAATTCCATTCTATCTAACGGTTTGATAATCACTTGCCATTCTATCTAACGGTTTGATACAACTGGACCATTCTATCTAACGGTTTGA	S7
synmiR-7_4×12bp	CTTTTCTAGGTAATTTAGCAATACAACAATTCTTTTCTAGGTAATTTAGCAATAATCACTTGCTTTTCTAGGTAATTTAGCAATACAACTGGACTTTTCTAGGTAATTTAGCAA	S8

synmiR-7_8×12bp	CTTTTCTAGGTAATTTAGCAAT ACAACAATTCTTTTCTAGGTA ATTTAGCAATAATCACTTGCTT TTCTAGGTAATTTAGCAATAC AACTGGACTTTTCTAGGTAAT TTAGCAATAACTAGTTCTTTG ACTTTTCTAGGTAATTTAGCA ATACAACAATTCTTTTCTAGG TAATTTAGCAATAATCACTTGC TTTTCTAGGTAATTTAGCAATA CAACTGGACTTTTCTAGGTAA TTTAGCAA	4F, S8
synmiR-8_4×8bp	GTAACGTCAAATTAGATCGAA TACAACAATTGTAACGTCAAA TTAGATCGAATAATCACTTGG TAACGTCAAATTAGATCGAAT ACAACTGGAGTAACGTCAAAT TAGATCGAA	S8
synmiR-8_4×9bp	GTAACGTCAAATAAGATCGAA TACAACAATTGTAACGTCAAA TAAGATCGAATAATCACTTGG TAACGTCAAATAAGATCGAAT ACAACTGGAGTAACGTCAAAT AAGATCGAA	S8
synmiR-8_8×8bp	GTAACGTCAAATTAGATCGAA TACAACAATTGTAACGTCAAA TTAGATCGAATAATCACTTGG TAACGTCAAATTAGATCGAAT ACAACTGGAGTAACGTCAAAT TAGATCGAATAACTAGTTCTT TGAGTAACGTCAAATTAGATC GAATACAACAATTGTAACGTC AAATTAGATCGAATAATCACT TGGTAACGTCAAATTAGATCG AATACAACTGGAGTAACGTCA AATTAGATCGAA	4F, S8
synmiR-9_4×16bp	TCCAGCCTTTCTTAAAACTTA TACAACAATTTCCAGCCTTTC TTAAAACTTATAATCACTTGTC CAGCCTTTCTTAAAACTTATA CAACTGGATCCAGCCTTTCTT AAAACTTA	S8
synmiR-9_4×17bp	TCCAACCTTTCTTAAAACTTA TACAACAATTTCCAACCTTTC TTAAAACTTATAATCACTTGTC CAACCTTTCTTAAAACTTATA CAACTGGATCCAACCTTTCTT AAAACTTA	4F, S8

synmiR-9_4×18bp	TCCCACCTTTCTTAAAACTTA TACAACAATTTCCCACCTTTC TTAAAACTTATAATCACTTGTC CCACCTTTCTTAAAACTTATA CAACTGGATCCCACCTTTCTT AAAACTTA	S8
synmiR-10_8×8bp	ATACCAAGACTCAATAGATGA TACAACAATTATACCAAGACT CAATAGATGATAATCACTTGA TACCAAGACTCAATAGATGAT ACAACTGGAATACCAAGACT CAATAGATGATAACTAGTTCT TTGAATACCAAGACTCAATAG ATGATACAACAATTATACCAA GACTCAATAGATGATAATCAC TTGATACCAAGACTCAATAGA TGATACAACTGGAATACCAAG ACTCAATAGATGA	4F, S8
synmiR-10_8×9bp	ATACCAAGACTCCATAGATGA TACAACAATTATACCAAGACT CCATAGATGATAATCACTTGA TACCAAGACTCCATAGATGAT ACAACTGGAATACCAAGACT CCATAGATGATAACTAGTTCT TTGAATACCAAGACTCCATAG ATGATACAACAATTATACCAA GACTCCATAGATGATAATCAC TTGATACCAAGACTCCATAGA TGATACAACTGGAATACCAAG ACTCCATAGATGA	S8
Note: the complementary sequences of the target are shaded with light blue.		

Table S2 Gene Annotation for the significantly differentially expressed genes suggested by bulk RNAseq

Shared differentially expressed genes were color-coded.

miRNA	**Differentially expressed genes**
synmiR-1	HSPA6, PCSK5, HSPA1A, HSPA1B, KRT17, LOLX4, BAG3, SSC4D, FADS2, GNB2, TUBA1A, TNNC1, CLPTM1L, FOXD1, CCN2
synmiR-2	HSPA6, PCSK5, HSPA1A, HSPA1B, RPL17
synmiR-3	HSPA6, PCSK5, HSPA1A, HSPA1B, SPARC
synmiR-4	HSPA6, PCSK5, HSPA1A, HSPA1B, CLPTM1L, POLR2L
synmiR-5	HSPA6, PCSK5, HSPA1A, HSPA1B, CLPTM1L,

	SSC4D, RPL17, TNNC1
synmiR-7	HSPA6, PCSK5, HSPA1A, HSPA1B, BAG3, GNB2, AP2S1, CLPTM1L, RHOC, HDLBP
synmiR-8	HSPA6, PCSK5, HSPA1A, HSPA1B, RPL17, LOXL4, FADS2, KRT17, RPS2, SSC4D, SPARC, DDX5, TNNC1
synmiR-9	HSPA6, LOXL4, FADS2, KRT17, PCSK5, RPL17, HSPA1B, SSC4D, DKK3, LOXL1, TPM1, TUBA1A, BCAM, SPARC, MAGED1, SEMA6B, TNNC1, AP2S1, RPS2, RPS19
synmiR-10	HSPA6, LOXL4, FADS2, KRT17, PCSK5, GNAS, RPL17, SSC4D, DKK3, BCAM, MAGED1, DDX5, SPARC, TUBA1A, YWHAZ, TNNC1, AP2S1, ACTG1, MRFAP1, RPS2, VIM
miR-L	HSPA6, GNAS, HSPA1B, RPL17, VIM

References

1. Franz, K., Singh, A. & Weinberger, L. S. Lentiviral vectors to study stochastic noise in gene expression. *Methods Enzymol.* **497**, 603–622 (2011).
2. Balasubramanian, S. *et al.* Rapid recombinant protein production from piggyBac transposon-mediated stable CHO cell pools. *J. Biotechnol.* **200**, 61–69 (2015).
3. Singer, Z. S. *et al.* Dynamic heterogeneity and DNA methylation in embryonic stem cells. *Mol. Cell* **55**, 319–331 (2014).
4. Maamar, H., Raj, A. & Dubnau, D. Noise in gene expression determines cell fate in Bacillus subtilis. *Science* **317**, 526–529 (2007).
5. Gam, J. J., DiAndreth, B., Jones, R. D., Huh, J. & Weiss, R. A 'poly-transfection' method for rapid, one-pot characterization and optimization of genetic systems. *Nucleic Acids Res.* **47**, e106 (2019).
6. Cabrera, A. *et al.* The sound of silence: Transgene silencing in mammalian cell engineering. *Cell Syst* **13**, 950–973 (2022).
7. Alon, U. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461 (2007).

8. Segall-Shapiro, T. H., Sontag, E. D. & Voigt, C. A. Engineered promoters enable constant gene expression at any copy number in bacteria. *Nat. Biotechnol.* **36**, 352–358 (2018).
9. Jones, R. D. *et al.* An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nat. Commun.* **11**, 5690 (2020).
10. Bleris, L. *et al.* Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Molecular Systems Biology* vol. 7 519 Preprint at https://doi.org/10.1038/msb.2011.49 (2011).
11. Strovas, T. J., Rosenberg, A. B., Kuypers, B. E., Muscat, R. A. & Seelig, G. MicroRNA-based single-gene circuits buffer protein synthesis rates against perturbations. *ACS Synth. Biol.* **3**, 324–331 (2014).
12. Yang, J. *et al.* A synthetic circuit for buffering gene dosage variation between individual mammalian cells. *Nat. Commun.* **12**, 4132 (2021).
13. Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20–51 (2018).
14. Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
15. Qiu, L., Wang, H., Xia, X., Zhou, H. & Xu, Z. A construct with fluorescent indicators for conditional expression of miRNA. *BMC Biotechnol.* **8**, 77 (2008).
16. Briskin, D., Wang, P. Y. & Bartel, D. P. The biochemical basis for the cooperative action of microRNAs. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 17764–17774 (2020).
17. Elkayam, E. *et al.* Multivalent Recruitment of Human Argonaute by GW182. *Mol. Cell* **67**, 646–658.e3 (2017).
18. Sheu-Gruttadauria, J. & MacRae, I. J. Phase Transitions in the Assembly and Function of Human miRISC. *Cell* **173**, 946–957.e16 (2018).
19. Kilikevicius, A., Meister, G. & Corey, D. R. Reexamining assumptions about miRNA-guided gene silencing. *Nucleic Acids Res.* **50**, 617–634 (2022).
20. La Rocca, G. *et al.* Inducible and reversible inhibition of miRNA-mediated gene repression in vivo. *Elife* **10**, (2021).

21. McGeary, S. E., Bisaria, N., Pham, T. M., Wang, P. Y. & Bartel, D. P. MicroRNA 3'-compensatory pairing occurs through two binding modes, with affinity shaped by nucleotide identity and position. *Elife* **11**, (2022).
22. Becker, W. R. *et al.* High-Throughput Analysis Reveals Rules for Target RNA Binding and Cleavage by AGO2. *Mol. Cell* **75**, 741–755.e11 (2019).
23. Michaels, Y. S. *et al.* Addendum: Precise tuning of gene expression levels in mammalian cells. *Nat. Commun.* **10**, 2622 (2019).
24. Ameres, S. L. *et al.* Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **328**, 1534–1539 (2010).
25. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).
26. McGeary, S. E. *et al.* The biochemical basis of microRNA targeting efficacy. *Science* **366**, (2019).
27. Pontén, J. & Saksela, E. Two established in vitro cell lines from human mesenchymal tumours. *Int. J. Cancer* **2**, 434–447 (1967).
28. Puck, T. T., Cieciura, S. J. & Robinson, A. Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects. *J. Exp. Med.* **108**, 945–956 (1958).
29. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**, 59–74 (1977).
30. Olmsted, J. B., Carlson, K., Klebe, R., Ruddle, F. & Rosenbaum, J. Isolation of microtubule protein from cultured mouse neuroblastoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **65**, 129–136 (1970).
31. Marchingo, J. M. & Cantrell, D. A. Protein synthesis, degradation, and energy metabolism in T cell immunity. *Cell. Mol. Immunol.* **19**, 303–315 (2022).
32. Lakatos, E., Salehi-Reyhani, A., Barclay, M., Stumpf, M. P. H. & Klug, D. R. Protein degradation rate is the dominant mechanism accounting for the differences in protein abundance of basal p53 in a human breast and colorectal cancer cell line. *PLoS One*

12, e0177336 (2017).

33. Liu, S. *et al.* Large cells activate global protein degradation to maintain cell size homeostasis. *bioRxiv* 2021.11.09.467936 (2021) doi:10.1101/2021.11.09.467936.
34. Swovick, K. *et al.* Interspecies Differences in Proteome Turnover Kinetics Are Correlated With Life Spans and Energetic Demands. *Mol. Cell. Proteomics* **20**, 100041 (2021).
35. Direct Readout of Neural Stem Cell Transgenesis with an Integration-Coupled Gene Expression Switch. *Neuron* **107**, 617–630.e6 (2020).
36. Andrews, B. *et al.* Imaging cell biology. *Nat. Cell Biol.* **24**, 1180–1185 (2022).
37. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
38. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nat. Protoc.* **12**, 1198–1228 (2017).
39. Reinhardt, S. C. M. *et al.* Ångström-resolution fluorescence microscopy. *Nature* **617**, 711–716 (2023).
40. Ma, H. *et al.* Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3002–3007 (2015).
41. Ma, H. *et al.* Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.* **34**, 528–530 (2016).
42. Wang, H. *et al.* CRISPR-mediated live imaging of genome editing and transcription. *Science* **365**, 1301–1305 (2019).
43. Cursons, J. *et al.* Combinatorial Targeting by MicroRNAs Co-ordinates Post-transcriptional Control of EMT. *Cell Syst* **7**, 77–91.e7 (2018).
44. Li, C.-J. *et al.* MicroRNA governs bistable cell differentiation and lineage segregation via a noncanonical feedback. *Mol. Syst. Biol.* **17**, e9945 (2021).
45. Liu, Z., Johnson, S. T., Zhang, Z. & Corey, D. R. Expression of TNRC6 (GW182) Proteins Is Not Necessary for Gene Silencing by Fully Complementary RNA Duplexes. *Nucleic Acid Ther.* **29**, 323–334 (2019).
46. Raser, J. M. & O'Shea, E. K. Noise in gene expression: origins, consequences, and

control. *Science* **309**, 2010–2013 (2005).

47. Pedraza, J. M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005).
48. Bashor, C. J., Hilton, I. B., Bandukwala, H., Smith, D. M. & Veiseh, O. Engineering the next generation of cell-based therapeutics. *Nat. Rev. Drug Discov.* **21**, 655–675 (2022).
49. Loh, Y.-H. *et al.* Reprogramming of T cells from human peripheral blood. *Cell Stem Cell* **7**, 15–19 (2010).
50. Wang, H., Yang, Y., Liu, J. & Qian, L. Direct cell reprogramming: approaches, mechanisms and progress. *Nat. Rev. Mol. Cell Biol.* **22**, 410–424 (2021).
51. Van Alstyne, M. *et al.* Gain of toxic function by long-term AAV9-mediated SMN overexpression in the sensorimotor circuit. *Nat. Neurosci.* **24**, 930–940 (2021).
52. Hinderer, C. *et al.* Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN. *Hum. Gene Ther.* **29**, 285–298 (2018).
53. Smith, S. E. P. *et al.* Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice. *Sci. Transl. Med.* **3**, 103ra97 (2011).
54. Collins, A. L. *et al.* Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.* **13**, 2679–2689 (2004).
55. Matagne, V. *et al.* Severe offtarget effects following intravenous delivery of AAV9-MECP2 in a female mouse model of Rett syndrome. *Neurobiol. Dis.* **149**, 105235 (2021).
56. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).
57. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–73 (2014).
58. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39**, D152–7 (2011).
59. Griffiths-Jones, S., Saini, H. K., van Dongen, S. & Enright, A. J. miRBase: tools for

microRNA genomics. *Nucleic Acids Res.* **36**, D154–8 (2008).

60. Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A. & Enright, A. J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34**, D140–4 (2006).
61. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Res.* **32**, D109–11 (2004).
62. Ambros, V. *et al.* A uniform system for microRNA annotation. *RNA* **9**, 277–279 (2003).
63. Meyers, B. C. *et al.* Criteria for annotation of plant MicroRNAs. *Plant Cell* **20**, 3186–3190 (2008).
64. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
65. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
66. Munsky, B., Neuert, G. & van Oudenaarden, A. Using gene expression noise to understand gene regulation. *Science* **336**, 183–187 (2012).
67. Raj, A. & van Oudenaarden, A. Single-molecule approaches to stochastic gene expression. *Annu. Rev. Biophys.* **38**, 255–270 (2009).
68. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
69. Strauss, S. & Jungmann, R. Up to 100-fold speed-up and multiplexing in optimized DNA-PAINT. *Nat. Methods* **17**, 789–791 (2020).

Supplementary Modeling

February 2, 2024

1 Modeling miRNA-based gene dosage compensation circuits

Here we develop a simplified mathematical model of miRNA-based incoherent feed-forward loop (IFFL) circuits, and use it to explore how various parameters control the scaling of target gene expression with gene dosage. The model is based on several reactions:

1. RISC production and removal. We assume miRNA is expressed, processed, and assembles with Argonaute (Ago) proteins to form an active (mature) RNA-Induced Silencing Complex (RISC)[1]. We denote the concentration of mature RISC (containing the miRNA) as r. We assume RISC is produced at a total rate of $D\beta_{\mathrm{r}}$, where D denotes gene copy number (gene dosage) and β_{r} denotes the rate of production RISC production per gene copy. This expression implicitly assumes that miRNA expression levels do not saturate available miRNA processing machinery, Ago, or other components. We also assume that the RISC complex is removed at total rate $\gamma_{\mathrm{r}} r$, where γ_{r} denotes a combined rate constant for dilution, degradation, and other removal processes.

2. The mRNA of the target gene, whose concentration we denote m, similarly undergoes production and removal. We assume mRNA is produced at a rate proportional to gene copy number, $D\beta_{\mathrm{m}}$, where β_{m} is the mRNA production rate per gene copy. mRNA can also be removed at a rate $\gamma_{\mathrm{m}} m$ due to dilution and degradation. We note that even though mRNA and miRNA are produced from the same engineered locus, the production rate constants β_{m} and β_{r} can differ, since the miRNA and mRNA are produced from distinct promoters, and because mRNA and RISC-miRNA complex production involve distinct biochemical steps.

3. RISC-mRNA complex formation and dissociation. We assume that the RISC and target mRNA associate to form a complex, whose concentration is denoted C, at a rate $k_{\mathrm{on}} r m$, following mass action kinetics with rate constant, k_{on}. Once formed, this complex can dissociate at rate $k_{\mathrm{off}} C$, undergo catalytic mRNA degradation, at rate $k_{\mathrm{c}} C$.

These chemical reactions can be summarized as follows, using G to denote

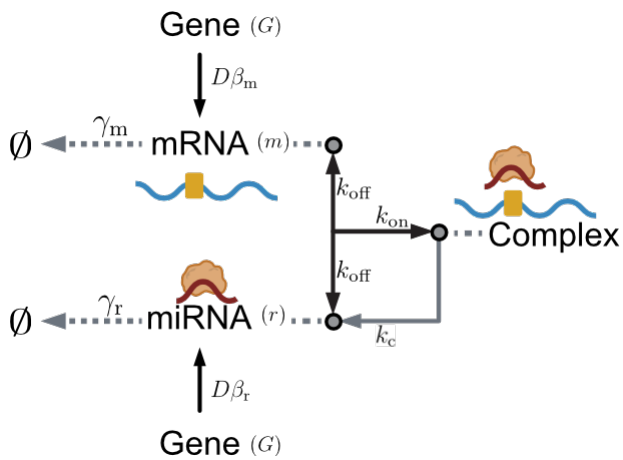

Figure 1: Reactions in model.

the gene, present at copy number (dosage) D:

$$
\begin{aligned}
G &\xrightarrow{D\beta_r} G + r \\
r &\xrightarrow{\gamma_r} \emptyset \\
G &\xrightarrow{D\beta_m} G + m \\
m &\xrightarrow{\gamma_m} \emptyset \\
r + m &\xrightarrow{k_{on}} C \\
C &\xrightarrow{k_{off}} r + m \\
C &\xrightarrow{k_c} r
\end{aligned}
$$

With these definitions and assumptions, we can write down a set of ordinary differential equations for the three variables, r, m, and C.

$$
\begin{aligned}
\frac{dr}{dt} &= \beta_r D - \gamma_r r - k_{on} r m + (k_c + k_{off}) C \\
\frac{dm}{dt} &= \beta_m D - \gamma_m m - k_{on} r m + k_{off} C \\
\frac{dC}{dt} &= k_{on} r m - (k_c + k_{off}) C
\end{aligned}
$$

These equations resemble previous modeling of regulation by small RNAs [2, 3], except for the explicit incorporation of gene dosage in the control of both mRNA and miRNA.

To gain insight into the possible behaviors of this system, we first non-dimensionalize it. We define a dimensionless time, $\widetilde{t} = \gamma_r t$, by rescaling time in

units of the RISC lifetime. We also define a dimensionless RISC concentration, $\widetilde{r} = r/(\beta_{\mathrm{r}}/\gamma_{\mathrm{r}})$. This effectively rescales r in units of the unregulated steady-state expression level produced by a single copy of the gene. We similarly define a dimensionless mRNA concentration, $\widetilde{m} = m/(\beta_{\mathrm{m}}/\gamma_{\mathrm{m}})$, normalizing m by its single copy steady state expression level. Finally, we define a dimensionless concentration of the RISC-miRNA-mRNA complex, $\widetilde{C} = C/(\beta_{\mathrm{m}}\beta_{\mathrm{r}}/\gamma_{\mathrm{m}}\gamma_{\mathrm{r}})$.

In addition, we define a set of convenient dimensionless parameter ratios:

$$
\begin{aligned}
K &= \frac{k_{\mathrm{c}} + k_{\mathrm{off}}}{k_{\mathrm{on}}} \\
\gamma &= \frac{\gamma_{\mathrm{m}}}{\gamma_{\mathrm{r}}} \\
\beta &= \frac{\beta_{\mathrm{m}}}{\beta_{\mathrm{r}}} \\
\widetilde{k}_{\mathrm{on}} &= \frac{k_{\mathrm{on}}\beta_{\mathrm{r}}}{\gamma_{\mathrm{m}}\gamma_{\mathrm{r}}} \\
\widetilde{k}_{\mathrm{off}} &= \frac{k_{\mathrm{off}}\beta_{\mathrm{r}}}{\gamma_{\mathrm{m}}\gamma_{\mathrm{r}}} \\
\widetilde{k}_{\mathrm{c}} &= \frac{k_{\mathrm{c}}\beta_{\mathrm{r}}}{\gamma_{\mathrm{m}}\gamma_{\mathrm{r}}}
\end{aligned}
$$

In the non-dimensionalized system, the differential equations can be written as,

$$
\begin{aligned}
\frac{d\widetilde{r}}{d\widetilde{t}} &= D - \widetilde{r} - \widetilde{k}_{\mathrm{on}}\beta\widetilde{m}\widetilde{r} + (\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}})\beta\widetilde{C} \\
\gamma^{-1}\frac{d\widetilde{m}}{d\widetilde{t}} &= D - \widetilde{m} - \widetilde{k}_{\mathrm{on}}\widetilde{m}\widetilde{r} + \widetilde{k}_{\mathrm{off}}\widetilde{C} \\
\gamma^{-1}\frac{d\widetilde{C}}{d\widetilde{t}} &= \frac{\gamma_r}{\beta_r}\widetilde{k}_{\mathrm{on}}(\widetilde{m}\widetilde{r} - K\widetilde{C})
\end{aligned}
$$

By definition, at steady state, the time derivatives all equal zero. Denoting steady values with a subscript s, we then have:

$$
\begin{aligned}
D - \widetilde{r}_{\mathrm{s}} - \widetilde{k}_{\mathrm{on}}\beta\widetilde{m}_{\mathrm{s}}\widetilde{r}_{\mathrm{s}} + (\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}})\beta\widetilde{C}_{\mathrm{s}} &= 0 \\
D - \widetilde{m}_{\mathrm{s}} - \widetilde{k}_{\mathrm{on}}\widetilde{m}_{\mathrm{s}}\widetilde{r}_{\mathrm{s}} + \widetilde{k}_{\mathrm{off}}\widetilde{C}_{\mathrm{s}} &= 0 \\
\widetilde{m}_{\mathrm{s}}\widetilde{r}_{\mathrm{s}} - K\widetilde{C}_{\mathrm{s}} &= 0
\end{aligned}
$$

Solving the equations above, we obtain an equation for steady-state mRNA concentration:

$$
\widetilde{m}_{\mathrm{s}} = \frac{D}{1 + \frac{\widetilde{k}_{\mathrm{c}}}{K}D}
$$

Henceforth, we will omit the tildes for notational convenience, and switch to the non-dimensionalized variables and parameters.

In the limit of large dosage, $\frac{k_c}{K}D \gg 1$. As a result, m_s approaches a limiting value, $m_s \to \frac{K}{k_c}$ independent of gene dosage. This is the regime that the constructs developed here are targeting.

Figure 2A plots this expression, and Figure S1A and B plot the expression in different paramters, i.e. k_{off}, and k_c.

Finally, to estimate biological parameter values, we assumed that the miRNA and mRNA were produced at similar rates, and selected values based on [4], as listed in the following table:

Non-dimensional parameter	**Dimensionless values**
β	1
k_{on}	200000
k_{off}	4000, 40000, 400000, 4000000
k_c	160, 1.6, 0.16

1.1 Incorporating ultrasensitivity

Alternatively, we also considered a more general, phenomenological model in which the inhibition of miRNA to the target was assumed to follow a Hill function, with the Hill coefficient n. This treatment omits intermediate steps, i.e., the RISC-mRNA complex formation and dissociation. With this assumption, we can write down a different set of ordinary differential equations for the two variables, r, and m.

$$\frac{dr}{dt} = \beta_r D - \gamma_r r$$
$$\frac{dm}{dt} = \frac{\beta_m D}{1 + (\frac{r}{\kappa})^n} - \gamma_m m$$

We non-dimensionalize this system by defining $\widetilde{t} = \gamma_r t$, $\widetilde{r} = r/(\beta_r/\gamma_r)$, and $\widetilde{m} = m/(\beta_m/\gamma_m)$. Additionally, we define $\widetilde{K} = \frac{\beta_r}{\gamma_r \kappa}$ for convenience. The steady state expression of mRNA concentration can then be written as:

$$\widetilde{m_s} = \frac{D}{1 + (\widetilde{K}D)^n}$$

Figure S1C plots the expression when $\widetilde{K}$ is set as 1, and $n = 0.5, 1, 2$. Critically, only when $n = 1$, does $\widetilde{m}_s$ approaches a limiting value. If $n < 1$, $\widetilde{m}_s$ shows a sublinear increase with D. If $n > 1$, $\widetilde{m}_s$shows a bandpass filter dependence with D.

References

1. Iwakawa, H.-o. & Tomari, Y. Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. *Molecular cell* **82,** 30–43 (2022).
2. Mehta, P., Goyal, S. & Wingreen, N. S. A quantitative comparison of sRNA-based and protein-based gene regulation. *Molecular systems biology* **4,** 221 (2008).
3. Mukherji, S. *et al.* MicroRNAs can generate thresholds in target gene expression. *Nature genetics* **43,** 854–859 (2011).
4. Nyayanit, D. & Gadgil, C. J. Mathematical modeling of combinatorial regulation suggests that apparent positive regulation of targets by miRNA could be an artifact resulting from competition for mRNA. *Rna* **21,** 307–319 (2015).