

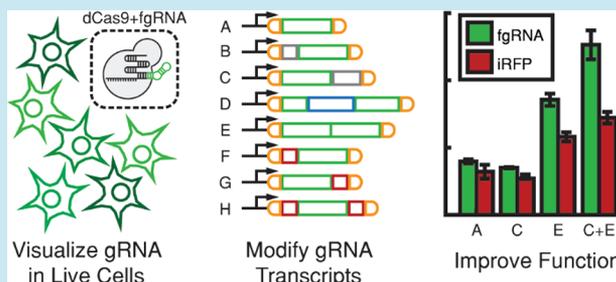
Fluorescent Guide RNAs Facilitate Development of Layered Pol II-Driven CRISPR Circuits

David J. Menn,[†] Swechchha Pradhan,[†] Samira Kiani,^{*,†} and Xiao Wang^{*,†}[†]School of Biological and Health Systems Engineering, Arizona State University, Tempe, Arizona 85281, United States

Supporting Information

ABSTRACT: Efficient clustered regularly interspaced short palindromic repeat (CRISPR) guide RNA (gRNA) expression from RNA Polymerase II (Pol II) promoters will aid in construction of complex CRISPR-based synthetic gene networks. Yet, we require tools to properly visualize gRNA directly to quantitatively study the corresponding network behavior. To address this need, we employed a fluorescent gRNA (fgRNA) to visualize synthetic CRISPR network dynamics without affecting gRNA functionality. We show that studying gRNA dynamics directly enables circuit modification and improvement of network function in Pol II-driven CRISPR circuits. This approach generates information necessary for optimizing the overall function of these networks and provides insight into the hurdles remaining in Pol II-regulated gRNA expression.

KEYWORDS: synthetic biology, CRISPR, fgRNA, genetic circuit, computational biology



Clustered regularly interspaced short palindromic repeats (CRISPR) technology has become a prime candidate for synthetic transcriptional regulation and creation of complex genetic networks due to its programmability, ease of design, and modularity.^{1–3} Originally a bacterial immune system, engineered CRISPR is composed of two parts: the protein Cas9 and a guide RNA (gRNA). The catalytically dead Cas9s (dCas9) have been used for transcriptional repression or activation in various organisms.^{3–5} Because dCas9 can be directed to nearly any region of DNA by changing the sequence of the gRNA, this technology allows for rapid construction of large libraries of activators and repressors which can act orthogonally to one another. Simple design and implementation allows dCas9-based circuits to fill the need for large libraries of components for network construction.

Transcriptional regulation of dCas9 can provide simple circuits with dynamic behavior. However, for circuits that incorporate more than one gRNA, dCas9 acts as a global variable, and its regulation cannot provide targeted modulation. In this case, expression of gRNA from RNA Polymerase type II (Pol II)-based promoters enables generation of sophisticated genetic circuits such as layered CRISPR-based genetic networks in which gRNAs act as individually controllable inputs and outputs.^{6,7} Such circuits will have tremendous value for stepwise or sequential modification of cell-fate or function in cases such as stem cell differentiation or tissue regenerative therapies. Through various RNA-editing methods which release gRNAs from primary RNA transcripts, the expression from Pol II promoters and composability problem of CRISPR-based components can be addressed.^{8–10} However, mRNA production rates from Pol II promoters tend

to be much lower than from Pol III promoters, which could lead to overall low efficiency or even nonfunctional Pol II-driven circuits in mammalian cells.^{7,11} There remains a need for methods to more accurately test, quantify, and optimize these systems by directly analyzing gRNA levels and assessing their impacts on circuit functionality.

gRNA levels in CRISPR circuits have previously been measured indirectly through evaluation of circuit output or a fluorescent protein coexpressed with the gRNA.^{6,7} Development of RNA binding fluorescent probes,¹² fluorescent protein binding RNA aptamers,¹³ and fluorophore-binding RNA aptamers^{14,15} have recently allowed for visualization of RNA and gRNA. It has been shown that modification of the gRNA transcript can be accomplished without destroying gRNA function, allowing insertion of fluorophore-binding RNA aptamers such as Spinach or Broccoli.^{14–18} Although these methods demonstrated feasibility to visualize gRNAs, their application for studying CRISPR-based synthetic gene circuits has not been explored.

Here we set out to address the need for more predictable and reliable Pol II-driven CRISPR circuits by employing a gRNA modified to include the green fluorescent aptamer Broccoli.¹⁵ We demonstrate this strategy can be used to analyze and model circuit behavior. Using gRNA constructs expressed from Pol II promoters, we show that the visualization and analysis of circuit components can be leveraged to improve the function of a layered CRISPR-based circuit composed of only Pol II-driven gRNAs.

Received: April 5, 2018

Published: July 18, 2018

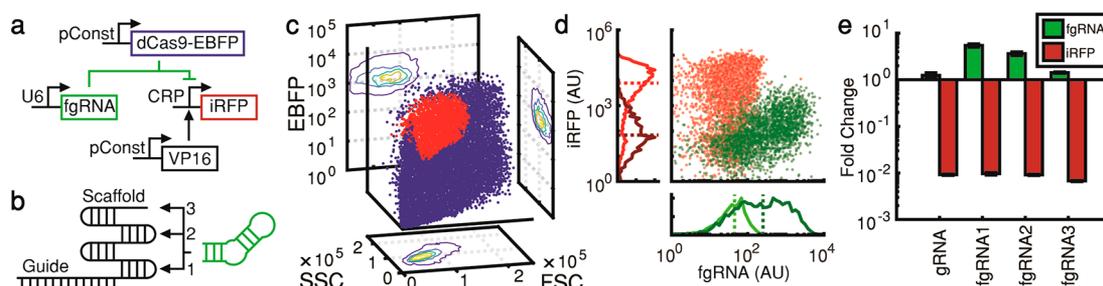


Figure 1. Fluorescent guide RNAs are visible without loss of function. (a) Diagram of the repressor circuit used to test Pol III-driven fgRNAs. The circuit was tested with and without the inclusion of the fgRNA component to calculate fold change in fgRNA and iRFP expression. (b) Guide RNA diagram showing locations into which the broccoli aptamer was inserted into the gRNA scaffold. DFHBI-1T binds to the broccoli aptamer and fluoresces green when excited. (c) Scatter plot of flow cytometry data showing cell size (front scatter; FSC), granularity (side scatter; SSC), and transfection level (EBFP) with the population gated via GMM clustering shown in red. Contour plots of the gated population are shown on the faces. (d) A representative repression experiment showing expression of fgRNA and iRFP. The scatterplot shows fluorescence levels in the absence (red) or presence (green) of fgRNA. Fluorescence channels are displayed independently as histograms on the axes, in the absence (light) or presence (dark) of fgRNA. Median values are indicated with dashed lines. To calculate fold change, medians of the repressed sample were divided by the medians of the unrepressed sample. (e) Fold changes of three fgRNA variants compared to a nonfluorescent gRNA control. Each bar is the mean of 4 flow cytometry replicates' medians \pm SD.

RESULTS AND DISCUSSION

To facilitate construction of diverse CRISPR based circuits, we first developed a fluorescing gRNA construct (fgRNA) without interfering with its downstream functionality. Toward this goal, various constructs were placed in a repressor circuit (Figure 1A), in which the infrared fluorescent protein (iRFP) gene was repressed by gRNA complexed with dCas9 fused to blue fluorescence protein (dCas9-EBFP). The selected fluorophores have wide spectral separation which requires minimal color compensation, though the long maturation time of iRFP necessitated slightly longer times between transfection and data collection. Previous studies suggest three candidate locations in a gRNA structure for additional sequences which do not disrupt guide function: the tetraloop, the second loop, and the 3' end¹⁵ (Figure 1B). As such, the broccoli sequence¹⁵ with a short hairpinning linker was inserted into each of these locations (see Table S1 for sequences).

Circuits were transfected into HEK293FT cells and assessed via flow cytometry after 72 h. We first employed a three-dimensional (3D) gating scheme, utilizing a Gaussian mixed model (GMM; see Methods for full description), to identify cells of interest using three channels: front scatter (FSC), side scatter (SSC), and blue fluorescence (EBFP, representing dCas9-EBFP). This allowed us to accurately separate cells from debris and choose a subset of cells which were both well-transfected (high blue) and of moderate size (Figure 1C). Selecting cells of moderate size allowed us to reduce population variability which might influence the analysis, as cell size is highly correlated with protein production.¹⁹ We then calculated median green and infrared fluorescent intensity in each circuit to analyze the fold change of both fgRNA and iRFP, a metric which we employ throughout this research (Figure 1D,E). Flow cytometry-based analysis reveals that repression of iRFP is strong across all three fgRNAs, with no significant difference between their effectiveness (Figure 1E). Broccoli fluorescence is strongest in fgRNA1 (tetraloop broccoli), with decreasing brightness in fgRNAs 2 (second loop broccoli) and 3 (3' tail broccoli). Because of its superior brightness, fgRNA1 is used in all further experiments, and all fgRNAs referenced hereafter are fgRNA1.

To characterize impacts of component abundances on circuit performances, we then generated dose–response curves

for each of the components within the circuit: fgRNA, dCas9, iRFP, the Gal4-VP16 regulator, as well as the total amount of DNA used in the transfection protocol (Figure 2A and Figure

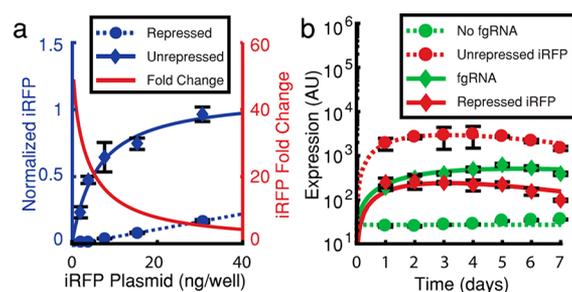


Figure 2. fgRNAs expressed from Pol III a promoter have a predictable response. (a) Normalized dose–response curves of iRFP plasmid (blue), while unrepressed by fgRNA (solid line, diamonds) and while repressed (dashed line, circles). Curves were fit with an exponential function and divided to determine total fold change (red), indicating that a lower amount of plasmid leads to greater differential expression. (b) Flow cytometry time course (points) and model fitting (lines) of the U6-fgRNA repressor circuit. fgRNA (green) and iRFP (red) expression are tracked while fgRNA is either expressed (solid lines, diamonds) or absent (dashed lines, circles). Data in all panels are the mean of 4 flow cytometry replicates' medians \pm SD.

S1). As expected, both Gal4-VP16 and iRFP display decreasing repressibility, showing larger fold changes at lower concentrations. More precisely, when titrating the iRFP plasmid, fold change drops from nearly 60 \times to <10 \times as concentration increases (Figure 2A). Both fgRNA and dCas9-EBFP responses saturate around 20 mM (Figure S1). Based on the dose–response experiments, plasmid concentrations were adjusted to minimal levels (Figure S2, Table S2). This resulted in circuits with much greater iRFP dynamic range despite utilizing smaller amounts of fgRNA repressor, while also decreasing the metabolic load on the cells.

We then analyzed the expression of the optimized circuit over time to quantify the underlying behavior of synthetic CRISPR-based repressors in mammalian cells (Figure 2B). Paired with these experiments, we developed a system of ordinary differential equations (ODEs) to model the network behavior (see Methods for model details and derivation). The

model consists of three ODEs representing overall plasmid levels, fgRNA, and iRFP abundances. Using these three equations, we demonstrate the relationship between fgRNA expression and iRFP regulation and scanned possible parameters governing this relationship (Figures S3 and S4). Analysis of fitted parameters indeed verified a few intuitions. For example, plasmid degradation (pDeg) and fgRNA degradation (fDeg) are inversely correlated (Figure S4), indicating a strict requirement for fgRNA abundance given specific dynamics. Therefore, a decrease of one parameter (pDeg) needs to be compensated by the increase of the other (fDeg). Alternately, positive correlations, such as those between iRFP production (rMax) and iRFP degradation (rDeg) illustrate a need for a ratio between certain paired parameters. Moreover, the analysis reveals that fgRNA production (fMax) has a narrow distribution, suggesting it as the most critical property to tune to achieve desired dynamic behaviors. Finally, this analysis also shows that the nonlinearity coefficient (b) is very close to 1. This lack of nonlinearity indicates weak cooperativity between fgRNA, dCas9, and DNA. This has ramifications for building circuits that demand nonlinearity for their function, such as noise reduction or multistability.^{20–23}

We next utilized inferred information on engineered fgRNA dynamics to develop Pol II-driven fgRNA production (Figure 3A). We evaluated three previously published RNA editing techniques to enable gRNA expression from Pol II promoters. The ribozyme-guide-ribozyme motif (RGR) is an fgRNA flanked by self-cleaving RNA sequences—a Hammerhead (HH) ribozyme on the 5' end and herpes delta virus (HDV) ribozyme on the 3' end—that excise the fgRNA shortly after transcription.⁸ The fgRNAs flanked by Csy4 editing sites (CGC) require exogenous expression of the Csy4 protein, which recognizes and cleaves a 20 nt hairpinning RNA sequencing inserted up- and downstream of the fgRNA.⁹ Because Csy4 cleaves on the 3' end of the hairpin, this method leaves one of the hairpins attached to the tail of the fgRNA transcript. The fgRNA flanked by functional tRNA sequences (TGT) utilizes endogenous tRNA editing proteins RNase P and RNase Z to cut around tRNA sequences placed up- and downstream of the fgRNA, leaving a 1 nt addition to the 5' end of the fgRNA and 6 nt on the 3' end.¹⁰

The circuits used to test these three techniques are designed to utilize tetracycline response element (TRE), a well characterized and widely used inducible Pol II promoter, for fgRNA expression (Figure 3B, only CGC method is shown for illustration). The TRE promoter has previously been shown to be a strong promoter for protein and gRNA expression in synthetic circuits.^{7,24,25} It requires the inclusion of a reverse tetracycline trans-activator (rtTA) protein which, when in the presence of doxycycline (Dox), activates expression of the fgRNA-containing transcript.

First, each of the three editing methods was tested using a Pol III (U6) promoter. It can be seen in Figure S5 that all three editing methods had no significant effects on transcriptional efficiency. Pol II (TRE) promoter-driven versions were then tested to characterize their Pol II expression and inducibility. Figure 3C shows that Pol II expression of CGC was both visible and caused downstream repression. RGR showed little expression, while TGT lacked inducibility, expressing at a high level regardless of Dox concentration. Therefore, CGC editing was used for further parametrization experiments. With the transition to Pol II expression, dose–

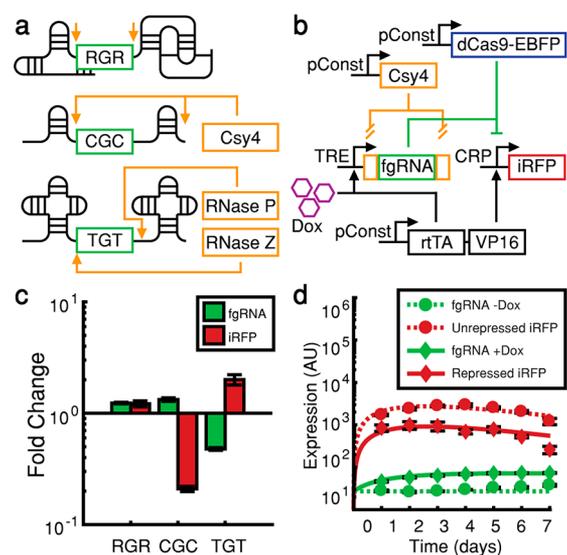


Figure 3. Pol II expression of fgRNAs show predictable dynamics. (a) Schematics for the three editing techniques employed in this work: self-cleaving with the hammerhead (HH) and herpes delta virus (HDV) ribozymes (RGR), targeting and cleavage of 20bp hairpins by the exogenous Csy4 protein (CGC), and excision of the guide by endogenous tRNA cleaving proteins RNase P and RNase Z (TGT). (b) Diagram of the doxycycline-inducible repressor circuit where the fgRNA is driven by a Pol II promoter. In this instance, Csy4 is shown editing the fgRNA transcript, though this component can be replaced with any of the editing methods shown in (a). (c) Fold change of the three editing methods when the fgRNA transcript is expressed from a Pol II (TRE) promoter. When expressed from a Pol II promoter, only the CGC construct shows repressive activity. (d) Flow cytometry time course (points) and model fitting (lines) of TRE-CGC repressor circuit. fgRNA (green) and iRFP (red) expression are tracked while fgRNA is either induced with dox (solid lines, diamonds) or without dox (dashed lines, circles). Data in all panels are the mean of 4 flow cytometry replicates' medians \pm SD.

response curves were generated to optimize the concentrations of Dox, CGC, and Csy4 components (Figure S6) for maximal repression. CGC effectiveness was shown to begin saturating around 75 mM. It was also observed that Csy4 is a very efficient editor, reaching peak effectiveness at 1.5–2 mM. Dox was most effective at a relatively high concentration (2 μ g/mL); however, increasing induction too far beyond this point resulted in cell sickness and network dysfunction. These experiments led us to select component concentrations yielding stronger output dynamic range (Table S2).

We then ran a time course to observe the direct relation between fgRNA and iRFP expression, fitting the results to our ODE model (Figure 3D) and quantifying the underlying parameters determining observed behaviors (Figures S7 and S8). For this application, the ODE model was expanded to include a fourth equation representing the mRNA transcript produced by the Pol II promoter. This transcript is then edited into the functional fgRNA which binds dCas9 and inhibits iRFP production (see Methods for details). Fit values of parameters shared between the Pol II and Pol III circuits—pDeg, fDeg, rMin, rMax, repression coefficient (rK), rDeg, b—are similar between experiments, verifying the model's applicability to both scenarios. Examining parameters for both models allows a quantitative comparison of the promoters used, revealing that the production rate (mMax) from TRE is roughly 10–100 \times weaker than that of U6 (fMax). We also

observed that the degradation rate of mRNA (mDeg) was almost 100-fold higher than fgRNA degradation (fDeg). In fact, the mRNA degradation rate was similar to the rate of editing into gRNA (mEdi), indicating that mRNAs are divided relatively equally between editing and export/degradation and become stabilized once edited into gRNA, possible through complexing with dCas9. Though one might assume that RNA-based regulation would result in fast dynamics, the slow degradation rate of gRNA indicates that the turnover is more similar to that of proteins, and similar expression dynamics should be expected. Engineering faster dynamic systems would require engineering destabilization of the gRNA while simultaneously increasing its expression to maintain a strong expression profile. Taken together, model and guided experiments provide detailed and quantitative characteristics of dCas9-based gene expression regulation dynamics, which are otherwise hard to acquire.

Cascades are a common motif in natural regulatory systems which have been shown to act as noise filters and as memory devices in synthetic networks.^{26–28} The use of fgRNAs in CRISPR-based circuits allows observation of previously hidden nodes whose activity could only be inferred from network inputs and outputs. This improved resolution allows us to more accurately characterize the network's behavior and to troubleshoot more effectively. CRISPR-based layered circuits enable us to leverage the power of CRISPR and combine it with logic-based design methods for sequential gene editing or epigenetic modulation, which will aid in more sophisticated and controllable therapies. However, synthetic layered Pol II CRISPR circuits previously failed to show functionality, so we set out to devise strategies to improve them.⁷

Network analysis revealed that increasing mRNA production or decreasing mRNA degradation were potential targets for improving network response. Therefore, we constructed and screened a number of fgRNA constructs modified to affect these areas (Figures S9 and S10). First, large portions of nontranslated RNA are often found on the 5' end of mRNA transcripts (5' UTR), and it is believed that it plays a regulatory role. It has been shown that the length of the 5' UTR can control the expression level from Pol II and may decrease nucleosome occupancy at the +1 position.²⁹ Therefore, we sought to unravel whether modified UTR length could influence gRNA expression from a Pol II promoter. For this, we inserted an additional random 20 nt sequence within the 5' end of the gRNA transcript, immediately after the transcriptional start, and we compared the efficiency with the original design. Second, the mRNA Poly-A tail is strongly associated with nuclear export.^{30–32} Therefore, we hypothesized that interfering with this component might lead to a larger fraction of mRNA transcripts being retained in the nucleus. A similar approach has been employed to optimize the efficiency of shRNA expression from Pol II promoters.³³ So, we incorporated a truncated minimal poly-A terminator (mPA) to provide a smaller poly-A tail to the transcript and compared the efficiency with transcripts harboring the original Pol II terminator. Additionally, introns have been shown to increase mRNA accumulation when compared to similar transcripts which lack introns.^{34,35} Along this line, a random, intronic 100 bp sequence was added into the middle of the fgRNA sequence. Next, we made several new designs to improve localization to the nucleus. An RNA sequence shown to impart nuclear localization in long noncoding RNAs (lncRNA) was added to either the 5' or 3' end of the mRNA transcript.³⁶

Because this sequence has been shown to reliably localize lncRNAs to the nucleus, it was hypothesized that it could have a similar effect on the fgRNA-containing mRNAs. Finally, gRNAs were multiplexed, placing between 2 and 8 copies of the same fgRNA plus editing sequences one after another in the transcript, each separated by a short linker sequence. We hypothesized this strategy would increase gRNA expression relative to the multiplex number per mRNA transcript.

Applied to both RGR and CGC constructs, the results of this screening are shown in Figures S9 and S10. While the CGC transcript was relatively functional to begin with, additional modifications had little effect. Conversely, initial screening of the RGR construct revealed that both expression and repression were minimal, but modification of the transcript resulted in a much more functional construct. The modifications yielding the greatest effect were addition of the mPA terminator, as well as multiplexing several copies of the fgRNA into a single transcript. These modifications were further analyzed using qPCR, which showed that fgRNA fluorescence is positively correlated with expression level (Figure S11). Sequential addition of guides increased performance up to 4 or 5 guides, at which point continued multiplexing did not appreciably increase expression. Other modifications—changing spacing between promoter and transcript, increasing availability using intronic sequences, and inclusion of lncRNA nuclear localization sequences—resulted in little improvement. As such, further efforts focused exclusively on multiplexing and terminator modification.

Finally, we sought to determine whether these improved Pol II-driven designs were capable of generating a functional two-layer, strictly Pol II gRNA transcriptional cascade. We transfected HEK293FT cells with the cascade circuit (Figure 4A) with and without Dox induction, and measured with flow cytometry 72 h post-transfection. An unmodified, non-fluorescent CGC construct repressed iRFP expression and was itself subsequently repressed by a Dox-responsive modified fgRNA RGR or CGC. To circumvent the lack of nonlinearity, twice as much fgRNA, relative to the middle-node repressor, was added to more efficiently derepress iRFP. Combinations of 3× multiplexing and the mPA terminator were then used for the Dox-responsive input node (Figure 4B). These yielded a moderately functional cascade that was not achievable in previous work using a similar circuit topology.⁷ As with the screening, modified CGC transcripts functioned similarly to unmodified CGCs, while modified RGRs showed marked improvement over the unmodified RGR, with iRFP fold change reaching a level similar to CGC with even greater fgRNA brightness. These results demonstrate that network improvements can be made via targeted re-engineering of circuit components directed by detailed analysis of network behavior. Even in the absence of cooperativity, components can be adjusted to function strongly enough to exhibit cascading behavior. Furthermore, the RGR results show that such modifications can be used to transform a circuit from nonfunctional to functional.

The ability to directly measure gRNA expression enables precise identification of single cell dynamical behaviors of CRISPR-based circuits, enabling informed optimization decisions to improve circuit functionality. We employed the fgRNA technology to interrogate the dynamics and function of otherwise hidden nodes within CRISPR-mediated synthetic gene circuits. Initial validation of fgRNA constructs shows that the placement of the fluorescent broccoli aptamer within the

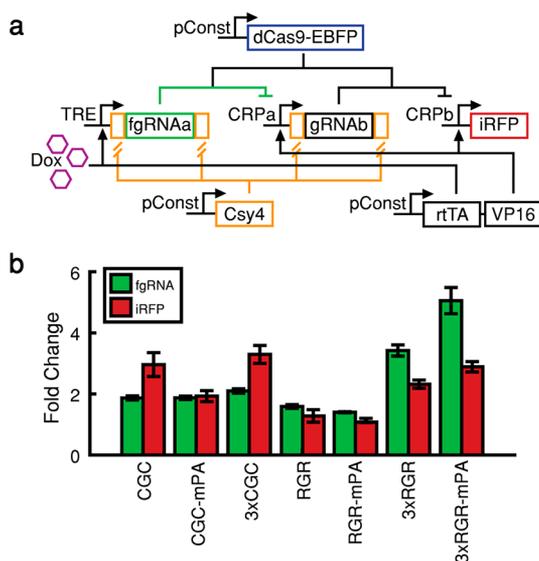


Figure 4. Targeted fgRNA modification improves function. (a) Diagram of the inducible, Pol II-driven two-tier repressor cascade. Csy4 is shown editing the gRNA transcripts in both positions into functional guides. In some experiments, the inducible first node was replaced with an RGR construct. Induction of the circuit with dox should increase fgRNA and iRFP expression. (b) Fold changes of the two-tier cascade with various fgRNA constructs in the first position. By modifying the fgRNA transcript via multiplexing (3×) or alterations to the terminator tail (mPA), fgRNA, and iRFP expression are improved using RGR constructs. Data are the mean of 4 flow cytometry replicates' medians \pm SD.

gRNA transcript has little effect on the function of the gRNA guided repression. The location of the insert only impacts overall fgRNA brightness. It was found that insertion of the aptamer into the gRNA tetraloop produces the highest fluorescence. We hypothesize that insertion into the second loop or tail may result in aptamer misfolding or prevent fluorophore binding. Additionally, inclusion at the tail may be hindered by premature transcriptional termination, resulting in a functional gRNA with an incomplete aptamer. This observation may indicate design constraints for other gRNA-aptamer systems.

gRNA regulated by RNA Pol II promoters provides an attractive platform to generate libraries of composable CRISPR-based gene networks, and thereby enable scaling to more sophisticated circuits. To identify the optimum strategy for gRNA expression from Pol II promoters, we compared three different RNA editing strategies RGR, CGC, and TGT. The CGC-based strategy is shown to be more efficient than the other two in our experiments. Furthermore, while TGT editing resulted in no loss of function from a Pol III promoter, when used in a Pol II context, it exhibited a lack of inducibility. This is likely because tRNAs themselves may act as promoters.³⁷ As tRNAs have been proposed as a means of efficiently multiplexing gRNAs in a single transcript, this is an important consideration for future studies.

As with Pol III, dynamics of the Pol II-driven repressor were evaluated mathematically, yielding several intriguing findings. First, though the editing sequences may interfere somewhat with transcription, CGC editing itself is highly efficient. Second, we confirmed mathematically that Pol II expression is 10–100× lower than Pol III expression. While some of this could be the result of the flanking editing sequences, it also

suggests that the Pol II promoters simply produce fewer transcripts than Pol III promoters. This is likely because mRNA transcripts can be upregulated during the translation stage in normally functioning Pol II expression systems.¹¹ Third, we show that while parameters shared between Pol II and Pol III circuits—pDeg, fDeg, rMin, rMax, rK, b—are centered around the same values as expected, the Pol II model exhibits increased variability around this center, suggesting that Pol II-driven gRNAs may be less well-regulated—in terms of production, degradation, and repression effectiveness—than their Pol III counterparts. This may be due to variability introduced by editing, as that is the primary point of difference between the networks; however, why this would be true mechanistically is unclear. Alternatively, this may be a mathematical artifact due to the more noticeable role of stochasticity within smaller populations. Smaller changes in output, like those seen in the Pol II network, may have more combinations of parameters that still fall within a physiologically relevant range, whereas the larger changes of the Pol III network, tend to group more clearly. If increased variability were to remain after further improving gRNA expression, identifying the source of this variability and ways to control it will be an interesting and necessary route for future experimentation. Finally, our analysis and experiments demonstrate a critical property of CRISPR circuits: dCas9 regulation lacks cooperativity, resulting in a linear relationship between the amount of gRNA-complexed dCas9 and the response of the circuit. Nonlinearity is an essential component of multistable networks and is crucial for noise reduction and maintaining signal fidelity in larger networks. Engineering cooperativity is still a challenge in CRISPR circuit construction and might be accomplished through dimerization of the Cas9 protein or through inclusion of RNA aptamers which allow gRNAs to recruit additional gRNAs.

With a mathematical understanding of the dynamics underlying the fgRNA repressor network, we identified areas—gRNA production and degradation—that we could alter to improve system function to produce a functional transcriptional cascade using only Pol II-driven components. We focused on improving gRNA availability by increasing gRNA production through multimerization and reducing gRNA nuclear export through terminator selection. As we demonstrate, lower Pol II production can be offset through multiplexing several identical gRNAs into the same transcript, though this method runs into limits from transcriptional falloff and plasmid instability due to highly repetitive sequences. Additionally, alterations to the terminator impart increased fgRNA brightness and downstream repression. Through these alterations, we were able to transform the previously nonfunctional Pol II-driven RGR-based CRISPR repressor into one capable of driving downstream derepression in a repressor cascade. Circuit component concentration optimization alone was enough to produce a functional CGC cascade. Thus, we succeeded in developing two editing methods that could produce functional Pol II gRNA transcriptional repression cascades, which was not previously achieved.⁷ While CGC requires coexpression of the Csy4 protein for editing, improved RGR efficiency provides an all-in-one system that works equally well and gives researchers additional flexibility to overcome experimental constraints. Many viral delivery methods, for example, impose a limit on the amount of DNA which can be packaged, so inclusion of an additional protein reduces the available space for therapeutics.³⁸ With a

process as complex as transcription, we are left with myriad angles for potential innovation in this area, such as the inclusion of enhancer sequences within the promoters, optimized RNA Pol II promoter sequences, optimized transcriptional start sites, or improved nuclear localization sequences. We used a mathematical approach to direct circuit modification, but the dissimilar response of RGR and CGC to similar modifications indicates that quantification and standardization of DNA sequence selection and assembly methods remains an area of importance for synthetic biology. There is still enough variability between research groups and experimental methods that approaches which yielded negative results in our hands cannot be rejected outright. Universal standardization of methods for similar studies will aid in better characterization of these networks.

Constructing reliable and predictable gene networks is a nontrivial undertaking. The recent prominence of CRISPR technology promises to improve the process by offering easier generation of unique, orthogonal components and by allowing easier engineering of interactivity between network parts. With the transition to RNA-based transcriptional regulation, however, additional areas require further exploration. Here, we present a tool for visualization of gRNA dynamics within cells and demonstrate how its proper implementation can allow for improved modeling, prediction, and functionality of CRISPR-based gene circuits.

METHODS

Cell Culture and Transfection. All experiments were performed in HEK293-rtTA3 cells (cell line generation detailed by Kiani, et al.⁷), a strain of HEK293FT cells with genomically integrated constitutively active rtTA activator. Cells were maintained in DMEM (Corning Life Sciences) supplemented with 10% FBS (Sigma-Aldrich), 1% nonessential amino acids (NEAA; Gibco), 1% L-glutamine–streptomycin–penicillin mix (Gibco), and 1% GlutaMax (Gibco). Transfections were performed using Polyethylenimine (PEI) as a transfection reagent (Polysciences, Inc.). Cells were seeded in 24-well plates the day before so that they were at ~80% confluence at the time of transfection. Masses of various plasmids used in each set of experiments are shown in Table S2. After DNA mixes were made, the volume was brought to 25 μ L by adding DMEM (no supplements) and then combined with an additional 25 μ L of DMEM (no supplements) with PEI equivalent to a 3:1 DNA:PEI ratio. This was vortexed twice for 1 s each and allowed to stand at room temperature for 30 min. While the DNA mixes were stationary, media was changed (DMEM with supplements above), and if necessary Dox was added to the wells as an inducer. All Dox inductions were performed at a concentration of 2 μ g/mL. After 30 min, 50 μ L of the DNA mixture was added to each well with micropipette, dipping the tip into the well's media and slowly ejecting while swirling inward, careful not to scrape the bottom of the well. Media and inducers were changed daily until analysis.

Plasmids. Plasmids were constructed using golden gate cloning methods, with pieces either copied from existing plasmids via PCR or de novo synthesis. All DNA components were purchased through Integrated DNA Technologies (IDT). The CRP-iRFP reporter plasmid was assembled using gateway cloning, combining the promoter and protein coding region in a gateway destination vector backbone. The Csy4 plasmid, PGK 1p-Csy4-pA (Construct 2), was a gift from Timothy Lu

(Addgene plasmid no. 55196).⁶ The dCas9-EBFP plasmid, pHR-SFFV-dCas9-BFP, was a gift from Stanley Qi and Jonathan Weissman (Addgene plasmid no. 46910).³⁹

Flow Cytometry. Prior to flow cytometry, wells were trypsinized with 100 μ L of 1 \times trypsin (Gibco) and then inactivated with 200 μ L of Hanks Balanced Salt Solution (HBSS; Corning Life Sciences) without calcium or magnesium but supplemented with 2% FBS. These were transferred to a 96-well plate and pelleted at 300g for 2 min at 4 $^{\circ}$ C. The supernatant was aspirated, and cells were resuspended in 200 μ L of phosphate-buffered saline (PBS; Corning Life Sciences) with 4% FBS and 40 mM DFHBI-1T (Lucerna), as recommended in prior literature.¹⁵ Flow cytometry was performed either daily (for time courses) or 72 h post-transfection, using a FACSCelesta flow cytometer (Becton Dickson) with HTS attachment. The cytometer was configured with violet (405 nm), blue (488 nm), and red (640 nm) lasers, used for excitation of EBFP (450/40 filter), Broccoli/DFHBI-1T (530/30 filter), and iRFP (780/60 filter), respectively. Samples were collected at 1.5 μ L/s to a total of 200 000 events.

qRT-PCR. 72 h after transfection, total RNA from cells was isolated using QIAGEN Rneasy Mini Plus kit. For cDNA synthesis, High Capacity Reverse Transcription kit (Applied Biosystems) was used. The reverse transcription conditions were 25 $^{\circ}$ C for 10 min followed by 37 $^{\circ}$ C for 120 min. The reaction was terminated at 85 $^{\circ}$ C for 5 min. qRT-PCR experiments were run on QuantStudio 3 Real Time PCR System (Applied Biosystems). First, 10 μ L of PowerUp SYBR Green Master Mix (Applied Biosystems) was used in each 20 μ L reaction along with 10 ng of cDNA and 1 μ L of each forward and reverse primers at 10uM concentration. fgRNA-specific primer sets used are F 5'-ACGGTCGGGTCCAGATATT-3' and R 5'-ACGGACTAGCCTTATTGAACTT-3'. The PCR thermal cycling conditions were as follows: 50 $^{\circ}$ C for 2 min followed by 95 $^{\circ}$ C for 10 min, and 40 cycles of 95 $^{\circ}$ C for 15 s, and 60 $^{\circ}$ C for 60 s. Relative fgRNA expression was normalized to 18S RNA and was calculated using $\Delta\Delta$ CT method.

Data Analysis. Data were analyzed using MATLAB (MathWorks, Inc.). Gates were generated against a test data set using a Gaussian Mixed Model (GMM) and then applied to all experimental data. The GMM used 6 clusters with 20 replicates, selecting the highest log likelihood. Channels used for gate generation were Front Scatter (FSC-A), Side Scatter (SSC-A), and EBFP (BV421-A) which was our transfection marker. Because the EBFP values were log distributed, we used a Log10 transform of the actual values to fit the GMM. This same transformation was also performed on all experimental data before clustering, then reversed to maintain the original values. Once gated, the median green (BB515-A) and infrared (APC-Cy7-A) fluorescence of all cells with expression >0 was calculated. Fold change of these fluorescences were calculated by dividing the expression with the addition of gRNA/Dox by the expression beforehand.

Modeling. Ordinary differential equation (ODE) models were solved and analyzed using MATLAB run on a personal computer. We designed a system of ODEs to describe the expression of important components in the system. For the U6-driven fgRNA repressor, we began with equations for fgRNA (F , eq 1) and iRFP (R , eq 2), following standard forms for production/degradation and Hill function repression.

$$\frac{dF}{dt} = \text{pol3} * f_{\text{Max}} - F * f_{\text{Deg}} \quad (1)$$

$$\frac{dR}{dt} = \text{reg} * i_{\text{RFP}} * \left(r_{\text{Min}} + \frac{r_{\text{Max}}}{1 + (rK * \text{cas} * F)^b} \right) - R * r_{\text{Deg}} \quad (2)$$

Because of the transient nature of the transfection protocol used, we added another equation to describe plasmid dilution with each subsequent cell division (P , eq 3), which was then integrated into the F and R eqs (eqs 4 and 5). These equations were used for fitting the Pol III experimental data.

$$\frac{dP}{dt} = -p_{\text{Deg}} * P \quad (3)$$

$$\frac{dF}{dt} = (P * \text{pol3}) * f_{\text{Max}} - F * f_{\text{Deg}} \quad (4)$$

$$\frac{dR}{dt} = (P * \text{reg}) * (P * i_{\text{RFP}}) * \left(r_{\text{Min}} + \frac{r_{\text{Max}}}{1 + (rK * (P * \text{cas}) * F)^b} \right) - R * r_{\text{Deg}} \quad (5)$$

To account for both Pol II as well as Pol III RNA production, a fourth equation was added representing mRNA expression (M , eq 6). This equation included an editing term, which converts some portion of M into F , requiring modification of the equation of F to account for this change (eq 7). The modified equations used for fitting the Pol II data were:

$$\frac{dM}{dt} = (P * \text{pol2}) * (m_{\text{Min}} + \text{dox} * (P * \text{reg}) * m_{\text{Max}}) - M * m_{\text{Edi}} - M * m_{\text{Deg}} \quad (6)$$

$$\frac{dF}{dt} = (P * \text{pol3}) * f_{\text{Max}} + M * m_{\text{Edi}} - F * f_{\text{Deg}} \quad (7)$$

Model fitting was performed in MATLAB software using a least-squares curve-fitting algorithm (lsqcurvefit). The fitting was first performed on the Pol III data using eqs 3, 4, and 5 to fit the following parameters: p_{Deg} , f_{Max} , f_{Deg} , r_{Min} , r_{Max} , rK , b , and r_{Deg} . The algorithm was run 500 times with randomly selected initial conditions within physiologically relevant bounds (determined empirically). The fitting rank and squared error (resnorm) are shown in Figure S3. Because there was no clear region of noticeably superior fit, we selected the 250 best fits, as well as a smaller subpopulation of the 50 best fits, for further analysis. Figure S4 shows these fit parameters' distributions as histograms and as scatter plots against one another.

When expanding to the 4-equation Pol II model (eqs 3, 6, 7, and 5) the same fitting algorithm was used on the 11 relevant parameters: p_{Deg} , m_{Min} , m_{Max} , m_{Deg} , m_{Edi} , f_{Deg} , r_{Min} , r_{Max} , rK , b , and r_{Deg} . Again, fittings were ranked by resnorm (Figure S7) and plotted against one another (Figure S8).

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00153.

Supplemental figures S1–S11, expounding on experimental optimization and model fitting methods, and tables S1–S2, detailing DNA sequences used and plasmid masses used for transfection (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: samira.kiani@asu.edu.

*E-mail: xiaowang@asu.edu.

ORCID

Samira Kiani: 0000-0003-2695-1501

Xiao Wang: 0000-0002-4056-0155

Author Contributions

D.M., S.K., and X.W. designed the research. D.M. performed experiments and developed computational models. S.P. performed qPCR experiments. D.M., S.K., and X.W. analyzed the data and wrote the manuscript.

Notes

The authors declare the following competing financial interest(s): a patent application has been filed relating to this work.

■ ACKNOWLEDGMENTS

We thank K. Haynes and M. Ebrahimkhani for helpful suggestions and discussion. D.M. was supported by the ASU Dean's Fellowship. This study was financially supported by an NIH grant (GM106081) (to X.W.) and the ASU School of Biological and Health Systems Engineering (S.K.).

■ REFERENCES

- (1) Ma, M., Ye, A. Y., Zheng, W., and Kong, L. (2013) A Guide RNA Sequence Design Platform for the CRISPR/Cas9 System for Model Organism Genomes. *BioMed Res. Int.* 2013, e270805.
- (2) Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., and Joung, J. K. (2014) Improving CRISPR-Cas Nuclease Specificity Using Truncated Guide RNAs. *Nat. Biotechnol.* 32 (3), 279–284.
- (3) Gander, M. W., Vrana, J. D., Voje, W. E., Carothers, J. M., and Klavins, E. (2017) Digital Logic Circuits in Yeast with CRISPR-DCas9 NOR Gates. *Nat. Commun.* 8, 15459.
- (4) Farzadfard, F., Perli, S. D., and Lu, T. K. (2013) Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. *ACS Synth. Biol.* 2 (10), 604–613.
- (5) Nielsen, A. A., and Voigt, C. A. (2014) Multi-Input CRISPR/Cas Genetic Circuits That Interface Host Regulatory Networks. *Mol. Syst. Biol.* 10 (11), 763.
- (6) Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P., and Lu, T. K. (2014) Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Mol. Cell* 54 (4), 698–710.
- (7) Kiani, S., Beal, J., Ebrahimkhani, M. R., Huh, J., Hall, R. N., Xie, Z., Li, Y., and Weiss, R. (2014) CRISPR Transcriptional Repression Devices and Layered Circuits in Mammalian Cells. *Nat. Nat. Methods* 11 (7), 723–726.
- (8) Gao, Y., and Zhao, Y. (2014) Self-Processing of Ribozyme-Flanked RNAs into Guide RNAs *In Vitro* and *In Vivo* for CRISPR-Mediated Genome Editing: Self-Processing of Ribozyme-Flanked RNAs into Guide RNAs. *J. Integr. Plant Biol.* 56 (4), 343–349.
- (9) Przybilski, R., Richter, C., Gristwood, T., Clulow, J. S., Vercoe, R. B., and Fineran, P. C. (2011) Csy4 Is Responsible for CRISPR RNA Processing in *Pectobacterium Atrosepticum*. *RNA Biol.* 8 (3), 517–528.
- (10) Xie, K., Minkenberg, B., and Yang, Y. (2015) Boosting CRISPR/Cas9 Multiplex Editing Capability with the Endogenous tRNA-Processing System. *Proc. Natl. Acad. Sci. U. S. A.* 112 (11), 3570–3575.

- (11) Brockmann, R., Beyer, A., Heinisch, J. J., and Wilhelm, T. (2007) Posttranscriptional Expression Regulation: What Determines Translation Rates? *PLoS Comput. Biol.* 3 (3), e57.
- (12) Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., and Tyagi, S. (2008) Imaging Individual mRNA Molecules Using Multiple Singly Labeled Probes. *Nat. Nat. Methods* 5 (10), 877–879.
- (13) Ma, H., Tu, L.-C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., and Pederson, T. (2016) Multiplexed Labeling of Genomic Loci with DCas9 and Engineered SgRNAs Using CRISPR-Rainbow. *Nat. Biotechnol.* 34 (5), 528–530.
- (14) Song, W., Strack, R. L., Svensen, N., and Jaffrey, S. R. (2014) Plug-and-Play Fluorophores Extend the Spectral Properties of Spinach. *J. Am. Chem. Soc.* 136 (4), 1198–1201.
- (15) Filonov, G. S., Moon, J. D., Svensen, N., and Jaffrey, S. R. (2014) Broccoli: Rapid Selection of an RNA Mimic of Green Fluorescent Protein by Fluorescence-Based Selection and Directed Evolution. *J. Am. Chem. Soc.* 136 (46), 16299–16308.
- (16) Paige, J. S., Wu, K. Y., and Jaffrey, S. R. (2011) RNA Mimics of Green Fluorescent Protein. *Science* 333 (6042), 642–646.
- (17) Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., Tsai, J. C., Weissman, J. S., Dueber, J. E., Qi, L. S., et al. (2015) Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell* 160 (1–2), 339–350.
- (18) Shechner, D. M., Hacisuleyman, E., Younger, S. T., and Rinn, J. L. (2015) Multiplexable, Locus-Specific Targeting of Long RNAs with CRISPR-Display. *Nat. Nat. Methods* 12 (7), 664–670.
- (19) Lloyd, D. R., Holmes, P., Jackson, L. P., Emery, A. N., and Al-Rubeai, M. (2000) Relationship between Cell Size, Cell Cycle and Specific Recombinant Protein Productivity. *Cytotechnology* 34 (1–2), 59–70.
- (20) Wu, M., Su, R.-Q., Li, X., Ellis, T., Lai, Y.-C., and Wang, X. (2013) Engineering of Regulated Stochastic Cell Fate Determination. *Proc. Natl. Acad. Sci. U. S. A.* 110 (26), 10610–10615.
- (21) Becskei, A., and Serrano, L. (2000) Engineering Stability in Gene Networks by Autoregulation. *Nature* 405 (6786), 590–593.
- (22) Wu, F., Menn, D. J., and Wang, X. (2014) Quorum-Sensing Crosstalk-Driven Synthetic Circuits: From Unimodality to Trimodality. *Chem. Biol.* 21 (12), 1629–1638.
- (23) Wu, F., Su, R.-Q., Lai, Y.-C., and Wang, X. (2017) Engineering of a Synthetic Quadrastable Gene Network to Approach Waddington Landscape and Cell Fate Determination. *eLife* 6, e23702.
- (24) Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., and Bujard, H. (1995) Transcriptional Activation by Tetracyclines in Mammalian Cells. *Science* 268 (5218), 1766–1769.
- (25) Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., Iyer, E. P. R., Lin, S., Kiani, S., Guzman, C. D., Wiegand, D. J., et al. (2015) Highly Efficient Cas9-Mediated Transcriptional Programming. *Nat. Methods* 12 (4), 326–328.
- (26) Nehlin, J. O., Carlberg, M., and Ronne, H. (1991) Control of Yeast GAL Genes by MIG1 Repressor: A Transcriptional Cascade in the Glucose Response. *EMBO J.* 10 (11), 3373–3377.
- (27) Hooshangi, S., Thiberge, S., and Weiss, R. (2005) Ultrasensitivity and Noise Propagation in a Synthetic Transcriptional Cascade. *Proc. Natl. Acad. Sci. U. S. A.* 102, 3581–3586.
- (28) Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., and Collins, J. J. (2009) Synthetic Gene Networks That Count. *Science* 324 (5931), 1199–1202.
- (29) Lin, Z., and Li, W.-H. (2012) Evolution of 5' Untranslated Region Length and Gene Expression Reprogramming in Yeasts. *Mol. Biol. Evol.* 29 (1), 81–89.
- (30) Zenklusen, D., and Stutz, F. (2001) Nuclear Export of mRNA. *FEBS Lett.* 498 (2), 150–156.
- (31) Vinciguerra, P., and Stutz, F. (2004) mRNA Export: An Assembly Line from Genes to Nuclear Pores. *Curr. Opin. Cell Biol.* 16 (3), 285–292.
- (32) Hector, R. E., Nykamp, K. R., Dheur, S., Anderson, J. T., Non, P. J., Urbinati, C. R., Wilson, S. M., Minvielle-Sebastia, L., and Swanson, M. S. (2002) Dual Requirement for Yeast HnRNP Nab2p in mRNA Poly(A) Tail Length Control and Nuclear Export. *EMBO J.* 21 (7), 1800–1810.
- (33) Giering, J. C., Grimm, D., Storm, T. A., and Kay, M. A. (2008) Expression of ShRNA from a Tissue-Specific Pol II Promoter Is an Effective and Safe RNAi Therapeutic. *Mol. Ther.* 16 (9), 1630–1636.
- (34) Xia, H., Mao, Q., Paulson, H. L., and Davidson, B. L. (2002) siRNA-Mediated Gene Silencing in Vitro and in Vivo. *Nat. Biotechnol.* 20 (10), 1006–1010.
- (35) Nott, A., Meislin, S. H., and Moore, M. J. (2003) A Quantitative Analysis of Intron Effects on Mammalian Gene Expression. *RNA* 9 (5), 607–617.
- (36) Zhang, B., Gunawardane, L., Niazi, F., Jahanbani, F., Chen, X., and Valadkhan, S. (2014) A Novel RNA Motif Mediates the Strict Nuclear Localization of a Long Noncoding RNA. *Mol. Cell. Biol.* 34 (12), 2318–2329.
- (37) Thompson, J. D., Ayers, D. F., Malmstrom, T. A., McKenzie, T. L., Ganousis, L., Chowrira, B. M., Couture, L., and Stinchcomb, D. T. (1995) Improved Accumulation and Activity of Ribozymes Expressed from a TRNA-Based RNA Polymerase III Promoter. *Nucleic Acids Res.* 23 (12), 2259–2268.
- (38) Grieger, J. C., and Samulski, R. J. (2005) Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *J. Virol.* 79 (15), 9933–9944.
- (39) Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., et al. (2013) CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154 (2), 442–451.