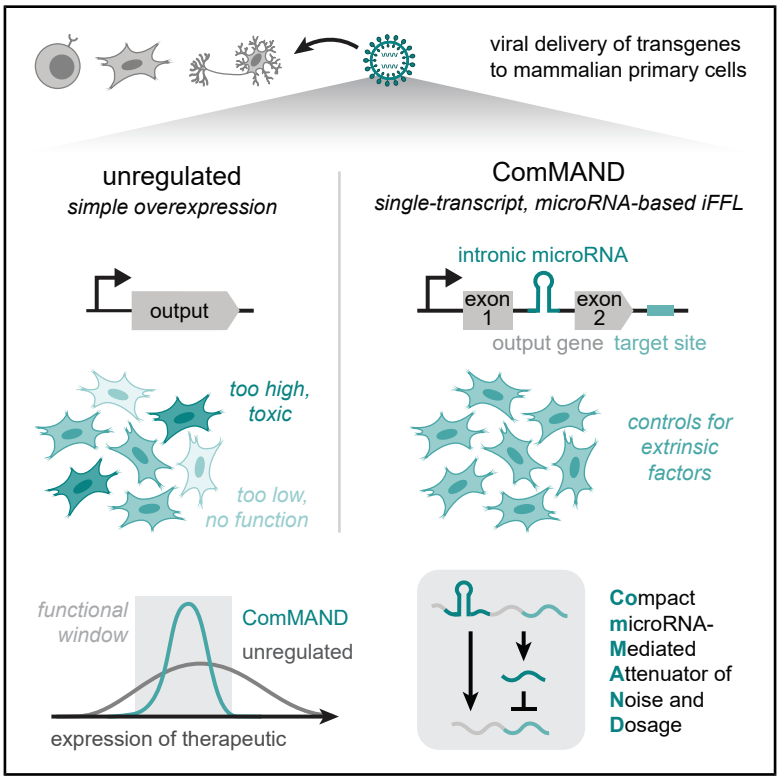
Article

Cell Systems

Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells

Graphical abstract



Authors

Kasey S. Love,
Christopher P. Johnstone,
Emma L. Peterman,
Stephanie Gaglione,
Michael E. Birnbaum, Kate E. Galloway

Correspondence

katiegal@mit.edu

In brief

Using a model to explore the design space, Love et al. developed ComMAND, a single-transcript, microRNA-mediated incoherent feedforward loop that mitigates variance due to gene dosage and tightly regulates transgene expression in primary cells, controlling cargoes relevant to the treatment of neurological disease.

Highlights

- ComMAND constrains the expression of output proteins across delivery methods
- ComMAND-regulated protein levels scale sublinearly with DNA dosage
- Modeling identifies physiological limits and tuning strategies for optimizing ComMAND
- ComMAND functions in primary cells and regulates clinically relevant genes

Love et al., 2025, Cell Systems *16*, 101269
June 18, 2025
https://doi.org/10.1016/j.cels.2025.101269

CellPress

Cell Systems

CellPress

Article

Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells

Kasey S. Love,[1] Christopher P. Johnstone,[2] Emma L. Peterman,[2] Stephanie Gaglione,[2] Michael E. Birnbaum,[1,3] and Kate E. Galloway[2,3,4,*]

[1]Department of Biological Engineering, MIT, 25 Ames St., Cambridge, MA 02139, USA
[2]Department of Chemical Engineering, MIT, 25 Ames St., Cambridge, MA 02139, USA
[3]Koch Institute for Integrative Cancer Research, Cambridge, MA, USA
[4]Lead contact
*Correspondence: katiegal@mit.edu
https://doi.org/10.1016/j.cels.2025.101269

SUMMARY

In a therapeutic context, supraphysiological expression of transgenes can compromise engineered phenotypes and lead to toxicity. To ensure a narrow range of transgene expression, we developed a single-transcript, microRNA-based incoherent feedforward loop called compact microRNA-mediated attenuator of noise and dosage (ComMAND). We experimentally tuned the ComMAND output profile, and we modeled the system to explore additional tuning strategies. By comparing ComMAND to two-gene implementations, we demonstrate the precise control afforded by the single-transcript architecture, particularly at low copy numbers. We show that ComMAND tightly regulates transgene expression from lentiviruses and precisely controls expression in primary human T cells, primary rat neurons, primary mouse embryonic fibroblasts, and human induced pluripotent stem cells. Finally, ComMAND effectively sets levels of the clinically relevant transgenes *frataxin* (*FXN*) and *fragile X messenger ribonucleoprotein 1* (*Fmr1*) within a narrow window. Overall, ComMAND is a compact tool well suited to precisely specify the expression of therapeutic cargoes. A record of this paper's transparent peer review process is included in the supplemental information.

INTRODUCTION

With massive advances in the development of delivery vectors, the tunable control of therapeutic cargoes remains the missing element in providing safe, effective *in vivo* and *ex vivo* gene therapies.[1–9] A large number of diseases result from mutations affecting a single gene.[10] Many of these monogenic disorders render the cell vulnerable by the insufficient production of an essential protein product.[11–15] For these haploinsufficiencies, delivery of functional copies of affected genes may restore essential cellular processes and rescue normal phenotypes. However, while improved vectors can efficiently deliver synthetic cargoes, excessive expression of transgenes can induce unforeseen neurological and cardiac disorders. For instance, in mouse models of neurological disorders that result from haploinsufficiency, gene replacement therapy generated mixed outcomes. When target proteins were overexpressed at 10–20 times their endogenous level, some mice showed improved function, but others suffered adverse events including cardiotoxicity and behavioral abnormalities.[16–18] For safety and efficacy, delivery of therapeutic cargoes must be tailored to ensure that transgenes are expressed within a "goldilocks" window (Figure 1A). Thus, there is a need for tools that can precisely set levels of transgene expression.

Gene dosage, including gene copy number, represents a major source of variance in gene expression.[19] While efficient at delivering genetic cargoes, random integration methods offer limited control of copy number. Adjusting vector dosage simultaneously—and unavoidably—affects both delivery efficiency and copy number, making it difficult to maintain high coverage of cells while also ensuring that most cells receive only one or a few copies of the transgene. Uneven biodistribution may further compound the challenge of tailoring expression by adjusting dosage.[20] Furthermore, for integrating vectors, differences in the site of integration introduce distinct genomic contexts for transgenes, which can generate cell-to-cell variability in transcription and thus expression.[21] Combined with wide variation in promoter activity across cell types,[22,23] local effects on transcriptional activity present challenges to robust, precise transgene expression.

Gene circuits offer a solution for implementing tight control of transgene expression. Gene circuits such as the incoherent feedforward loop (iFFL) can reduce expression variation caused by DNA copy number.[24–30] To do so, the iFFL senses and compensates for changes in the input by both positively (directly) and negatively (indirectly via an additional species) regulating the output. Mathematical analysis of the iFFL demonstrates that this topology—in the absence of resource limitations—enables

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

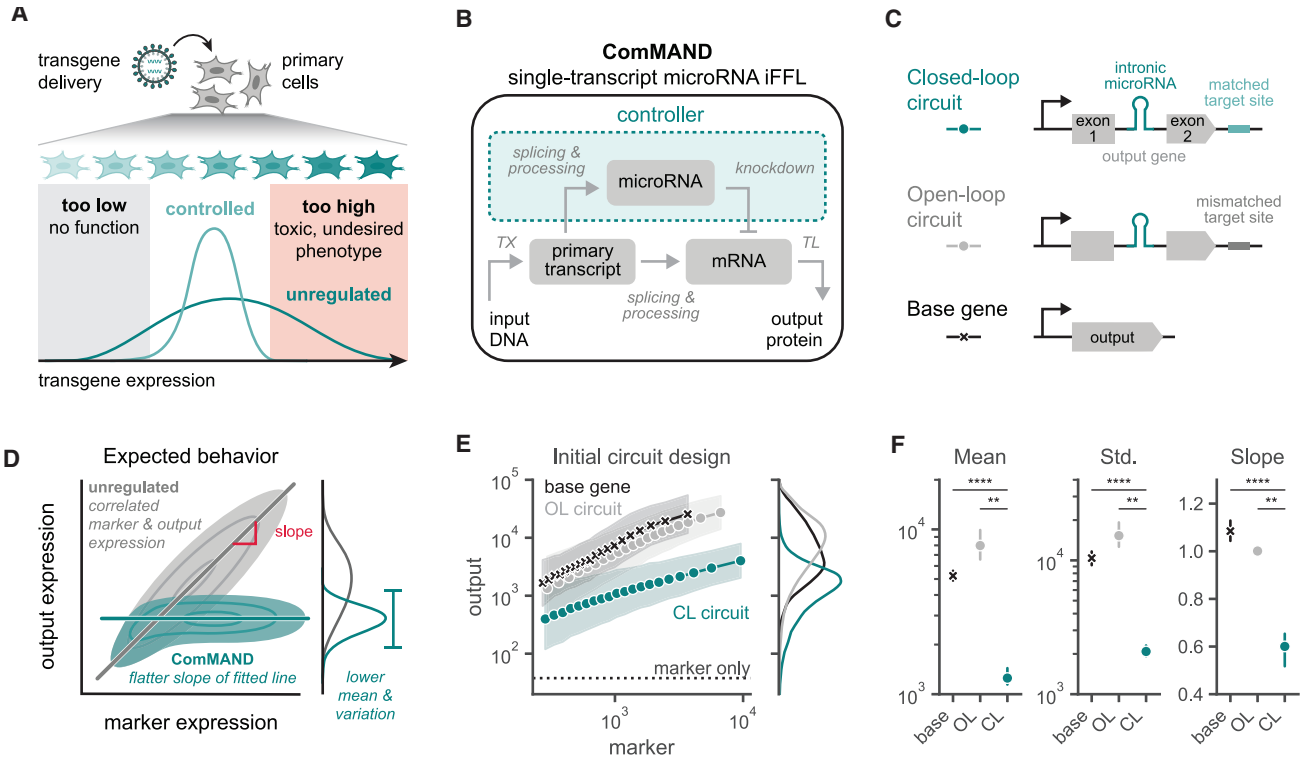


Figure 1. ComMAND, a single-transcript, microRNA-based iFFL, reduces output mean and variability compared with unregulated genes

(A) Delivery of transgenes to primary cells results in high expression variance. Unregulated transgene expression levels may be too low, where the transgene does not function, or too high, which can lead to toxicity or other undesired phenotypes. Controlled transgene expression has the potential to narrow this expression distribution to within the functional range.

(B) Block diagram of ComMAND, a single-transcript, microRNA-based iFFL. DNA, the input, is transcribed (TX) into a primary transcript, which is then spliced and processed into mature mRNA and microRNA. The microRNA can knock down the mRNA; microRNA production and function constitute the controller (teal). Alternatively, the mRNA can be translated (TL) into protein, the output of the circuit.

(C) DNA construct diagrams of ComMAND. The closed-loop (CL) circuit consists of an intronic microRNA between two exons of the output gene and a 22-bp complementary microRNA target site in the 3′ UTR. In the open-loop (OL) circuit, the microRNA target site is orthogonal to the microRNA sequence. The base gene includes only the output coding sequence, lacking both the microRNA and the target site.

(D) Expected behavior of ComMAND across a population of cells when co-delivered with a constitutively expressed marker gene. For an unregulated gene (base gene or OL circuit, gray), the marker and output expression should be well-correlated, giving rise to a wide output expression distribution with a positive slope. In contrast, the CL circuit (teal) is expected to display relatively constant output expression levels across a range of marker expression levels, leading to a lower output expression mean, a narrower output distribution, and a flatter slope.

(E) Left: output expression as a function of marker expression for constructs in (C) co-transfected with a marker gene in HEK293T cells. Flow cytometry measurements for one representative biological replicate are binned by marker expression into 20 equal-quantile groups per condition. Points represent geometric means of output expression for cells in each bin, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Bins are plotted at their median marker value. Dashed line represents the output geometric mean of cells transfected only with the marker gene. Right: histograms of output expression for cells in each condition. The output gene, mRuby2, was expressed via the EF1α promoter, and circuits include miR-FF5.

(F) Summary statistics of output expression for populations in (E). The plotted mean values use the geometric mean. SD refers to the standard deviation, and the slope represents the slope of the line fitted to the binned, log-transformed marker-output points in (E). Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$, independent t test. All units are arbitrary units from a flow cytometer.

perfect adaptation, in which the system response returns to its original level after a disturbance.[31–37] Previous work has implemented iFFLs in mammalian cells via transcriptional[25,38] and posttranscriptional[1,25–28,30,33] control and demonstrated experimentally that this circuit architecture enables robustness to resource competition and sources of extrinsic noise such as plasmid copy number. A similar circuit has also been used to mitigate toxic transgene expression during adeno-associated virus production.[39] Constructing an iFFL using microRNA elements is particularly advantageous because these components are small and may require fewer cellular resources to function. However, until recently,[1,30] it remained unclear if a microRNA-based implementation of this circuit could reduce variability introduced by random integration methods or could function effectively in primary cells. Furthermore, there is a need to articulate design rules to build circuits that function across diverse delivery methods and to identify optimal performance regimes. A deeper understanding of the principles governing circuit function in application-relevant cell types would lay the foundation for the development of more effective therapeutics.

To this end, we constructed a genetic control system called compact microRNA-mediated attenuator of noise and dosage

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

(ComMAND). ComMAND comprises a microRNA-based iFFL encoded on a single transcript (Figure 1B). Instead of relying on protein components, ComMAND requires only the addition of an intronic microRNA and a corresponding target site, allowing these circuits to remain extremely compact with minimal resource burden and low immunogenicity. The single-transcript architecture requires fewer genetic elements than multi-transcript or multi-vector designs, making it advantageous for use in therapeutic applications where delivery poses a key challenge. Furthermore, encoding the microRNA in an intron of the output gene closely links microRNA and output mRNA production, since transcription and proper splicing are required to produce molecules of each of these species. The single-transcript design thus more tightly couples the direct and indirect effects of the iFFL on the output, affording more precise control. Additionally, this intronic design provides an inherently fail-safe mechanism of control. Unlike non-intronic designs, if splicing of the microRNA fails, the regulated transgene cannot be correctly translated from that transcript. Finally, because ComMAND functions downstream of transcription, it is compatible with a wide range of expression methods, including cell-type-specific or small-molecule-inducible promoters, enabling additional layers of regulation. Here, we characterize ComMAND and demonstrate its generalizability across cell types, delivery methods, and regulated genes, paving the way for therapeutic applications.

RESULTS

A single-transcript, microRNA-based iFFL reduces the expression mean and variability of an output protein

To take advantage of the compact size of microRNA components, we constructed ComMAND, a single-transcript version of an iFFL (Figure 1B). To prevent crosstalk with endogenous transcripts, we selected a panel of previously developed synthetic microRNAs and cognate target sequences derived from firefly luciferase.[40] These microRNAs are expressed within an intron containing the human miR-30a scaffold and have targeting sequences orthogonal to the human genome. We first verified that this panel of microRNAs efficiently knocks down their targets in transfections of HEK293T cells (Figure S1A). Of the four reported sequences, we found that three sequences (FF4, FF5, and FF6) are orthogonal and effective at target knockdown (Figures S1B and S1C). Additionally, the presence of the microRNA or target site alone does not alter gene expression (Figures S1D and S1E). To construct the iFFL, we inserted an intron bearing the FF5 microRNA (miR-FF5) at a 5′-AGGT-3′ sequence within the output gene, mRuby2, to generate favorable splice donor and acceptor sites. The intronic design of the iFFL adds only ~450 bp to the total transcript size, ensuring ComMAND remains sufficiently compact for therapeutic applications with strict cargo limits. For the closed-loop (CL) circuit, we included a perfectly complementary 22-bp FF5 target sequence in the 3′ untranslated region (UTR) of the transcript (Figure 1C). For comparison, we constructed an open-loop (OL) version of the circuit that replaces the complementary target site with one of the other orthogonal sequences from our panel. We also included a "base gene," a second unregulated control that lacks both the microRNA-containing intron and the microRNA target site (i.e., the output gene alone). Together, these constructs enable us to parse the effects of the circuit on output expression.

To examine the ability of ComMAND to control variation introduced by differences in copy number and other extrinsic factors, we co-delivered the circuit with a separate fluorescent marker gene. Expression of the fluorescent marker varies across a population of cells as a function of DNA dosage and cellular physiology. For unregulated genes, the output and marker expression should be strongly correlated (Figure 1D, gray). We quantify correlation by the slope of the marker-output line in logarithmic space, which indicates how much the output expression varies as marker expression changes. A slope below one indicates sublinear scaling. For unregulated genes, we expect marker and output expression to co-vary, resulting in a slope around one. In contrast, we expect that the CL circuit will have relatively lower output expression variation across a range of marker levels (Figure 1D, teal). As ComMAND approaches optimal performance, the slope will approach zero. The CL circuit should also have a lower mean output level than an unregulated gene due to knockdown of the output mRNA. Indeed, when we co-transfected the base gene, OL circuit, or CL circuit with a marker gene in HEK293T cells, we observed a reduction in output expression level, output standard deviation, and slope of the marker-output line for the CL circuit, compared with the unregulated genes (Figures 1E and 1F). Although previous work[29,33] demonstrates microRNA-based iFFLs can reduce the coefficient of variation, a relative metric of variability in protein expression, we did not observe differences in this metric for the CL circuit (Figure S1F). Theory predicts that microRNA-based iFFL circuits can increase or decrease noise at the protein level.[41] Thus, we focused on characterizing ComMAND by absolute measures of variability that are relevant for therapeutic applications, such as standard deviation, which does decrease for the CL circuit. Together, ComMAND significantly reduces the absolute variability of the output gene and achieves sublinear scaling but does not provide perfect adaptation to extrinsic noise. To better understand the system and improve performance, we next examined how to tune circuit output.

Selection of circuit components enables tuning of output expression

In order to define design principles for implementing ComMAND, we characterized output expression profiles of constructs with different sets of genetic parts. We sought to understand how part identity contributes to circuit performance and tunes output expression levels. First, we investigated how properties of the microRNA affect output expression. Increasing the strength of the repression arm of the iFFL should reduce output variability.[28] As FF4 showed the greatest target knockdown (Figure S1C), we hypothesized that replacing the FF5 sequences with FF4 sequences in ComMAND would improve regulation of the output gene. Indeed, introduction of FF4 reduces output mean and standard deviation relative to FF5 for the CL circuit in transfection of HEK293T cells (Figures 2A, 2C, and S1G). To further increase knockdown, we next modified the microRNA scaffold. Previously, rational sequence modifications to enhance pri-microRNA processing led to the development of the Mv3 (miRE) scaffold, which improves AND gate function in transfection of HEK293T cells.[42] Therefore, we replaced the miR-30a-based

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

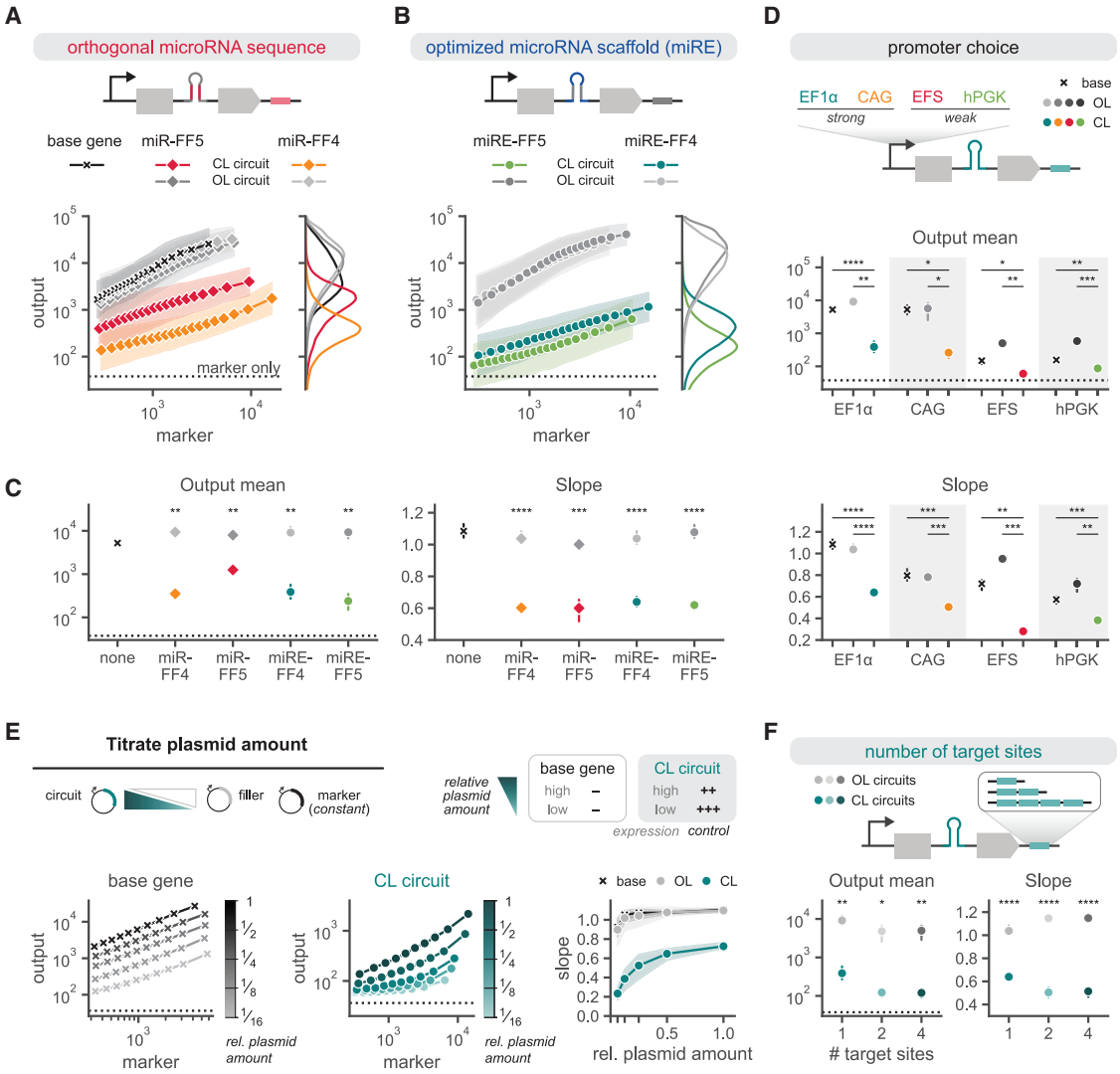


Figure 2. Selection of circuit components enables tuning of output expression

Three constructs with various circuit components are quantified. Base genes (x markers) include only the output gene, without intronic microRNA and target sites. OL (gray) and CL (colored) circuits combine the indicated intronic microRNA with one copy of an orthogonal (OL) or matched (CL) target site.

(A and B) Left: output expression as a function of marker expression for constructs co-transfected with a marker gene in HEK293T cells. Flow cytometry measurements for one representative biological replicate are binned by marker expression into 20 equal-quantile groups per condition. Points represent geometric means of cells in each bin, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Right: histograms of output expression for cells in each condition. Circuits are expressed from an EF1α promoter.

(C) Summary statistics of output expression for populations in (A) and (B). Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval.

(D) Summary statistics of output expression for HEK293T cells co-transfected with a marker gene and constructs expressed from the indicated promoters. OL and CL circuits contain miRE-FF4. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval.

(E) As plasmid amount decreases, expression level decreases for both the base gene and CL circuit. The base gene offers no control over output expression, while control afforded by the CL circuit increases as plasmid amount decreases. Left, middle: output expression as a function of marker expression for varying doses of base gene (left) or CL circuit (middle) co-transfected with a constant amount of marker gene in HEK293T cells. Cells were also transfected with a "filler" plasmid expressing a different fluorescent protein cargo to maintain the same total DNA amount in each condition. Data depict one representative biological replicate and are analyzed as in (A), without error bars for clarity. Plasmid amount is represented as a fraction relative to the standard dosage. Color bars depict relative plasmid amount on a $\log_2$ scale. Right: slope of the marker-output curves as a function of relative plasmid amount. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Circuits contain miRE-FF4 and are expressed from an EF1α promoter.

(F) Summary statistics of output expression for HEK293T cells transfected with EF1α-driven miRE-FF4 circuits containing 1, 2, or 4 copies of matched (CL) or orthogonal (OL) target sites. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval.

Dashed lines represent geometric mean output level for cells transfected only with the marker gene. All unnormalized values use arbitrary units from a flow cytometer. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$, independent t test. Unless otherwise indicated, tests compare CL and OL conditions, and CL and base conditions are also significantly different.

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

scaffold with the miRE scaffold in ComMAND. The miRE scaffold improves the performance of the FF5 CL circuit, approaching the performance of the FF4 CL circuit (Figures 2B, 2C, and S1G). However, both miRE-FF4 and miRE-FF5 circuits perform similarly to the miR-FF4 design. Because miRE-FF4 putatively has the strongest target knockdown, we moved forward with miRE-FF4 in ComMAND.

Since altering the microRNA sequence has only a small effect on output expression profiles, we considered other methods to tune the circuit. As ComMAND acts at the posttranscriptional level (Figure 1B), it is compatible with diverse promoters. Tuning promoter strength can alter the setpoint of microRNA-based iFFLs,[30] but it remains unclear how this property affects single-transcript architectures. To explore how promoter selection affects ComMAND output expression, we transfected circuits driven by a panel of strong and weak promoters (Figure 2D). Output levels are higher with the strong promoters, compared with the weaker promoters as expected. In all cases, the CL circuit reduces expression level and variability, compared with the base gene and OL circuit with the same promoter (Figures 2D and S1H). For the CL circuits, weaker promoters produce smaller slopes, compared with stronger promoters. Because transfected cells receive high plasmid copy numbers, ComMAND may operate in a resource-limited regime in transfection. In such a regime, we expect lower expression levels to generate less competition for resources, affording better control. Potentially, ComMAND performance may improve at lower DNA copy numbers than those achieved in transfection.

To test this hypothesis, we varied the plasmid copy number of the circuit in transfection. We reasoned that at lower DNA doses, the CL circuit would exist in a regime unconstrained by cellular resources, leading to greater control of output expression. In contrast, we expect unregulated genes to lack control at all plasmid amounts. Indeed, when we transfected the circuits driven by a strong human elongation factor 1 alpha (EF1α) promoter with varying DNA dosage in HEK293T cells, we found that the slopes of the base gene and OL circuit remain constant across DNA amount (Figure 2E). However, for the CL circuit, the slope decreases as plasmid amount decreases. Additionally, at lower plasmid doses, the coefficient of variation decreases for ComMAND but increases for the base gene and OL circuit (Figure S1I). These data support the hypothesis that ComMAND affords greater control of output expression at lower copy numbers, putatively due to reductions in resource limitations.

Next, we explored whether changing the number of target sites could tune output expression. We first verified that changes in target site number could alter microRNA-mediated knockdown outside of an iFFL context, as shown previously.[43] When we co-transfected an intronic microRNA and a separate target gene containing varying numbers of target sites, we found that target gene expression decreases as the number of target sites increases (Figures S1J and S1K). Therefore, we expected that ComMAND would more tightly control output expression with additional target sites. However, with a second target site, the output mean and slope only slightly decrease (Figures 2F and S1L). Increasing to four target sites does not change output expression further for the CL circuit. Our observation of limited tuning via the number of target sites matches the behavior of recently reported circuits containing microRNAs paired with fully complementary target sites.[30]

By varying the genetic elements in ComMAND, we were able to determine strategies to shift the output setpoint. In particular, we find that microRNA sequence, microRNA scaffold, and promoter strength can set the output expression level. Across these varied tuning strategies, we then looked to identify the underlying biochemical constraints at work in ComMAND. We turned to a mathematical model of ComMAND function to explain our experimental data, to unify our understanding of the constraints, and to illuminate further opportunities to modulate controller function.

A model of ComMAND activity identifies physiological limits and tuning strategies

To achieve a better understanding of circuit properties and to more rapidly explore the wide design space, we modeled the reactions involved in ComMAND. From previously developed microRNA-based models,[29] we chose a detailed set of reactions in order to account for the potentially diverse couplings of both our genetically encoded components and available cellular resources (Figure 3A; Table S1). To better understand the trends we observed experimentally, we derived a steady-state analytical solution of protein output (STAR Methods). Using this solution, we computed protein levels as a function of DNA copy number.

First, we sought to understand why output expression increases as marker expression increases, instead of remaining constant as predicted for an ideal iFFL (Figures 1D and 1E). The model predicts that protein levels remain at the circuit setpoint at lower DNA copy numbers. Above a particular copy-number threshold where RNA-induced silencing complex (RISC) becomes fully saturated with bound microRNA, protein levels exceed this setpoint (Figure 3B, left).[29] Above this threshold, the circuit exists in a RISC-limited regime where microRNA-mediated knockdown cannot fully control output expression. In this regime, protein levels increase linearly with copy number, and the total amount of RISC affects the rate of this increase. By changing the amount of RISC in the model, we could shift the threshold at which RISC becomes limiting as well as the slope in the RISC-limited regime (Figures 3B, right and S2A, top). Because cells take up thousands of plasmids in transfection,[44] our transfections may generate a RISC-limited state. Alternatively, saturation of splicing machinery could reduce ComMAND output. While it is difficult to disentangle the various effects of splicing on circuit output, in the case of splicing saturation, we would expect lower expression for the OL circuit, compared with the base gene at high marker levels. We do not observe this pattern, as the slopes of the base gene and OL circuits are similar (Figures 1E and 1F). Agreeing with our experimental results (Figure 2E), these resource limitations suggest that ComMAND may better regulate output expression at DNA copy numbers lower than those we obtained in transfection.

Transcriptional activity influences the rate of RISC saturation and thus the performance of ComMAND. RISC saturation is governed by the amounts of both RISC and microRNA, so we expect changes in production of primary transcript—and thus changes in microRNA levels—to shift properties of saturation. Our model predicts that promoter activity shapes the RISC-limited regime

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

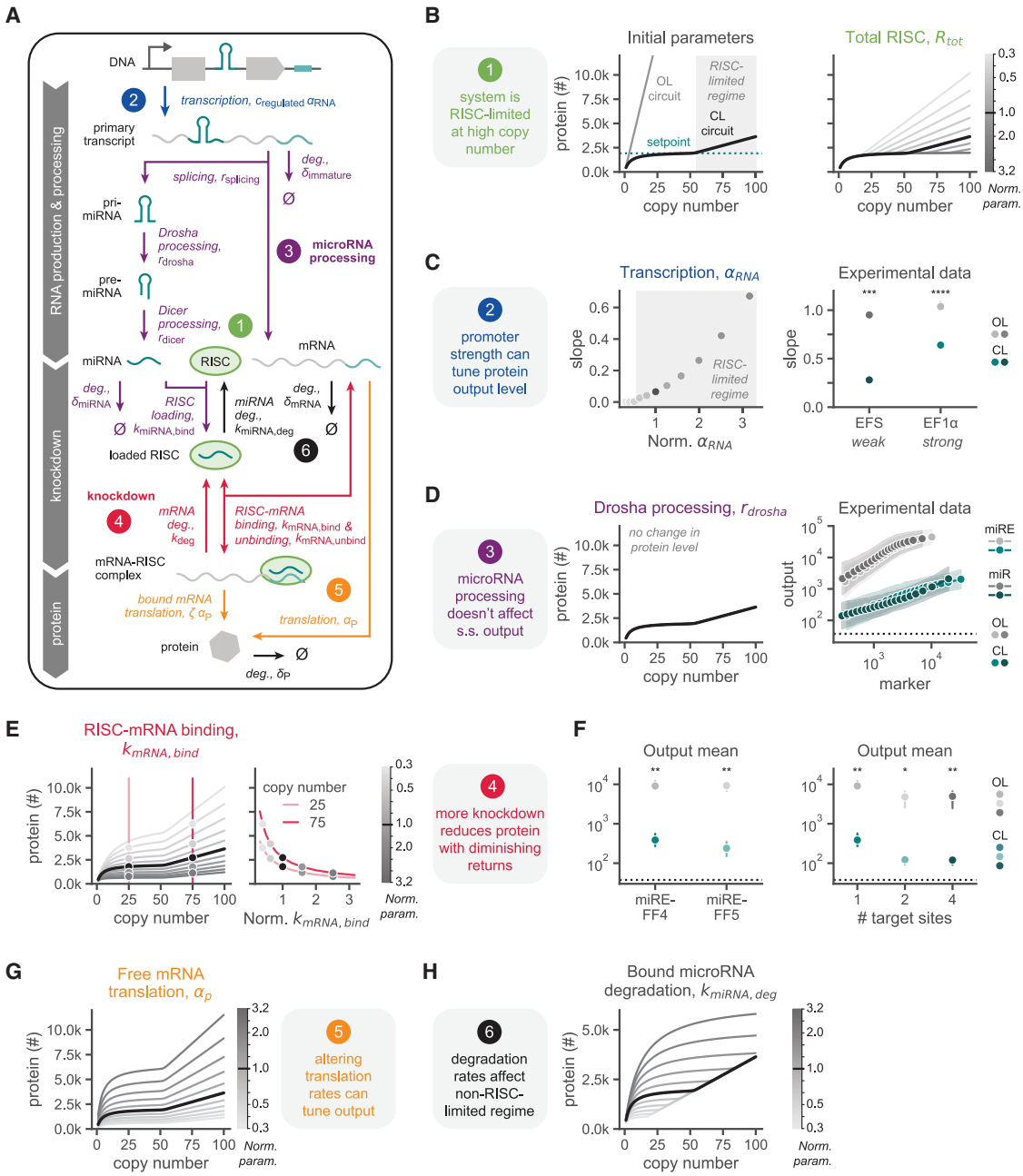


Figure 3. A model of ComMAND activity identifies physiological limits and tuning strategies

(A) Schematic of a reaction network model of ComMAND function. This diagram includes all species, reactions, and parameters that comprise RNA production and processing, microRNA-mediated knockdown, and protein production steps. Exact reactions and steady-state analysis can be found in STAR Methods. Reactions and parameters are color-coded by design principles highlighted in the remaining panels.

(B) Left: output protein (in molecules) plotted as a function of DNA copy number, $c_{regulated}$, for the steady-state analytical solution of the model of ComMAND (CL circuit, black) with the base parameters in Table S3. Gray line indicates the solution for an unregulated gene (OL circuit). Gray shading highlights the DNA copy numbers over which the system is RISC limited. Dashed teal line represents the system setpoint or the protein level in the non-resource-limited regime that remains constant even as copy-number changes. Right: output protein (in number of molecules) as a function of DNA copy number, $c_{regulated}$, for the model with a sweep of values of R_{tot}, the total amount of RISC. Chosen parameter values are evenly log distributed over an order of magnitude centered on the original parameter value. Color bar depicts parameter values normalized to the original value on a $\log_{10}$ scale. The thick black line in this and subsequent panels represents the output of the model with the original parameter values.

(C) Left: slope of the copy number-protein curve at $c_{regulated} = 100$ as a function of transcription rate, α_{RNA}. See STAR Methods for details on the slope calculation. Values of α_{RNA} are normalized to the original value (black dot). Shaded region indicates the values of α_{RNA} at which the system is RISC limited. Right: slope of the marker-output expression distribution for HEK293T cells co-transfected with a marker gene and OL (gray) or CL (teal) circuits expressed from EF1α or EFS

(legend continued on next page)

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

(Figures 3C, left and S2B). Namely, lower transcriptional activity increases the DNA copy-number threshold at which RISC becomes limiting and decreases the slope beyond this threshold. This coupling between promoter strength and RISC limitations may explain why we observe a lower slope for the CL circuit when ComMAND is expressed from weak promoters (Figures 2D and 3C, right). The fact that we observe reduced output variability across promoters in transfection highlights that ComMAND offers control even under physiological resource-limited conditions.

Next, we explored how splicing and processing of the single primary transcript impact circuit function. Varying these parameters within one order of magnitude has little or no effect on steady-state protein output (Figures 3D, left and S2C). Putatively, the generation of mature microRNA from the primary transcript is not limiting at steady state for ComMAND. This may explain why the optimized miRE scaffold has no effect on output expression profiles for the FF4 circuit (Figures 2A, 2C, and 3D, right).

While tuning the circuit in transfection, we hypothesized that directly increasing microRNA-mediated knockdown would enhance ComMAND performance. Indeed, in our model, increasing the binding affinity of the loaded RISC for the mature mRNA decreases the output protein setpoint and the slope in the RISC-limited region (Figure 3E, left). However, this effect has diminishing returns on an absolute scale; if binding affinity is already high, further increases lead to only small decreases in protein level and slope (Figure 3E, right). This may explain why, experimentally, we find that changing the microRNA targeting sequence and the number of target sites—both potentially related to RISC-mRNA binding—has a limited effect on ComMAND performance (Figures 2B, 2C, 2F, and 3F). Altering knockdown activity by increasing the rate of degradation of the RISC-bound mRNA demonstrates similar diminishing returns for decreasing output protein levels and slope (Figure S2D).

We next investigated how translation affects circuit output. As expected, changing the translation rate of the free, mature mRNA shifts the output setpoint without changing the shape of the curve (Figure 3G). We expect that the RISC-bound mRNA can be translated, albeit at a lower rate than free mRNA. Accordingly, we assumed that the rate of bound mRNA translation is smaller than that of free mRNA. As the ratio of translation rates between the bound and free mRNA species approaches one, the slope of the curve increases somewhat, even in the region not limited by RISC (Figure S2E). However, we observe that the "leakiness" of expression from bound mRNAs does not qualitatively change the behavior of ComMAND.

Finally, we explored how the rate of RISC-mediated degradation of the two principal species, RISC-bound microRNA and the mature mRNA, affects ComMAND function. Although we observe little impact of free microRNA degradation rate on output levels (Figure S2C), it is also possible for degradation to occur when the microRNA is loaded in RISC.[45] As the rate of degradation of bound microRNA increases, the output setpoint increases, and the RISC saturation point shifts to higher DNA copy numbers while maintaining the same slope above the inflection point (Figure 3H). Greater degradation of RISC-bound microRNA increases the turnover of RISC, freeing up this cellular resource. Lastly, the degradation rate of the free mRNA has only a small effect on circuit behavior (Figure S2F).

Altogether, the model helps explain our experimental results and suggests additional avenues for tuning the behavior of ComMAND. Moreover, it offers insights into circuit behavior in unconstrained and limiting resource regimes, which are largely dictated by DNA copy number, relative transcription rate, and total amount of RISC.

The single-transcript circuit matches or exceeds performance of two-gene implementations

The single-transcript architecture of ComMAND offers the advantage of compactness; however, alternative implementations may

promoters as in Figure 2D. Slope represents the slope of the line fitted to the binned, log-transformed marker-output points. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Units are arbitrary units from a flow cytometer.

(D) Left: output protein (in number of molecules) as a function of DNA copy number for the model with a sweep of values of r_{drosha}, the rate of Drosha processing, over one order of magnitude. All curves lie under the black line, which represents the output of the model with the original parameter values as in (B). Right: output expression as a function of marker expression for HEK293T cells co-transfected with a marker gene and OL (gray) or CL (teal) circuits with miR-FF4 or miRE-FF4 as in Figure 2B. Flow cytometry measurements for one representative biological replicate are binned by marker expression into 20 equal-quantile groups per condition. Points represent geometric means of cells in each bin, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Dashed line represents geometric mean output level for cells transfected only with the marker gene. Units are arbitrary units from a flow cytometer.

(E) Left: output protein (in number of molecules) as a function of DNA copy number for the model with a sweep of values of $k_{mRNA,bind}$, the RISC-mRNA binding rate. Chosen parameter values are evenly log distributed over an order of magnitude centered on the original parameter value. Red lines highlight $c_{regulated} = 25$ and $c_{regulated} = 75$, and points indicate the values of several of the curves at these copy numbers. Right: output protein (in molecules) as a function of the RISC-mRNA binding rate at two copy numbers, $c_{regulated} = 25$ and $c_{regulated} = 75$. Values of $k_{mRNA,bind}$ are normalized to the original value. Color bar depicts normalized parameter values on a $\log_{10}$ scale.

(F) Geometric mean output expression of constructs co-transfected with a marker gene in HEK293T cells. Left: OL and CL circuits with miRE-FF4 or miRE-FF5 as in Figure 2C. Right: circuits with miRE-FF4 and varying numbers of target sites as in Figure 2F. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Dashed lines represent values for cells transfected only with the marker gene. Units are arbitrary units from a flow cytometer.

(G) Output protein (in number of molecules) as a function of DNA copy number for the model with a sweep of values of α_p, the free mRNA translation rate. Chosen parameter values are evenly log distributed over an order of magnitude centered on the original parameter value. Color bar depicts parameter values normalized to the original value on a $\log_{10}$ scale.

(H) Output protein (in number of molecules) as a function of DNA copy number for the model with a sweep of values of $k_{miRNA,deg}$, the RISC-bound microRNA degradation rate. Chosen parameter values are evenly log distributed over an order of magnitude centered on the original parameter value. Color bar depicts parameter values normalized to the original value on a $\log_{10}$ scale.

$^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$, independent t test. Tests compare CL and OL conditions.

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

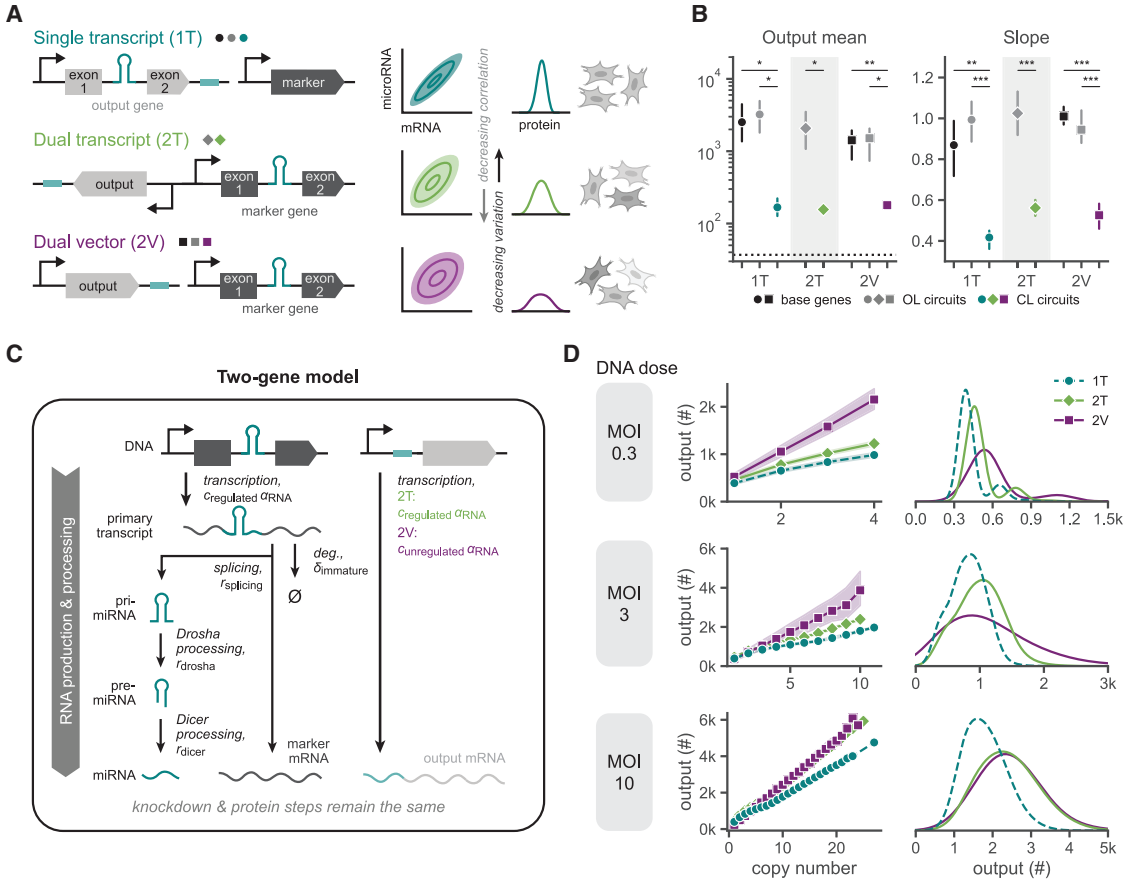


Figure 4. The single-transcript circuit matches or exceeds performance of two-gene implementations

(A) Left: DNA construct diagrams of single-transcript and two-gene circuit implementations. The single-transcript circuit (ComMAND, 1T) consists of an output gene with an intronic microRNA and matched (CL) or orthogonal (OL) target sites, co-delivered with a separate marker gene. The dual-transcript circuit (2T) consists of a single vector with two divergently oriented genes: the marker gene with an intronic microRNA and the output gene with matched (CL) or orthogonal (OL) target sites. The dual-vector circuit (2V) consists of the same two genes as in the dual-transcript circuit, each on a separate vector. All genes are expressed from the EF1α promoter, and all circuits use miRE-FF4. Right: delivering the circuit on a single transcript leads to strongly correlated, stoichiometrically equal expression of microRNA and mRNA species and thus to a narrow distribution of output protein expression. Dual-transcript or dual-vector architectures introduce more intrinsic noise, decreasing the correlation between microRNA and mRNA levels and increasing variation of the output protein.

(B) Summary statistics of output expression for HEK293T cells transfected with the base genes, CL circuits, or OL circuits depicted in (A). Presented mean values use the geometric mean. Slope represents the slope of the line fitted to the binned, log-transformed marker-output points. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Dashed line represents geometric mean output level for cells transfected only with a marker gene lacking an intronic microRNA. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Units are arbitrary units from a flow cytometer. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, independent t test.

(C) Schematic of RNA production and processing reactions modeled for two-gene circuit implementations. Unlike the single-transcript model depicted in Figure 3A, in these systems, mature output mRNA is not produced by transcription and splicing of the primary transcript. Instead, it is produced directly from transcription of a second gene with transcription rate α_{RNA}. The copy number of this second gene is either equivalent to ($c_{regulated}$, dual-transcript) or different from ($c_{unregulated}$, dual-vector) that of the microRNA gene. All other reactions remain the same as in Figure 3A. The spliced mature mRNA from the primary transcript represents marker gene transcripts and is ignored. See STAR Methods for additional details.

(D) Left: output protein (in number of molecules) as a function of DNA copy number, $c_{regulated}$, from stochastic simulations of single-transcript (1T), dual-transcript (2T), and dual-vector (2V) models depicted in (C) and Figure 3A. The plot shows summary statistics of output protein levels for simulations binned by copy number, where points represent the geometric mean of each bin, and shaded regions illustrate the geometric mean divided or multiplied by the geometric standard deviation. Right: histograms of output protein levels (in number of molecules) for the same simulations. For each simulation, DNA copy numbers (DNA doses) were chosen from a Poisson distribution with mean equal to the given MOI, the effective "multiplicity of infection," as illustrated in Figure S3C. 10,000 simulations were run for each condition. See STAR Methods for more details.

provide additional properties. Similar iFFL circuits have been constructed using two genes, in which the microRNA and output mRNA are transcribed separately.[26,27,30,46] These architectures support varying ratios of microRNA and mRNA, which can tune knockdown and may improve circuit function. To benchmark ComMAND, we compared its performance with several alternative circuit architectures (Figure 4A). In the dual-transcript implementation, the output gene containing a microRNA target site is divergently expressed with a second gene containing the intronic microRNA. In a dual-vector implementation, these genes are

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

separately encoded on two co-delivered vectors. While we expect the dual-transcript and dual-vector implementations to have the same average expression, the dual-vector system may introduce additional extrinsic noise due to copy-number variation between the microRNA and output components (Figure 4A, right). When we transfected these circuits in HEK293T cells, we found that the single-transcript CL circuit controls output expression as well as or better than both two-gene architectures (Figures 4B, S3A, and S3B). We observed minimal differences in expression of the base genes and OL circuits across all architectures, indicating that the improved control is specific to circuit activity.

Therapeutic applications often require orders of magnitude lower DNA copy numbers than those in transfection. As copy number decreases, the effect of noise becomes more pronounced, potentially magnifying differences between circuit implementations. We expect ComMAND to outperform two-gene architectures at low DNA copy numbers because the expression of circuit components is more tightly coupled. To explore circuit behavior in different copy-number regimes, we turned to our model. For the two-gene architectures, we modeled output mRNA transcription separately from production of the microRNA and allowed the associated DNA copy number to vary in the dual-vector case (Figure 4C). We performed stochastic simulations using these models to account for the effects of noise, choosing DNA copy-number values for each run from a distribution defined by an effective multiplicity of infection (MOI) (Figure S3C; STAR Methods). As expected, mean output protein level increases with MOI and is higher for an unregulated gene than for the CL circuits (Figures 4D and S3D). Furthermore, output variability is lower for the single-transcript architecture than for two-gene models, especially at lower DNA copy numbers (Figures 4D, S3D, and S3E). Thus, the stochastic simulations highlight the advantage of ComMAND's single-transcript implementation in noisy contexts.

We designed ComMAND to have a fail-safe mechanism, where errors in splicing of the primary transcript likely produce neither correct mRNA nor correct microRNA. In this case, the entire transcript is unproductive, and mRNA and microRNA levels remain tightly coupled. In contrast, for non-intronic designs or two-gene architectures, splicing errors or transcriptional differences in one component of the circuit may alter the ratio of mRNA and microRNA molecules, negatively impacting circuit control. While splicing errors for common transgenic promoters are rare in HEK293T cells,[47] we nevertheless sought to explore ComMAND performance in the context of mis-splicing, which may increase in disease.[48] We modified our original model to include mis-splicing reactions that produce only microRNA or mRNA from the primary transcript and performed stochastic simulations (Figure S3F; STAR Methods). The output variability in each mis-splicing model is lower than that for the two-gene models (Figures S3G and S3H). These mis-splicing scenarios are extreme (50% splicing errors), so they offer further evidence that the single-transcript architecture may outperform two-gene implementations in physiological regimes.

Comparing ComMAND with two-gene, microRNA-based iFFLs, we show experimentally and via stochastic modeling that ComMAND maintains performance comparable or superior to two-gene architectures. Thus, we demonstrate that ComMAND maintains tight regulation across DNA copy-number distributions. These results suggest that ComMAND may effectively control output expression for delivery methods with variable delivery efficiency, such as low-dosage vectors in therapeutically relevant, physiological contexts.

Multiple designs of the single-transcript circuit regulate output expression across delivery methods

After establishing the performance of ComMAND's single-transcript architecture, we investigated how the placement of genetic parts along the transcript affects output expression profiles. Previous work demonstrated that microRNA-mediated knockdown of mRNA is greater when target sites are located in the 5′ UTR compared with in the 3′ UTR.[43,49] Additionally, other RNA polymerase II (RNA Pol II)-driven microRNA expression systems have placed the intronic microRNA in the 3′ UTR rather than within a protein-coding sequence.[28,29] Therefore, we constructed two additional single-transcript circuit designs (Figure 5A). In the new designs, designated designs 2 and 3, we moved the microRNA target site to the 5′ UTR. In design 3, we additionally moved the intronic microRNA to the 3′ UTR. Transfecting these circuits in HEK293T cells, we observed that both CL circuits reduce output expression mean and variability, compared with the corresponding OL circuits (Figures 5B and S4A). Output expression of both OL and CL circuits is lower for designs 2 and 3 than for the original design. Because output genes including only a microRNA or target sites also have reduced expression levels (Figures S4B and S4C), we reasoned that the location of genetic elements affects circuit activity in part by altering processes other than knockdown, such as translation of the mature, unbound mRNA (Figure 3G). Location of the microRNA target sites may also affect the mechanism of microRNA-mediated knockdown: 3′ UTR target sites are thought to lead to direct microRNA degradation, while 5′ UTR target sites may mainly inhibit translation, altering usage of cellular resources.[49]

Next, we explored whether these positional effects are element specific or generalizable across microRNAs. To test this, we constructed circuit designs 2 and 3 with our previous panel of microRNAs. Trends in expression profiles for design 3 relative to the original design remain nearly constant across microRNAs, while those for design 2 vary across microRNAs (Figures S4D–S4F). Additionally, designs 2 and 3 perform as well as or better than corresponding two-gene architectures with target sites in the 5′ UTR (Figures S4G and S4H). However, design 3 CL circuits have very low output levels, just above background. Thus, the original design of ComMAND may be optimal for its consistent results across parts, although design 3 could alternatively be used to set a very low level of expression. Together, these results indicate that positional and element-specific effects combine to influence output expression profiles for ComMAND.

To identify optimal designs for future therapeutic applications, we encoded ComMAND in a lentiviral vector. To maintain high viral titer, we expressed the circuit from a small-molecule-inducible promoter on the antisense strand of the virus, oriented divergently from a constitutive marker gene (Figure S5A). We transduced HEK293T cells with these viruses in the presence of inducer and found that output expression levels and variation decrease relative to unregulated genes for all three circuit

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

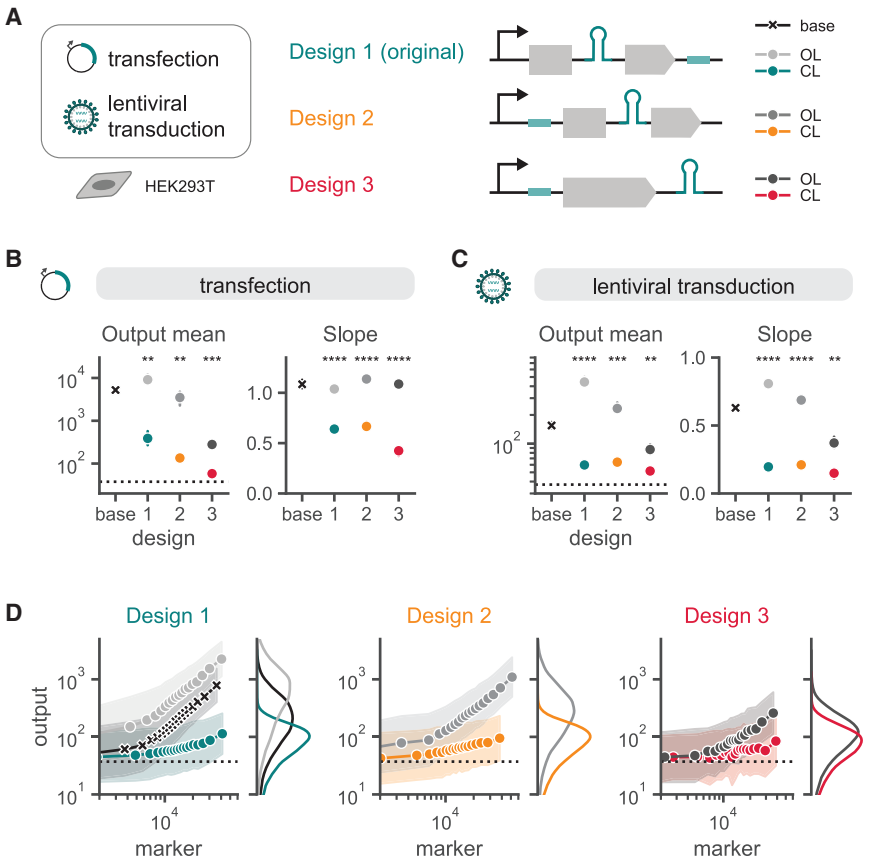


Figure 5. Multiple designs of the single-transcript circuit regulate output expression across delivery methods

(A) DNA construct diagrams for three designs of the single-transcript OL and CL circuits delivered to HEK293T cells via transfection or lentiviral transduction. Design 1 is the original circuit design, with an intronic microRNA within the output coding sequence and target sites in the 3′ UTR. In design 2, the target site is located in the 5′ UTR. In design 3, the target site is located in the 5′ UTR, and additionally, the intronic microRNA is moved to the 3′ UTR. Full vectors for lentiviral transduction are depicted in Figure S5A.

(B and C) Summary statistics of output expression in HEK293T cells transfected (B) or lentivirally transduced (C) with constructs in (A). Plotted mean values use the geometric mean, and slope represents the slope of the line fitted to the binned, log-transformed marker-output points. Dashed lines represent geometric mean output levels for cells transfected only with the marker gene (B) or for untransduced cells (C). Design "base" refers to the base gene construct that does not contain an intronic microRNA or target sites. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$, independent t test. Tests compare CL and OL conditions, and CL and base conditions are also significantly different.

(D) Output expression as a function of marker expression for one representative biological replicate of data in (C). Points represent the geometric means of equal-quantile bins, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Dashed line represents the output geometric mean of untransduced cells. Histograms depict output expression for cells in each condition.

All units are arbitrary units from a flow cytometer.

designs (Figures 5C, 5D, and S5B–S5D). In fact, these vectors elicit the tightest control yet for ComMAND. We attribute this to low transgene copy number in lentiviral transduction, allowing the circuit to operate in a regime not limited by cellular resources. Viral delivery also preserves the trends in output expression between designs, further suggesting that these patterns are generalizable.

Transposase-based integration is another common transgene delivery method that results in variable payload copy number.[19,50] To test whether ComMAND can regulate output expression in this context, we delivered PiggyBac transposase and vectors expressing design 1 of the circuit to HEK293T cells (Figure S5E). The CL circuit effectively reduces output variation relative to unregulated genes (Figure S5F). Thus, ComMAND functions when expressed via diverse delivery and integration methods.

In sum, ComMAND effectively regulates output expression in transfection, lentiviral transduction, and transposase-based integration of HEK293T cells. While multiple configurations of genetic elements enable circuit function, our original design regulates output expression most consistently across contexts and thus may be used for many applications. Having established the generalizability of ComMAND across delivery methods and having demonstrated high performance in lentiviral vectors, we next looked to deliver the circuit to therapeutically relevant cell types.

ComMAND functions in primary cells and can regulate clinically relevant output genes

With the ability to deliver ComMAND cargoes via lentivirus, we tested the efficacy of our gene circuit in primary cells (Figure 6A). For therapeutic applications, expression must be precisely tailored. High, unregulated expression of transgenes or heterogeneity across a population can compromise normal cell functions.[16–18] For example, overexpression of *frataxin* (*FXN*) in a mouse model of Friedrich's ataxia shows limited toxicity when *FXN* expression remains within a range of one order of magnitude of endogenous levels but is toxic when expressed at higher levels.[17,18]

We investigated the performance of ComMAND circuits in application-relevant cells by transducing the base gene, OL circuit, and CL circuit in multiple primary cell types: rat cortical neurons (Figures 6B, 6C, S6A, and S6B), mouse embryonic fibroblasts (Figures 6D, S6C, and S6E), and human T cells (Figures 6E, S6F, and S6G). Rat cortical neurons and mouse embryonic fibroblasts are important model cell types for neurological gene therapy delivery and for cellular reprogramming, respectively, while T cells are used for a variety of immunotherapies. Observing clusters of transduced neurons with microscopy (Figures 6B and S6B), we find that ComMAND generates more uniform expression across and between clusters of neurons.

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

CellPress

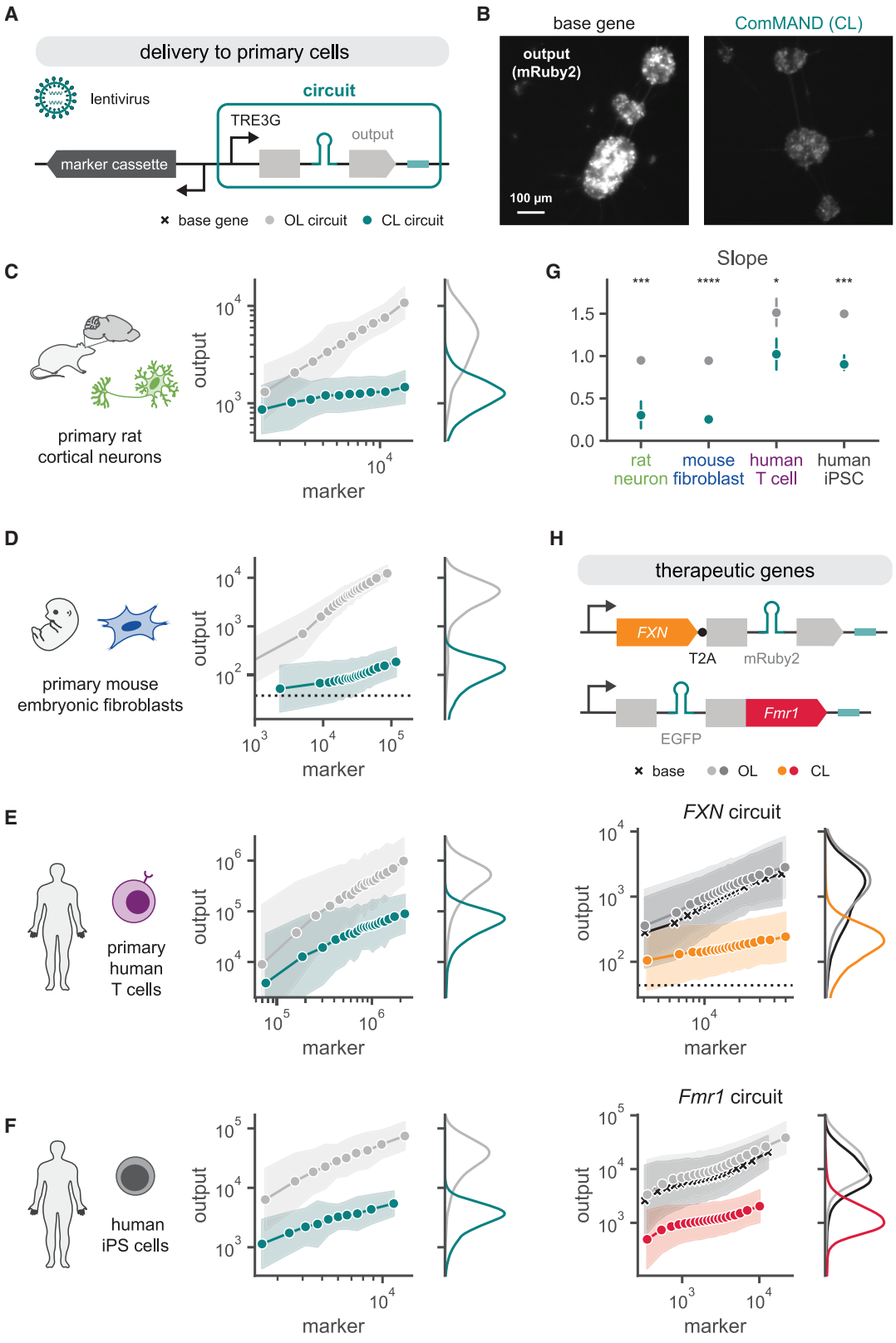


Figure 6. ComMAND functions in primary cells and can regulate clinically relevant output genes

(A) Schematic of lentiviral ComMAND vectors expressing the base gene, OL circuit, or CL circuit for delivery to primary cells. The regulated gene is expressed via the doxycycline-inducible TRE3G promoter with a divergently oriented marker cassette. See Figure S5A for more details.

(legend continued on next page)

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

We also tested ComMAND activity in delivery to human induced pluripotent stem cells (iPSCs). iPSCs are increasingly used as a starting cell type for disease modeling and to produce diverse cell therapies.[19,51,52] Similarly, we find that ComMAND reduces output variability in iPSCs for both lentiviral transduction (Figures 6F, S6H, and S6I) and transfection (Figures S7A–S7C).

For each cell type, we quantified the slope of the output distribution and observed strongly sublinear scaling with respect to marker expression, indicating tight control (Figure 6G). While the absolute slope varies across cell types, the slope for the CL circuit is lower than for the OL circuit in all cases. To investigate these differences in slope, we returned to our model. As ComMAND cannot buffer noise arising during translation, we hypothesized that changes in the translation rate could alter output protein expression. Our model predicts that as the translation rate increases, the slope of the output-marker curve in the resource-limited regime increases for both the OL circuit and the CL circuit (Figure S6J). This may explain the differences we observe in output protein profiles across cell types. Altogether, we find that ComMAND reduces population heterogeneity in application-relevant cell types, successfully narrowing the expression distribution of a model fluorescent protein cargo.

With the ability to control transgenes in therapeutically relevant delivery contexts and cell types, we sought to use ComMAND to regulate genes affected in two monogenic neurological disorders that have a narrow window of therapeutic efficacy. Friedrich's ataxia is a progressive neurodegenerative disease that results from the loss of frataxin (FXN).[53] Fragile X syndrome results from loss of expression of fragile X messenger ribonucleoprotein (FMRP) and is the most common inherited form of mental retardation.[54] To model gene therapies for Friedrich's ataxia and fragile X syndrome, we sought to control expression of human *FXN* and mouse *fragile X messenger ribonucleoprotein 1* (*Fmr1*) in transfection of HEK293T cells as a proof of concept. We expressed *FXN* alongside a fluorescent protein encoding the microRNA-containing intron using a "self-cleaving" 2A peptide sequence (Figure 6H, top). For *Fmr1*, we added the microRNA-containing intron to a fluorescent protein directly linked to the gene, generating a fusion protein. In both cases, ComMAND reduces the mean and slope of the output protein expression (Figures S7D and S7E). With these promising initial results, we next delivered the circuits regulating *FXN* and *Fmr1* expression using lentiviral vectors in HEK293T cells. Here too, the CL circuit exhibits lower output mean and variability, compared with unregulated genes (Figures 6H and S7F–S7I). Because FMRP is directly fused to a fluorescent protein, fluorescence measurements accurately reflect levels of the therapeutic cargo. Additionally, the fluorescence distribution of the fusion protein is consistent with localization in cytoplasmic granules (Figure S7J), suggesting that EGFP-FMRP may remain functional.[55]

To investigate how the expression of *FXN* and *Fmr1* in transduced cells compares to endogenous levels, we quantified RNA transcripts using quantitative reverse-transcription PCR (RT-qPCR). We selected primer sets that amplify either (1) both the endogenous and transgenic gene (*FXN*) or (2) only the transgenic transcripts (*Fmr1* and fluorescent proteins). We quantified expression in HEK293T cells lentivirally transduced with ComMAND circuits and calculated expression levels relative to the base gene (STAR Methods). Trends in relative expression across base genes, OL circuits, and CL circuits align with those for protein levels measured by flow cytometry (Figures S7F, S7H, S7K, and S7L). *FXN* transcript levels for cells transduced with the base gene are ~56-fold greater than for endogenous expression in untransduced cells (Figure S7K). In contrast, cells transduced with the CL circuit express *FXN* mRNA at levels only ~7.5 times those for untransduced cells.

Altogether, we find that ComMAND effectively regulates output expression in primary cells and of the therapeutic cargoes *FXN* and *Fmr1*. By increasing the proportion of cells with low levels of expression, ComMAND may achieve control of functional proteins within physiological regimes. This proof-of-concept work lays the foundation for translation of ComMAND to therapeutically relevant contexts where tight control of transgene expression is essential for safety and efficacy.

DISCUSSION

To build gene circuits that provide controlled dosages of transgenic cargoes, we developed ComMAND, a single-transcript, microRNA-mediated iFFL (Figure 1). As an RNA-based control system, ComMAND provides compact and modular control, minimal immunogenicity, low resource burden, orthogonality to native gene networks, and programmability via selection of

(B) Representative images of rat cortical neurons transduced (MOI = 1) with lentiviral vectors expressing a base gene or the ComMAND CL circuit. Images depict fluorescence of the output gene, mRuby2, for ~200 representative cells, and the scale bar represents 100 μm.

(C–F) Output expression as a function of marker expression for primary rat cortical neurons, primary mouse embryonic fibroblasts, primary human T cells, and human iPSCs, respectively, transduced with the lentiviral vectors in Figure S5A. Cells were transduced at an MOI of 7, 1, 1, and 0.3, respectively, and cultured in the presence of inducer (1 μg/mL doxycycline). Flow cytometry measurements are binned by marker expression into equal-quantile groups per condition. Points represent geometric means of cells in each bin, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Plots depict one representative biological replicate.

(G) Slopes of the marker-output curves for the OL and CL circuits in (C)–(F). Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$, independent t test.

(H) Lentiviral delivery of ComMAND circuits controlling the therapeutically relevant genes *FXN* and *Fmr1*. For the *FXN* circuit, the microRNA is encoded in an intron within the *mRuby2* gene, and mRuby2 is co-expressed with the human FXN protein via a T2A "self-cleaving" peptide. For the *Fmr1* circuit, the intronic microRNA is located within the *egfp* coding sequence, and EGFP is directly fused to mouse FMRP. The full lentiviral vector also contains a divergently oriented marker cassette, as in (A). Plots depict output expression as a function of marker expression for HEK293T cells transduced (MOI = 1) with these circuits in the presence of inducer (1 μg/mL doxycycline). Flow cytometry measurements are binned by marker expression into equal-quantile groups per condition. Points represent geometric means of cells in each bin, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Plots depict one representative biological replicate. Dashed lines represent geometric mean output levels for untransduced cells. All units are arbitrary units from a flow cytometer.

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

sequences.[56] While other microRNA-mediated iFFLs have been developed,[1,25–28,30,33] ComMAND combines compact, single-transcript design with a fail-safe mechanism achieved by encoding the microRNA within an intron in the transgene. This design prevents translation of the primary mRNA if splicing of the mRNA fails. Thus, ComMAND ensures that both the mRNA and microRNA components of the iFFL are produced in equal amounts, supporting robust performance and failing into a non-toxic state of low expression. ComMAND achieves sublinear scaling of transgene dosage, approaching the theoretical limit of control at low, clinically relevant DNA doses (Figure 2). As ComMAND operates through posttranscriptional regulation, the dosage of the transgenic cargo delivered by ComMAND can be orthogonally tuned via promoter selection. Following rational tuning of components, we used mathematical modeling to examine physiological constraints on ComMAND performance and to identify design rules (Figure 3). Further, we used stochastic modeling coupled with experiments to compare ComMAND with alternative two-gene architectures and to identify optimal regimes of performance (Figure 4). Notably, we predicted that ComMAND performance would improve at the low DNA doses relevant for therapies. In lentiviral delivery, we demonstrate that ComMAND tightly controls expression across diverse cell types including human iPSCs, primary mouse embryonic fibroblasts, primary rat neurons, and primary human T cells (Figures 5 and 6). We show that ComMAND constrains expression of the therapeutically relevant transgenes *FXN* and *Fmr1* within a narrow window, supporting translational therapies for two neurological disorders. Altogether, these data provide compelling evidence that ComMAND can support safe, effective control in regimes that approach therapeutic relevance.

Transgene dosage from ComMAND can be rationally tuned by selection of genetic parts. Stronger promoters increase expression of transgenes from ComMAND, and optimization of the microRNA sequence and scaffold can improve mean expression and reduce variability, albeit with diminishing returns (Figures 2A–2C). While the addition of multiple target sites tunes the output in other microRNA-mediated iFFLs, we observe only a small effect from increasing the number of target sites in ComMAND (Figure 2F). However, our target sites are perfectly complementary to the microRNA. Because perfectly and imperfectly complementary sequences are processed via different mechanisms, ComMAND may have a relatively small range of tunability with respect to target site number.[1,30,33,46] As previously observed for other iFFLs,[27,28,49] ComMAND reduces resource burden, extending the range of co-expression for the marker and the output, compared with the base gene (Figure 2A).

We used our mathematical model to understand the parameters that physiologically constrain the performance of ComMAND. We identify the ratio of the transcription rate and the total pool of RISC as a sensitive parameter that defines the regime of controllability for ComMAND (Figures 3B and 3C). Specifically, saturation of RISC sets a physiological limit on the control afforded by microRNA-mediated circuits. As total levels of RISC are regulated through a myriad of loading and degradation mechanisms that can vary by cell type and genetic background,[57,58] it is more feasible to tune ComMAND to reside in the regime of controllability through promoter selection. We also identify that most parameters affecting microRNA processing do not significantly impact output expression, suggesting that optimization of these steps may not be fruitful (Figures 3D and S2C).

We designed ComMAND as a single transcript for optimal compactness and tight coupling between the production of mRNA and microRNA species. In theory, a single-transcript, microRNA-mediated iFFL provides optimal noise suppression.[41] Dosage control from ComMAND matches or exceeds the performance of dual-transcript and dual-vector architectures in transfection of HEK293T cells (Figure 4B). Transfections deliver thousands of copies of plasmids, potentially obscuring differences in designs that would manifest at lower copy numbers. Examining a range of low copy numbers with stochastic simulations, we find that ComMAND outperforms two-gene implementations in this clinically relevant regime (Figure 4D). Thus, we conclude that single-transcript designs will provide optimal control of therapeutic cargoes.

Although ComMAND consistently decreases absolute variability in protein expression across cell types, we did not observe reductions in coefficient of variation as others have reported for microRNA-based iFFLs.[29,33] Recent theory work predicts that noise in expression from microRNA-based iFFLs can increase or decrease depending on the parameter regime,[41] which may explain why we did not observe differences in coefficient of variation. It remains possible that ComMAND does not reduce variation in all cell types or conditions. Nevertheless, the consistent decrease in slope for ComMAND, compared with the unregulated genes, matches the trends observed in recent microRNA-based and protein-based iFFLs delivered to application-relevant cells.[1,30,59]

While ComMAND provides precise control of transgene expression by mitigating variance inherent to methods of delivery, there remain limits to this control and opportunities to further define design parameters. As a form of negative regulation, we expect that ComMAND will decrease expression of the output gene. In general, we observe reductions in the mean expression relative to the OL circuit and base gene. However, introduction of the intronic microRNA increases expression from ComMAND, compared with the base gene, in mouse embryonic fibroblasts, human T cells, and human iPSCs (Figures 6D–6F). Addition of an intron may increase expression through splicing-mediated transcriptional activation and may be differentially regulated across cell types.[60,61] As the position of introns influences transcription, there remains an opportunity to understand how placement of the intronic microRNA along ComMAND's single transcript tunes expression for therapeutically relevant genes and varies across cell types. In theory, ComMAND controls for DNA copy number as well as transcriptional variance to reduce noise in gene expression.[41] However, ComMAND may not offer effective control when the sources of cell-to-cell variation arise from processes other than transcription. Changes in translation rates or significant variances introduced in translation may obscure noise suppression at the transcriptional level. Translation rates may vary across genes, cell types, and cell states,[47,62,63] potentially explaining the differences in slope we observed for lentiviral delivery to primary cell

types. Other forms of regulation such as translational control and protein-based circuits may be coupled with ComMAND to control expression across the central dogma.[59,64–67]

Here, we show that ComMAND controls transgene dosage in a range of mammalian primary cells including primary human T cells and iPSCs, demonstrating the enormous potential to apply ComMAND to control transgenes for therapeutic applications and basic research. For haploinsufficiency disorders such as Friedrich's ataxia and fragile X syndrome, ComMAND-regulated gene supplementation of FXN and FMRP may be developed to offer safer, effective gene therapies. While further quantification of protein levels in disease-relevant cell types is needed, *FXN* transcript expression measured by RT-qPCR for ComMAND is within one order of magnitude of endogenous levels (Figure S7K). With further tuning, ComMAND may be able to achieve expression levels within the physiological regime. As a posttranscriptional regulator, ComMAND can putatively control cargoes from cell-type-specific promoters for targeted expression. Additionally, integration of cell-state and pathway-responsive promoters may scale the expression of ComMAND for cell-autonomous feedback. As transgenes are increasingly used to augment cellular functions and program cell fate, ComMAND may support well-controlled expression of transgenes for diverse therapeutic applications.[68–70] By increasing the predictability of transgene expression, we expect ComMAND will broadly improve the performance of gene circuits in therapeutic contexts.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Kate E. Galloway, PhD (katiegal@mit.edu).

Materials availability

Most plasmids used in this work have been deposited at Addgene, see Tables S4 and S5. Plasmids not available through Addgene are available from the lead contact upon request. Sequences for all plasmids have been deposited at Zenodo, and the DOI is listed in the key resources table.

Data and code availability

- All raw data, analyzed data, and modeling simulations have been deposited at Zenodo and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- All code for data analysis, modeling, and figure generation is publicly available on GitHub at https://github.com/GallowayLabMIT/ComMAND. The code has also been deposited at Zenodo, and the DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

Research reported in this manuscript was supported by the National Institute of General Medical Sciences of the National Institutes of Health (under award numbers R35-GM143033 and DP2-AI158126), the National Science Foundation (with NSF-CAREER under award number 2339986), the Institute for Collaborative Biotechnologies (W911NF-19-2-0026), the Air Force Research Laboratory MURI (FA9550-22-1-03t16), and the DoD (W81XWH2210300). K.S.L. and E.L.P. are supported by the National Science Foundation Graduate Research Fellowship Program under grant number 1745302. Work completed at the Koch Institute's Robert A. Swanson (1969) Biotechnology Center Flow Cytometry Core and the MIT BioMicro Center was supported in part by Koch Institute Support (core) grant P30-CA014051 from the National Cancer Institute. We thank Jack Toppen, Albert Blanch Asensio, Adam Beitz, Deon Ploessl, and Diya Godavarti for their help with experiments. We thank Sneha Kabaria, Nathan Wang, Deon Ploessl, and Yunbeen Bae for their feedback on the development of the manuscript. iPS11 cells and the hyperactive PiggyBac plasmid were gifts from the Weiss lab.

AUTHOR CONTRIBUTIONS

K.S.L. and K.E.G. conceived and outlined the project. K.S.L., C.P.J., E.L.P., and S.G. performed experiments. K.S.L. and C.P.J. analyzed experiments. C.P.J. and E.L.P. composed the model. K.S.L., K.E.G., C.P.J., and E.L.P. wrote and edited the manuscript. M.E.B supervised the human T cell experiments. K.E.G. supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL DETAILS
  - HEK293T cells
  - Human induced pluripotent stem cells
  - Primary mouse embryonic fibroblasts
  - Primary rat cortical neurons
  - Primary human T cells
- METHOD DETAILS
  - Cloning
  - HEK293T transfection
  - PiggyBac integration
  - iPSC transfection
  - Lentivirus production
  - Lentiviral transduction
  - Of primary rat cortical neurons
  - RT-qPCR
  - Fluorescent imaging
  - Steady-state modeling
  - Stochastic simulations
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Flow cytometry
  - Viral titer calculation
  - RT-qPCR analysis
  - Statistics

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cels.2025.101269.

Received: July 8, 2024
Revised: December 12, 2024
Accepted: April 2, 2025
Published: April 28, 2025

REFERENCES

1. Flynn, M.J., Mayfield, A.M.H., Du, R., Gradinaru, V., and Elowitz, M.B. (2024). Synthetic dosage-compensating miRNA circuits allow precision gene therapy for Rett syndrome. Preprint at bioRxiv. https://doi.org/10.1101/2024.03.13.584179.
2. Adusumilli, P.S., Cherkassky, L., Villena-Vargas, J., Colovos, C., Servais, E., Plotkin, J., Jones, D.R., and Sadelain, M. (2014). Regional delivery of mesothelin-targeted CAR T cell therapy generates potent and long-lasting

CD4-dependent tumor immunity. Sci. Transl. Med. *6*, 261ra151. https://doi.org/10.1126/scitranslmed.3010162.

3. Huang, Q., Chan, K.Y., Wu, J., Botticello-Romero, N.R., Zheng, Q., Lou, S., Keyes, C., Svanbergsson, A., Johnston, J., Mills, A., et al. (2024). An AAV capsid reprogrammed to bind human transferrin receptor mediates brain-wide gene delivery. Science *384*, 1220–1227. https://doi.org/10.1126/science.adm8386.
4. Tervo, D.G.R., Hwang, B.-Y., Viswanathan, S., Gaj, T., Lavzin, M., Ritola, K.D., Lindo, S., Michael, S., Kuleshova, E., Ojala, D., et al. (2016). A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. Neuron *92*, 372–382. https://doi.org/10.1016/j.neuron.2016.09.021.
5. Kotterman, M.A., and Schaffer, D.V. (2014). Engineering adeno-associated viruses for clinical gene therapy. Nat. Rev. Genet. *15*, 445–451. https://doi.org/10.1038/nrg3742.
6. Joung, J., Ma, S., Tay, T., Geiger-Schuller, K.R., Kirchgatterer, P.C., Verdine, V.K., Guo, B., Arias-Garcia, M.A., Allen, W.E., Singh, A., et al. (2023). A transcription factor atlas of directed differentiation. Cell *186*, 209–229.e26. https://doi.org/10.1016/j.cell.2022.11.026.
7. Veerapandian, V., Ackermann, J.O., Srivastava, Y., Malik, V., Weng, M., Yang, X., and Jauch, R. (2018). Directed Evolution of Reprogramming Factors by Cell Selection and Sequencing. Stem Cell Rep. *11*, 593–606. https://doi.org/10.1016/j.stemcr.2018.07.002.
8. Eguchi, A., Wleklinski, M.J., Spurgat, M.C., Heiderscheit, E.A., Kropornicka, A.S., Vu, C.K., Bhimsaria, D., Swanson, S.A., Stewart, R., Ramanathan, P., et al. (2016). Reprogramming cell fate with a genome-scale library of artificial transcription factors. Proc. Natl. Acad. Sci. USA *113*, E8257–E8266. https://doi.org/10.1073/pnas.1611142114.
9. Legut, M., Gajic, Z., Guarino, M., Daniloski, Z., Rahman, J.A., Xue, X., Lu, C., Lu, L., Mimitou, E.P., Hao, S., et al. (2022). A genome-scale screen for synthetic drivers of T cell proliferation. Nature *603*, 728–735. https://doi.org/10.1038/s41586-022-04494-7.
10. Apgar, T.L., and Sanders, C.R. (2022). Compendium of causative genes and their encoded proteins for common monogenic disorders. Protein Sci. *31*, 75–91. https://doi.org/10.1002/pro.4183.
11. Grabczyk, E., Kumari, D., and Usdin, K. (2001). Fragile X syndrome and Friedreich's ataxia: two different paradigms for repeat induced transcript insufficiency. Brain Res. Bull. *56*, 367–373. https://doi.org/10.1016/S0361-9230(01)00572-X.
12. Shi, Y., Lin, S., Staats, K.A., Li, Y., Chang, W.-H., Hung, S.-T., Hendricks, E., Linares, G.R., Wang, Y., Son, E.Y., et al. (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. Nat. Med. *24*, 313–325. https://doi.org/10.1038/nm.4490.
13. Han, X., Chen, S., Flynn, E., Wu, S., Wintner, D., and Shen, Y. (2018). Distinct epigenomic patterns are associated with haploinsufficiency and predict risk genes of developmental disorders. Nat. Commun. *9*, 2138. https://doi.org/10.1038/s41467-018-04552-7.
14. Chénier, S., Yoon, G., Argiropoulos, B., Lauzon, J., Laframboise, R., Ahn, J.W., Ogilvie, C.M., Lionel, A.C., Marshall, C.R., Vaags, A.K., et al. (2014). CHD2 haploinsufficiency is associated with developmental delay, intellectual disability, epilepsy and neurobehavioural problems. J. Neurodev. Disord. *6*, 9. https://doi.org/10.1186/1866-1955-6-9.
15. Rom, A., Melamed, L., Gil, N., Goldrich, M.J., Kadir, R., Golan, M., Biton, I., Perry, R.B.-T., and Ulitsky, I. (2019). Regulation of CHD2 expression by the Chaserr long noncoding RNA gene is essential for viability. Nat. Commun. *10*, 5092. https://doi.org/10.1038/s41467-019-13075-8.
16. Peier, A.M., McIlwain, K.L., Kenneson, A., Warren, S.T., Paylor, R., and Nelson, D.L. (2000). (Over)correction of FMR1 deficiency with YAC transgenics: behavioral and physical features. Hum. Mol. Genet. *9*, 1145–1159. https://doi.org/10.1093/hmg/9.8.1145.
17. Huichalaf, C., Perfitt, T.L., Kuperman, A., Gooch, R., Kovi, R.C., Brenneman, K.A., Chen, X., Hirenallur-Shanthappa, D., Ma, T., Assaf, B. T., et al. (2022). In vivo overexpression of frataxin causes toxicity mediated by iron-sulfur cluster deficiency. Mol. Ther. Methods Clin. Dev. *24*, 367–378. https://doi.org/10.1016/j.omtm.2022.02.002.
18. Belbellaa, B., Reutenauer, L., Messaddeq, N., Monassier, L., and Puccio, H. (2020). High Levels of Frataxin Overexpression Lead to Mitochondrial and Cardiac Toxicity in Mouse Models. Mol. Ther. Methods Clin. Dev. *19*, 120–138. https://doi.org/10.1016/j.omtm.2020.08.018.
19. Peterman, E.L., Ploessl, D.S., and Galloway, K.E. (2024). Accelerating Diverse Cell-Based Therapies Through Scalable Design. Annu. Rev. Chem. Biomol. Eng. *15*, 267–292. https://doi.org/10.1146/annurev-chembioeng-100722-121610.
20. Chen, X., Lim, D.A., Lawlor, M.W., Dimmock, D., Vite, C.H., Lester, T., Tavakkoli, F., Sadhu, C., Prasad, S., and Gray, S.J. (2023). Biodistribution of Adeno-Associated Virus Gene Therapy Following Cerebrospinal Fluid-Directed Administration. Hum. Gene Ther. *34*, 94–111. https://doi.org/10.1089/hum.2022.163.
21. Hong, C.K.Y., and Cohen, B.A. (2022). Genomic environments scale the activities of diverse core promoters. Genome Res. *32*, 85–96. https://doi.org/10.1101/gr.276025.121.
22. Ede, C., Chen, X., Lin, M.-Y., and Chen, Y.Y. (2016). Quantitative Analyses of Core Promoters Enable Precise Engineering of Regulated Gene Expression in Mammalian Cells. ACS Synth. Biol. *5*, 395–404. https://doi.org/10.1021/acssynbio.5b00266.
23. Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.-Z., and Lahn, B.T. (2010). Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. PLoS One *5*, e10611. https://doi.org/10.1371/journal.pone.0010611.
24. Segall-Shapiro, T.H., Sontag, E.D., and Voigt, C.A. (2018). Engineered promoters enable constant gene expression at any copy number in bacteria. Nat. Biotechnol. *36*, 352–358. https://doi.org/10.1038/nbt.4111.
25. Bleris, L., Xie, Z., Glass, D., Adadey, A., Sontag, E., and Benenson, Y. (2011). Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. Mol. Syst. Biol. *7*, 519. https://doi.org/10.1038/msb.2011.49.
26. Lillacci, G., Benenson, Y., and Khammash, M. (2018). Synthetic control systems for high performance gene expression in mammalian cells. Nucleic Acids Res. *46*, 9855–9863. https://doi.org/10.1093/nar/gky795.
27. Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., and Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. Nat. Commun. *11*, 4641. https://doi.org/10.1038/s41467-020-18392-x.
28. Jones, R.D., Qian, Y., Siciliano, V., DiAndreth, B., Huh, J., Weiss, R., and Del Vecchio, D. (2020). An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. Nat. Commun. *11*, 5690. https://doi.org/10.1038/s41467-020-19126-9.
29. Yang, J., Lee, J., Land, M.A., Lai, S., Igoshin, O.A., and St-Pierre, F. (2021). A synthetic circuit for buffering gene dosage variation between individual mammalian cells. Nat. Commun. *12*, 4132. https://doi.org/10.1038/s41467-021-23889-0.
30. Du, R., Flynn, M.J., Honsa, M., Jungmann, R., and Elowitz, M.B. (2024). miRNA circuit modules for precise, tunable control of gene expression. Preprint at bioRxiv. https://doi.org/10.1101/2024.03.12.583048.
31. Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. USA *100*, 11980–11985. https://doi.org/10.1073/pnas.2133841100.
32. Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining Network Topologies that Can Achieve Biochemical Adaptation. Cell *138*, 760–773. https://doi.org/10.1016/j.cell.2009.06.013.
33. Strovas, T.J., Rosenberg, A.B., Kuypers, B.E., Muscat, R.A., and Seelig, G. (2014). MicroRNA-based single-gene circuits buffer protein synthesis rates against perturbations. ACS Synth. Biol. *3*, 324–331. https://doi.org/10.1021/sb4001867.
34. Aoki, S.K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., and Khammash, M. (2019). A universal biomolecular integral feedback controller for robust perfect adaptation. Nature *570*, 533–537. https://doi.org/10.1038/s41586-019-1321-1.

35. Khammash, M.H. (2021). Perfect adaptation in biology. Cell Syst. *12*, 509–521. https://doi.org/10.1016/j.cels.2021.05.020.

36. Blanchini, F., Franco, E., and Giordano, G. (2012). Determining the structural properties of a class of biological models. In 51st IEEE Conference on Decision and Control (CDC) (IEEE), pp. 5505–5510. https://doi.org/10.1109/CDC.2012.6427037.

37. Osella, M., Bosia, C., Corá, D., and Caselle, M. (2011). The Role of Incoherent MicroRNA-Mediated Feedforward Loops in Noise Buffering. PLoS Comput. Biol. *7*, e1001101. https://doi.org/10.1371/journal.pcbi.1001101.

38. Shimoga, V., White, J.T., Li, Y., Sontag, E., and Bleris, L. (2013). Synthetic mammalian transgene negative autoregulation. Mol. Syst. Biol. *9*, 670. https://doi.org/10.1038/msb.2013.27.

39. Blahetek, G., Mayer, C., Zuber, J., Lenter, M., and Strobel, B. (2024). Suppression of toxic transgene expression by optimized artificial miRNAs increases AAV vector yields in HEK-293 cells. Mol. Ther. Methods Clin. Dev. *32*, 101280. https://doi.org/10.1016/j.omtm.2024.101280.

40. Leisner, M., Bleris, L., Lohmueller, J., Xie, Z., and Benenson, Y. (2010). Rationally designed logic integration of regulatory signals in mammalian cells. Nat. Nanotechnol. *5*, 666–670. https://doi.org/10.1038/nnano.2010.135.

41. Carignano, A., Mukherjee, S., Singh, A., and Seelig, G. (2018). Extrinsic Noise Suppression in micro RNA mediated Incoherent feedforward Loops. Preprint at bioRxiv. https://doi.org/10.1101/422394.

42. Nissim, L., Wu, M.-R., Pery, E., Binder-Nissim, A., Suzuki, H.I., Stupp, D., Wehrspaun, C., Tabach, Y., Sharp, P.A., and Lu, T.K. (2017). Synthetic RNA-Based Immunomodulatory Gene Circuits for Cancer Immunotherapy. Cell *171*, 1138–1150.e15. https://doi.org/10.1016/j.cell.2017.09.049.

43. Gam, J.J., Babb, J., and Weiss, R. (2018). A mixed antagonistic/synergistic miRNA repression model enables accurate predictions of multi-input miRNA sensor activity. Nat. Commun. *9*, 2430. https://doi.org/10.1038/s41467-018-04575-0.

44. Cohen, R.N., van der Aa, M.A.E.M., Macaraeg, N., Lee, A.P., and Szoka, F.C. (2009). Quantification of Plasmid DNA Copies in the Nucleus after Lipoplex and Polyplex Transfection. J. Control. Release *135*, 166–174. https://doi.org/10.1016/j.jconrel.2008.12.016.

45. Zhang, Z., Qin, Y.-W., Brewer, G., and Jing, Q. (2012). MicroRNA degradation and turnover: regulating the regulators. Wiley Interdiscip. Rev. RNA *3*, 593–600. https://doi.org/10.1002/wrna.1114.

46. Ilia, K., Shakiba, N., Bingham, T., Jones, R.D., Kaminski, M.M., Aravera, E., Bruno, S., Palacios, S., Weiss, R., Collins, J.J., et al. (2023). Synthetic genetic circuits to uncover the OCT4 trajectories of successful reprogramming of human fibroblasts. Sci. Adv. *9*, eadg8495. https://doi.org/10.1126/sciadv.adg8495.

47. Peterman, E.L., Ploessl, D.S., Love, K.S., Sanabria, V., Daniels, R.F., Johnstone, C.P., Godavarti, D.R., Kabaria, S.R., Oakes, C.G., Pai, A.A., et al. (2024). High-resolution profiling reveals coupled transcriptional and translational regulation of transgenes. Preprint at bioRxiv. https://doi.org/10.1101/2024.11.26.625483.

48. Garcia-Cabau, C., Bartomeu, A., Tesei, G., Cheung, K.C., Pose-Utrilla, J., Picó, S., Balaceanu, A., Duran-Arqué, B., Fernández-Alfara, M., Martín, J., et al. (2025). Mis-splicing of a neuronal microexon promotes CPEB4 aggregation in ASD. Nature *637*, 496–503. https://doi.org/10.1038/s41586-024-08289-w.

49. Cella, F., Perrino, G., Tedeschi, F., Viero, G., Bosia, C., Stan, G.-B., and Siciliano, V. (2023). MIRELLA: a mathematical model explains the effect of microRNA-mediated synthetic genes regulation on intracellular resource allocation. Nucleic Acids Res. *51*, 3452–3464. https://doi.org/10.1093/nar/gkad151.

50. Kettlun, C., Galvan, D.L., George, A.L., Jr., Kaja, A., and Wilson, M.H. (2011). Manipulating piggyBac Transposon Chromosomal Integration Site Selection in Human Cells. Mol. Ther. *19*, 1636–1644. https://doi.org/10.1038/mt.2011.129.

51. Rowe, R.G., and Daley, G.Q. (2019). Induced pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Genet. *20*, 377–388. https://doi.org/10.1038/s41576-019-0100-z.

52. Shi, Y., Inoue, H., Wu, J.C., and Yamanaka, S. (2017). Induced pluripotent stem cell technology: a decade of progress. Nat. Rev. Drug Discov. *16*, 115–130. https://doi.org/10.1038/nrd.2016.245.

53. Payne, R.M., and Wagner, G.R. (2012). Cardiomyopathy in Friedreich Ataxia: Clinical Findings and Research. J. Child Neurol. *27*, 1179–1186. https://doi.org/10.1177/0883073812448535.

54. Lee, H.-G., Imaichi, S., Kraeutler, E., Aguilar, R., Lee, Y.-W., Sheridan, S.D., and Lee, J.T. (2023). Site-specific R-loops induce CGG repeat contraction and fragile X gene reactivation. Cell *186*, 2593–2609.e18. https://doi.org/10.1016/j.cell.2023.04.035.

55. Okray, Z., de Esch, C.E.F., Van Esch, H., Devriendt, K., Claeys, A., Yan, J., Verbeeck, J., Froyen, G., Willemsen, R., de Vrij, F.M.S., et al. (2015). A novel fragile X syndrome mutation reveals a conserved role for the carboxy-terminus in FMRP localization and function. EMBO Mol. Med. *7*, 423–437. https://doi.org/10.15252/emmm.201404576.

56. Takahashi, K., and Galloway, K.E. (2024). RNA-based controllers for engineering gene and cell therapies. Curr. Opin. Biotechnol. *85*, 103026. https://doi.org/10.1016/j.copbio.2023.103026.

57. Kawamata, T., and Tomari, Y. (2010). Making RISC. Trends Biochem. Sci. *35*, 368–376. https://doi.org/10.1016/j.tibs.2010.03.009.

58. Iwakawa, H.-O., and Tomari, Y. (2022). Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. Mol. Cell *82*, 30–43. https://doi.org/10.1016/j.molcel.2021.11.026.

59. Katz, N., An, C., Lee, Y.-J., Tycko, J., Zhang, M., Kang, J., Bintu, L., Bassik, M.C., Huang, W.-H., and Gao, X.J. (2024). Tunable, self-contained gene dosage control via proteolytic cleavage of CRISPR-Cas systems. Preprint at bioRxiv. https://doi.org/10.1101/2024.10.09.617463.

60. Fiszbein, A., Krick, K.S., Begg, B.E., and Burge, C.B. (2019). Exon-Mediated Activation of Transcription Starts. Cell *179*, 1551–1565.e17. https://doi.org/10.1016/j.cell.2019.11.002.

61. Uriostegui-Arcos, M., Mick, S.T., Shi, Z., Rahman, R., and Fiszbein, A. (2023). Splicing activates transcription from weak promoters upstream of alternative exons. Nat. Commun. *14*, 3435. https://doi.org/10.1038/s41467-023-39200-2.

62. Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. Cell *157*, 624–635. https://doi.org/10.1016/j.cell.2014.02.033.

63. Riba, A., Di Nanni, N., Mittal, N., Arhné, E., Schmidt, A., and Zavolan, M. (2019). Protein synthesis rates and ribosome occupancies reveal determinants of translation elongation rates. Proc. Natl. Acad. Sci. USA *116*, 15023–15032. https://doi.org/10.1073/pnas.1817299116.

64. Gao, X.J., Chong, L.S., Kim, M.S., and Elowitz, M.B. (2018). Programmable protein circuits in living cells. Science *361*, 1252–1258. https://doi.org/10.1126/science.aat5062.

65. Vlahos, A.E., Kang, J., Aldrete, C.A., Zhu, R., Chong, L.S., Elowitz, M.B., and Gao, X.J. (2022). Protease-controlled secretion and display of intercellular signals. Nat. Commun. *13*, 912. https://doi.org/10.1038/s41467-022-28623-y.

66. Duan, M., Dev, I., Lu, A., You, M.Y., and Shapiro, M.G. (2023). Stoichiometric expression of messenger polycistrons by eukaryotic ribosomes (SEMPER) for compact, ratio-tunable multi-gene expression from single mRNAs. Cell Syst. *15*, 597–609.e4. https://doi.org/10.1101/2023.05.26.541240.

67. Barrington, C.L., Galindo, G., Koch, A.L., Horton, E.R., Morrison, E.J., Tisa, S., Stasevich, T.J., and Rissland, O.S. (2023). Synonymous codon usage regulates translation initiation. Cell Rep. *42*, 113413. https://doi.org/10.1016/j.celrep.2023.113413.

68. Chmielewski, M., and Abken, H. (2017). CAR T Cells Releasing IL-18 Convert to T-Bethigh FoxO1low Effectors that Exhibit Augmented

Activity against Advanced Solid Tumors. Cell Rep. *21*, 3205–3219. https://doi.org/10.1016/j.celrep.2017.11.063.

69. Christodoulou, I., Ho, W.J., Marple, A., Ravich, J.W., Tam, A., Rahnama, R., Fearnow, A., Rietberg, C., Yanik, S., Solomou, E.E., et al. (2021). Engineering CAR-NK cells to secrete IL-15 sustains their anti-AML functionality but is associated with systemic toxicities. J. Immunother. Cancer *9*, e003894. https://doi.org/10.1136/jitc-2021-003894.

70. Zhang, L., Morgan, R.A., Beane, J.D., Zheng, Z., Dudley, M.E., Kassim, S. H., Nahvi, A.V., Ngo, L.T., Sherry, R.M., Phan, G.Q., et al. (2015). Tumor-Infiltrating Lymphocytes Genetically Engineered with an Inducible Gene Encoding Interleukin-12 for the Immunotherapy of Metastatic Melanoma. Clin. Cancer Res. *21*, 2278–2288. https://doi.org/10.1158/1078-0432.CCR-14-2085.

71. Li, X., Meng, B., Zhang, Z., Wei, L., Chang, W., Wang, Y., Zhang, K., Li, T., and Lu, K. (2025). qPrimerDB 2.0: an updated comprehensive gene-specific qPCR primer database for 1172 organisms. Nucleic Acids Res. *53*, D205–D210. https://doi.org/10.1093/nar/gkae684.

72. Wang, N.B., Lende-Dorn, B.A., Adewumi, H.O., Beitz, A.M., Han, P., O'Shea, T.M., and Galloway, K.E. (2023). Proliferation history and transcription factor levels drive direct conversion. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.26.568736.

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human Peripheral Blood Leukopak, Fresh	STEMCELL Technologies	Cat#70500
Primary Rat Cortex Neurons, Sprague Dawley	Thermo Scientific	Cat#A36511
Chemicals, peptides, and recombinant proteins		
B-27™ Supplement (50X), serum free	Thermo Scientific	Cat#17-504-044
Diethylaminoethyl-Dextran	Sigma-Aldrich	Cat#D9885
DMEM, High Glucose, with L-Glutamine, with Sodium Pyruvate	Genesee Scientific	Cat#25-500
DNase	Worthington Biochemical	Cat#9003-98-9
DNase I-XT	New England Biolabs	Cat#M0570
Doxycycline hydrochloride	Sigma-Aldrich	Cat#D3447
Dulbecco's Modified Eagle's Medium (DMEM)/ Ham's F-12 1:1, With L-glutamine	Corning	Cat#10-090-CV
Dynabeads™ Human T-Activator CD3/CD28 for T Cell Expansion and Activation	Thermo Fisher Scientific	Cat#11131D
Fetal Bovine Serum	Genesee Scientific	Cat#25-514H
FuGENE® HD Transfection Reagent	FuGENE	Cat#HD-1000
Gelatin	Sigma-Aldrich	Cat#G1890
Geltrex™ LDEV-Free, hESC-Qualified, Reduced Growth Factor Basement Membrane Matrix	Thermo Fisher Scientific	Cat#A1413302
Gentle Cell Dissociation Reagent	STEMCELL Technologies	Cat#100-1077
GlutaMAX™ Supplement	Thermo Fisher Scientific	Cat#35050061
HEPES	Sigma-Aldrich	Cat#H3375
Hexadimethrine bromide	Sigma-Aldrich	Cat#H9268
IL-2	R&D Systems	Cat#202-IL-050
KAPA SYBR® FAST qPCR Kit Master Mix (2X) Universal	Kapa Biosystems	Cat#KK4600
KnockOut™ DMEM	Fisher Scientific	Cat#10-829-01
Laminin, Mouse	Corning	Cat#354232
mTeSR™ Plus	STEMCELL Technologies	Cat#100-1130
Neurobasal™ Medium	Thermo Scientific	Cat#21103049
Opti-MEM™ I Reduced Serum Medium	Thermo Fisher Scientific	Cat#31985070
Papain	Worthington Biochemical	Cat#9001-73-4
Penicillin Streptomycin (10,000 U/mL)	Fisher Scientific	Cat#15-140-122
Phosphate buffered saline	Sigma-Aldrich	Cat#P4417
Poly-D-Lysine	Thermo Scientific	Cat#A3890401
Polyethyleneimine, linear	Fisher Scientific	Cat#AA4389603
Puromycin	Invivogen	Cat#ant-pr-1
ReLeSR™	STEMCELL Technologies	Cat#100-0484
RPMI 1640 Medium	Thermo Fisher Scientific	Cat#11875093
Trypsin-EDTA, 0.25% 1X without Calcium and Magnesium	Genesee Scientific	Cat#25-510
Y-27632 dihydrochloride	Millipore Sigma	Cat#Y0503
Critical commercial assays		
EasySep™ Direct Human PBMC Isolation Kit	STEMCELL Technologies	Cat#19654
EasySep™ Human CD8+ T Cell Enrichment Kit	STEMCELL Technologies	Cat#19053
LIVE/DEAD™ Fixable Near-IR Dead Cell Stain Kit	Thermo Fisher Scientific	Cat#L34976

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
LIVE/DEAD™ Fixable Violet Dead Cell Stain Kit	Thermo Fisher Scientific	Cat#L34964
Monarch Total RNA Miniprep Kit	New England Biolabs	Cat#T2010
ProtoScript First Strand cDNA Synthesis Kit	New England Biolabs	Cat#E6300
Deposited data		
Raw data, analyzed data, and modeling simulations	This work	https://doi.org/10.5281/zenodo.14946133
Experimental models: Cell lines		
Human: HEK293T	ATCC	Cat#CRL-3216; RRID:CVCL_0063
Human: Human iPSC Line (Episomal, HFF)	Alstem	Cat#iPS11
Human: Lenti-X™ 293T Cell Line	Takara Bio	Cat#632180; RRID:CVCL_4401
Experimental models: Organisms/strains		
Mouse: C57BL/6J	The Jackson Laboratory	Cat#000664; RRID:IMSR_JAX:000664
Oligonucleotides		
Primers for qPCR, see Table S7	qPrimerDB[71]	N/A
Recombinant DNA		
Plasmids constructed in this work, see Table S5 and deposited plasmid sequences	This work	Addgene #235253– 235327; DOI: https://doi.org/10.5281/zenodo.14946133
pMD2.G (lentiviral envelope plasmid)	Addgene	Addgene #12259
psPAX2 (lentiviral packaging plasmid)	Addgene	Addgene #12260
pRJ0084 (EF1α-hyperactive.PiggyBac-syn.pA)	Gift from the Weiss lab	N/A
Software and algorithms		
Code for data analysis, modeling, and figure generation	This work	https://github.com/GallowayLabMIT/ComMAND; DOI: https://doi.org/10.5281/zenodo.14962012
Python 3.12	Python Software Foundation	https://www.python.org; RRID:SCR_008394
Julia 1.10	The Julia Programming Language	https://julialang.org/; RRID:SCR_021666
Mathematica 14.0	Wolfram	https://www.wolfram.com/mathematica/; RRID:SCR_014448
FlowJo v10	BD	https://www.flowjo.com/solutions/flowjo; RRID:SCR_008520
Adobe Illustrator	Adobe Systems	http://www.adobe.com/products/illustrator.html; RRID:SCR_010279

EXPERIMENTAL MODEL DETAILS

HEK293T cells

HEK293T cells (ATCC, CRL-3216) and Lenti-X HEK293T cells (Takara Bio, 632180) were cultured using DMEM (Genesee Scientific, 25-500) plus 10% FBS (Genesee Scientific, 25-514H) on plates with 0.1% gelatin coating (Sigma-Aldrich, G1890-100G) and incubated at 37° C with 5% CO2. For routine passaging, seeding, and preparation for flow cytometry, HEK293T cells were dissociated using 0.25% Trypsin-EDTA (Genesee Scientific, 25-510) diluted in PBS (Sigma-Aldrich, P4417-100TAB) for four minutes then quenched with an equal volume of DMEM + 10% FBS. Cells were passaged every three to four days to maintain consistent growth before experiments.

Human induced pluripotent stem cells

iPS11 cells (Alstem, iPS11) were cultured using mTeSR™ Plus (STEMCELL Technologies, 100-1130) on plates coated with Geltrex™ (Thermo Fisher Scientific, A1413302) and incubated at 37° C with 5% CO2. For passaging, iPS11 cells were dissociated in clumps using ReLeSR™ (STEMCELL Technologies, 100-0484) according to manufacturer's instructions. For experiments, iPS11 cells were dissociated as single cells using Gentle Cell Dissociation Reagent (STEMCELL Technologies, 100-1077) according to manufacturer's instructions and counted using a hemocytometer. Cells were plated in mTeSR™ Plus with 10 μM ROCK inhibitor (Millipore Sigma, Y0503-5MG) and 100 U/mL penicillin-streptomycin (Fisher Scientific, 15-140-122).

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

Primary mouse embryonic fibroblasts

Primary mouse embryonic fibroblasts were isolated as described in Wang et al.[72] C57BL/6 mice were mated, and embryos were harvested at E12.5-E14.5 under a dissection scope. Embryo heads and internal organs were removed, then razors were used to break up the tissue with the addition of 0.25% Trypsin-EDTA. One or two embryos were processed simultaneously. After five minutes, the solution was quenched with DMEM + 10% FBS, spun down, and resuspended in fresh Trypsin-EDTA. After triturating, the solution was again quenched DMEM + 10% FBS and spun down. The resulting cells were resuspended in DMEM + 10% FBS and passed through a 40-μm filter, then plated on 10-cm dishes coated with 0.1% gelatin, one dish per embryo (passage 0). Cells were incubated at 37° C with 5% CO2. Once cells reached ~80% confluence (after two to three days), they were dissociated using 0.25% Trypsin-EDTA and passaged 1:3 onto fresh, gelatin-coated 10-cm dishes (passage 1). During passaging, a subset of cells were removed, expanded, and tested for mycoplasma. Once confluent (after two to four days), cells were dissociated using 0.25% Trypsin-EDTA, cryopreserved in 10% DMSO + 90% FBS, and stored in liquid nitrogen. For experiments, vials of cryopreserved passage 1 cells were thawed into DMEM + 10% FBS in T75 flasks coated with 0.1% gelatin and allowed to recover for one to two days before seeding.

Primary rat cortical neurons

Rat cortical neurons (Thermo Scientific, A36511) were recovered from cryopreservation following manufacturer's instructions. A 48-well plate was coated first with 0.1% gelatin followed by 10 mg/mL laminin (Corning, 354232), or coated with 4.5 μg/cm2 poly-D-lysine (Thermo Scientific, A3890401). Thawed neurons were counted and plated at 80,000 cells per well in Neurobasal Medium (Thermo Scientific, 21103049) supplemented with 0.5 mM GlutaMAX™ Supplement (Thermo Fisher Scientific, 35050061) and 20 mL/L B-27 (Thermo Scientific, 17504044). Cells were cultured at 37° C with 5% CO2. Six hours after plating, a half-media change was performed, where half the media volume was removed and replaced with fresh, pre-warmed media. After this point, a half-media change was performed every day. Cells were allowed to recover for two to seven days before use in experiments.

Primary human T cells

Peripheral blood mononuclear cells from healthy donors were purified from a leukopak (STEMCELL Technologies, 70500) using an EasySep Direct Human PBMC Isolation Kit (STEMCELL Technologies, 19654) according to manufacturer's instructions. Primary CD8+ T cells were isolated using EasySep Human CD8+ T cell Enrichment Kits (STEMCELL Technologies, 19053) and cultured in RPMI-1640 (ATCC, 30-2001) supplemented with 10% FBS, penicillin-streptomycin, and 30 IU/mL recombinant human IL-2 (R&D Systems, 202-IL-050). Cells were incubated at 37° C with 5% CO2.

METHOD DETAILS

Cloning

The sequences for FF3, FF4, FF5, and FF6 microRNA hairpins and cognate target sites were obtained from Leisner et al.[40] To clone the microRNAs, oligonucleotides were ordered from Azenta/Genewiz, phosphorylated, annealed, and ligated into a microRNA scaffold digested by XhoI and EcoRI. The miR-30a scaffold was PCR amplified from Addgene #25748, and the miR-E scaffold was ordered as a gBlock from Azenta/Genewiz based on the sequence in Nissim et al.[42] The intronic microRNAs were then introduced into *mRuby2*, *mGreenLantern*, and *egfp* coding sequences at a 5′-AGGT-3′ site via Gibson assembly using Hifi DNA Assembly Master Mix (NEB, M5520) according to manufacturer's instructions.

To clone the microRNA target sites, oligonucleotides were ordered from Azenta/Genewiz, phosphorylated, annealed, and ligated into a "part vector" backbone digested by HindIII and AvrII.

The therapeutically relevant genes *Fmr1* and *FXN* were ordered from Addgene (#87929 and #23620) and cloned into "part vectors" via Gibson assembly using Hifi DNA Assembly Master Mix according to manufacturer's instructions.

All plasmids for transfection were constructed via BsaI (NEB, R3733L) Golden Gate cloning using "part vectors" corresponding to promoter, coding sequence, UTR, and polyadenylation signal. Sources for additional parts can be found in Peterman et al.[47] These single transcriptional units were then combined with a backbone vector and a 200-bp spacer via PaqCI (NEB, R0745L) Golden Gate cloning to construct the lentiviral and PiggyBac integration vectors. The lentiviral backbone is a third-generation vector derived from Addgene #17297. The PiggyBac backbone was derived from Addgene #63800.

HEK293T transfection

For transfection experiments, HEK293T cells were counted using a hemocytometer and plated at a density of 25,000-35,000 cells per well in a 96-well plate 24 hours before transfection. Transfection was performed using linear polyethylenimine, PEI (Fisher Scientific, AA4389603). Transfection mixes were prepared using a ratio of 4 μg PEI to 1 μg DNA. First, a master mix of PEI and KnockOut™ DMEM (Fisher Scientific, 10-829-018) was prepared and incubated for a minimum of ten minutes. This mixture was then added to DNA mixes containing 112.5 ng of output plasmid and 56.25 ng of marker plasmid per well. These conditions mixes were incubated for an additional 10 to 15 minutes and then added on top of the growth media in the 96-well plate. After 24 hours, media was replaced with fresh DMEM + 10% FBS. At two days post-transfection, cells were prepared for flow cytometry by dissociating with Trypsin-EDTA. After centrifuging at 500x*g* for five minutes, cells were resuspended in PBS and transferred to a v-bottom plate for flow cytometry. Biological replicates represent independent transfections ($n \geq 3$).

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

For the plasmid titration experiment, the "standard" (1x) dose of output plasmid was 112.5 ng per well. Lower dosages included a "filler" plasmid expressing a different fluorescent protein, such that the total amount of output and filler plasmid remained constant at 112.5 ng per well. All conditions also contained 56.25 ng of marker plasmid per well.

PiggyBac integration

For PiggyBac integration, 50,000 HEK293T cells were plated per well of a 24-well plate coated with 0.1% gelatin. The following day, 315 ng of the integration vector, 135 ng of a plasmid expressing the hyperactive PiggyBac transposase (pRJ0084, gift from the Weiss lab), and 135 ng of a transfection reporter plasmid per well were co-transfected using PEI. The day after transfection, media was replaced with DMEM + 10% FBS plus 1 μg/mL puromycin (InvivoGen, ant-pr-1). Media was replaced with puromycin-containing media for two more days. On the fourth day post-transfection, media was changed to fresh DMEM + 10% FBS without puromycin. The following day, cells were passaged to 6-well plates.

Two days later, cells were prepared for fluorescence-activated cell sorting (FACS) by dissociating with Trypsin-EDTA. Cells were resuspended in fresh DMEM + 10% FBS and sorted into polyclonal populations on a Sony MA900 Cell Sorter at the Koch Institute Flow Cytometry Core. Cells expressing the marker and/or output genes were collected using gates set manually based on an untransfected population. The total numbers of single cells sorted per condition were ∼10,000 (replicate 1), ∼50,000 (replicate 2), and ∼200,000 (replicate 3). After sorting, cells were transferred to plates coated with 0.1% gelatin and allowed to recover for several days in DMEM + 10% FBS with penicillin-streptomycin. Once confluent, cells were dissociated with Trypsin-EDTA. A fraction of the cells were passaged to a new plate, and the remaining cells were assessed via flow cytometry. Several days later, cells were again passaged and assessed via flow cytometry. Biological replicates represenT cells analyzed from two passages each of three separate integrations ($n = 6$).

iPSC transfection

For transfection experiments, 15,000 iPS11 cells were plated per well in 96-well plates 48 hours before transfection. 24 hours before transfection, media containing ROCK inhibitor was removed and replaced with fresh mTeSR™ Plus with penicillin-streptomycin. On the day of transfection, the media was changed to Opti-MEM™ (ThermoFisher Scientific, 31985062) and transfection mixes were prepared with FuGENE® HD (FuGENE, HD-1000) according to the manufacturer's instructions using a ratio of 3 μL reagent to 1 μg DNA. Each well was transfected with 66 ng output plasmid and 33 ng marker plasmid. Four hours after transfection, mTeSR™ Plus with penicillin-streptomycin was added to the wells. 24 hours after transfection, the media was changed to mTeSR™ Plus with penicillin-streptomycin. Two days post-transfection, cells were detached from wells using Gentle Cell Dissocation Reagent and centrifuged at 500x*g* for five minutes. Cells were resuspended in PBS and transferred to a v-bottom plate for flow cytometry. Biological replicates represent independent transfections ($n = 3$).

Lentivirus production

Lenti-X HEK293T cells were seeded at one million cells per well in 6-well plates coated with 0.1% gelatin. The following day (day one), 1 μg of the third-generation lentiviral expression plasmid (containing the circuit), 1 μg of the packaging plasmid (psPAX2, Addgene #12260), and 2 μg of the envelope plasmid (pMD2.G, Addgene #12259) per well were co-transfected using PEI. After six hours, the media was replaced with 1.25 mL of DMEM + 10% FBS with 25 mM HEPES (Sigma-Aldrich, H3375). On the following day (day two), the media was collected, stored at 4° C, and replaced with HEPES-buffered DMEM + 10% FBS. On day three, the media was again collected. The collected media was filtered through a 0.45-μm PES filter.

To the filtered virus-containing media, Lenti-X Concentrator (Takara Bio, 631232) was added in a 3 parts media: 1 part concentrator volume ratio, mixed gently, and left overnight at 4° C. On day four, the media was centrifuged at 1500x*g* at 4° C for 45 minutes. The supernatant was aspirated, and the resulting pellet was resuspended to a total volume of 200 μL in HEPES-buffered DMEM + 10% FBS. Virus was used immediately or stored at -80° C.

For transductions of primary rat cortical neurons, primary human T cells, and human iPSCs, virus was produced as above at a 10-cm dish scale, using 6 μg/dish of the lentiviral expression plasmid, 6 μg/dish of the packaging plasmid, and 12 μg/dish of the envelope plasmid. On successive days, 6.5 mL of HEPES-buffered DMEM + 10% FBS per dish was collected, and the final pellet was resuspended to a total volume of 500-600 μL. All other virus production steps were the same as for the 6-well scale.

Lentiviral transduction

Of HEK293T cells

On the day of transduction, HEK293T cells were dissociated using Trypsin-EDTA, counted using a hemocytometer, and diluted to a concentration of 20,000 cells per well in DMEM + 10% FBS. Cells were combined with 5 μg/mL polybrene (hexadimethrine bromide, Sigma-Aldrich, H9268-5G) and either a constant amount of virus (1.0 μL per well) or a two-fold serial dilution of virus to compute viral titer (highest concentration: 1.0 μL concentrated virus per well). Additional DMEM + 10% FBS was added for a total volume of 100 μL per well. The resulting cell, polybrene, and virus mixture was plated onto 96-well plates coated with 0.1% gelatin.

Six hours later, the media was replaced with fresh DMEM + 10% FBS containing 1 μg/mL doxycycline (dox, Sigma-Aldrich, D3447-500MG). The following day, the media was replaced with fresh dox-containing media. Cells were prepared for flow cytometry three days later (four days post-transduction) by dissociating with Trypsin-EDTA. After centrifuging at 500x*g* for five minutes, cells were

resuspended in PBS and transferred to a v-bottom plate for flow cytometry. Biological replicates were obtained by using six different batches of virus ($n = 6$).

For transductions with vectors expressing the *FXN* and *Fmr1* circuits, HEK293T cells were transduced at an MOI of 1. Biological replicates for flow cytometry experiments were obtained by using three different batches of virus ($n = 3$). For the RT-qPCR analysis, 500,000 cells per well were transduced and plated on 12-well plates coated with 0.1% gelatin. One and three days post-transduction, media was replaced fresh DMEM + 10% FBS with 1 μg/mL doxycycline. Four days post-transduction, cells were dissociated with Trypsin-EDTA for RNA isolation. Biological replicates were obtained by using three batches of virus in four different transductions ($n = 5$).

Of primary mouse embryonic fibroblasts

After recovery from thawing, cells were dissociated using 0.25% Trypsin-EDTA diluted in PBS and counted using a hemocytometer. Cells were plated onto 96-well plates coated with 0.1% gelatin at 10,000 cells per well. The following day, the media was replaced with DMEM + 10% FBS plus 5 μg/mL polybrene and either a constant amount of virus (10.0 μL per well) or a two-fold serial dilution of virus (highest concentration: 20 μL concentrated virus per well). The cells were spinfected by centrifuging at 1500x*g* for 90 minutes at 32° C.

Six hours after spinfection, the media was replaced with fresh DMEM + 10% FBS containing 1 μg/mL doxycycline. The cells were changed to fresh dox-containing media the following day. Cells were prepared for flow cytometry three days later (four days post-transduction) by dissociating with Trypsin-EDTA. After centrifuging at 500x*g* for five minutes, cells were resuspended in PBS and transferred to a v-bottom plate for flow cytometry. Biological replicates were obtained by using cells from four separate isolations with six different batches of virus ($n = 9$).

Of primary rat cortical neurons

After recovering for two to seven days post-thawing, the cells were transduced with virus added at an MOI of 1, 5, or 7 with 5 μg/mL of polybrene. MOI was estimated from viral titers computed from the primary mouse embryonic fibroblast transduction. Six hours after transduction, the cells were half-media changed twice into Neurobasal Medium supplemented with 0.5 mM GlutaMAX™ Supplement, 20 mL/L B-27, and 1 μg/mL doxycycline. This double media change was performed again on the two subsequent days. Three to seven days after transduction, the cells were imaged and prepared for flow cytometry. The neurons were dissociated gently using DMEM/F12 (Corning, 10-090-CV) with 17 U/mL DNase (Worthington Biochemical, LK003172) and 167 U/mL papain (Worthington Biochemical, LK003178) at 37◦ C for 20-40 minutes. The resulting clusters of neurons were centrifuged at 400x*g* for four minutes, resuspended in PBS, and transferred to a v-bottom plate for flow cytometry. Biological replicates were obtained by using four separate vials of cells with four different batches of virus ($n = 4$).

Of primary human T cells

Following isolation, cells were activated with DynaBeads Human T-Activator CD3/CD28 (Thermo Fisher, 11131D) at a 1:1 cell-to-bead ratio, plated in 24-well plates at 1 million cells per mL, and cultured overnight at 37° C and 5% CO2. For transduction the following day, cells were re-plated at 1 million cells per mL (500,000 cells per condition), and concentrated lentiviruses were added to cells at an MOI of 1, 5, or 7 with 8 μg/mL diethylaminoethyl-dextran (Sigma-Aldrich, D9885). MOI was estimated from viral titers computed from the primary mouse embryonic fibroblast transduction. On day four post-isolation, DynaBeads were removed according to manufacturer's instructions and cells were expanded in fresh IL-2-containing media. On day 7, cells were treated with 1 ug/mL doxycycline. After 24 hours, cells were stained with LIVE/DEAD fixable near-IR or violet dye (ThermoFisher, L34976 or L34964) according to manufacturer's instructions and analyzed using a Cytoflex S flow cytometer. Biological replicates were obtained by using cells from four separate donors with three batches of virus ($n = 4$).

Of human induced pluripotent stem cells

For lentiviral transductions, iPS11 cells were plated in mTeSR™ Plus at 15,000 cells per well in 96-well plates coated with Geltrex™. The next day, the media was replaced with concentrated lentivirus diluted in fresh mTeSR™ Plus with 5 μg/mL of polybrene. To compute viral titer, cells were transduced with a two-fold serial dilution of each virus. For subsequent experiments, cells were transduced at an MOI of 0.3. The next day, media was replaced with fresh mTeSR™ Plus containing 1 μg/mL doxycycline. Two days later, cells were detached using Gentle Cell Dissocation Reagent and centrifuged at 500x*g* for five minutes. Cells were resuspended in PBS and transferred to a v-bottom plate for flow cytometry. Biological replicates were obtained using three batches of virus ($n = 3$).

RT-qPCR

Bulk RNA samples were isolated for RT-qPCR analysis using the Monarch Total RNA Miniprep Kit (New England Biolabs, T2010). An additional DNase I treatment (New England Biolabs, M0570) was performed, and RNA samples were eluted into 50 μL of nuclease-free water. For cDNA synthesis, 6 μL of each RNA sample was used with the ProtoScript First Strand cDNA Synthesis Kit (New England Biolabs, E6300). To capture only polyadenylated RNA, oligo-dT primers were used for cDNA synthesis, and a control reaction without the reverse transcriptase enzyme was included. cDNA samples were stored at -20° C until qPCR.

qPCR was performed using a Roche LightCycler 480 at the MIT BioMicro Center. Each reaction mix was run in quadruplicate with the following components: 2.5 μL KAPA SYBR FAST qPCR Master Mix (2X) Universal (Kapa Biosystems, KK4600), 0.5 μL each 2 μM forward and reverse primers, 0.5 μL cDNA product, and 1.5 μL nuclease-free water. Primer sequences were obtained from qPrimerDB[71] and are reported in Table S7. C_t values were called using the "High Sensitivity" analysis mode.

Fluorescent imaging

Images of the primary rat cortical neurons were taken three days post-transduction on a Nikon ECLIPSE Ti2-E fluorescent microscope. Images of HEK293T cells transduced with circuits regulating *Fmr1* were taken three days post-transduction on a Keyence BZ-X800 fluorescent microscope.

Steady-state modeling

We modeled the ComMAND system using the set of reactions shown in Figure 3A. Reactions were modeled as mass-action reactions with the stoichiometries and rate constants listed in Table S1. In this system, the microRNA is encoded in an intron within the output gene. These components are transcribed as a single immature primary transcript ($mRNA_i$, "Transcription" reaction), which is then spliced into pri-miRNA (pri) and mature output mRNA (mRNA, "Splicing" reaction) and which can be degraded ("Immature degradation" reaction). The pri-miRNA is processed by endogenous Drosha into pre-miRNA (pre, "Drosha" reaction), which in turn is processed by endogenous Dicer into mature microRNA (miR, "Dicer" reaction). Both the mature microRNA and the mature output mRNA can be degraded ("microRNA degradation" and "mRNA degradation" reactions). The mature microRNA can be irreversibly loaded onto RISC ("RISC loading" reaction), and RISC-bound microRNA can be degraded ("Loaded degradation" reaction). The RISC-bound microRNA can reversibly bind output mRNA ("RISC-mRNA binding" and "RISC-mRNA unbinding" reactions). From this complex, the output mRNA is consumed by RISC knockdown, while the RISC-bound microRNA is recycled ("mRNA knockdown"). Both the free output mRNA and the RISC-bound mRNA can be translated ("Translation" and "Bound translation" reactions), and the resulting output protein can be degraded ("Protein degradation" reaction). Together, DNA produces output protein through the activity of the circuit.

To analyze the behavior of the dual-transcript, dual-vector, and non-intronic dual-vector systems presented in Figure 4, we replaced the first two reactions ("Transcription" and "Splicing") with the reactions listed in Table S2. In the two-gene systems, microRNA and output mRNA are encoded on separate transcripts with or without introns, leading to different reactions as described below. To understand these circuits, we computed an analytical steady-state solution for each system.

ComMAND single-transcript system

For the single-transcript case, we have the following ODEs defining the production of the pre-miRNA (pre) from the pri-miRNA (pri) and the immature primary transcript ($mRNA_i$):

$$\frac{d\mathrm{mRNA}_i(t)}{dt} = -r_{\mathrm{splicing}}\mathrm{mRNA}_i(t) + c_{\mathrm{regulated}}\alpha_{\mathrm{RNA}} - \mathrm{mRNA}_i(t)\delta_{\mathrm{immature}}$$

$$\frac{d\mathrm{pri}(t)}{dt} = -r_{\mathrm{drosha}}\mathrm{pri}(t) + r_{\mathrm{splicing}}\mathrm{mRNA}_i(t)$$

$$\frac{d\mathrm{pre}(t)}{dt} = -r_{\mathrm{dicer}}\mathrm{pre}(t) + r_{\mathrm{drosha}}\mathrm{pri}(t)$$

At steady state, this can be solved:

$$\mathrm{mRNA}_i = \frac{c_{\mathrm{regulated}}\alpha_{\mathrm{RNA}}}{\delta_{\mathrm{immature}} + r_{\mathrm{splicing}}} \quad \text{(Equation 1)}$$

$$\mathrm{pri} = \frac{r_{\mathrm{splicing}}}{r_{\mathrm{drosha}}} \cdot \frac{c_{\mathrm{regulated}}\alpha_{\mathrm{RNA}}}{\delta_{\mathrm{immature}} + r_{\mathrm{splicing}}} \quad \text{(Equation 2)}$$

$$\mathrm{pre} = \frac{r_{\mathrm{splicing}}}{r_{\mathrm{dicer}}} \cdot \frac{c_{\mathrm{regulated}}\alpha_{\mathrm{RNA}}}{\delta_{\mathrm{immature}} + r_{\mathrm{splicing}}} \quad \text{(Equation 3)}$$

Additionally, we can define the effective production rate of the output mRNA:

$$\tilde{r}_{\mathrm{controlled}} = r_{\mathrm{splicing}}\mathrm{mRNA}_i \quad \text{(Equation 4)}$$

The effective production rate differs across the single-transcript and two-gene systems, so it is helpful to define it separately and then later substitute it into the solution for the shared system.

Dual-transcript system

In the dual-transcript case, the microRNA is encoded in an intron within a gene that is not the output gene. The immature primary transcript produces mature microRNA through the same Drosha and Dicer processing steps as in the single-transcript system, so the steady-state solution is the same as in Equations 2 and 3. These reactions produce a second mature mRNA transcript ($mRNA_{ignore}$) that does not interact with the rest of the system and thus can be ignored.

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

Additionally, here the output mRNA does not contain an intron and thus is not spliced. The effective production rate is instead directly dependent on the copy number and the transcription rate:

$$\tilde{r}_{\text{controlled}} = c_{\text{regulated}}\alpha_{\text{RNA}} \quad \text{(Equation 5)}$$

Dual-vector system

In the dual-vector case, the steady-state solution for the pri-miRNA and pre-miRNA is also the same as in Equations 2 and 3. However, the output mRNA and microRNA are encoded on separate DNA vectors. Therefore, the effective production rate of the output mRNA depends on a different copy number, $c_{\text{unregulated}}$:

$$\tilde{r}_{\text{controlled}} = c_{\text{unregulated}}\alpha_{\text{RNA}} \quad \text{(Equation 6)}$$

Non-intronic dual-vector system

In this simplest case, there is no splicing involved in the production of the microRNA, such as when expression is driven from a RNA Polii promoter (e.g., U6). Instead, microRNA is produced directly by transcription of the U6 vector, which has copy number c_{U6} and transcription rate α_{U6}:

$$\frac{\text{dpri}(t)}{dt} = -r_{\text{drosha}}\text{pri}(t) + c_{U6}\alpha_{U6}$$

$$\frac{\text{dpre}(t)}{dt} = -r_{\text{dicer}}\text{pre}(t) + r_{\text{drosha}}\text{pri}(t)$$

which can be solved at steady state:

$$\text{pri} = \frac{c_{U6}\alpha_{U6}}{r_{\text{drosha}}} \quad \text{(Equation 7)}$$

$$\text{pre} = \frac{c_{U6}\alpha_{U6}}{r_{\text{dicer}}} \quad \text{(Equation 8)}$$

The effective production rate of the output mRNA is the same as in the dual-transcript case:

$$\tilde{r}_{\text{controlled}} = c_{\text{regulated}}\alpha_{\text{RNA}} \quad \text{(Equation 9)}$$

Shared system

All implementations of the microRNA-based iFFL share a common set of reactions, as listed below the line in Table S1 (“Immature degradation” and on). We can think of the set of common reactions as a network that takes in two species—the *output mRNA* and the *pre-miRNA*—and returns a final protein (and mRNA) concentration.

The specific form of the output mRNA term differs across systems, so we use the effective production rate $\tilde{r}_{\text{controlled}}$ computed earlier. These reactions are:

$$\frac{\text{miR}(t)}{dt} = r_{\text{dicer}}\text{pre}(t) - \text{miR}(t)\delta_{\text{miRNA}} - k_{\text{miRNA,bind}}\text{miR}(t)R(t) \quad \text{(Equation 10)}$$

$$\frac{\text{dR}(t)}{dt} = k_{\text{miRNA,deg}}\text{RI}(t) - k_{\text{miRNA,bind}}\text{miR}(t)R(t) \quad \text{(Equation 11)}$$

$$\frac{\text{dRI}(t)}{dt} = +k_{\text{deg}}\text{RIM}(t) + k_{\text{mRNA,unbind}}\text{RIM}(t) - k_{\text{mRNA,bind}}\text{RI}(t)\text{mRNA}(t) - k_{\text{miRNA,deg}}\text{RI}(t) + k_{\text{miRNA,bind}}\text{miR}(t)\text{R}(t) \quad \text{(Equation 12)}$$

$$\frac{\text{dRIM}(t)}{dt} = -k_{\text{deg}}\text{RIM}(t) - k_{\text{mRNA,unbind}}\text{RIM}(t) + k_{\text{mRNA,bind}}\text{RI}(t)\text{mRNA}(t) \quad \text{(Equation 13)}$$

$$\frac{\text{dmRNA}(t)}{dt} = k_{\text{mRNA,unbind}}\text{RIM}(t) + \tilde{r}_{\text{controlled}} - \delta_{\text{mRNA}}\text{mRNA}(t) - k_{\text{mRNA,bind}}\text{RI}(t)\text{mRNA}(t) \quad \text{(Equation 14)}$$

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

$$\frac{dP(t)}{dt} = -\delta_p P(t) + \alpha_p \mathrm{mRNA}(t) + \zeta\alpha_p \mathrm{RIM}(t) \quad \text{(Equation 15)}$$

Immediately, we can solve Equations 10 and 15 at steady state to get:

$$\mathrm{miR} = \frac{r_{\mathrm{dicer}} \cdot \mathrm{pre}}{k_{\mathrm{miRNA,bind}} \cdot R + \delta_{\mathrm{miRNA}}} \quad \text{(Equation 16)}$$

$$P = \frac{\alpha_p}{\delta_p}(\mathrm{mRNA} + \zeta \cdot \mathrm{RIM}) \quad \text{(Equation 17)}$$

To solve the rest, note that equations Equations 11, 12, and 13 sum to zero at steady state; this means that we will need to introduce the RISC conservation equation and substitute to fully solve this system. We assume that the total amount of RISC, R_{tot}, remains constant:

$$R_{\mathrm{tot}} = R + \mathrm{RI} + \mathrm{RIM} \quad \text{(Equation 18)}$$

We can then solve for RI using Equation 11:

$$\mathrm{RI} = \frac{k_{\mathrm{miRNA,bind}} r_{\mathrm{dicer}} \cdot \mathrm{pre} \cdot R}{k_{\mathrm{miRNA,deg}}\left(\delta_{\mathrm{miRNA}} + k_{\mathrm{miRNA,bind}} \cdot R\right)} = \frac{R}{\kappa_1 + \kappa_2 R} \quad \text{(Equation 19)}$$

for

$$\kappa_1 = \frac{k_{\mathrm{miRNA,deg}}\delta_{\mathrm{miRNA}}}{k_{\mathrm{miRNA,bind}} r_{\mathrm{dicer}} \cdot \mathrm{pre}} \quad \kappa_2 = \frac{k_{\mathrm{miRNA,deg}} k_{\mathrm{miRNA,bind}}}{k_{\mathrm{miRNA,bind}} r_{\mathrm{dicer}} \cdot \mathrm{pre}}$$

Similarly, by plugging this result back in and using Equation 13, we can define:

$$\kappa_3 = \frac{\delta_{\mathrm{mRNA}}\delta_{\mathrm{miRNA}} k_{\mathrm{miRNA,deg}}\left(k_{\mathrm{deg}} + k_{\mathrm{mRNA,unbind}}\right)}{k_{\mathrm{mRNA,bind}} k_{\mathrm{miRNA,bind}} r_{\mathrm{dicer}} \cdot \mathrm{pre} \cdot \tilde{r}_{\mathrm{controlled}}}$$

$$\kappa_4 = \frac{\delta_{\mathrm{mRNA}} k_{\mathrm{miRNA,deg}}\left(k_{\mathrm{deg}} + k_{\mathrm{mRNA,unbind}}\right)}{k_{\mathrm{mRNA,bind}} r_{\mathrm{dicer}} \cdot \mathrm{pre} \cdot \tilde{r}_{\mathrm{controlled}}} + \frac{k_{\mathrm{deg}}}{\tilde{r}_{\mathrm{controlled}}}$$

so that

$$\mathrm{RIM} = \frac{R}{\kappa_3 + \kappa_4 R} \quad \text{(Equation 20)}$$

Then, the RISC mass closure equation can be written:

$$R_{\mathrm{tot}} = R + \frac{R}{\kappa_1 + \kappa_2 R} + \frac{R}{\kappa_3 + \kappa_4 R}$$

$$0 = (R - R_{\mathrm{tot}})(\kappa_1 + \kappa_2 R)(\kappa_3 + \kappa_4 R) + (\kappa_3 + \kappa_4 R)R + (\kappa_1 + \kappa_2 R)R$$

This is a cubic equation in R:

$$0 = (\kappa_2\kappa_4)R^3 + (\kappa_1\kappa_4 + \kappa_2\kappa_3 + \kappa_4 + \kappa_2 - R_{\mathrm{tot}}\kappa_2\kappa_4)R^2 + (\kappa_1\kappa_3 + \kappa_3 + \kappa_1 - R_{\mathrm{tot}}(\kappa_1\kappa_4 + \kappa_2\kappa_3))R - R_{\mathrm{tot}}\kappa_1\kappa_3 \quad \text{(Equation 21)}$$

While this represents an analytic steady-state solution for R, it is only tractable when solved numerically. For the parameter regimes chosen here, the cubic discriminant is positive, which means that there are three real roots. In practice—and at every parameter range included in this work—two of the roots are unphysical (negative), so we choose the positive root.

With this positive root, we can solve for the steady-state mRNA concentration:

$$\mathrm{mRNA} = \tilde{r}_{\mathrm{controlled}}\left(\delta_{\mathrm{mRNA}} + \frac{k_{\mathrm{deg}} k_{\mathrm{mRNA,bind}} k_{\mathrm{miRNA,bind}} r_{\mathrm{dicer}} \cdot \mathrm{pre} \cdot R}{k_{\mathrm{miRNA,deg}}\left(k_{\mathrm{deg}} + k_{\mathrm{mRNA,unbind}}\right)\left(\delta_{\mathrm{miRNA}} + k_{\mathrm{miRNA,bind}} \cdot R\right)}\right)^{-1} \quad \text{(Equation 22)}$$

By back-substituting this result and Equations 20 into Equation 17, we can compute the steady-state protein concentration. This gives the steady-state output protein values shown throughout this work.

Some of the equation algebra to find the steady-state solution was performed using Mathematica (14.0) and can be found in steady_state.nb. Parameter values used for the steady-state analysis are shown in Table S3. Solutions using these parameters were computed in Julia (1.10) in steady_state.jl. See data and code availability for access to these scripts.

Parameter sweeps

To assess the impact of the parameters on the model predictions, we solved the steady-state protein equation with different values of each parameter. Ten different values evenly distributed in logspace around the original value were chosen for each parameter, except

for the RISC-bound mRNA translation factor, ζ, where values were evenly distributed in linear space between zero and one. All other parameters were held constant at their base values (Table S3) for each parameter sweep. Solutions were computed in Julia in steady_state_all_param_sweep.jl (see data and code availability).

Slope calculation

To calculate the slope in the RISC-limited regime for the modeling predictions, we first defined a system for expression of the simplest unregulated gene, equivalent to the experimental marker gene. Namely, this system includes only transcription, translation, and degradation reactions, defined by the equations

$$\frac{\mathrm{dmRNA}(t)}{dt} = -\delta_{\mathrm{mRNA}}\mathrm{mRNA}(t) + c_{\mathrm{regulated}}\alpha_{\mathrm{RNA}}$$

$$\frac{dP(t)}{dt} = -\delta_p P(t) + \alpha_p \mathrm{mRNA}(t)$$

which can be solved at steady state:

$$P = \frac{c_{\mathrm{regulated}}\alpha_{\mathrm{RNA}}\alpha_p}{\delta_{\mathrm{mRNA}}\delta_p}$$

We then computed the slope of the copy number–protein curve for both the ComMAND output gene and the marker gene using the protein values at the two highest copy numbers for each simulation. Selecting these very large copy numbers ensured that we were firmly in the RISC-limited regime. We then normalized the ComMAND slope s by the marker slope, which is simply multiplication by a normalization factor:

$$\text{normalized slope} = \frac{\alpha_{\mathrm{RNA}}\alpha_p}{\delta_{\mathrm{mRNA}}\delta_p} \cdot s$$

This normalized slope calculation allows us to simulate the slope of the marker–output curve for the steady-state solution of the model.

Stochastic simulations

We performed stochastic simulations for both the single-transcript (ComMAND) and two-gene systems in order to better understand the effects of noise on protein expression. In particular, stochastic simulations are useful to investigate differences between the dual-transcript and dual-vector systems, which differ only in the copy number of the gene expressing the output mRNA. The steady-state analytical solutions including Equations 5 and 6 do not provide much insight into differences between the systems, since they do not account for effects of noise on copy number or transcription. Instead, we can use stochastic simulations where in the dual-vector case, the DNA copy number of the output gene is allowed to randomly vary from the that of the microRNA gene to simulate imperfect co-delivery of the two vectors. Furthermore, the stochastic simulations can account for noise in the production and processing steps for the microRNA and output mRNA, which are tightly coupled in the single-transcript case but not in the two-gene systems.

We modeled the systems at three different effective multiplicity of infections (MOIs), a common parameter in viral transductions that specifies the number of viral transducing units added per cell. In this context, the copy number of the transduced cell population can be modeled as a Poisson distribution with a mean equal to the MOI. To recapitulate this variation in delivery in our model, we chose DNA copy numbers for each simulation run from a Poisson distribution with one of three different means: 0.3, to simulate single-copy integration; 3, to simulate a low-copy regime; and 10, to simulate a high-copy regime (Figure S3C).

The resulting systems were simulated using a Gillespie algorithm in Julia, using the packages Cataylst.jl, DifferentialEquations.jl, and JumpProblems.jl. Parameter values used for the stochastic simulations are the same as for the steady-state analytical solution and are shown in Table S3. 10,000 simulation runs were performed for each set of conditions.

Stochastic mis-splicing simulations

To investigate the effects of mis-splicing on circuit performance, we analyzed scenarios where splicing of the microRNA-containing intron yields only the mRNA molecule (Model A) or only the microRNA molecule (Model B), as depicted in Figure S3F. Each of these scenarios was simulated as described above with the correct splicing reaction and incorrect splicing reaction occurring with equal probability (i.e., $r_{\mathrm{mRNA-only}} = r_{\mathrm{splicing}}$ and $r_{\mathrm{miRNA-only}} = r_{\mathrm{splicing}}$). Therefore, the results shown in Figures S3G and S3H represent the scenario where 50% of transcripts are spliced incorrectly.

QUANTIFICATION AND STATISTICAL ANALYSIS

Flow cytometry

For all experiments except those with primary human T cells, flow was performed with an Attune NxT flow cytometer. The primary human T cells were analyzed using a Cytoflex S flow cytometer. See Table S6 for channel mappings, lasers, filters, and voltage settings. Data were analyzed by selecting single cells using forward and side scatter gates in FlowJo (BD, v10). Raw data (fcs files), FlowJo workspaces (wsp files), and gated single cell populations (csv files) have been deposited at Zenodo (see key resources table).

Please cite this article in press as: Love et al., Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101269

Next, populations were gated for expressing cells in Python (3.12). For primary human T cell experiments, live cells were gated as the negative population of the live/dead dye using a manually drawn gate, and >96% of cells were viable. For transfections, cells were gated on marker expression based on the 99.9th percentile of the untransfected population within each biological replicate. For transductions, the gate for marker expression was drawn manually to best separate expressing cells from the untransduced population.

After gating the marker-expressing cells, populations were further analyzed in Python using the pandas and scipy packages. Throughout the paper, marker–output plots depict one representative biological replicate, where data are binned by marker expression into 10-20 equal-quantile groups per condition. Points represent geometric means of output expression for cells in each bin, and shaded regions represent this value multiplied or divided by the geometric standard deviation of the bin. Geometric mean and geometric standard deviation were used to decrease the impact of outlier points since the expression data is log-distributed. Bins are plotted at their median marker value. Expression values are in arbitrary units from a flow cytometer.

All other flow cytometry data plots show summary statistics of output gene expression, as calculated in Python. The plotted mean values use the geometric mean (arbitrary units), Std. refers to the standard deviation (arbitrary units), and CV is the coefficient of variation (unitless). The slope represents the slope of the least-squares regression line fitted to the binned marker–output points in log-space (arbitrary units). Points represent means of $n \geq 3$ biological replicates; see the relevant experimental sections in method details for more information about replicates. A few conditions contained outlier points with aberrantly low marker expression; these points were removed. Error bars show the 95% confidence interval.

Viral titer calculation

Lentiviral titer was calculated by analyzing the cells transduced with the two-fold serial dilution of each virus. From the flow cytometry data, the fraction of single cells expressing the marker gene was computed for each dilution. This data was fit to a Poisson distribution to calculate u, the number of transducing units (TUs) per cell per volume:

$$f = 1 - e^{-u \cdot v} \quad \text{(Equation 23)}$$

where f is the fraction infected and v is the volume of virus added. By multiplying u by the total number of cells per well at the time of transduction, the viral titer in TUs per volume is obtained. This value is specific to each batch of virus and transduced cell type, and it is used to calculate the volume of virus needed for a given multiplicity of infection (MOI), the ratio of viral TUs to cells in a transduction. The viral titers calculated with primary mouse embryonic fibroblasts were used to estimate MOIs for the transductions in primary rat cortical neurons and primary human T cells.

RT-qPCR analysis

To analyze the RT-qPCR data, technical replicates with no amplification ($C_t = 35$) were excluded. C_t values for each reaction were computed as the median of the four technical replicates. Reactions with C_t values at or above those for control reactions (lacking reverse transcriptase or cDNA) for each biological replicate were removed. Conditions (independent cDNA samples) with low *GAPDH* expression ($C_t \geq 21$) were also excluded. C_t values were then normalized to those for *GAPDH* for each condition and biological replicate (ΔC_t). Normalized C_t values were then further normalized to those for the base gene conditions within each biological replicate ($\Delta\Delta C_t$). Expression is reported as $2^{-\Delta\Delta C_t}$.

Statistics

All statistical tests are two-sided independent t-tests where ns $p > 0.05$ is not significant, and * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$ are significant.

Cell Systems, Volume *16*

Supplemental information

Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells

Kasey S. Love, Christopher P. Johnstone, Emma L. Peterman, Stephanie Gaglione, Michael E. Birnbaum, and Kate E. Galloway

Supplementary Figures

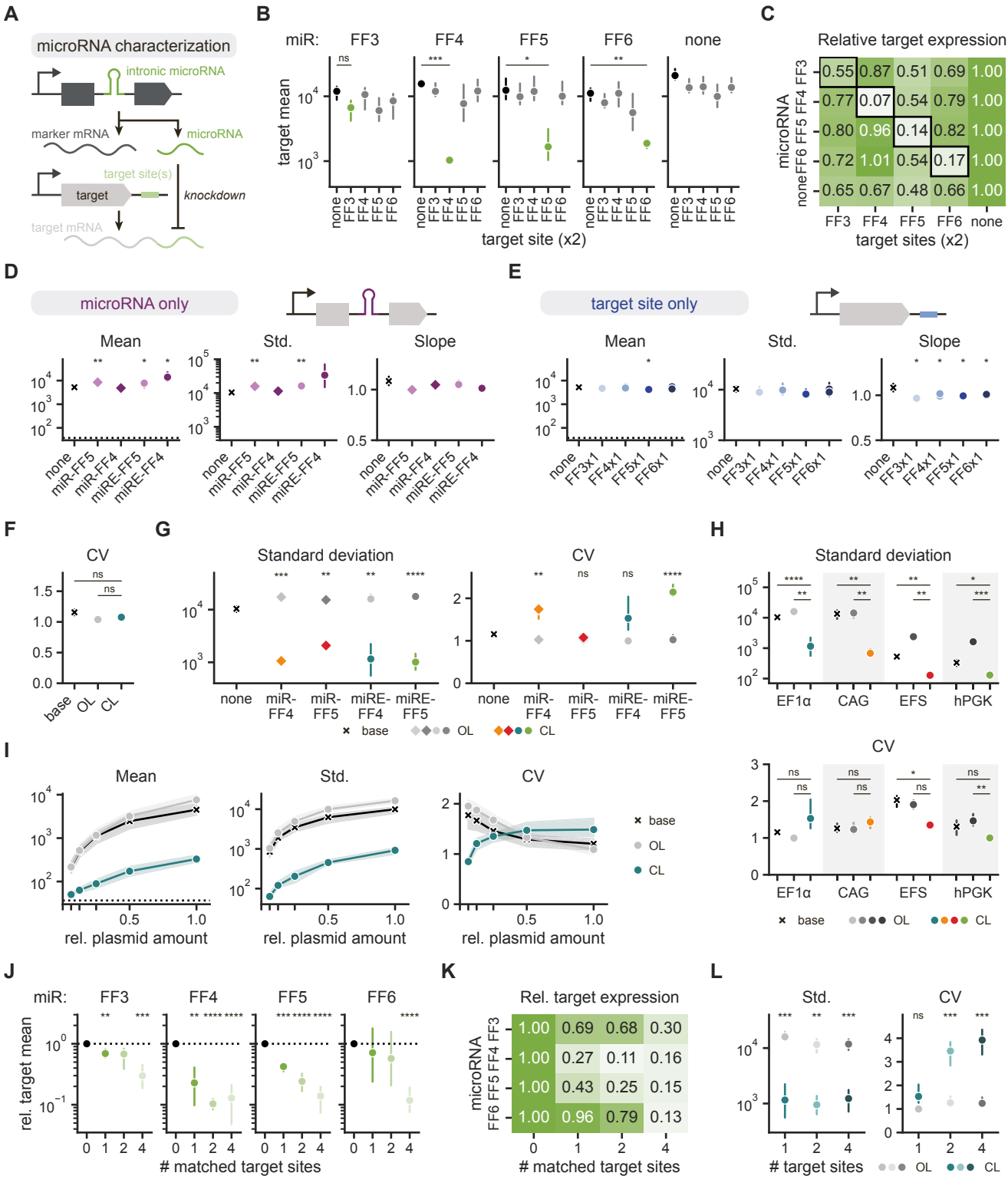


Figure S1. Separate microRNA and target site characterization and additional circuit tuning data, related to Figures 1 and 2.

A. Schematic of constructs for microRNA and target site characterization. A construct with a miR-30a–based intronic microRNA (miR-FF3, miR-FF4, miR-FF5, or miR-FF6) is transcribed, spliced, and processed into mature marker mRNA and mature microRNA. A co-delivered construct containing a target gene with 1, 2, or 4 target sites (FF3, FF4, FF5, or FF6 target sequences) is transcribed into mature target mRNA, which can be knocked down by the cognate microRNA.

B. Target gene expression in HEK293T cells co-transfected with microRNA and target site sequence combinations as shown in **A**. Each marker-microRNA construct was paired with a target gene containing two copies of the indicated target site sequence. Target gene geometric means are plotted for each condition.

C. Heatmap of relative target gene expression for conditions in **B**. Within each biological replicate, target gene expression levels were normalized for each microRNA (rows) to the no target site condition (rightmost column). Final values represent averages of $n = 3$ biological replicates. Black outlines indicate conditions with matched microRNA and target site sequences.

D, **E**. Summary statistics for HEK293T cells transfected with output genes containing only an intronic microRNA (**D**) or only a 3' UTR target site (**E**). All constructs were expressed by an EF1α promoter. Dashed lines represent geometric mean output levels for cells transfected only with the marker gene. Statistical tests compare conditions to the base gene (none) condition.

F. Coefficient of variation (CV) of output expression for populations in Figure 1E.

G. Output standard deviation and CV for populations in Figures 2A, 2B. Statistical tests compare CL and OL conditions, and CL and base conditions also have significantly different standard deviations.

H. Output standard deviation and CV for populations in Figure 2D.

I. Output mean, standard deviation, and CV as a function of plasmid dose for populations in Figure 2E. Plasmid amount is represented as a fraction relative to the standard dosage. Dashed line represents geometric mean output level for cells transfected only with the marker gene.

J. Target gene expression in HEK293T cells co-transfected with microRNA and target site sequence combinations as shown in **A**. Each marker-microRNA construct was paired with a target gene containing 0, 1, 2, or 4 copies of the matched target site sequence. Target gene geometric means were calculated for each condition, and these values were then normalized within biological replicates to the no target site (0) condition for each microRNA to obtain relative target expression. Dashed lines depict a relative target expression of 1. Statistical tests compare conditions to the no target site condition for each microRNA.

K. Heatmap of relative target gene expression for conditions in **J**. Target gene expression levels were normalized for each microRNA (rows) to the no target site condition (leftmost column) for each biological replicate. Final values represent averages of $n = 3$ biological replicates.

L. Output standard deviation and CV for populations in Figure 2F. Statistical tests compare CL and OL conditions, and CL and base conditions also have significantly different standard deviations.

The plotted mean values use the geometric mean. Std. refers to the standard deviation, and the slope represents the slope of the line fitted to the binned, log-transformed marker–output points. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. All unnormalized values are in arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, independent t-test. Unlabeled comparisons are not statistically significant.

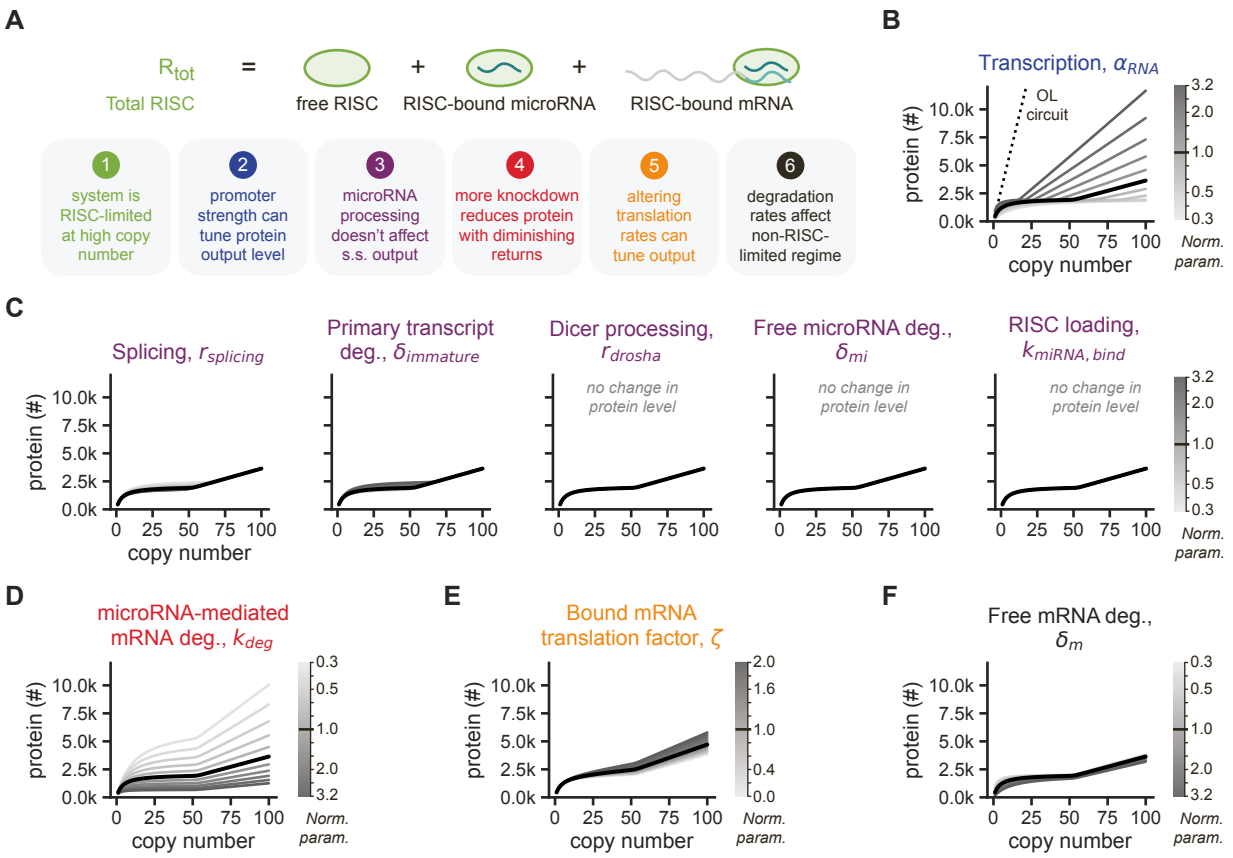


Figure S2. Additional model parameter sweeps, related to Figure 3.

A. Top: The total amount of RISC in the system, R_{tot}, is the sum of the amount of free RISC, RISC bound to microRNA, and RISC-microRNA bound to mRNA. The model assumes that the total amount of RISC remains constant. Bottom: The six design principles revealed by the model, as illustrated in Figure 3.

B. Output protein level (in number of molecules) as a function of DNA copy number, $c_{\text{regulated}}$, with a sweep of values of the transcription rate of the primary transcript, α_{RNA}. Chosen parameter values are evenly log-distributed over an order of magnitude centered on the original parameter value. Color bar depicts parameter values normalized to the original value on a $\log_{10}$ scale. The thick black line in this panel and subsequent panels represents the output of the model with the original parameter values. Dashed black line indicates the solution for an unregulated gene (OL circuit).

C, **D**. Output protein level (in number of molecules) as a function of DNA copy number with a sweep of the values of the following parameters: splicing rate, r_{splicing}; primary transcript degradation rate, δ_{immature}; Dicer processing rate, r_{dicer}; free microRNA degradation rate, δ_{mi}; microRNA loading in RISC, $k_{\text{miRNA,bind}}$; the microRNA-mediated mRNA degradation rate, k_{deg}. Chosen parameter values are evenly log-distributed over an order of magnitude centered on the original parameter value. Color bars depict parameter values normalized to the original value on a $\log_{10}$ scale. Output curves that do not change as the given parameter changes are noted on the plot.

E. Output protein level (in number of molecules) as a function of DNA copy number with a sweep of the RISC-bound mRNA translation factor, ζ. Chosen parameter values are linearly spaced between zero and one, where zero represents no translation of the bound mRNA and one represents translation at a rate equivalent to that of free mRNA. Color bar depicts parameter values normalized to the original value on a linear scale.

F. Output protein level (in number of molecules) as a function of DNA copy number with a sweep of values of the free mRNA degradation rate, δ_{m}. Chosen parameter values are evenly log-distributed over an order of magnitude centered on the original parameter value. Color bar depicts parameter values normalized to the original value on a $\log_{10}$ scale. See also model schematic in Figure 3A. Exact reactions, steady-state analysis, and description of parameter sweeps can be found in STAR Methods. Base parameter values are listed in Table S3.

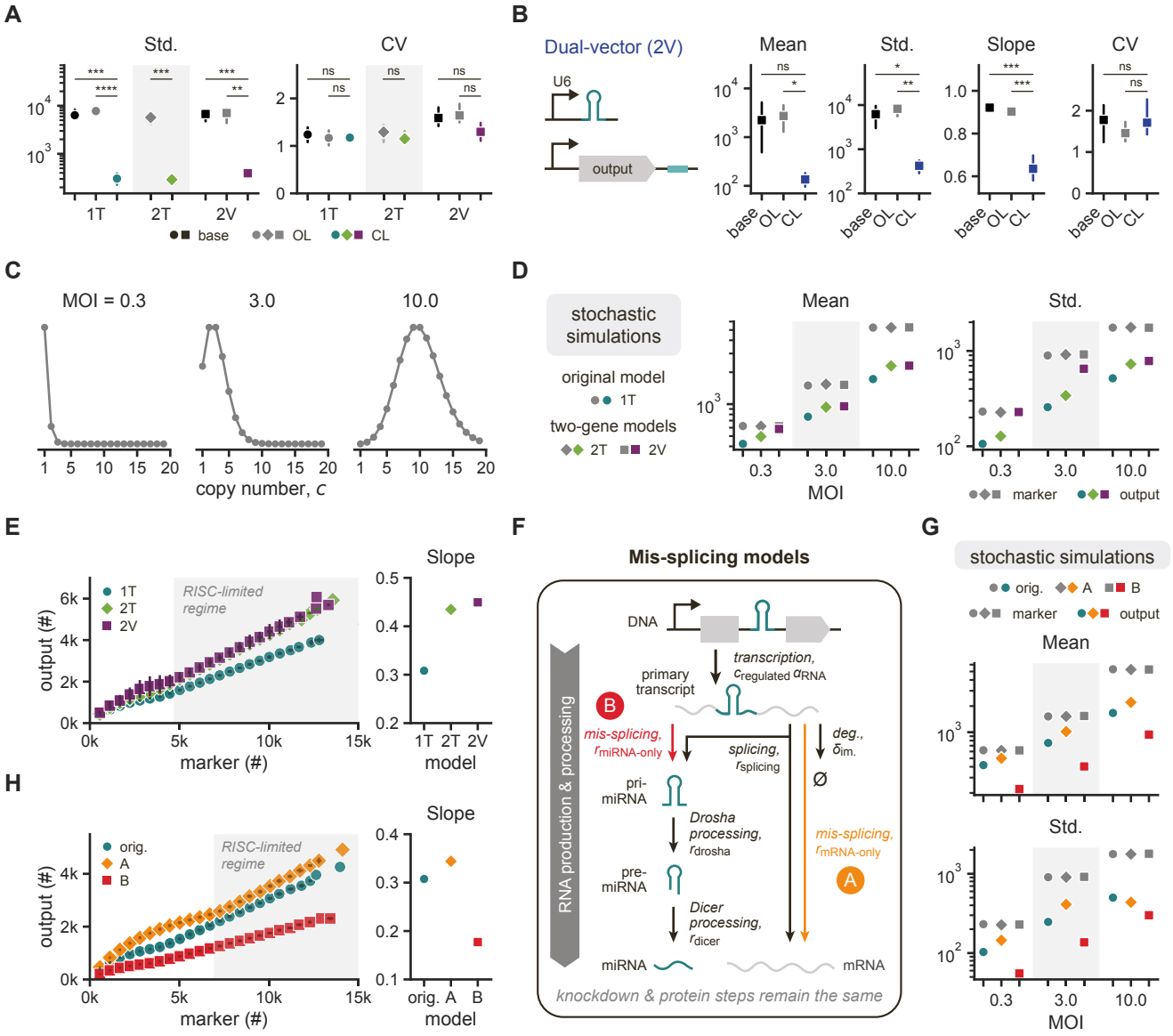

A
Std.
CV
1T
2T
2V
base
OL
CL
B
Dual-vector (2V)
U6
output
Mean
Std.
Slope
CV
C
MOI = 0.3
3.0
10.0
copy number, c
D
stochastic simulations
original model
1T
two-gene models
2T
2V
Mean
Std.
MOI
marker
output
E
RISC-limited regime
output (#)
marker (#)
Slope
model
F
Mis-splicing models
DNA
primary transcript
transcription,
RNA production & processing
mis-splicing,
splicing,
deg.,
pri-miRNA
Drosha processing,
pre-miRNA
Dicer processing,
miRNA
mRNA
knockdown & protein steps remain the same
G
stochastic simulations
orig.
A
B
marker
output
Mean
Std.
MOI
H
RISC-limited regime
output (#)
marker (#)
Slope
model

Figure S3. Additional two-gene circuit implementations and simulations, related to Figure 4.
A. Output standard deviation and coefficient of variation (CV) for HEK293T cells transfected with base genes, open-loop (OL) circuits, or closed-loop (CL) circuits depicted in Figure 4A. All circuits use miRE-FF4 and are expressed by the EF1α promoter. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Units are arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, independent t-test.
B. Left: DNA construct diagram of a non-intronic dual-vector (2V) circuit implementation, where microRNA expression is driven by the U6 promoter. Right: Summary statistics of output expression for HEK293T cells transfected with the base gene, OL circuit, or CL circuit for the system depicted to the left. Presented mean values use the geometric mean. Std. refers to the standard deviation, and slope represents the slope of the line fitted to the binned, log-transformed marker–output points. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Dashed line represents geometric mean output level for cells transfected only with a marker gene. Units are arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, independent t-test.
C. Probability density function (PDF) for Poisson distributions with means 0.3, 3, and 10 (left to right). These distributions are truncated such that all values are greater than or equal to one (i.e., the zero point of the PDF is removed). The distributions correspond to the MOI, the effective "multiplicity of infection," in **D**, **G**, and Figure 4D. See STAR Methods for more information.
D. Summary statistics for 10,000 stochastic simulations run for the single-transcript (1T, original), dual-transcript (2T), and dual-vector (2V) models with varying MOIs as described in STAR Methods. Plots depict protein geometric means and protein standard deviations (in molecules) for the distributions presented in Figure 4D.
E. Left: The same simulations in **D** and Figure 4D are combined across all MOIs. These data are binned by copy number (points), and the mean marker and output protein levels (in number of molecules) are plotted. Error bars represent standard deviations of the marker and output proteins for each bin. Gray shading highlights the DNA copy numbers at or above which the system becomes RISC-limited. Right: Slope of the data from left in the RISC-limited regime. Slope represents the slope of the line fitted to the binned marker–output points.
F. Schematic of RNA production and processing reactions modeled for two mis-splicing scenarios. Model A represents mis-splicing that results in unproductive microRNA, encoded by an additional reaction where primary transcript produces only mature mRNA at a rate $r_{\text{mRNA-only}}$. Model B represents mis-splicing that results in unproductive mRNA, encoded by an additional reaction where primary transcript produces only pri-miRNA at a rate $r_{\text{miRNA-only}}$. All other reactions remain the same as in Figure 3A.
G. Summary statistics for 10,000 stochastic simulations run for the original model and for mis-splicing models A and B depicted in **F**. Plots depict protein geometric means and protein standard deviations (in molecules).
H. Simulations from **G**, presented as in **E**.

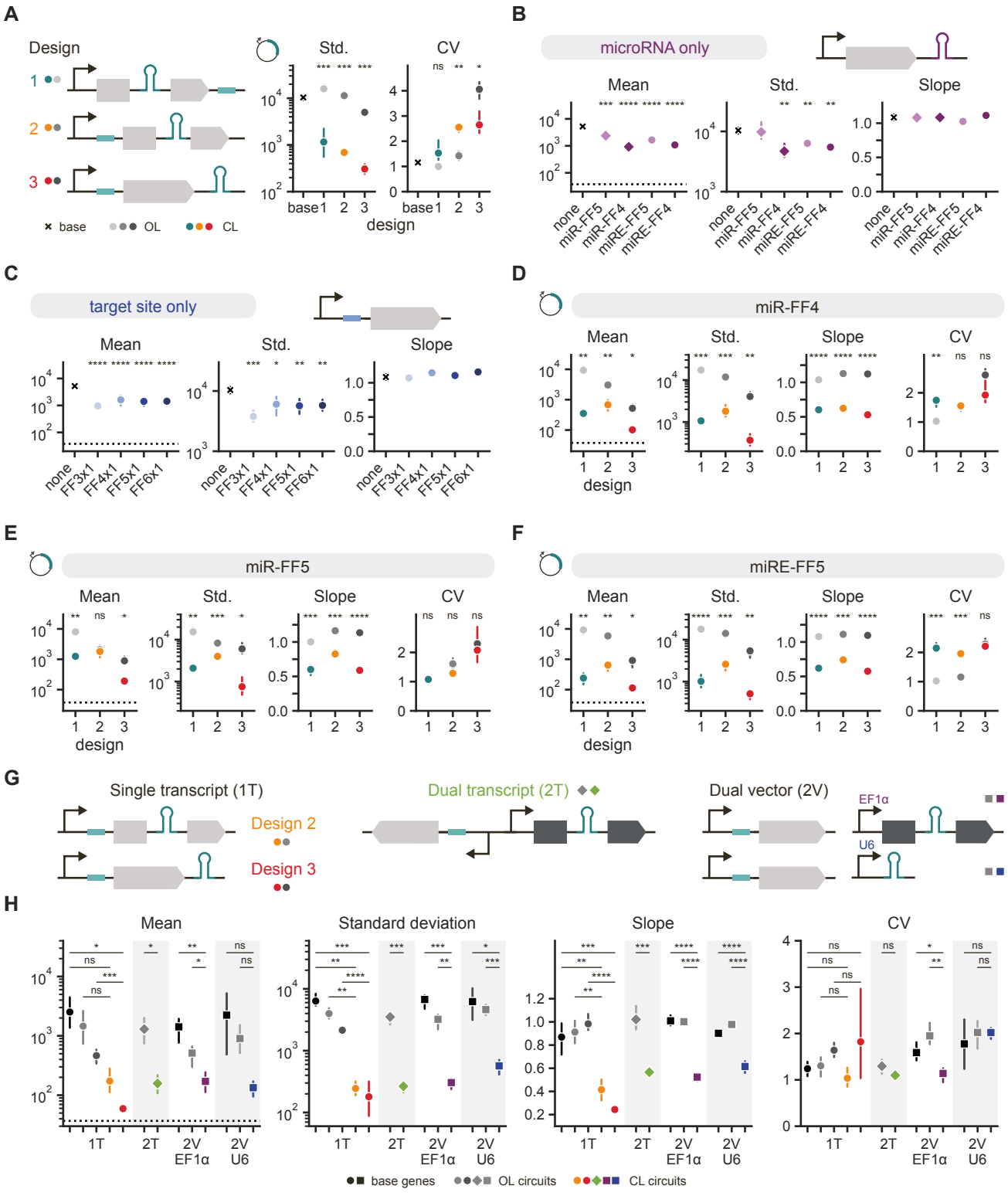

A
Design
1
2
3
base
OL
CL
Std.
CV
design
B
microRNA only
Mean
Std.
Slope
none
miR-FF5
miR-FF4
miRE-FF5
miRE-FF4
C
target site only
Mean
Std.
Slope
none
FF3x1
FF4x1
FF5x1
FF6x1
D
miR-FF4
Mean
Std.
Slope
CV
design
E
miR-FF5
Mean
Std.
Slope
CV
design
F
miRE-FF5
Mean
Std.
Slope
CV
design
G
Single transcript (1T)
Design 2
Design 3
Dual transcript (2T)
Dual vector (2V)
EF1α
U6
H
Mean
Standard deviation
Slope
CV
1T
2T
2V EF1α
2V U6
base genes
OL circuits
CL circuits

Figure S4. Additional single-transcript and two-gene circuit designs in transfection of HEK293T cells, related to Figure 5.

A. Output standard deviation and coefficient of variation (CV) for HEK293T cells transfected with circuit designs (left) as in Figure 5B. Design "base" refers to the base gene construct that does not contain an intronic microRNA or target sites. Statistical tests compare CL and OL conditions, and CL and base condition comparisons have the same significance trends.

B, **C**. Summary statistics for HEK293T cells transfected with output genes containing only a microRNA in the 3' UTR (**B**) or only a target site in the 5' UTR (**C**). Statistical tests compare conditions to the base gene, and unlabeled comparisons are not significant.

D–**F**. Summary statistics of output expression in HEK293T cells transfected with circuit designs in Figure 5A using miR-FF4 (**D**), miR-FF5 (**E**), or miRE-FF5 (**F**). Statistical tests compare CL and OL conditions, and CL and base conditions also have significantly different means, standard deviations, and slopes.

G. DNA construct diagrams for single-transcript (1T), dual-transcript (2T), and dual-vector (2V) circuits with microRNA target sites in the 5' UTR of the output gene.

H. Summary statistics of output expression in HEK293T cells transfected with the base gene, OL circuit, or CL circuit for systems shown in **G**.

All data depict expression of output genes driven by an EF1α promoter. The plotted mean values use the geometric mean. Std. refers to the standard deviation, and the slope represents the slope of the line fitted to the binned, log-transformed marker–output points. Dashed lines represent geometric mean output levels for cells transfected only with a marker gene lacking an intronic microRNA. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. All units are arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, independent t-test.

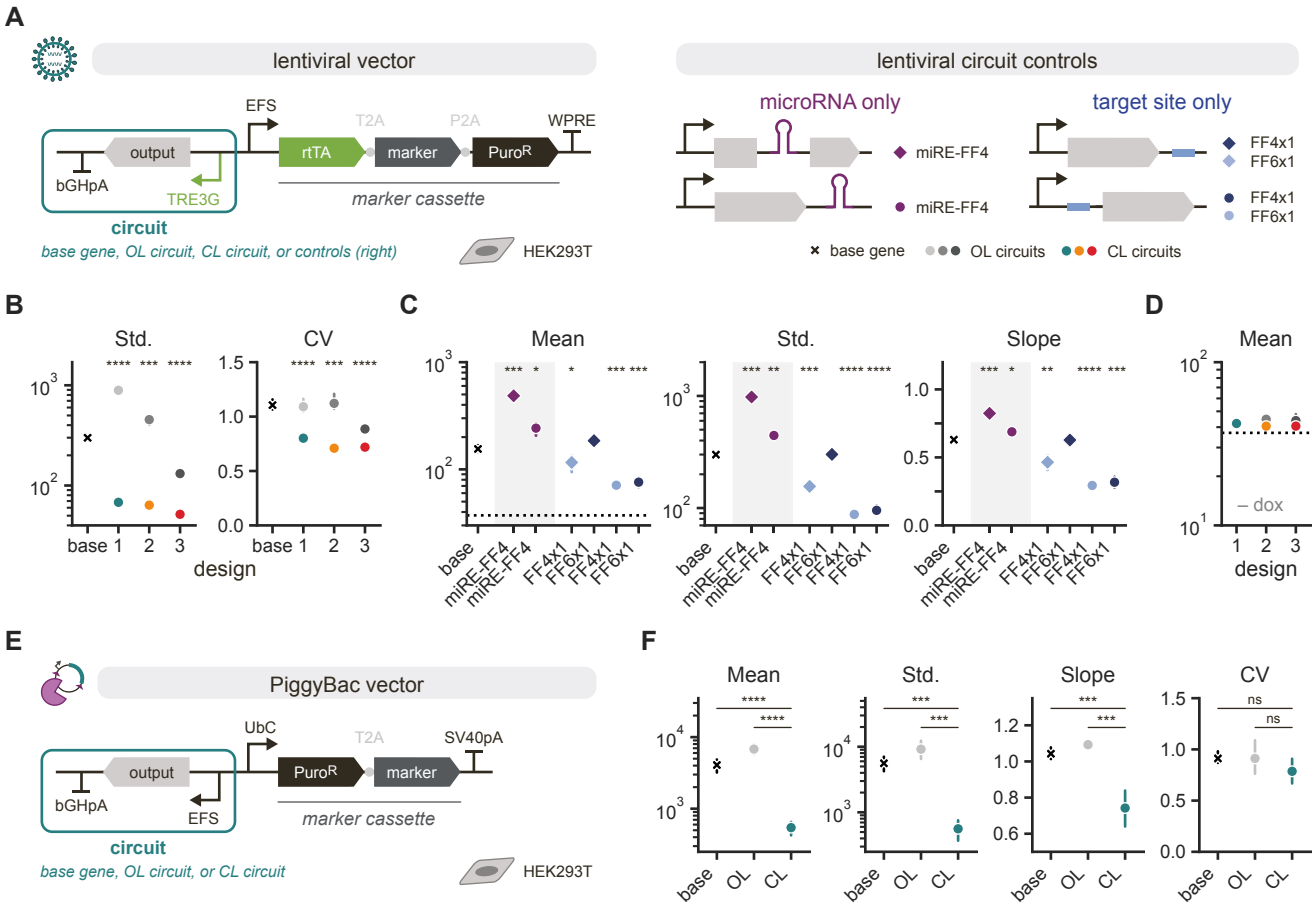

A
lentiviral vector
EFS
T2A
P2A
WPRE
output
rtTA
marker
PuroR
bGHpA
TRE3G
marker cassette
circuit
base gene, OL circuit, CL circuit, or controls (right)
HEK293T
lentiviral circuit controls
microRNA only
target site only
miRE-FF4
miRE-FF4
FF4x1
FF6x1
FF4x1
FF6x1
base gene
OL circuits
CL circuits
B
Std.
CV
design
C
Mean
Std.
Slope
D
Mean
– dox
design
E
PiggyBac vector
UbC
T2A
SV40pA
output
PuroR
marker
bGHpA
EFS
marker cassette
circuit
base gene, OL circuit, or CL circuit
HEK293T
F
Mean
Std.
Slope
CV
base
OL
CL

Figure S5. Lentiviral and PiggyBac vectors delivered to HEK293T cells, related to Figure 5.
A. Left: DNA construct diagram of the lentiviral vectors with base gene, open-loop (OL) circuit, or closed-loop (CL) circuit regulating expression of the output gene. The output gene is placed on the antisense strand of the viral genome under the control of a doxycycline-inducible TRE3G promoter. The marker cassette is expressed divergently from an EFS promoter and consists of the doxycycline-inducible activator rtTA, a marker gene, and a puromycin-resistance gene separated by 2A "self-cleaving" peptides. Right: DNA construct diagrams for unregulated circuit controls consisting of the output gene with either only a microRNA or only a target site. These output genes were expressed from a TRE3G promoter with a bGH polyadenylation signal in the same lentiviral vectors (left) in place of the circuit.
B. Output standard deviation and coefficient of variation (CV) for populations in Figures 5C, 5D. Design "base" refers to the base gene construct that does not contain an intronic microRNA or target sites. Statistical tests compare CL and OL conditions, and CL and base conditions are also significantly different.
C. Summary statistics of output expression in HEK293T cells lentivirally transduced with the unregulated control circuits in **A** (right) in the presence of inducer. Dashed line represents the geometric mean output level for untransduced cells. Statistical tests compare conditions to the base gene, and unlabeled conditions are not statistically significant.
D. Output expression for HEK293T cells lentivirally transduced with OL or CL circuits in Figure 5A in the absence of the inducer doxycycline (– dox). Dashed line represents the geometric mean output level for untransduced cells.
E. DNA construct diagram of the PiggyBac vectors used for transposase-mediated integration in HEK293T cells. Expression of the base gene, OL circuit, or CL circuit is driven by an EFS promoter. The marker cassette is expressed divergently from a UbC promoter and consists of a puromycin-resistance gene and a marker gene separated by a 2A "self-cleaving" peptide.
F. Summary statistics of output expression in HEK293T cells integrated with the PiggyBac vectors in **E**.
The plotted mean values use the geometric mean. Std. refers to the standard deviation, and the slope represents the slope of the line fitted to the binned, log-transformed marker–output points. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. All units are arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, independent t-test.

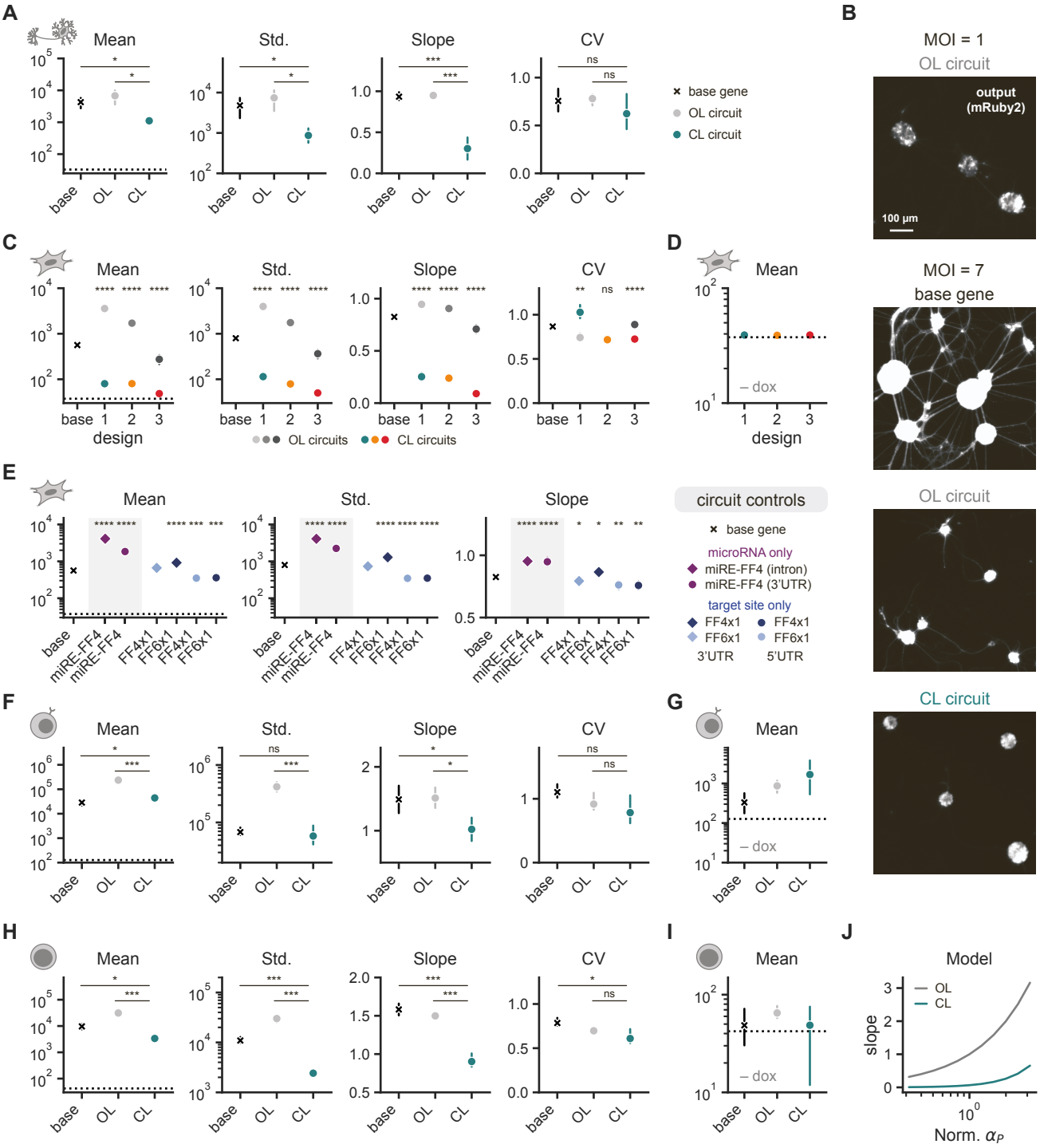

A
Mean
Std.
Slope
CV
base
OL
CL
base gene
OL circuit
CL circuit
B
MOI = 1
OL circuit
output (mRuby2)
100 µm
MOI = 7
base gene
OL circuit
CL circuit
C
design
OL circuits
CL circuits
D
– dox
E
circuit controls
microRNA only
miRE-FF4 (intron)
miRE-FF4 (3'UTR)
target site only
FF4x1
FF6x1
3'UTR
5'UTR
F
G
H
I
J
Model
slope
Norm. α_P

Figure S6. Additional data and controls for lentiviral delivery to primary cells, related to Figure 6.
A. Summary statistics of output expression for primary rat cortical neurons lentivirally transduced with the base gene, open-loop (OL) circuit, or closed-loop (CL) circuit in the presence of inducer as in Figure 6C.
B. Representative images of primary rat cortical neurons transduced with lentiviral vectors as in Figure 6B. Top image: Open-loop (OL) circuit transduced at an MOI of 1. Bottom three images: Base gene, open-loop (OL) circuit, and closed-loop (CL) circuit transduced at an MOI of 7. Images depict fluorescence of the output gene, mRuby2, for $\sim$200 representative cells, and the scale bar represents 100 µm.
C. Summary statistics of output expression for primary mouse embryonic fibroblasts lentivirally transduced with the base gene or different designs of the OL or CL circuits in the presence of inducer as in Figure 6D. Statistical tests compare CL and OL conditions, and CL and base conditions are all significantly different.
D. Output expression for primary mouse embryonic fibroblasts lentivirally transduced with OL or CL circuits in the absence of the inducer doxycycline (– dox).
E. Summary statistics of output expression for primary mouse embryonic fibroblasts lentivirally transduced with unregulated control circuits in the presence of inducer. Statistical tests compare conditions to the base gene, and unlabeled conditions are not statistically significant.
F. Summary statistics of output expression for primary human T cells lentivirally transduced with the base gene, OL circuit, or CL circuit in the presence of inducer as in Figure 6E.
G. Output expression for primary human T cells lentivirally transduced with the base gene, OL circuit, or CL circuit in the absence of the inducer doxycycline (– dox).
H. Summary statistics of output expression for human induced pluripotent stem cells lentivirally transduced with the base gene, OL circuit, or CL circuit in the presence of inducer as in Figure 6F.
I. Output expression for human induced pluripotent stem cells lentivirally transduced with the base gene, OL circuit, or CL circuit in the absence of the inducer doxycycline (– dox).
A–**I**. The plotted mean values use the geometric mean. Std. refers to the standard deviation, and CV is the coefficient of variation. Slope represents the slope of the line fitted to the binned, log-transformed marker–output points. Dashed lines indicate the geometric mean output level for untransduced cells. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. All units are arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, independent t-test.
J. Slope of the copy number–protein curve in the RISC-limited regime of the model depicted in Figure 3 as a function of translation rate, α_p, for the open-loop (OL) and closed-loop (CL) circuits. The slopes are normalized by that of an unregulated gene to model the marker–output relationship. Values of α_p are normalized to the original value. See STAR Methods for more details on the slope calculation.

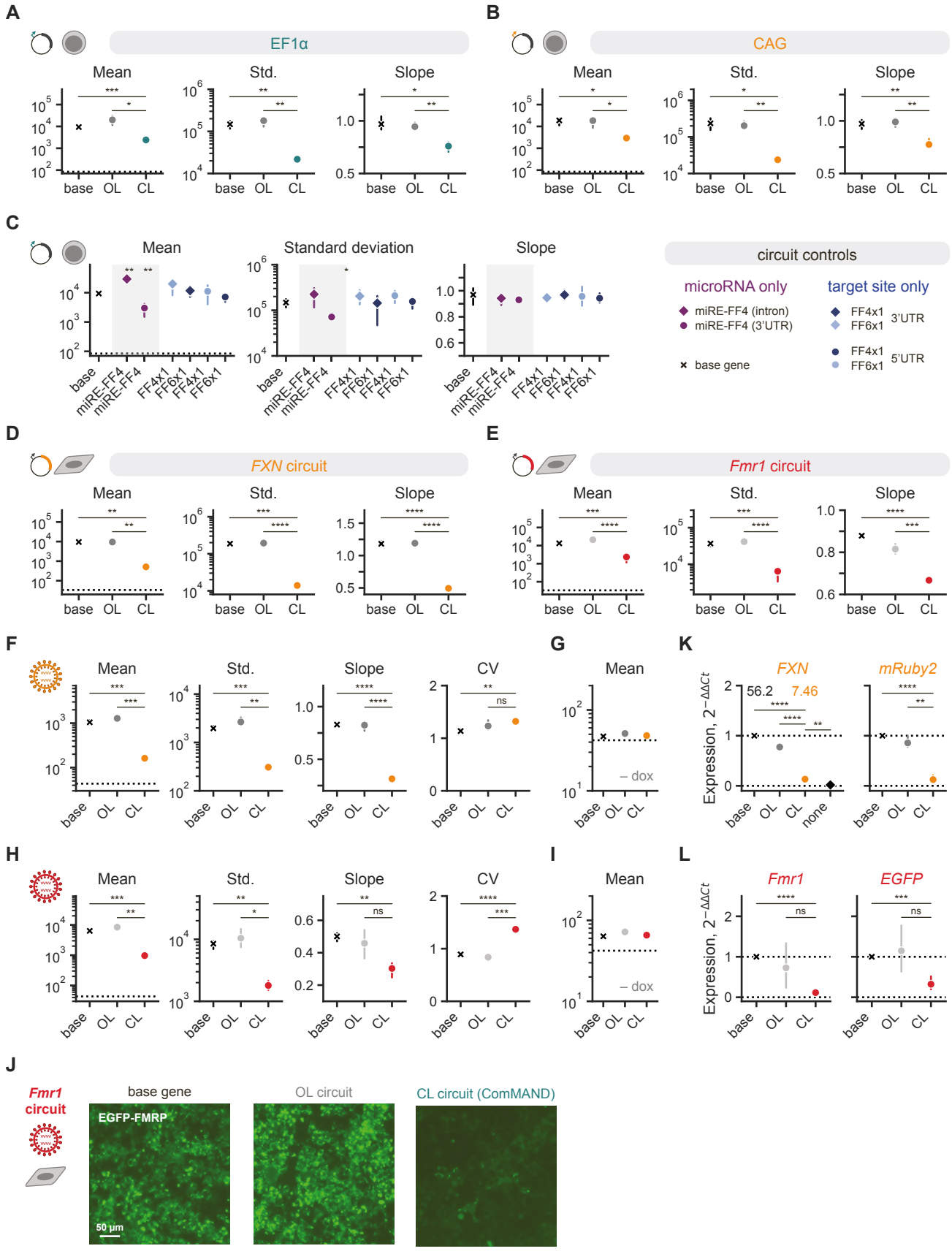

A
EF1α
Mean
Std.
Slope
base OL CL
B
CAG
C
Mean
Standard deviation
Slope
base
miRE-FF4
miRE-FF4
FF4x1
FF6x1
FF4x1
FF6x1
circuit controls
microRNA only
miRE-FF4 (intron)
miRE-FF4 (3'UTR)
base gene
target site only
FF4x1
FF6x1
3'UTR
FF4x1
FF6x1
5'UTR
D
FXN circuit
E
Fmr1 circuit
F
CV
G
– dox
K
FXN
mRuby2
Expression, 2−ΔΔCt
56.2
7.46
none
H
I
L
Fmr1
EGFP
J
Fmr1 circuit
base gene
OL circuit
CL circuit (ComMAND)
EGFP-FMRP
50 µm

Figure S7. Additional data and controls for iPSC transfections and for delivery of therapeutically relevant genes, related to Figure 6.

A, **B**. Summary statistics of output expression for induced pluripotent stem cells (iPSCs) co-transfected with a marker gene and the base gene, open-loop (OL) circuit, or closed-loop (CL) circuit expressed by the EF1α promoter (**A**) or CAG promoter (**B**). Circuits use miRE-FF4. Dashed lines indicate the geometric mean output level for cells transfected only with the marker gene.

C. Summary statistics of output expression for iPSCs transfected with unregulated circuit controls consisting of the output gene with either only a microRNA or only a target site. Circuit controls were expressed by the EF1α promoter. Dashed line indicates the geometric mean output level for cells transfected only with the marker gene. Statistical tests compare conditions to the base gene, and unlabeled conditions are not statistically significant.

D, **E**. Summary statistics of output expression for HEK293T cells transfected with circuits shown in Figure 6H regulating the therapeutically relevant genes *FXN* (**D**) or *Fmr1* (**D**). Circuits were expressed by the EF1α promoter. Dashed lines indicate the geometric mean output level for cells transfected only with the marker gene.

F. Summary statistics of output expression for HEK293T cells lentivirally transduced with the base gene, OL circuit, or CL circuit regulating *FXN* in the presence of inducer as in Figure 6H.

G. Output expression for HEK293T cells lentivirally transduced with the base gene, OL circuit, or CL circuit regulating *FXN* in the absence of the inducer doxycycline (– dox).

H. Summary statistics of output expression for HEK293T cells lentivirally transduced with the base gene, OL circuit, or CL circuit regulating *Fmr1* in the presence of inducer as in Figure 6H.

I. Output expression for HEK293T cells lentivirally transduced with the base gene, OL circuit, or CL circuit regulating *Fmr1* in the absence of the inducer doxycycline (– dox).

J. Representative images of conditions in **H**. Images depict fluorescence of the output fusion protein, EGFP-FMRP, for $\sim$400 representative cells, and the scale bar represents 50 µm.

K. Relative expression of output mRNA measured via RT-qPCR for conditions in **F** amplified with primers targeting *FXN* (left) or *mRuby2* (right). Expression was first normalized to expression of the housekeeping gene *GAPDH* for each condition (ΔC_t), then normalized to the expression of the corresponding base gene condition within each biological replicate ($\Delta\Delta C_t$). Plots display relative expression as $2^{-\Delta\Delta C_t}$. "None" refers to untransduced cells. Dashed lines indicate relative expression levels 0 and 1. Annotations describe the fold-change relative expression of the base gene (56.2x) and CL circuit (7.46x) compared to the untransduced population.

L. Relative expression of output mRNA measured via RT-qPCR for conditions in **H** amplified with primers targeting *Fmr1* (left) or *EGFP* (right). Expression was first normalized to expression of the housekeeping gene *GAPDH* for each condition (ΔC_t), then normalized to the expression of the corresponding base gene condition within each biological replicate ($\Delta\Delta C_t$). Plots display relative expression as $2^{-\Delta\Delta C_t}$. Dashed lines indicate relative expression levels 0 and 1.

The plotted mean values use the geometric mean. Std. refers to the standard deviation, and CV is the coefficient of variation. Slope represents the slope of the line fitted to the binned, log-transformed marker–output points. Points represent means of $n \geq 3$ biological replicates, and error bars show the 95% confidence interval. Units for plots in **A**–**I** are arbitrary units from a flow cytometer. ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, independent t-test.

Supplementary Tables

Reaction	Rate constant	mRNA_i	pri	pre	miR	mRNA	R	RI	RIM	protein
Transcription	$c_{\text{regulated}}\alpha_{\text{RNA}}$	1	0	0	0	0	0	0	0	0
Splicing	r_{splicing}	-1	1	0	0	1	0	0	0	0
Immature degradation	δ_{immature}	-1	0	0	0	0	0	0	0	0
Drosha	r_{drosha}	0	-1	1	0	0	0	0	0	0
Dicer	r_{dicer}	0	0	-1	1	0	0	0	0	0
microRNA degradation	δ_{miRNA}	0	0	0	-1	0	0	0	0	0
mRNA degradation	δ_{mRNA}	0	0	0	0	-1	0	0	0	0
RISC loading	$k_{\text{miRNA,bind}}$	0	0	0	-1	0	-1	1	0	0
Loaded degradation	$k_{\text{miRNA,deg}}$	0	0	0	0	0	1	-1	0	0
RISC-mRNA binding	$k_{\text{mRNA,bind}}$	0	0	0	0	-1	0	-1	1	0
RISC-mRNA unbinding	$k_{\text{mRNA,unbind}}$	0	0	0	0	1	0	1	-1	0
mRNA knockdown	k_{deg}	0	0	0	0	0	0	1	-1	0
Translation	$\alpha_p[\text{mRNA}]$	0	0	0	0	0	0	0	0	1
Bound translation	$\zeta\alpha_p[\text{RIM}]$	0	0	0	0	0	0	0	0	1
Protein degradation	δ_p	0	0	0	0	0	0	0	0	-1

Table S1. Reaction stoichiometries and rate constants for our model of the ComMAND circuit, related to Figure 3 and STAR Methods. "mRNA_i" is the immature primary transcript, "pre" is the pre-miRNA, and "pri" is the pri-miRNA. "miR" represents the microRNA, and "mRNA" represents the output (regulated) mRNA. "R" refers to unbound RISC, "RI" refers to RISC loaded with microRNA, and "RIM" refers to loaded RISC bound to mRNA. The rate of a reaction with reaction rate constant r and stoichiometric coefficients ζ_i for species s_i is $r\prod_{\zeta_i<0} s_i^{|\zeta_i|}$, i.e., a mass action equation in terms of the reactants (species with negative stoichiometric coefficients).

System	Reaction	Rate constant	mRNA_i	pri	mRNA	$\text{mRNA}_{\text{ignore}}$
Dual-transcript (Figure 4)	microRNA transcription	$c_{\text{regulated}}\alpha_{\text{RNA}}$	1	0	0	0
	Splicing	r_{splicing}	-1	1	0	1
	Output mRNA transcription	$c_{\text{regulated}}\alpha_{\text{RNA}}$	0	0	1	0
Dual-vector (Figure 4)	microRNA transcription	$c_{\text{regulated}}\alpha_{\text{RNA}}$	1	0	0	0
	Splicing	r_{splicing}	-1	1	0	1
	Output mRNA transcription	$c_{\text{unregulated}}\alpha_{\text{RNA}}$	0	0	1	0
Dual-vector, non-intronic	microRNA transcription	$c_{\text{U6}}\alpha_{\text{U6}}$	0	1	0	0
	Output mRNA transcription	$c_{\text{regulated}}\alpha_{\text{RNA}}$	0	0	1	0

Table S2. Reaction stoichiometries and rate constants for alternate reactions used to simulate the dual-transcript and dual-vector systems, related to Figure 4 and STAR Methods. In the dual-transcript and dual-vector cases, splicing of the intronic microRNA creates an unregulated mRNA transcript, $\text{mRNA}_{\text{ignore}}$ that does interact with the rest of the system. The dual-transcript and dual-vector cases then only differ in the copy number variable used in the transcription reactions. For the dual-vector non-intronic (U6-driven) system, the transcription steps directly produce pri-miRNA or the output mRNA.

Parameter	Value	Units	Source
α_{RNA}	$4.67 \cdot 10^{-2}$	1 / s	Tigges *et al.* [1]
r_{splicing}	$2.0 \cdot 10^{-3}$	1 / s	Bleris *et al.* [2]
δ_{immature}	$2.88 \cdot 10^{-4}$	1 / s	Tigges *et al.* [1]
r_{drosha}	$1.0 \cdot 10^{-2}$	1 / s	Bleris *et al.* [2]
r_{dicer}	$1.0 \cdot 10^{-3}$	1 / s	Tigges *et al.* [3]
δ_{miRNA}	$2.88 \cdot 10^{-4}$	1 / s	Tigges *et al.* [1]
δ_{mRNA}	$2.88 \cdot 10^{-4}$	1 / s	Tigges *et al.* [1]
$k_{\text{miRNA,bind}}$	$1.0 \cdot 10^{-5}$	1 / (molecules · s)	Tigges *et al.* [3]
$k_{\text{miRNA,deg}}$	$2.16 \cdot 10^{-5}$	1 / s	Tigges *et al.* [3]
$k_{\text{mRNA,bind}}$	$1.84 \cdot 10^{-6}$	1 / (molecules · s)	Tigges *et al.* [3]
$k_{\text{mRNA,unbind}}$	0.303	1 / s	Yang *et al.* [4]
k_{deg}	$7.0 \cdot 10^{-3}$	1 / s	Tigges *et al.* [3]
α_p	$3.33 \cdot 10^{-4}$	1 / s	Tigges *et al.* [1, 3]
ζ	0.0	dimensionless	n/a
δ_p	$9.67 \cdot 10^{-5}$	1 / s	Tigges *et al.* [1]

Table S3. Base parameter values used in the model, related to Figures 3, 4 and STAR Methods. Both the steady-state solution and stochastic simulations use these parameters. In Figure 3, parameter values are varied from these base values while keeping non-modified parameters constant.

Figure	Main plasmids	Additional plasmid(s)
Figures 1E, 1F, S1F	pKG3107 (base), pKG3147 (OL), pKG3148 (CL)	pKG2352 (marker)
Figures 2A–2C, 3D (right), 3F (left), S1G	pKG3107 (base) miR-FF5: pKG3147 (OL), pKG3148 (CL) miR-FF4: pKG3166 (OL), pKG3167 (CL) miRE-FF5: pKG2998 (OL), pKG2999 (CL) miRE-FF4: pKG3109 (OL), pKG3110 (CL)	pKG2352 (marker)
Figures 2D, 3C (right), S1H	EF1α: pKG3107 (base), pKG3109 (OL), pKG3110 (CL) CAG: pKG3185 (base), pKG3186 (OL), pKG3187 (CL) EFS: pKG1850 (base), pKG3781 (OL), pKG3782 (CL) hPGK: pKG1851 (base), pKG3783 (OL), pKG3784 (CL)	pKG2352 (marker)
Figure 2E, S1I	pKG3107 (base), pKG3109 (OL), pKG3110 (CL)	pKG2352 (marker), pKG2898 (filler)
Figures 2F, 3F (right), S1L	1 target site: pKG3109 (OL), pKG3110 (CL) 2 target sites: pKG3181 (OL), pKG3182 (CL) 4 target sites: pKG3183 (OL), pKG3184 (CL)	pKG2352 (marker)
Figure 4B, S3A	1T: pKG3188 + {pKG3107 (base), pKG3109 (OL), pKG3110 (CL)} 2T: pKG3010 (OL), pKG3011 (CL) 2V: pKG3009 + {pKG3107 (base), pKG3127 (OL), pKG3129 (CL)}	pKG2908 (marker)
Figure 5B, S4A	pKG3107 (base) design 1: pKG3109 (OL), pKG3110 (CL) design 2: pKG3173 (OL), pKG3174 (CL) design 3: pKG3175 (OL), pKG3176 (CL)	pKG2352 (marker)
Figures 5C, 5D, S5B, S5D, S6C, S6D	pKG3097 (base) design 1: pKG3800 (OL), pKG3801 (CL) design 2: pKG3802 (OL), pKG3803 (CL) design 3: pKG3804 (OL), pKG3805 (CL)	psPAX2 (viral), pMD2.G (viral)
Figures 6B–6G, S6A, S6B, S6F–S6I	pKG3097 (base), pKG3800 (OL), pKG3801 (CL)	psPAX2 (viral), pMD2.G (viral)
Figure 6H, S7F–S7L	*FXN*: pKG3669 (base), pKG3670 (OL), pKG3671 (CL) *Fmr1*: pKG3672 (base), pKG3673 (OL), pKG3674 (CL)	psPAX2 (viral), pMD2.G (viral)
Figures S1B, S1C, S1J, S1K	microRNA plasmids: pKG3809, pKG3810, pKG3811, pKG3812, pKG3813 target site plasmids: pKG3814, pKG3815, pKG3816, pKG3817, pKG3818, pKG3819, pKG3820, pKG3821, pKG3822, pKG3823, pKG3824, pKG3825, pKG3826	n/a
Figures S1D, S1E	pKG3107 (base) microRNA only: pKG3108, pKG3145, pKG3164, pKG2997 target site only: pKG3111, pKG3112, pKG3126, pKG3127, pKG3128, pKG3129	pKG2352 (marker)
Figure S3B	pKG2600 + {pKG3107 (base), pKG3127 (OL), pKG3129 (CL)}	pKG2908 (marker)

Table S4. *Continued on the next page.*

Figure	Main plasmids	Additional plasmid(s)
Figures S4B, S4C	pKG3107 (base) microRNA only: pKG3146, pKG3165, pKG3172, pKG2974 target site only: pKG3130, pKG3131, pKG3132, pKG3133	pKG2352 (marker)
Figure S4D	design 1: pKG3166 (OL), pKG3167 (CL) design 2: pKG3168 (OL), pKG3169 (CL) design 3: pKG3170 (OL), pKG3171 (CL)	pKG2352 (marker)
Figure S4E	design 1: pKG3147 (OL), pKG3148 (CL) design 2: pKG3150 (OL), pKG3151 (CL) design 3: pKG3153 (OL), pKG3154 (CL)	pKG2352 (marker)
Figure S4F	design 1: pKG2998 (OL), pKG2999 (CL) design 2: pKG3000 (OL), pKG3001 (CL) design 3: pKG2975 (OL), pKG2976 (CL)	pKG2352 (marker)
Figure S4H	1T: pKG3188 + pKG3107 (base) 1T design 2: pKG3188 + {pKG3173 (OL), pKG3174 (CL)} 1T design 3: pKG3188 + {pKG3175 (OL), pKG3176 (CL)} 2T: pKG3012 (OL), pKG3013 (CL) 2V EF1α: pKG3009 + {pKG3107 (base), pKG3131 (OL), pKG3133 (CL)} 2V U6: pKG2600 + {pKG3107 (base), pKG3131 (OL), pKG3133 (CL)}	pKG2908 (marker)
Figures S5C, S6E	pKG3097 (base) microRNA only: pKG3794, pKG3795 target site only: pKG3796, pKG3797, pKG3798, pKG3799	psPAX2 (viral), pMD2.G (viral)
Figure S5F	pKG3806 (base), pKG3807 (OL), pKG3808 (CL)	pRJ0084 (transposase), pKG2898 (reporter)
Figures S7A, S7B	EF1α: pKG3107 (base), pKG3109 (OL), pKG3110 (CL) CAG: pKG3185 (base), pKG3186 (OL), pKG3187 (CL)	pKG2027 (marker)
Figure S7C	pKG3107 (base) microRNA only: pKG3108, pKG3172 target site only: pKG3111, pKG3112, pKG3131, pKG3133	pKG2027 (marker)
Figures S7D, S7E	*FXN*: pKG3189 (base), pKG3190 (OL), pKG3191 (CL) *Fmr1*: pKG3076 (base), pKG3077 (OL), pKG3078 (CL)	pKG2905 (marker)

Table S4. Plasmids used in each figure, related to STAR Methods. Base, base gene; OL, open-loop circuit; CL, closed-loop circuit. For information on the contents of the plasmids constructed in this work (beginning pKG), see Table S5. For information on plasmids used but not constructed in this work, see the Key Resources Table.

Name	Plasmid	Addgene
pKG1850	EFS-mRuby2-bGH	235253
pKG1851	hPGK-mRuby2-bGH	235254
pKG2027	EF1α-mGL-bGH	–
pKG2352	EFS-mGL-bGH	235255
pKG2600	U6-miR.FF4	235256
pKG2898	EF1α-tagBFP-bGH	235257
pKG2905	CAG-SNAP-iRFP670-bGH	–
pKG2908	EF1α-SNAP-iRFP670-bGH	–
pKG2974	EF1α-mRuby2-miRE.FF5-bGH	–
pKG2975	EF1α-TS.FF6x1-mRuby2-miRE.FF5-bGH	235258
pKG2976	EF1α-TS.FF5x1-mRuby2-miRE.FF5-bGH	235259
pKG2997	EF1α-mRuby2(miRE.FF5)-bGH	–
pKG2998	EF1α-mRuby2(miRE.FF5)-TS.FF6x1-bGH	235260
pKG2999	EF1α-mRuby2(miRE.FF5)-TS.FF5x1-bGH	235261
pKG3000	EF1α-TS.FF6x1-mRuby2(miRE.FF5)-bGH	235262
pKG3001	EF1α-TS.FF5x1-mRuby2(miRE.FF5)-bGH	235263
pKG3009	EF1α-mGL(miRE.FF4)-bGH	235264
pKG3010	EF1α-mRuby2-TS.FF4x1-bGH	–
pKG3011	EF1α-mRuby2-TS.FF6x1-bGH	–
pKG3012	EF1α-TS.FF4x1-mRuby2-bGH	–
pKG3013	EF1α-TS.FF6x1-mRuby2-bGH	–
pKG3076	EF1α-EGFP-FMRP-bGH	235265
pKG3077	EF1α-EGFP(miRE.FF4)-FMRP-TS.FF6x1-bGH	235266
pKG3078	EF1α-EGFP(miRE.FF4)-FMRP-TS.FF4x1-bGH	235267
pKG3097	[TRE3G-mRuby2-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235268
pKG3107	EF1α-mRuby2-bGH	235269
pKG3108	EF1α-mRuby2(miRE.FF4)-bGH	–
pKG3109	EF1α-mRuby2(miRE.FF4)-TS.FF6x1-bGH	235270
pKG3110	EF1α-mRuby2(miRE.FF4)-TS.FF4x1-bGH	235271
pKG3111	EF1α-mRuby2-TS.FF6x1-bGH	–
pKG3112	EF1α-mRuby2-TS.FF4x1-bGH	–
pKG3126	EF1α-mRuby2-TS.FF3x1-bGH	–
pKG3127	EF1α-mRuby2-TS.FF4x1-bGH	235272
pKG3128	EF1α-mRuby2-TS.FF5x1-bGH	–
pKG3129	EF1α-mRuby2-TS.FF6x1-bGH	235273
pKG3130	EF1α-TS.FF3x1-mRuby2-bGH	–
pKG3131	EF1α-TS.FF4x1-mRuby2-bGH	235274
pKG3132	EF1α-TS.FF5x1-mRuby2-bGH	–
pKG3133	EF1α-TS.FF6x1-mRuby2-bGH	235275
pKG3145	EF1α-mRuby2(miR.FF5)-bGH	–
pKG3146	EF1α-mRuby2-miR.FF5-bGH	–
pKG3147	EF1α-mRuby2(miR.FF5)-TS.FF6x1-bGH	235276
pKG3148	EF1α-mRuby2(miR.FF5)-TS.FF5x1-bGH	235277
pKG3150	EF1α-TS.FF6x1-mRuby2(miR.FF5)-bGH	235278
pKG3151	EF1α-TS.FF5x1-mRuby2(miR.FF5)-bGH	235279
pKG3153	EF1α-TS.FF6x1-mRuby2-miR.FF5-bGH	235280
pKG3154	EF1α-TS.FF5x1-mRuby2-miR.FF5-bGH	235281
pKG3164	EF1α-mRuby2(miR.FF4)-bGH	–
pKG3165	EF1α-mRuby2-miR.FF4-bGH	–
pKG3166	EF1α-mRuby2(miR.FF4)-TS.FF6x1-bGH	235282

Table S5. *Continued on the next page.*

Name	Plasmid	Addgene
pKG3167	EF1α-mRuby2(miR.FF4)-TS.FF4x1-bGH	235283
pKG3168	EF1α-TS.FF6x1-mRuby2(miR.FF4)-bGH	235284
pKG3169	EF1α-TS.FF4x1-mRuby2(miR.FF4)-bGH	235285
pKG3170	EF1α-TS.FF6x1-mRuby2-miR.FF4-bGH	235286
pKG3171	EF1α-TS.FF4x1-mRuby2-miR.FF4-bGH	235287
pKG3172	EF1α-mRuby2-miRE.FF4-bGH	–
pKG3173	EF1α-TS.FF6x1-mRuby2(miRE.FF4)-bGH	235288
pKG3174	EF1α-TS.FF4x1-mRuby2(miRE.FF4)-bGH	235289
pKG3175	EF1α-TS.FF6x1-mRuby2-miRE.FF4-bGH	235290
pKG3176	EF1α-TS.FF4x1-mRuby2-miRE.FF4-bGH	235291
pKG3181	EF1α-mRuby2(miRE.FF4)-TS.FF6x2-bGH	235292
pKG3182	EF1α-mRuby2(miRE.FF4)-TS.FF6x4-bGH	235293
pKG3183	EF1α-mRuby2(miRE.FF4)-TS.FF4x2-bGH	235294
pKG3184	EF1α-mRuby2(miRE.FF4)-TS.FF4x4-bGH	235295
pKG3185	CAG-mRuby2-bGH	235296
pKG3186	CAG-mRuby2(miRE.FF4)-TS.FF6x1-bGH	235297
pKG3187	CAG-mRuby2(miRE.FF4)-TS.FF4x1-bGH	235298
pKG3188	EF1α-mGL-bGH	235299
pKG3189	EF1α-FXN-P2A-mRuby2-bGH	235300
pKG3190	EF1α-FXN-P2A-mRuby2(miRE.FF4)-TS.FF6x1-bGH	235301
pKG3191	EF1α-FXN-P2A-mRuby2(miRE.FF4)-TS.FF4x1-bGH	235302
pKG3669	[TRE3G-FXN-P2A-mRuby2-bGH] - EFS-rtTA-P2A-mGL-WPRE	235303
pKG3670	[TRE3G-FXN-P2A-mRuby2(miRE.FF4)-TS.FF6x1-bGH] - EFS-rtTA-P2A-mGL-WPRE	235304
pKG3671	[TRE3G-FXN-P2A-mRuby2(miRE.FF4)-TS.FF4x1-bGH] - EFS-rtTA-P2A-mGL-WPRE	235305
pKG3672	[TRE3G-EGFP-FMRP-bGH] - EFS-rtTA-P2A-mRuby2-WPRE	235306
pKG3673	[TRE3G-EGFP(miRE.FF4)-FMRP-TS.FF6x1-bGH] - EFS-rtTA-P2A-mRuby2-WPRE	235307
pKG3674	[TRE3G-EGFP(miRE.FF4)-FMRP-TS.FF4x1-bGH] - EFS-rtTA-P2A-mRuby2-WPRE	235308
pKG3781	EFS-mRuby2(miRE.FF4)-TS.FF6x1-bGH	235309
pKG3782	EFS-mRuby2(miRE.FF4)-TS.FF4x1-bGH	235310
pKG3783	hPGK-mRuby2(miRE.FF4)-TS.FF6x1-bGH	235311
pKG3784	hPGK-mRuby2(miRE.FF4)-TS.FF4x1-bGH	235312
pKG3794	[TRE3G-mRuby2-TS.FF6x1-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235313
pKG3795	[TRE3G-mRuby2-TS-FF4x1-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235314
pKG3796	[TRE3G-TS.FF6x1-mRuby2-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235315
pKG3797	[TRE3G-TS.FF4x1-mRuby2-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235316
pKG3798	[TRE3G-mRuby2(miRE.FF4)-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235317
pKG3799	[TRE3G-mRuby2-miRE.FF4-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235318
pKG3800	[TRE3G-mRuby2(miRE.FF4)-TS.FF6x1-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235319

Table S5. *Continued on the next page.*

Name	Plasmid	Addgene
pKG3801	[TRE3G-mRuby2-(miRE.FF4)-TS.FF4x1-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235320
pKG3802	[TRE3G-TS.FF6x1-mRuby2(miRE.FF4)-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235321
pKG3803	[TRE3G-TS.FF4x1-mRuby2(miRE.FF4)-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235322
pKG3804	[TRE3G-TS.FF6x1-mRuby2-miRE.FF4-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235323
pKG3805	[TRE3G-TS.FF4x1-mRuby2-miRE.FF4-bGH] - EFS-rtTA-T2A-mGL-P2A-PuroR-WPRE	235324
pKG3806	[EF1α-mRuby2-bGH] - UbC-PuroR-T2A-mGL-SV40pA	235325
pKG3807	[EF1α-mRuby2(miRE.FF4)-TS.FF6x1-bGH] - UbC-PuroR-T2A-mGL-SV40pA	235326
pKG3808	[EF1α-mRuby2(miRE.FF4)-TS.FF4x1-bGH] - UbC-PuroR-T2A-mGL-SV40pA	235327
pKG3809	hPGK-mRuby2-bGH	–
pKG3810	hPGK-mRuby2(miR.FF3)-bGH	–
pKG3811	hPGK-mRuby2(miR.FF4)-bGH	–
pKG3812	hPGK-mRuby2(miR.FF5)-bGH	–
pKG3813	hPGK-mRuby2(miR.FF6)-bGH	–
pKG3814	hPGK-mGL-bGH	–
pKG3815	hPGK-mGL-TS.FF3x1-bGH	–
pKG3816	hPGK-mGL-TS.FF3x2-bGH	–
pKG3817	hPGK-mGL-TS.FF3x4-bGH	–
pKG3818	hPGK-mGL-TS.FF4x1-bGH	–
pKG3819	hPGK-mGL-TS.FF4x2-bGH	–
pKG3820	hPGK-mGL-TS.FF4x4-bGH	–
pKG3821	hPGK-mGL-TS.FF5x1-bGH	–
pKG3822	hPGK-mGL-TS.FF5x2-bGH	–
pKG3823	hPGK-mGL-TS.FF5x4-bGH	–
pKG3824	hPGK-mGL-TS.FF6x1-bGH	–
pKG3825	hPGK-mGL-TS.FF6x2-bGH	–
pKG3826	hPGK-mGL-TS.FF6x4-bGH	–

Table S5. List of plasmids constructed in this work, related to Table S4 and STAR Methods. Parentheses indicate sequences in introns, and sequences in square brackets are inverted. For information on plasmids used but not constructed in this work, see the Key Resources Table. Sequence maps for all plasmids have been deposited at Zenodo (see Key Resources Table).

Figure	Cytometer	Channel	Laser	Filter	PMTV
Figures 1–5, 6D, 6H; Figures S1D–S1I, S1L, S3A, S3B, S4, S5, S6C–S6E, S7A–S7C, S7F–S7I	Attune NxT	mGL/EGFP mRuby2	488 561	530 / 30 585 / 16	220 200
Figure 6C; Figure S6A	Attune NxT	mGL mRuby2	488 561	530 / 30 585 / 16	220 260
Figure 6E; Figures S6F, S6G	Cytoflex S	LIVE/DEAD Violet mGL mRuby2 LIVE/DEAD near-IR	405 488 561 638	450 / 45 525 / 40 585 / 42 780 / 60	178 211 263 150
Figure 6F; Figures S6H, S6I	Attune NxT	mGL mRuby2	488 561	530 / 30 620 / 15	220 320
Figures S1B, S1C, S1J, S1K	Attune NxT	mGL mRuby2	488 561	530 / 30 620 / 15	240 300
Figures S7D, S7E	Attune NxT	EGFP mRuby2 iRFP670	488 561 637	530 / 30 615 / 25 670 / 14	220 300 320

Table S6. Cytometer, channel assignments, and settings for all flow cytometry experiments, related to STAR Methods. Laser and filter wavelengths are in nm, and PMTV values are voltages.

Gene	Forward primer	Reverse primer
EGFP	AGGAGCGCACCATCTTCTTC	TTGTACTCCAGCTTGTGCCC
Fmr1	AGGTGCCAGAAGATTTACGACA	CTCGCTTTGAGGTGACTTCATT
FXN	GGAAACGCTGGACTCTTTAGC	CCAGTTTGACAGTTAAGACACCA
GAPDH	GTATCGTGGAAGGACTCATGAC	ACCACCTTCTTGATGTCATCAT
mRuby2	ATGGTCCCGTGATGCAGAAG	AGAGCAAGACAGATGGCCAC

Table S7. Primer sequences used in RT-qPCR experiments, related to Figures S7K, S7L and STAR Methods.

Supplemental References

1. Tigges, M., Marquez-Lago, T.T., Stelling, J., and Fussenegger, M. (2009). A tunable synthetic mammalian oscillator. *Nature* **457**, 309–312. 10.1038/nature07616.
2. Bleris, L., Xie, Z., Glass, D., Adadey, A., Sontag, E., and Benenson, Y. (2011). Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Mol Syst Biol* **7**, 519. 10.1038/msb.2011.49.
3. Tigges, M., Dénervaud, N., Greber, D., Stelling, J., and Fussenegger, M. (2010). A synthetic low-frequency mammalian oscillator. *Nucleic Acids Research* **38**, 2702–2711. 10.1093/nar/gkq121.
4. Yang, J., Lee, J., Land, M.A., Lai, S., Igoshin, O.A., and St-Pierre, F. (2021). A synthetic circuit for buffering gene dosage variation between individual mammalian cells. *Nat Commun* **12**, 4132. 10.1038/s41467-021-23889-0.