

# Engineering high-titer lentiviral vectors for robust expression of RNA-based gene circuits

Kasey S. Love<sup>1</sup>, Brittany A. Lende-Dorn<sup>2</sup>, and Kate E. Galloway<sup>2,3,4,\*</sup>

<sup>1</sup>Department of Biological Engineering, MIT, 25 Ames St., Cambridge, MA 02139, USA

<sup>2</sup>Department of Chemical Engineering, MIT, 25 Ames St., Cambridge, MA 02139, USA

<sup>3</sup>Broad Institute of MIT and Harvard, 415 Main St., Cambridge, MA 02142, USA

<sup>4</sup>Koch Institute for Integrative Cancer Research, 500 Main St., Cambridge, MA 02139, USA

\*Corresponding author and lead contact, [katiegal@mit.edu](mailto:katiegal@mit.edu)

## Abstract

Lentiviral vectors enable efficient delivery of genetic cargoes for gene and cell therapies. With their ~10-kb packaging limit, lentiviral vectors can encode multiple transcription units, supporting delivery of compact gene circuits. RNA-based devices offer highly compact control including ligand-responsive induction and closed-loop regulation. However, RNA devices such as ribozymes and splicing switches may interfere with vector production via activity on the single-stranded RNA genome. Here, we examine the impact of gene syntax and genetic parts to define design strategies for two-gene vectors encoding RNA devices. We find that titer decreases with genetic parts that interfere with transcription or processing of the viral transcript during production. Compared to initial vectors, our best-performing design boosts titer more than 30-fold, enabling fine-scale tuning of expression to optimize cell-fate conversion within a non-monotonic landscape. Together, this work illuminates principles for constructing two-gene lentiviral vectors with both high titer and robust expression, enhancing efficacy for downstream applications.

## Introduction

Lentiviral vectors are effective delivery methods for *ex vivo* engineering of cell therapies and *in vivo* gene therapies. Eleven lentiviral-mediated therapies are currently approved by the United States Food and Drug Administration<sup>1</sup>, and hundreds more have been registered in clinical trials<sup>2</sup>. These vectors have broad or programmable tropism and can transduce both dividing and non-dividing cells, enabling targeting to therapeutic cell types<sup>2-6</sup>. Additionally, lentiviral vectors offer stable, long-term expression, minimizing required dosages and enhancing efficacy<sup>5</sup>. Furthermore, the relatively large cargo capacity (~10 kilobases) of these vectors allows for inclusion of multiple transcriptional units to deliver several cargoes or gene circuits<sup>2,7,8</sup>. Gene circuits afford sophisticated control, including autonomous regulation and response to exogenous inputs, that can contribute to greater safety and efficacy of therapeutic vectors.

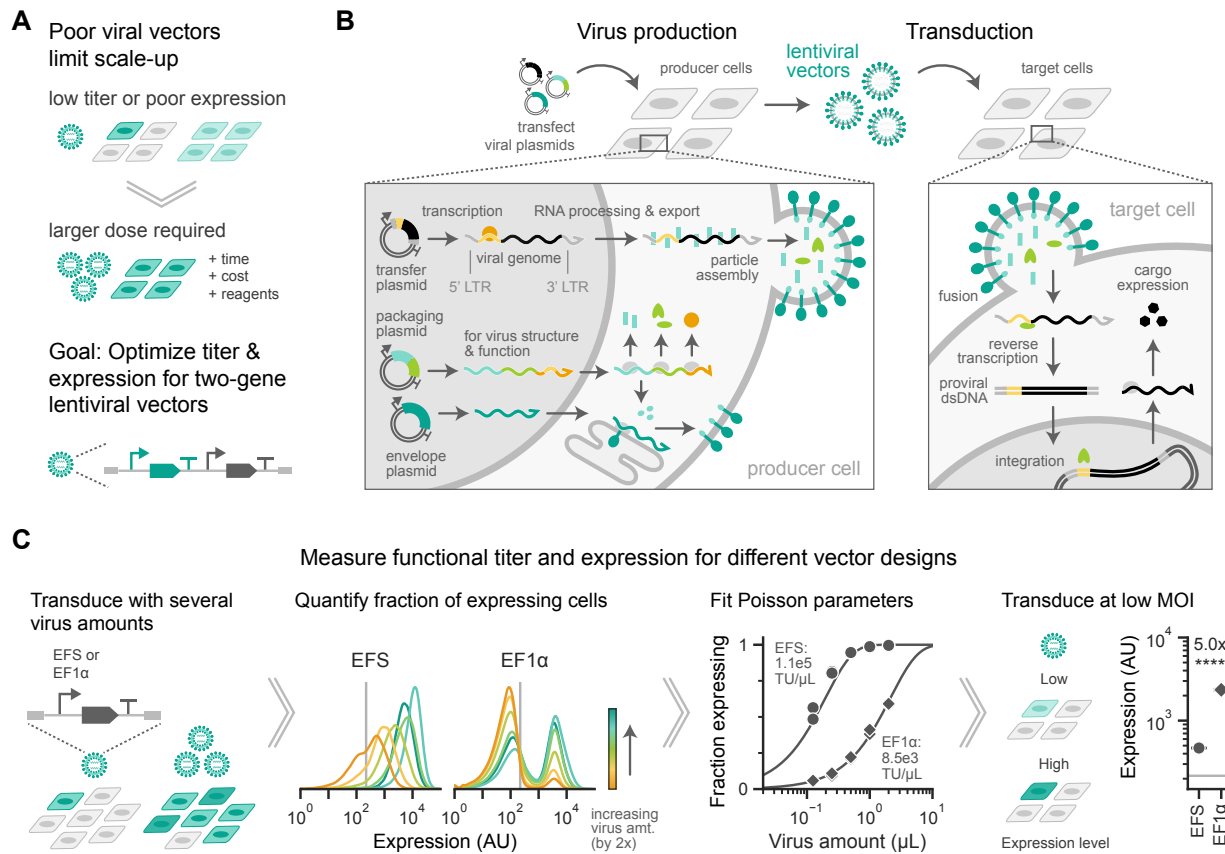
Composing gene circuits with RNA devices is particularly advantageous, as these elements are highly compact, portable, and programmable<sup>9</sup>. RNA devices like microRNAs and ribozyme

switches are typically no more than a few hundred base pairs, contributing a small footprint in size-limited delivery vectors. As these elements may not require exogenous protein components to function, they minimize immunogenicity and toxicity when expressed in target cells. RNA components may be sequence-programmable to interface with or act orthogonally to native networks, and ligand-responsive aptamers can be composed with functional domains via high-throughput design techniques<sup>10-12</sup>. RNA-based circuits with cell type-specificity<sup>13,14</sup> or small-molecule control<sup>15-23</sup> are increasingly being applied in therapeutic contexts. However, RNA-based regulatory elements that include splice sites, strong secondary or tertiary structures, or sequences targeted by machinery in producer cells may disrupt production of viral vectors used for delivery. Thus, there is an opportunity to define principles for designing lentiviral vectors containing these elements to unlock their clinical potential.

Effective lentiviral vectors require both efficient production and robust expression of their cargo<sup>24,25</sup>. If viral titer is low or the cargo expresses poorly, a larger dose of the vector is required to elicit the desired expression level in transduced target cells, limiting vector scale-up (Fig. 1A). Effective vector design may be able to optimize both titer and expression simultaneously, or designs may illuminate tradeoffs for application-specific decisions. These design choices may impact production and expression at several steps during lentiviral vector manufacturing and delivery<sup>26</sup>. Correct processing and packaging of the single-stranded RNA (ssRNA) viral genome in producer cells is essential for reverse transcription, integration, and expression of the desired cargo in transduced target cells (Fig. 1B). Cargo sequences that disrupt any of these steps will reduce production and packaging of full-length genomes<sup>27</sup>. For instance, early termination during transcription of the viral genome in producer cells will generate transcripts that lack components required for integration of the proviral DNA in target cells<sup>28-30</sup>. Cryptic splicing between viral sequences and transgenic elements may produce transcripts that are packaged less efficiently or that integrate an unwanted, modified sequence in target cells<sup>31-35</sup>. Potential integration of unintended sequences is particularly problematic for *in vivo* gene therapy, as aberrant sequences cannot be easily removed. Alternatively, encoding disruptive sequences on the antisense strand of the cargo may limit unintended activity of strand-specific RNA elements on the ssRNA viral genome<sup>31,33</sup>. However, inverting an entire transcriptional unit may interfere with virus production. Strong expression of transcripts that are antisense to the primary viral transcript may inhibit transcription and processing of the desired viral genome or disrupt packaging into viral particles via formation of double-stranded RNA species<sup>36-39</sup>.

To support transgene expression in target cells, vector design must balance tradeoffs between viral titer and expression after integration. For vectors with two transcriptional units, expression is affected by gene syntax, the order and orientation of adjacent genes. We previously demonstrated that divergent syntax can amplify expression from both genes, while tandem syntax can reduce expression of the downstream gene due to DNA supercoiling-mediated feedback<sup>40</sup>. Additionally, readthrough transcription from the upstream gene in tandem syntax can reduce expression of one or both genes<sup>24,41</sup>. On the other hand, inclusion of a polyadenylation signal (PAS) on the upstream gene can mitigate readthrough<sup>42,43</sup> and boost expression levels<sup>44</sup>. As expression of virally integrated transgenes can silence over time, some applications may require inclusion of sequences that help resist silencing<sup>45-47</sup>, even if these reduce viral titer. Thus, identifying tradeoffs in titer and expression is important for negotiating competing design objectives in downstream applications.

Here, we aimed to uncover how design choices affect both titer and expression. In order to optimize these properties, we systematically varied genetic elements and gene syntax in lentiviral vectors containing two transcriptional units. By quantifying both titer and expression, we identified vector designs that function well across cargoes and support expression of RNA devices in gene circuits, including a microRNA-based dosage controller, a ligand-responsive ribozyme switch, and a



**Figure 1. Promoter identity sets tradeoffs between viral titer and expression level of single genes.**

- A.** Viral vectors with low titer or poor expression require larger doses to elicit the desired behavior in target cells. Optimized vectors can improve production scale-up.
- B.** Lentiviral vectors are packaged in producer cell lines transfected with a transfer plasmid to express the viral genome. The viral genome is a single-stranded RNA species spanning from the 5' long terminal repeat sequence (LTR) to the 3' LTR that contains the transgenic cargo along with other viral sequences that facilitate packaging. This transcript is exported from the nucleus and packaged with viral proteins into particles that bud from the membrane of producer cells. The viral particles are collected and then applied to target cells. During transduction, the envelope protein mediates fusion of the particle with the target cell membrane, releasing the particle contents into the cytoplasm. The viral genome is reverse transcribed into proviral double-stranded DNA (dsDNA), which is randomly integrated into the genome of the target cell. The target cell then expresses the transgenic cargo.
- C.** Workflow to measure functional titer and expression for various vector designs. Single-gene vectors containing a fluorescent protein expressed by either the human elongation factor 1 alpha ( $EF1\alpha$ ) or  $EF1\alpha$  short (EFS) promoter were transduced at varying amounts into HEK293T cells. **Left:** Expression of the single-gene vectors measured via fluorescence in arbitrary units (AU, log scale) from a flow cytometer. Light gray vertical line depicts the expression gate, used to calculate the fraction of expressing cells in each condition. Distributions show one representative batch of virus (one biological replicate). **Middle:** For the same vectors, fraction expressing as a function of virus volume was fit to a Poisson distribution to obtain titer in transducing units (TU) per  $\mu\text{L}$  virus. Points represent technical replicates for a representative batch of virus, and lines depict the Poisson curve fits. The computed titer is annotated on the plot, and virus volume is plotted on a log scale. **Right:** To quantify cargo expression, HEK293T cells were transduced at a low multiplicity of infection (MOI). Expression is the geometric mean of expressing cells (AU, log scale). Solid light gray line depicts the expression gate. Points represent means  $\pm$  standard error for  $n = 3$  biological replicates. Statistic is a two-sided Student's t-test, \*\*\*\*  $p < 0.0001$ . Annotation shows fold change between indicated points.

ligand-responsive splicing switch. Our best-performing vector design encodes the RNA devices under an inducible promoter antisense to the viral genome with divergent syntax, increasing titer more than 30-fold and maintaining functionality compared to alternatives arranged in tandem. This design enables fine-scale titration of functional genes in primary cells. By tuning expression of a biphasic cell-fate regulator, we identified the expression level at which we could optimize cell-fate conversion in primary mouse cells, highlighting the utility of our design principles in downstream applications.

## Results

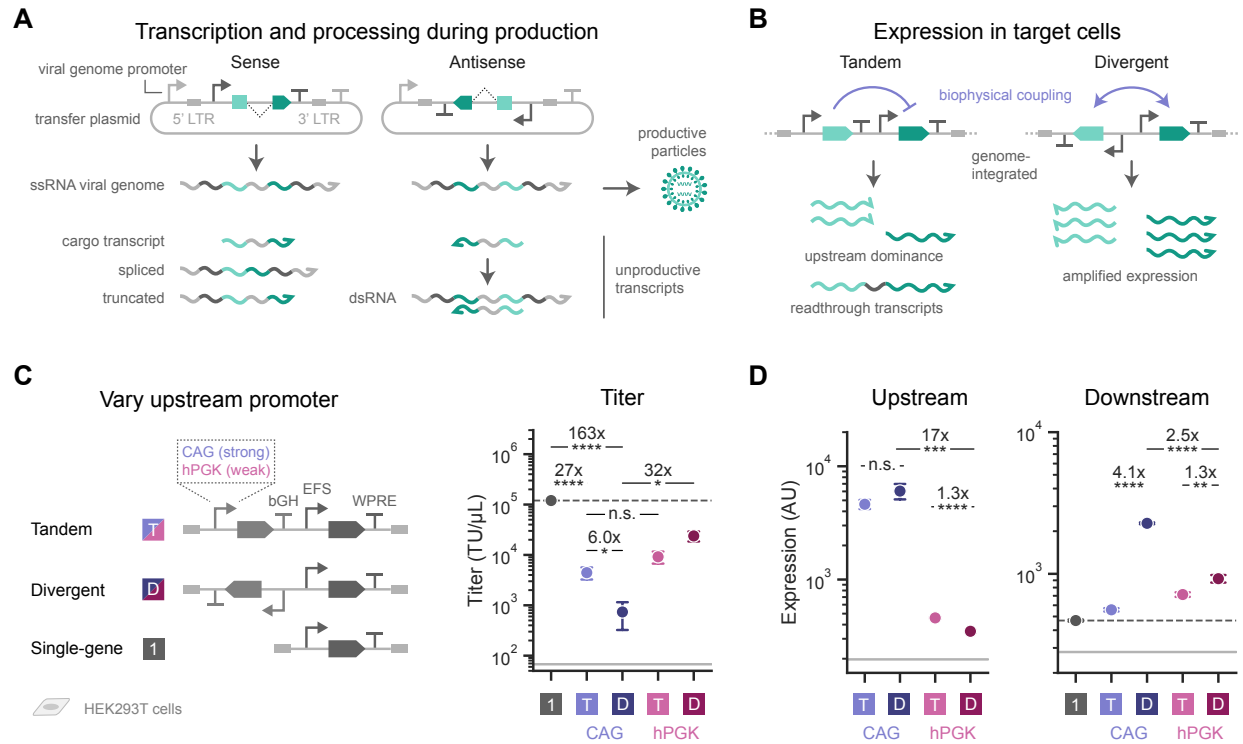
### Promoter identity sets tradeoffs between viral titer and expression level of single genes

We set out to understand design rules for composing two transcriptional units on a single lentiviral vector, beginning with vectors containing a single transcriptional unit to establish a baseline. To quantify the effects of design choices on vector production and expression, we measured both functional titer and expression after single-copy integration in target cells (Fig. 1C). Specifically, following standard virus production in HEK293T cells, we transduced fresh HEK293T cells with varying amounts of concentrated virus and computed the fraction of expressing cells three days later using flow cytometry. Assuming a Poisson process for transduction events, we fit the data to calculate titer in transducing units per volume of concentrated virus. We chose this functional metric for titer to capture both the particle formation and integration processes, as these together are relevant for downstream applications. Then, using these titer values, we transduced new HEK293T cells at a standard multiplicity of infection (MOI) of 0.3 to ensure single-copy integration and quantified expression level in transduced cells. Using fluorescent reporters, we could compare both viral titer and cargo expression across vector designs.

Promoter choice represents an important design consideration for specifying gene expression at the RNA and protein levels<sup>48</sup>. We first compared two commonly used promoters, measuring their effects on titer and expression. For single-gene vectors expressing a fluorescent protein from the human elongation factor 1 alpha (EF1 $\alpha$ ) promoter or a shorter, intronless variant (EF1 $\alpha$  short, EFS), we observed a more than 10-fold difference in functional titer between the designs (Fig. 1C). Potentially, the intron sequence associated with EF1 $\alpha$  interferes with processing of the viral genome. Despite its lower titer, the vector encoding EF1 $\alpha$  led to five-fold higher expression compared to the EFS design at the same MOI (Fig. 1C), reflecting previously reported differences in expression strength between these two promoters<sup>48</sup>. Thus, we find that promoter identity can set tradeoffs between viral titer and expression level of single genes.

### Promoter strength and gene syntax define the tradeoffs in two-gene vectors

For two-gene vectors, gene syntax and genetic parts affect the RNA species generated in both producer and target cells. The orientation of cargo genes relative to the viral genome affects transcription and RNA processing during virus production (Fig. 2A). In producer cells, a promoter on the transfer plasmid upstream of the 5' LTR initiates transcription of the viral genome, which can be packaged into productive particles<sup>49</sup>. However, genetic elements on the sense strand can result in splicing, truncation, or other aberrant processing of the viral transcript, decreasing production of particles carrying the correct sequence<sup>28-35</sup>. In addition to the full-length, ssRNA viral genome (5' LTR to 3' LTR), promoters within the cargo of the vector may generate other transcripts. Antisense transcripts can interfere with transcription or processing of the viral genome or form double-stranded RNA species that inhibit packaging<sup>36-38</sup>. For two-gene vectors, the orientation and



**Figure 2. Promoter strength and gene syntax define tradeoffs in two-gene vectors.**

**A.** Schematic of the RNA species transcribed in producer cells for vectors with cargo on the sense or antisense strand relative to the viral genome. The transfer plasmid expresses the viral genome, as well as cargo transcripts and spliced or truncated forms of the viral transcript. The viral genome can be packaged into productive particles, while the other species are unproductive, meaning they do not form viral particles that integrate the intended sequence into target cells. Additionally, antisense cargo transcripts can form double-stranded RNA (dsRNA) by binding to the viral genome, potentially inhibiting proper packaging.

**B.** Expression in target cells is influenced by biophysical coupling defined by the syntax of the encoded genes, resulting in attenuated or amplified expression.

**C.** Schematic of two-gene vectors with tandem or divergent syntax that express a fluorescent protein from the strong, synthetic CAG promoter or from the relatively weak human phosphoglycerate kinase (hPGK) promoter, and a second fluorescent protein from a downstream EFS promoter. These vectors are compared to a single-gene vector containing only the downstream gene from Fig. 1. The upstream gene includes the bovine growth hormone (bGH) polyadenylation signal (PAS), and the downstream gene includes the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) in the 3' UTR. Viral titer is in transducing units (TU) per  $\mu$ L (log scale).

**D.** Expression of the upstream and downstream genes for each vector in HEK293T cells. Expression is the geometric mean fluorescence in arbitrary units (AU, log scale).

Solid light gray lines represent titers calculated for untransduced cells or expression gates. Dashed dark gray lines depict values for the single-gene vector for reference. Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points.

order of cargo genes (i.e., gene syntax) affects expression after integration (Fig. 2B). In target cells, genes with tandem syntax may exhibit supercoiling-mediated upstream dominance<sup>40</sup> or readthrough transcription<sup>24,42,43</sup>, decreasing expression of the downstream gene. In contrast, divergent syntax may amplify expression of both genes through biophysical coupling<sup>40</sup>. For both syntaxes, choice of genetic elements alters transcript processing, stability, and translation to impact expression<sup>42,43,48</sup>. Given these diverse processes, we sought to identify how gene syntax and components together influence both vector production and expression.

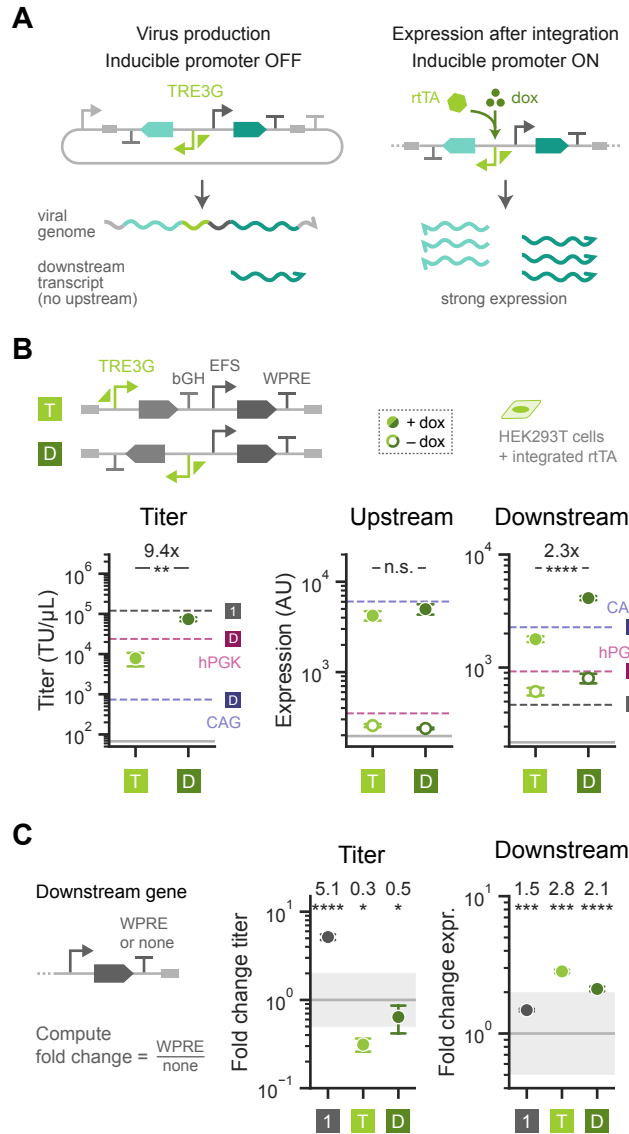
To build two-gene vectors, we added to the single-gene vectors an upstream gene expressing a fluorescent protein under control of the strong, synthetic CAG promoter. We encoded the two genes with tandem or divergent syntax and measured both titer and expression (Figs. 2C, S1A, S1B). Compared to the best-producing single-gene vector, the two-gene tandem vector generates 27-fold lower titer with either the EFS or EF1 $\alpha$  promoter in the downstream gene (Figs. 2C, S1B). For divergent syntax, titer drops an additional six-fold, a decrease of more than two orders of magnitude compared to the best single-gene vector (Fig. 2C). We hypothesize that with divergent syntax, strong antisense transcription from the upstream gene interferes with production of the viral genome.

In contrast, when transduced at constant MOI, the divergent vector results in four-fold higher mean expression of the downstream gene compared to the tandem vector and at least 1.8-fold higher expression than the single-gene vectors (Figs. 2D, S1B), matching predictions of transcriptional amplification. For all designs, the EF1 $\alpha$  promoter produces higher expression of the downstream gene compared to the EFS promoter, as expected (Fig. S1B). Finally, expression of the CAG-driven upstream gene is similar for the divergent and tandem vectors (Figs. 2D, S1B). Thus, there exists a tradeoff between titer and expression level for these vectors, with the divergent design expressing most strongly but producing least efficiently.

We hypothesized that replacing CAG with a weaker upstream promoter could restore titer for the divergent design. Indeed, using a promoter derived from the human phosphoglycerate kinase gene (hPGK), which drives lower expression<sup>48</sup>, titers of divergent two-gene vectors increase by more than an order of magnitude (Figs. 2C, S1B). Meanwhile, the titer of tandem designs do not increase, and none of the two-gene designs are produced as efficiently as the best single-gene vector. However, as expected, replacing CAG with hPGK reduces expression of the upstream gene for both syntaxes (Figs. 2D, S1B), further underscoring the tradeoff between viral titer and upstream expression.

## An inducible upstream promoter resolves tradeoffs in vector performance

To overcome these tradeoffs, an ideal vector would provide strong expression in target cells while limiting interference in producer cells. An inducible promoter fulfills these criteria. In the absence of inducer, the upstream gene generates low, basal levels of transcription in producer cells, minimally disrupting production of the viral transcript (Fig. 3A). In target cells, addition of the inducer turns on the gene of interest after transduction. We chose the widely used Tet-ON system for the inducible gene, in which the TRE3G promoter is activated by the small molecule doxycycline (dox). Binding of dox to the reverse tetracycline transactivator (rtTA) protein allows rtTA to bind to the TRE3G promoter, inducing transcription. To measure performance of these vectors, we separately integrated an rtTA-expressing cassette into HEK293T cells. For the divergent vector, replacing the hPGK promoter with TRE3G increases the viral titer three-fold, nearing the level of the single-gene vector and remaining higher than the titer for the tandem vector (Fig. 3B). Expression of the upstream gene is an order of magnitude larger with TRE3G upon induction than with hPGK for both syntaxes, reaching the level achieved by the CAG promoter (Fig. 3B). Across all



**Figure 3. An inducible upstream promoter can resolve tradeoffs in vector performance.**

**A.** Schematic of transcripts produced from a divergent vector with the upstream, antisense gene driven by a Tet-ON inducible promoter (TRE3G). In producer cells cultured without inducer, the viral genome is transcribed from the transfer plasmid, and some unproductive downstream transcripts may also be produced. After integration in target cells, addition of the inducer doxycycline (dox) and the reverse tetracycline transactivator (rtTA) protein leads to strong expression of the inducible gene, along with strong expression of the downstream gene due to biophysical coupling.

**B.** Two-gene vectors with an inducible upstream gene and tandem (T) or divergent (D) syntax were delivered to HEK293T cells containing a separately integrated rtTA cassette. Titer is in transducing units (TU) per  $\mu\text{L}$  (log scale), and expression is the geometric mean fluorescence in arbitrary units (AU, log scale). Cells were treated with 1  $\mu\text{g}/\text{mL}$  dox (filled points) or untreated (open points). Solid light gray lines represent titers calculated for untransduced cells or expression gates. Dashed lines depict values for several vectors from Fig. 2 for reference. Annotations show the fold change between indicated points.

**C.** Performance of vectors in B is compared to vectors lacking WPRE in the 3' UTR of the downstream gene. Titer and expression of the downstream gene for the vectors with WPRE are normalized to values for vectors without WPRE, plotted on a log scale. Solid gray line indicates a fold change of 1, and gray shading spans a two-fold change in either direction (0.5 to 2) for reference. Annotations show the values of the points, and statistics compare unnormalized values.

Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ .

metrics, the inducible divergent vector matches or exceeds performance of the corresponding single-gene and tandem vectors, resolving the tradeoff between titer and expression.

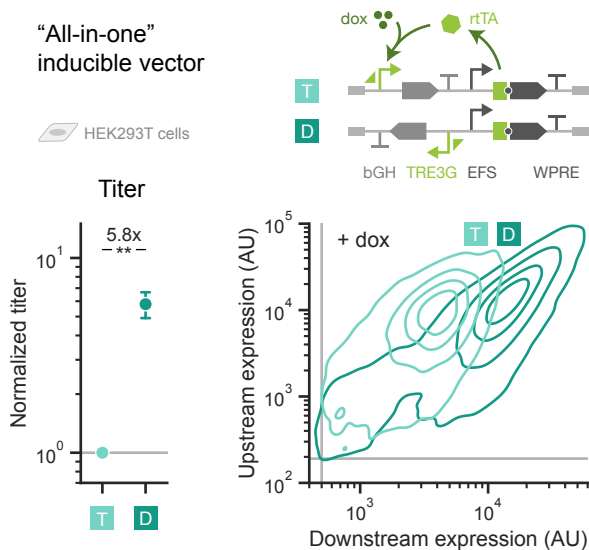
Having identified a promoter combination that optimizes viral titer and expression of both cargoes, we next sought to understand how elements in the 3' UTR such as a PAS affect vector performance. In our initial testing, we found that for tandem syntax, the choice of upstream PAS unexpectedly has little impact on viral titer and upstream expression (Fig. S2A). However, for the divergent vectors, including the bGH PAS moderately improves both metrics, though the presence of the bGH PAS on both the upstream and downstream genes abolishes virus production (Fig. S2A). For the downstream gene, our initial vector designs contained the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) in the 3' UTR without a subsequent PAS. WPRE is commonly included in viral vectors to enhance expression, supporting several RNA processing steps including nuclear export and transcript stability<sup>50,51</sup>. Consistent with these findings, addition of the WPRE results in about a two-fold increase in expression of the downstream gene for both single-gene and inducible two-gene vectors (Figs. 3C, S2A, S2B). Across all vectors, WPRE in

combination with a subsequent PAS has mixed effects on titer (Figs. 3C, S2A, S2B). In particular, the wide variation in titer for vectors with the EF1 $\alpha$  promoter may suggest the appearance of emergent effects between 3' elements and the EF1 $\alpha$  intron. Overall, WPRE consistently affords a small boost in expression of the downstream gene with small effects on titer, and addition of a subsequent PAS does not broadly improve vector performance.

### “All-in-one” vectors with divergent syntax offer high titer and inducible control

While an inducible upstream gene elicits both high titer and strong expression, this design relies on the presence of the associated activator protein. Instead of co-transducing a separate vector containing the activator, which may reduce overall delivery efficiency, the activator can be encoded within the constitutively expressed gene on the same vector to generate an “all-in-one” design<sup>52</sup>. We constructed “all-in-one” Tet-ON vectors with tandem and divergent syntaxes, expressing rtTA alongside the downstream fluorescent protein separated by a “self-cleaving” 2A peptide (Fig. 4). As expected, both titer and expression of each gene match values for the inducible two-gene vectors lacking the rtTA (Fig. S3). In the absence of inducer, expression of the upstream gene is minimal (Fig. S3B). Both syntaxes support robust induction (Figs. 4, S3B), consistent with our prior investigation of gene syntax in diverse cell types<sup>40</sup>. However, with these genetic parts, the divergent design increases titer five-fold over the tandem vector (Figs. 4, S3A).

To explore the generality of the “all-in-one” design, we varied the upstream PAS and downstream promoter. Divergent vectors have higher upstream expression but similar titer with a bGH PAS versus no PAS, while tandem designs have higher titer but similar expression with no PAS for downstream EFS and hPGK promoters (Fig. S3). With a downstream EF1 $\alpha$  promoter, trends are less clear for the tandem vectors, similar to the variation observed for the single-gene vectors (Fig. S2A). Additionally, for both divergent and tandem vectors, expression of the downstream gene is similar with either the EFS or EF1 $\alpha$  promoter (Fig. S3B). This stands in contrast to the previous and expected observation of higher expression with EF1 $\alpha$  in single-gene and two-gene vectors lacking the rtTA (Figs. 1C, S1B, S2A, S2B). These results suggest that the effects of the EF1 $\alpha$  promoter may be less predictable than those of promoters lacking an intron (e.g., EFS, hPGK). Nevertheless, the divergent vector with the bGH PAS in the upstream gene is the best-performing design for each downstream promoter.



**Figure 4. “All-in-one” vectors with divergent syntax offer high titer and inducible control.**

An “all-in-one” inducible vector contains both an inducible gene and a constitutive gene that expresses the required activator, which can be arranged with tandem (T) or divergent (D) syntax. Here, the dox-inducible TRE3G promoter (Tet-ON system) controls expression of a fluorescent protein in the upstream gene, and in the downstream position EFS drives expression of rtTA and a second fluorescent protein, separated by a “self-cleaving” 2A peptide. The upstream gene includes the bGH PAS, and the downstream gene includes the tandem vector to the tandem design. **Left:** Viral titer is shown normalized to the tandem vector for each batch of virus. Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistic is a two-sided Student’s t-test, \*\*  $0.05 \leq p < 0.01$ . **Right:** 2D density distributions depict upstream and downstream expression via fluorescence in arbitrary units (AU, log scale) for a representative biological replicate of HEK293T cells transduced with each vector and treated with 1  $\mu\text{g}/\text{mL}$  dox. Populations are gated on cells expressing both genes. Solid gray lines show expression gates.

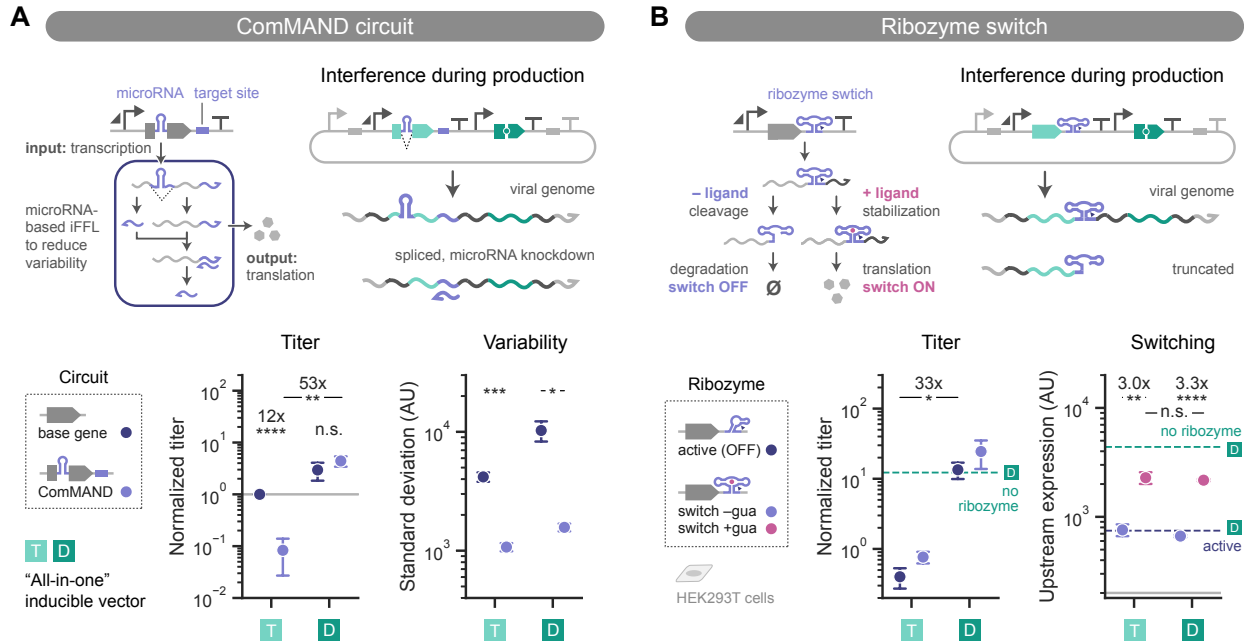
## The divergent “all-in-one” design effectively encodes diverse RNA-based gene circuits

Although some tandem “all-in-one” designs perform well, we hypothesized that inclusion of RNA devices such as microRNAs, ribozymes, and splicing switches in the sense direction of the viral transcript could alter RNA processing and decrease production efficiency. In contrast, encoding these elements on the antisense strand (i.e., in the upstream gene with divergent syntax) would preclude device activity during production. To test this hypothesis, we first delivered ComMAND, a post-transcriptional circuit that controls for gene dosage<sup>53</sup>. ComMAND uses a synthetic microRNA within an intron of the output coding sequence to target a matched microRNA target site in the 3' UTR of the same transcript, resulting in an incoherent feedforward loop that buffers variation in expression. We expressed the circuit in the upstream gene of the best-performing divergent and tandem “all-in-one” vectors and compared these designs to vectors containing only the output coding sequence (denoted as “base gene”) and to an open-loop circuit containing the microRNA and an orthogonal target site (Figs. 5A, S4). Compared to the base gene vector, both the open-loop circuit and ComMAND decrease titer for vectors with tandem syntax (Figs. 5A, S4A). In contrast, for divergent syntax, titer remains the same across circuits, resulting in a 50-fold greater titer for the ComMAND vector with divergent syntax over the tandem design. Meanwhile, syntax minimally impacts circuit function: both the tandem and divergent designs show comparable reduction in expression variability with ComMAND relative to the base gene and open-loop circuit (Figs. 5A, S4B–S4D).

To assess the generality of this syntax-specific performance, we constructed tandem and divergent “all-in-one” vectors expressing a ligand-responsive ribozyme switch<sup>54</sup> in the 3' UTR of the upstream gene (Figs. 5B, S5). In the absence of ligand, the ribozyme element in the switch self-cleaves, decreasing transcript stability and thus protein expression. Binding of the ligand to an aptamer in the switch alters the conformation of the ribozyme, reducing cleavage rates and increasing protein expression. As controls, we replaced the ribozyme switch with a constitutively active ribozyme (protein expression OFF) or a mutant, minimally cleaving ribozyme (protein expression ON). Analogous to our observations for ComMAND, titer remains similar across vectors with divergent syntax but decreases with increasing ribozyme activity for tandem syntax (Figs. 5B, S5A). Furthermore, upstream expression and switch performance are the same between syntaxes (Figs. 5B, S5C). Together, divergent syntax improves production of lentiviral vectors containing RNA devices such as microRNAs and ribozymes while maintaining function of the RNA-based circuits in target cells.

## Divergent syntax supports small-molecule tuning of expression from a splicing switch

Observing that both syntaxes lead to strong circuit performance for ComMAND and the ribozyme switch, we next explored how the performance of a splicing switch is affected by lentiviral delivery. Encoded between exons of a gene of interest, the splicing switch consists of introns flanking an alternative exon with a premature stop codon. In the absence of the ligand guanine, splicing includes the alternative exon in the mature mRNA, leading to truncated protein (expression OFF). Binding of guanine to an aptamer in one of the introns alters the RNA structure and changes the splicing pattern to exclude the alternative exon, resulting in production of full-length protein (expression ON) (Fig. 6A). The amount of guanine sets the ratio of these splice products in cells to tune protein levels. We found that “all-in-one” vectors encoding this RNA device have similar titers for the best tandem and divergent designs, though addition of an upstream bGH PAS with tandem syntax abolishes virus production (Figs. 6A, S6A). However, the tandem vector without an upstream PAS has no switch function, despite its efficient delivery as evidenced by expression of the



**Figure 5. The divergent “all-in-one” inducible vector effectively delivers RNA-based circuits.**

**A.** ComMAND is a compact, post-transcriptional circuit that reduces dosage-associated variability in expression. ComMAND uses a synthetic microRNA within an intron of the output coding sequence to target a matched microRNA target site in the 3' UTR of the same transcript, resulting in an incoherent feedforward loop. Because ComMAND operates at the RNA level, circuit activity can interfere with virus production when encoded on the sense strand of the viral genome. “All-in-one” inducible vectors as in Fig. 4 were constructed with ComMAND (light purple) or the output gene alone (base gene, dark blue) as the upstream cargo. **Left:** Titer is normalized to the base gene vector with tandem syntax for each batch of virus (log scale). **Right:** Standard deviation of circuit (upstream gene) expression measured via fluorescence in arbitrary units (AU, log scale).

**B.** A ribozyme switch, consisting of a ligand-responsive RNA aptamer fused with a self-cleaving ribozyme, controls expression of an output gene when placed in the 3' UTR. In the absence of ligand, the ribozyme switch cleaves, leading to degradation of the mRNA and a reduction in protein levels (OFF). Binding of the ligand changes the folding of the switch to minimize ribozyme cleavage and restore protein expression (ON). When encoded on the sense strand of the viral genome, ribozymes can interfere with virus production by cleaving and truncating transcripts. “All-in-one” inducible vectors as in Fig. 4 were constructed with a guanine (gua)-responsive ribozyme (not shown) or a constitutively active ribozyme added to the upstream gene. **Left:** Titer is normalized to a vector lacking a ribozyme (not shown) for each batch of virus (log scale). **Right:** Upstream expression is the geometric mean fluorescence (AU, log scale) of cells transduced with vectors containing the ribozyme switch and treated without (light purple) or with (pink) 100  $\mu$ M guanine. Solid light gray line shows the expression gate. Dashed teal lines depict values for a divergent vector lacking a ribozyme, and dashed dark blue line represents the value for the divergent vector with the constitutively active ribozyme.

To compare optimal designs for each syntax, tandem vectors lack a PAS on the upstream gene, while divergent vectors include a bGH PAS. In all plots, HEK293T cells were transduced and treated with 1  $\mu$ g/mL dox. Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points.

downstream gene (Fig. 6A, Figs. S6A–S6C). Target cells transduced with this vector express the upstream gene in the absence of guanine, and addition of guanine does not change these levels. In contrast, the divergent vector results in upstream expression that is minimal in the absence of guanine and that increases 100-fold upon induction (Figs. 6A, S6A, S6C), as intended for the switch.

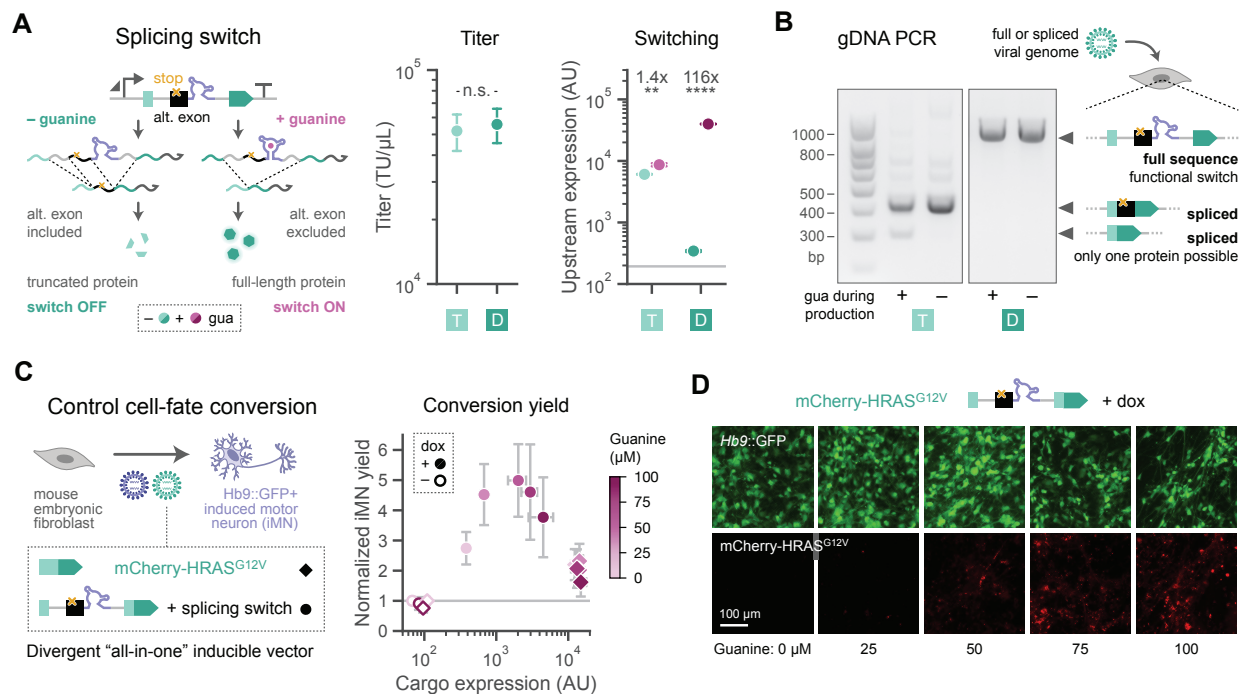
We hypothesized that the lack of inducibility for the tandem design results from integration of an incomplete splicing switch (Fig. S6D). To investigate the integrated sequence, we used PCR to amplify genomic DNA isolated from cells transduced with each vector. The single amplicon for the divergent vector matches the expected size for the construct containing the full-length splicing switch, while conditions with the tandem vector have multiple amplicons (Figs. 6B, S6E). For the tandem vector, the amplicon at the full-length size is absent, and the most prominent band on the

gel matches the splice product containing the alternative exon but lacking the intronic regions. This suggests that the viral transcript undergoes splicing in producer cells, packaging sequences that lack essential switch components. To test this hypothesis, we treated producer cells with guanine during virus production, reasoning that this would activate the splicing switch and generate viral transcripts without the alternative exon (Fig. S6D). Indeed, performing the same PCR on gDNA from cells transduced with these vectors resulted in an amplicon corresponding to the protein coding sequence without the entire switch (Figs. 6B, S6E). Further, for these conditions, expression increases in target cells without restoring responsiveness to guanine (Figs. S6A–S6C). Thus, during production of tandem vectors, the splicing switch is active and changes the splice products of the viral transcripts that are packaged into particles. Conversely, addition of guanine during production of the divergent vector has no impact on gDNA PCR amplicons, expression, or switch activity (Figs. 6A, 6B, S6). Thus, divergent syntax supports production and delivery of a functional splicing switch on a lentiviral vector.

Having identified a high-titer, guanine-responsive vector design, we used the splicing switch to titrate expression of a conversion factor during direct conversion of primary mouse embryonic fibroblasts to induced motor neurons (iMNs) (Figs. 6C, S7A, S7B). High rates of conversion rely on delivery of a high-efficiency cocktail that combines with neuronal transcription factors to drive cells to adopt a motor neuron identity<sup>55–57</sup>. Within this cocktail, we had previously identified that an intermediate level of the HRAS mutant HRAS<sup>G12V</sup> is optimal for conversion<sup>58</sup>. To titrate expression, we had used an array of techniques including varying lentiviral dosage, cassette design, promoter usage, and promoter setpoint editing<sup>58,59</sup>. In contrast, we hypothesized that the splicing switch would allow us to easily modulate HRAS<sup>G12V</sup> expression via guanine addition while maintaining identical transduction conditions and a constant MOI. As a control, we also constructed a vector expressing HRAS<sup>G12V</sup> without the splicing switch. As desired, titrating the amount of guanine unimodally shifts expression of the cargo over an order of magnitude in converting cells (Figs. 6C, 6D, S7C–S7E). For the same conditions, we measured conversion yield by counting the number of cells expressing *Hb9::GFP* per cell seeded and normalizing to controls without overexpression of HRAS<sup>G12V</sup>. Matching previous findings<sup>58</sup>, we determined that intermediate concentrations of guanine and levels of HRAS<sup>G12V</sup> expression optimize iMN yield (Figs. 6C, 6D, S7E). Together, the divergent “all-in-one” vector incorporating the splicing switch provides a robust and practical approach for tuning expression of a functional cargo in primary cells.

## Discussion

In this work, we sought to define design rules for constructing lentiviral vectors to enhance delivery and control of genetic cargoes in mammalian cells. We focused on vectors containing two transcriptional units that incorporate RNA devices and circuits, as these tools offer sophisticated control over cargo expression. Measuring functional titer reveals effects of vector design across the virus lifecycle, from production and packaging of the viral transcript to integration in target cells. Production efficiency is crucial for downstream applications of lentiviral vectors, as it impacts manufacturing scale-up. We observed tradeoffs between viral titer and functional expression for some designs. We addressed this challenge by expressing one of the genes from an inducible promoter on the antisense strand of the viral transcript (i.e., with divergent syntax), and constructed effective “all-in-one” vectors that also express the required transactivator. These vectors are similar to some reported in the literature<sup>60</sup>; however, previous studies do not compare gene syntax in combination with genetic parts. Studies examining the effects of genetic parts primarily focus on expression in target cells and often lack characterization of production efficiency. We



**Figure 6. Divergent syntax supports small-molecule tuning of expression from a splicing switch.**

**A.** A guanine-responsive splicing switch consists of two introns and an alternative exon (alt. exon) containing a stop codon. Addition of the splicing switch within the coding sequence of a gene of interest enables tunable control of protein expression. In the absence of guanine (gua), the main spliced isoform includes the alternative exon, leading to truncated protein product. In the presence of guanine, the RNA aptamer changes structure and alters the splicing pattern to exclude the alternative exon and enable production of full-length protein. Switches are encoded within the upstream gene of the best "all-in-one" inducible vectors identified previously (Fig. 4). The tandem vector (T) does not include a PAS on the upstream gene while the divergent vector (D) includes the bGH PAS. **Left:** Titer is measured in transducing units (TU) per  $\mu$ L (log scale). **Right:** Upstream expression is geometric mean fluorescence in arbitrary units (AU, log scale) for HEK293T cells treated with 1  $\mu$ g/mL dox and without (teal) or with (pink) 100  $\mu$ M guanine. Solid light gray line shows the expression gate.

**B.** Gel depicts PCR products amplified from genomic DNA (gDNA) of cells transduced with vectors containing the splicing switch. Producer cells were cultured with or without 100  $\mu$ M guanine during virus production. The amplicon spans the site where the splicing switch was added. Lengths of the ladder bands are labeled in base pairs (bp), and the most prominent bands in the conditions are annotated with their predicted splice products. The uncropped gel is displayed in Fig. S6E.

**C.** Schematic of direct conversion of primary mouse embryonic fibroblasts to *Hb9::GFP*+ induced motor neurons (iMNs) via viral delivery of conversion factors. *Hb9::GFP* is a transgenic reporter of iMN fate used to identify converted cells. The splicing switch regulates expression of the oncogenic HRAS mutant HRAS<sup>G12V</sup>, which is fused to the fluorescent protein mCherry and delivered via the the best-performing divergent "all-in-one" inducible vector. Vectors contain the splicing switch within the mCherry coding sequence (circles) or lack the switch (diamonds) and were transduced into mouse embryonic fibroblasts alongside the remaining conversion factors. Cells were treated with 1  $\mu$ g/mL dox (filled points) or without (open points), and additionally treated with 0–100  $\mu$ M guanine for the duration of the conversion process. At 14 days post-transduction, iMN yield was quantified as the number of cells expressing *Hb9::GFP* per cell seeded, normalized to the same count for the untreated condition (neither dox nor guanine added) for each vector for each biological replicate. Cargo expression is the geometric mean of mCherry fluorescence for the iMN population (AU, log scale).

**D.** Images show fluorescence of *Hb9::GFP* (top) and mCherry-HRAS<sup>G12V</sup> (bottom) at 14 days post-transduction for cells transduced with the vector containing the splicing switch, treated with dox, and treated with varying concentrations of guanine. Scale bar represents 100  $\mu$ m.

Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points.

aimed to systematically fill this knowledge gap to facilitate predictable vector design.

In the best-performing vectors, we encoded a microRNA-based circuit, ribozyme switch, and splicing switch, finding that the “all-in-one” design with divergent syntax affords both high titer and functionality for these cargoes, while vectors with tandem syntax do not. Our results suggest that when encoded on the sense strand, these RNA devices primarily reduce titer by preventing formation of particles with intact viral transcripts. Similarly, these designs may compromise functionality by altering the sequence of packaged transcripts. As lentiviral vectors are increasingly used to deliver gene therapies *in vivo*, integration of aberrant sequences presents clinical risks. In contrast, when encoded on the antisense strand with divergent syntax, we find that these RNA devices minimally impact virus production, express well in target cells, and retain their characteristic functions. We anticipate that our “all-in-one” vectors offer an effective framework for delivering other sequences that may interfere with processing of the viral transcript.

We expect our best-performing vector design to generalize beyond the cargo, genetic parts, and cell types tested here. As a demonstration, our best-performing design expresses a functional cargo, a dose-dependent cell-fate conversion factor, suggesting that the vector can deliver other therapeutic payloads. Nevertheless, further testing across diverse primary human cell types is required for downstream applications. Encoding native genes in the vector may affect titer by introduction of sequences that affect RNA processing, such as cryptic splice sites<sup>31-35</sup>. Comparison to well-defined lentiviral vectors may reveal hidden features within cargo sequences and offer a design alternatives. Further, cargo size may also influence vector performance, warranting future work. As we expand our designs to other sequences, we envision replacing the bacterial-derived Tet-ON system with humanized components<sup>61,62</sup> and FDA-approved inducers, such as the synZiFTR platform<sup>63</sup>, for greater clinical relevance. We may also consider exchanging other elements for human-derived sequences. Though viral in origin, the WPRE sequence is often included in lentiviral vectors, as it reliably boosts expression<sup>50,51</sup> with minimal effect on titer, which our work supports, and mitigates transgene silencing<sup>47</sup>. Although we did not test vectors encoding this element in the 3' UTR of the upstream gene, we hypothesize that two copies of WPRE in divergent syntax may abolish vector production, as observed for two bGH PAS sequences. Finally, while inclusion of an upstream PAS consistently increases titer and expression for the divergent vectors we tested, tandem vectors performed as well or better without an upstream PAS, though it is possible that other PAS elements may behave differently.

Fundamentally, our approach leverages an inducible promoter that does not express in producer cells to maintain high titer, and that elicits strong expression in target cells upon induction. Beyond the Tet-ON system used here, we expect that cell state-responsive promoters that are transcriptionally active in target cells but not producer cells would fulfill these criteria<sup>64-67</sup>. Using promoters that recruit endogenous transcriptional machinery would eliminate the need for an exogenous transactivator, reducing cargo size and avoiding potential side effects of synthetic systems. One potential challenge to this approach is that cell state-responsive promoters may have more “leaky” expression outside of the intended cell state, which may be exacerbated by nearby transcription in divergent syntax<sup>37</sup>. We did not observe expression in uninduced conditions for genes regulated by the tightly controlled Tet-ON system. However, even with tandem syntax, we observed an unexpected increase in expression of the downstream EFS-driven gene upon induction of the adjacent TRE3G promoter, though not to the same degree as with divergent syntax. Potentially, the high local concentration of transcriptional machinery recruited to the inducible gene enhances expression of the neighboring gene regardless of syntax. This effect may only be apparent for very strong inducible promoters like TRE3G used here<sup>48</sup>, or it may be specific to EF1 $\alpha$ -derived promoters, as vectors with hPGK did not exhibit this behavior. Nevertheless, effects of the inducible gene on the neighboring constitutive gene could disrupt expression of functional cargoes with

narrow expression windows. As yet, the easiest way to robustly “insulate” expression of one gene from another is to deliver them on separate vectors, at the cost of delivery efficiency. Future studies could determine the generality of this coupling and devise new insulation approaches.

Together, this work illuminates key design principles for constructing two-gene lentiviral vectors encoding RNA-based circuits. Our results highlight the pitfalls of common designs and offer strategies for robust manufacturability and functionality. Optimizing for both of these properties will ensure that vectors are efficacious and support scalable production, advancing toward feasible clinical applications. Delivering compact, tunable, and modular RNA-based circuits on lentiviral vectors will unlock avenues for precise control in gene and cell therapies.

## Materials and methods

### Cloning

All lentiviral transfer plasmids were constructed using a multi-level Golden Gate cloning scheme described in Peterman *et al.*<sup>48</sup>, with the third-generation Lenti-X1 lentiviral backbone derived from Addgene #17297. ComMAND circuit<sup>53</sup> plasmids include Addgene #235268, #235319, and #235320, and additional plasmids were derived from these. The plasmids containing the constitutively active and mutant ribozymes<sup>10</sup> were derived from Addgene #131744 and #131745, and the ribozyme switch was ordered as a gBlock from Azenta/Genewiz using the sequence from Mustafina *et al.*<sup>54</sup>. The splicing switch was a gift from Shiva Razavi, Jonathan C. Chen, and James J. Collins at MIT. Plasmids containing the splicing switch were derived from this vector. The plasmids containing HRAS<sup>G12V</sup> were derived from Addgene #246340 (ref.<sup>59</sup>). Sources for other genetic parts can be found in Peterman *et al.*<sup>48</sup>. Key plasmids will be deposited at Addgene, and other plasmids will be made available upon request.

### Cell culture

#### Cell lines

HEK293T cells (ATCC, CRL-3216), Lenti-X HEK293T cells (Takara Bio, 632180), and Plat-E retroviral packaging HEK293T cells (Cell Biolabs, RV-101) were cultured using DMEM (Genesee Scientific, 25-500) plus 10% FBS (Genesee Scientific, 25-514H) and incubated at 37°C with 5% CO<sub>2</sub>. 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) and incubated at 37°C for four minutes, then quenched with an equal volume of DMEM + 10% FBS. Cells were passaged about every three days to maintain consistent growth before experiments. Plat-E cells were selected using 10 µg/mL blasticidin and 1 µg/mL puromycin every three passages. Cells were tested for mycoplasma.

#### Primary mouse embryonic fibroblasts

Primary mouse embryonic fibroblasts were isolated as described in Wang *et al.*<sup>56</sup>. C57BL/6 mice (The Jackson Laboratory, 000664) were mated with B6CBAF1 mice bearing the transgenic *Hb9::EGFP* reporter (The Jackson Laboratory, 005029), 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 (Sigma-Aldrich, G1890-100G), one dish per embryo (passage 0). Cells were incubated at 37°C with 5% CO<sub>2</sub>. 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.

## Lentiviral vector production

Lenti-X HEK293T cells were seeded at 7.5 million cells per well in 10-cm dishes coated with 0.1% gelatin. The following day (day one), the transfer plasmid, packaging plasmid (psPAX2, Addgene #12260), and envelope plasmid (pMD2.G, Addgene #12259) were co-transfected using PEI (linear polyethylenimine, Fisher Scientific, AA4389603). Transfection mixes were prepared using a ratio of 4  $\mu\text{g}$  PEI to 1  $\mu\text{g}$  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 prepared DNA mixes containing 6  $\mu\text{g}$  of the transfer plasmid, 6  $\mu\text{g}$  of the packaging plasmid, and 12  $\mu\text{g}$  of the envelope plasmid per dish. These condition mixes were incubated for an additional 10 to 15 minutes and then added dropwise on top of the growth media in the 10-cm dishes. After six hours, the media was replaced with 6.5 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 fresh 6.5 mL HEPES-buffered DMEM + 10% FBS. On day three, the media was again collected. The collected media was filtered through a 0.45- $\mu\text{m}$  polyethersulfone (PES) filter (Thermo Scientific, CH4525-PES). To concentrate the virus, Lenti-X Concentrator (Takara Bio, 631232) was added to the filtered virus-containing media in a volume ratio of 3 parts media to 1 part concentrator, mixed gently, and left overnight at 4°C. On day four, the media was centrifuged at 1500 $\times g$  at 4°C for 45 minutes. The supernatant was aspirated, and the resulting pellet was resuspended in 200  $\mu\text{L}$  of DMEM + 10% FBS. Virus was stored at -80°C before use.

## Lentiviral vector transduction

On the day of transduction, HEK293T cells were dissociated and diluted to a concentration of 20,000 cells per well (50  $\mu\text{L}$ /well) in DMEM + 10% FBS. Cells were combined with 5  $\mu\text{g}/\text{mL}$  polybrene (hexadimethrine bromide, Sigma-Aldrich, H9268-5G) and the specified volume of virus (see below). Additional DMEM + 10% FBS was added to reach a total volume of 100  $\mu\text{L}$  per well. The resulting cell, polybrene, and virus mixture was plated onto 96-well plates coated with 0.1% gelatin. The following day, the media was replaced with fresh DMEM + 10% FBS, with or without inducers according to the condition. Cells were prepared for flow cytometry two days later (three days post-transduction) by dissociating with Trypsin-EDTA. After centrifuging the dissociated cells at 500 $\times g$  for five minutes, the cells were resuspended in PBS and transferred to a v-bottom plate for flow cytometry.

For transduction experiments to calculate functional titer, a two-fold serial dilution of concentrated virus was used, with six conditions per vector and 4  $\mu\text{L}$  of concentrated virus per well in the highest condition. For transduction experiments using a constant multiplicity of infection (MOI), the volume of virus to use was calculated for each vector for each batch of virus from the titer data to achieve a MOI of 0.3. Namely, virus volume in  $\mu\text{L}$ ,  $v$ , is a function of the number of cells being transduced,  $c$ , the MOI in transducing units (TU) per cell,  $\lambda$ , and the viral titer in TU/ $\mu\text{L}$ ,  $T$ :

$$v = \frac{c \lambda}{T} \quad (1)$$

For vectors containing a Tet-ON inducible gene, 1  $\mu\text{g}/\text{mL}$  dox (doxycycline, Sigma-Aldrich, D3447-500MG) dissolved in water was added to the media at one day post-transduction both for experiments to calculate titer and for experiments with a constant MOI. Experiments with a constant MOI also included conditions without dox. For vectors containing the ribozyme switch or the splicing switch, experiments with a constant MOI also included conditions treated with 100  $\mu\text{M}$  guanine (from stocks of 25 mM dissolved in 0.2 N NaOH), with or without dox. Conditions without guanine were also included, using an equal volume of 0.2 N NaOH as a vehicle control instead.

## Direct conversion of primary mouse embryonic fibroblasts to induced motor neurons

### Retroviral vector production

Conversion factors other than HRAS<sup>G12V</sup> were delivered via retroviral vectors. These include LNI (Addgene #233154), a polycistronic cassette of Lhx3, Ngn2, and Isl1 neuronal transcription factors separated by 2A “self-cleaving” peptides, and SNAP-p53DD (Addgene #244168), a fusion of the SNAPtag and an oncogenic p53 mutant lacking the DNA binding domain. To produce the retrovirus, Plat-E cells were seeded at 850,000 cells per well in 6-well plates coated with 0.1% gelatin. The next day, cells were transfected with 1.8 µg of transfer plasmid per well using PEI, then 18 hours later the media was replaced with 1.25 mL fresh HEPES-buffered DMEM + 10% FBS. On the following day, viral supernatant was collected and filtered through a 0.45-µm PES filter, then immediately used to transduce mouse embryonic fibroblasts.

### Transduction of primary mouse embryonic fibroblasts

Primary mouse embryonic fibroblasts were seeded at 5,000 cells per well in 96-well plates coated with 0.1% gelatin. The next day, media was replaced with fresh DMEM + 10% FBS containing 5 µg/mL polybrene and 11 µL of each retrovirus (LNI and SNAP-p53DD). Cells were spininfected by centrifuging at 1500xg for 30 minutes. On the following day (defined as 0 days post-transduction), media was again replaced with fresh DMEM + 10% FBS plus 5 µg/mL polybrene and 1 µL of concentrated HRAS<sup>G12V</sup> lentivirus. Cells were again spininfected. After 24 hours, media was replaced with fresh DMEM + 10% FBS containing 1 µg/mL dox, 0–100 µM guanine, and/or vehicle controls, according to the condition. Conditions with inducer included these for the remainder of the experiment. At three days post-transduction, cells were switched into N3 media: DMEM/F-12 (Fisher Scientific, 21-331-020) supplemented with N-2 (Thermo Fisher Scientific, 17-502-048), B-27 (Thermo Scientific, 17-504-044), 1% GlutaMAX (Thermo Fisher Scientific, 35-050-061), 7.5 µM RepSox (a TGF-β inhibitor, Selleck Chemicals, S7223), and 10 ng/mL each of recombinant human BDNF (R&D Systems, 248-BDB-050), GDNF (R&D Systems, 212-GD-050), CNTF (R&D Systems, 257-NT-050), and FGF (Peprotech, 100-18B). Media was replaced with fresh N3 media every 2 to 3 days. At 14 days post-transduction, cells were imaged then dissociated for flow cytometry by incubating with DMEM/F12 plus 20 U/mL DNase (STEMCELL Technologies, 07470), 2000 U/mL papain (Fisher Scientific, AAJ61875MC), 0.5 mM EDTA (Millipore Sigma, E5134), and 1 mM L-cysteine (Thermo Scientific, J63745-14) at 37°C for 15 minutes. The cells were then centrifuged at 400xg for four minutes, resuspended in PBS, and transferred to a v-bottom plate for flow cytometry.

### Amplification of genomic DNA from transduced cells

HEK293T cells were transduced with 4 µL of concentrated lentivirus per well in a 96-well plate as for transductions to calculate functional titer. At three days post-transduction, cells were dissociated using Trypsin-EDTA, spun down at 500xg for 5 minutes, and combined with 50 µL Cell Lysis Buffer (Cell Signaling Technology, 9803S) and 1 µL Proteinase K (New England Biolabs, P8107S) per well. Conditions were incubated for 45 minutes at 85°C to lyse the cells. Polymerase chain reaction (PCR) mixes were prepared for each sample in a 10 µL reaction volume using Apex Taq RED 2X Master Mix (Genesee Scientific, 42-138B), 1 µL of lysate, and 0.5 µL each of 10 µM stocks of the primers in [Table S1](#). PCR was performed in a thermocycler with the following protocol: 95°C for 30 seconds for an initial denaturation; 30 cycles of 15 seconds at 95°C (denature), 15 seconds at 61°C (anneal), and 60 seconds at 68°C (extension); and 5 minutes at 68°C for the final extension. Then, 5 µL of

each PCR product was run on a 2% agarose gel (Genesee Scientific, 20-102GP) at 110 V for 45 to 60 minutes and imaged using the ChemiDoc MP Imaging System (BioRad, 12003154).

## Fluorescence imaging

Fluorescence images were taken on a Keyence All-in-One Fluorescence Microscope (Keyence, BZ-X800) at 10x magnification. Scale bars represent 100  $\mu\text{m}$ .

## Quantification and statistical analysis

### Analysis of flow cytometry data

Flow cytometry was performed on an Attune NxT Flow Cytometer (Thermo Fisher Scientific, A24858). See [Table S2](#) for corresponding fluorophores, lasers, and filters. Data were first analyzed in FlowJo (BD Biosciences, v10) to select cells using forward and side scatter gates (FSC-A, SSC-A) and single cells using forward scatter gates (FSC-A, FSC-H). The resulting single cell populations were analyzed in Python using the `matplotlib` (v3.10.3), `numpy` (v2.2.6), `pandas` (v2.2.3), `rushd` (v0.5.1), `scipy` (v1.15.3), and `seaborn` (v0.13.2) packages.

Fluorescence values are in arbitrary units plotted on a log scale and are shown as 1D or 2D kernel density estimates or as geometric means of gated populations. Cells were gated on expression of the downstream gene (for single-gene vectors) or on expression of both genes (for two-gene vectors), using gates defined by the 99.5<sup>th</sup> percentile of untransduced cells or defined manually. Expression variability for the ComMAND vectors is shown as standard deviation of fluorescence (arbitrary units) of the gated population. Fold change comparing similar vectors with one different genetic part (vectors with a downstream EF1 $\alpha$  promoter versus otherwise identical designs with a downstream EFS promoter, vectors with WPRE versus without WPRE) was computed by averaging the geometric means across biological replicates for the second vector, then dividing the geometric means for each biological replicate of the first vector by this value. Fold change comparing conditions treated with or without guanine was computed by dividing the geometric mean expression for the condition with guanine by the value for the condition without guanine within biological replicates.

Points represent means  $\pm$  standard error across biological replicates. For transductions at a constant MOI, biological replicates are defined as independent transductions of separate batches of virus, for a total of  $n \geq 3$  for each vector. All statistics are two-sided Student's t-tests. Annotations show fold changes computed by dividing the values for the indicated conditions averaged across biological replicates. For vectors with a Tet-ON inducible gene, the annotations always represent comparisons between conditions treated with dox.

### Calculating functional titer

The fluorescence data from transduction of HEK293T cells with a two-fold serial dilution of virus was used to calculate a functional metric of viral titer. Specifically, transduction events were assumed to follow a Poisson process, where cells can be transduced by multiple viral particles and transduction events are independent. The expected number of transduction events per cell is the MOI,  $\lambda$ , meaning the probability of  $k$  events in a cell is

$$P(x = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (2)$$

Across a population, this value is reflected by the fraction of cells with  $k$  transduction events. We can use the fraction of expressing cells (i.e., cells with at least one transduction event) to solve for MOI:

$$P(x > 0) = 1 - P(x = 0) \quad (3)$$

$$= 1 - \frac{\lambda^0 e^{-\lambda}}{0!} \quad (4)$$

$$= 1 - e^{-\lambda} \quad (5)$$

Combining [Equations 1](#) and [5](#), we have

$$\text{Fraction(expressing)} = 1 - e^{-T v/c} \quad (6)$$

Using this equation, we fitted our data—the fraction of expressing cells as a function of virus volume ( $v$ ) for six volumes of virus and a known number of cells ( $c$ )—to this equation for each batch of each vector to calculate titer ( $T$ , in TU/ $\mu$ L). Fraction expressing was calculated from expression gates defined by the 99.5<sup>th</sup> percentile of untransduced cells. For curve fitting, we used the mean of two technical replicates per condition, and we excluded points where the fraction of expressing cells decreased for larger virus volumes. Titer values were calculated for untransduced cells (i.e., fraction expressing of 0.005 by definition) to serve as a reference. Normalized titer was computed relative to one vector per batch of virus to account for batch-to-batch variability. Titer is plotted on a log scale in TU/ $\mu$ L, or is unitless for normalized titer. Points represent means  $\pm$  standard error across biological replicates. For titer calculations, biological replicates are defined as transductions of separate batches of virus, with  $n \geq 3$  batches per vector. All statistics are two-sided Student's t-tests.

### Calculating direct conversion yield

To calculate yield for direct conversion of mouse embryonic fibroblasts to induced motor neurons (iMNs), the iMN population was first defined as cells with high *Hb9*::GFP expression measured via fluorescence. The number of such cells was divided by the number of cells seeded at the beginning of the experiment, then normalized to the untreated condition (neither dox nor guanine added) for each lentiviral vector (i.e., vectors expressing HRAS<sup>G12V</sup> with or without the splicing switch) for each biological replicate. Points represent means  $\pm$  standard error across biological replicates. Biological replicates are defined as separate batches of donor cells in three independent conversion experiments, for a total of  $n = 5$ .

## Acknowledgments

We thank Shiva Razavi, Jonathan C. Chen, and James J. Collins at MIT for the splicing switch. A manuscript detailing the splicing switch, titled “A framework for programmable RNA-level gene control,” is currently in preparation. We thank Maria Castellanos, Mary Ehmann, Christopher Johnstone, Sneha Kabaria, Emma Peterman, and Nathan Wang for feedback on the manuscript.

## Funding

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health [R35-GM143033]; the National Science Foundation CAREER [2339986]; the Pershing Square Foundation MIND Prize; the National Science Foundation Graduate Research Fellowship [1745302 to K.S.L. and B.A.L.-D.]; and the MIT Health and Life Sciences Collaborative Graduate Fellowship [to K.S.L.].

## Author contributions

Following the CRediT taxonomy, K.S.L. contributed to conceptualization, methodology, investigation, validation, data curation, formal analysis, visualization, and writing (original draft, review & editing). B.A.L.-D. contributed to methodology, investigation, validation, formal analysis, and writing (review & editing). K.E.G. contributed to conceptualization, methodology, writing (review & editing), resources, supervision, funding acquisition, and project administration.

## Declaration of interests

There are no competing interests to declare.

## References

1. Geng, G., Xu, Y., Hu, Z., Wang, H., Chen, X., Yuan, W. & Shu, Y. Viral and non-viral vectors in gene therapy: current state and clinical perspectives. *eBioMedicine* **118**. doi:[10.1016/j.ebiom.2025.105834](https://doi.org/10.1016/j.ebiom.2025.105834) (Aug. 1, 2025).
2. Jargalsaikhan, B.-E., Muto, M. & Ema, M. The Era of Gene Therapy: The Advancement of Lentiviral Vectors and Their Pseudotyping. *Viruses* **17**, 1036. doi:[10.3390/v17081036](https://doi.org/10.3390/v17081036) (Aug. 2025).
3. Lewis, P., Hensel, M. & Emerman, M. Human immunodeficiency virus infection of cells arrested in the cell cycle. *The EMBO Journal* **11**, 3053–3058. doi:[10.1002/j.1460-2075.1992.tb05376.x](https://doi.org/10.1002/j.1460-2075.1992.tb05376.x) (Aug. 1, 1992).
4. Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M. & Trono, D. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–267. doi:[10.1126/science.272.5259.263](https://doi.org/10.1126/science.272.5259.263) (Apr. 12, 1996).
5. Michels, A., Ho, N. & Buchholz, C. J. Precision medicine: In vivo CAR therapy as a showcase for receptor-targeted vector platforms. *Molecular Therapy* **30**, 2401–2415. doi:[10.1016/j.ymthe.2022.05.018](https://doi.org/10.1016/j.ymthe.2022.05.018) (July 6, 2022).

6. Duvergé, A. & Negroni, M. Pseudotyping Lentiviral Vectors: When the Clothes Make the Virus. *Viruses* **12**, 1311. doi:[10.3390/v12111311](https://doi.org/10.3390/v12111311) (Nov. 16, 2020).
7. Wolff, J. H. & Mikkelsen, J. G. Delivering genes with human immunodeficiency virus-derived vehicles: still state-of-the-art after 25 years. *Journal of Biomedical Science* **29**, 79. doi:[10.1186/s12929-022-00865-4](https://doi.org/10.1186/s12929-022-00865-4) (Oct. 9, 2022).
8. Kumar, M., Keller, B., Makalou, N. & Sutton, R. E. Systematic Determination of the Packaging Limit of Lentiviral Vectors. *Human Gene Therapy* **12**, 1893–1905. doi:[10.1089/104303401753153947](https://doi.org/10.1089/104303401753153947) (Oct. 1, 2001).
9. Takahashi, K. & Galloway, K. E. RNA-based controllers for engineering gene and cell therapies. *Current Opinion in Biotechnology* **85**, 103026. doi:[10.1016/j.copbio.2023.103026](https://doi.org/10.1016/j.copbio.2023.103026) (Feb. 1, 2024).
10. Xiang, J. S., Kaplan, M., Dykstra, P., Hinks, M., McKeague, M. & Smolke, C. D. Massively parallel RNA device engineering in mammalian cells with RNA-Seq. *Nature Communications* **10**, 4327. doi:[10.1038/s41467-019-12334-y](https://doi.org/10.1038/s41467-019-12334-y) (Sept. 23, 2019).
11. Dykstra, P. B., Kaplan, M. & Smolke, C. D. Engineering synthetic RNA devices for cell control. *Nature Reviews Genetics*, 1–14. doi:[10.1038/s41576-021-00436-7](https://doi.org/10.1038/s41576-021-00436-7) (Jan. 4, 2022).
12. Mukherjee, S. & Barash, D. Designing RNA switches for synthetic biology using inverse-RNA-folding. *Trends in Biotechnology* **42**, 517–521. doi:[10.1016/j.tibtech.2023.11.005](https://doi.org/10.1016/j.tibtech.2023.11.005) (May 1, 2024).
13. Ling, J. P., Bygrave, A. M., Santiago, C. P., Carmen-Orozco, R. P., Trinh, V. T., Yu, M., Li, Y., Liu, Y., Bowden, K. D., Duncan, L. H., *et al.* Cell-specific regulation of gene expression using splicing-dependent frameshifting. *Nature Communications* **13**, 5773. doi:[10.1038/s41467-022-33523-2](https://doi.org/10.1038/s41467-022-33523-2) (Oct. 1, 2022).
14. Han, S. R., Lee, C. H., Im, J. Y., Kim, J. H., Kim, J. H., Kim, S. J., Cho, Y. W., Kim, E., Kim, Y., Ryu, J.-H., *et al.* Targeted suicide gene therapy for liver cancer based on ribozyme-mediated RNA replacement through post-transcriptional regulation. *Molecular Therapy Nucleic Acids* **23**, 154–168. doi:[10.1016/j.omtn.2020.10.036](https://doi.org/10.1016/j.omtn.2020.10.036) (Mar. 5, 2021).
15. Chen, Z., Yang, L., Zhang, Y., Li, J., Yang, Y., Li, Y., Fan, L., Chen, W., Miao, L., Liu, J., *et al.* The drug-elicitable alternative splicing module for tunable vector expression in the heart. *Nature Cardiovascular Research* **4**, 938–955. doi:[10.1038/s44161-025-00665-7](https://doi.org/10.1038/s44161-025-00665-7) (July 2025).
16. Monteys, A. M., Hundley, A. A., Ranum, P. T., Tecedor, L., Muehlmann, A., Lim, E., Lukashev, D., Sivasankaran, R. & Davidson, B. L. Regulated control of gene therapies by drug-induced splicing. *Nature* **596**, 291–295. doi:[10.1038/s41586-021-03770-2](https://doi.org/10.1038/s41586-021-03770-2) (Aug. 2021).
17. Stagno, J. R., Deme, J. C., Dwivedi, V., Lee, Y.-T., Lee, H. K., Yu, P., Chen, S.-Y., Fan, L., Degenhardt, M. F. S., Chari, R., *et al.* Structural investigation of an RNA device that regulates PD-1 expression in mammalian cells. *Nucleic Acids Research* **53**, gkaf156. doi:[10.1093/nar/gkaf156](https://doi.org/10.1093/nar/gkaf156) (Mar. 24, 2025).
18. Fukunaga, K., Dhamodharan, V., Miyahira, N., Nomura, Y., Mustafina, K., Oosumi, Y., Takayama, K., Kanai, A. & Yokobayashi, Y. Small-Molecule Aptamer for Regulating RNA Functions in Mammalian Cells and Animals. *Journal of the American Chemical Society* **145**, 7820–7828. doi:[10.1021/jacs.2c12332](https://doi.org/10.1021/jacs.2c12332) (Apr. 12, 2023).
19. Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W. & Salis, H. M. Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic Acids Research* **44**, 1–13. doi:[10.1093/nar/gkv1289](https://doi.org/10.1093/nar/gkv1289) (Jan. 8, 2016).

20. Chen, Y. Y., Jensen, M. C. & Smolke, C. D. Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proceedings of the National Academy of Sciences* **107**, 8531–8536. doi:[10.1073/pnas.1001721107](https://doi.org/10.1073/pnas.1001721107) (May 11, 2010).
21. Mou, H., Zhong, G., Gardner, M. R., Wang, H., Wang, Y.-W., Cheng, D. & Farzan, M. Conditional Regulation of Gene Expression by Ligand-Induced Occlusion of a MicroRNA Target Sequence. *Molecular Therapy* **26**, 1277–1286. doi:[10.1016/j.ymthe.2018.02.021](https://doi.org/10.1016/j.ymthe.2018.02.021) (May 2, 2018).
22. Strobel, B., Düchs, M. J., Blazevic, D., Rechtsteiner, P., Braun, C., Baum-Kroker, K. S., Schmid, B., Ciossek, T., Gottschling, D., Hartig, J. S., *et al.* A Small-Molecule-Responsive Riboswitch Enables Conditional Induction of Viral Vector-Mediated Gene Expression in Mice. *ACS Synthetic Biology* **9**, 1292–1305. doi:[10.1021/acssynbio.9b00410](https://doi.org/10.1021/acssynbio.9b00410) (June 19, 2020).
23. Eriksson, R. A. E., Nieminen, T., Galibert, L., Peltola, S. K., Tikkanen, P., Käyhty, P., Leinonen, H. M., Oruetebarria, I., Lepola, S., Valkama, A. J., *et al.* Optimized riboswitch-regulated AAV vector for VEGF-B gene therapy. *Frontiers in Medicine* **9**. doi:[10.3389/fmed.2022.1052318](https://doi.org/10.3389/fmed.2022.1052318) (Dec. 13, 2022).
24. Yu, X., Zhan, X., D’Costa, J., Tanavde, V. M., Ye, Z., Peng, T., Malehorn, M. T., Yang, X., Civin, C. I. & Cheng, L. Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Molecular Therapy* **7**, 827–838. doi:[10.1016/S1525-0016\(03\)00104-7](https://doi.org/10.1016/S1525-0016(03)00104-7) (June 1, 2003).
25. Merten, O.-W., Hebben, M. & Bovolenta, C. Production of lentiviral vectors. *Molecular Therapy - Methods & Clinical Development* **3**, 16017. doi:[10.1038/mtm.2016.17](https://doi.org/10.1038/mtm.2016.17) (Jan. 1, 2016).
26. Hanawa, H., Yamamoto, M., Zhao, H., Shimada, T. & Persons, D. A. Optimized Lentiviral Vector Design Improves Titer and Transgene Expression of Vectors Containing the Chicken  $\beta$ -Globin Locus HS4 Insulator Element. *Molecular Therapy* **17**, 667–674. doi:[10.1038/mt.2009.1](https://doi.org/10.1038/mt.2009.1) (Apr. 1, 2009).
27. Nielsen, T. T., Jakobsson, J., Rosenqvist, N. & Lundberg, C. Incorporating double copies of a chromatin insulator into lentiviral vectors results in less viral integrants. *BMC biotechnology* **9**, 13. doi:[10.1186/1472-6750-9-13](https://doi.org/10.1186/1472-6750-9-13) (Feb. 24, 2009).
28. Tolmachov, O., Tolmachova, T., Al-Allaf, F. A., Tolmachov, O., Tolmachova, T. & Al-Allaf, F. A. in *Viral Gene Therapy* (IntechOpen, July 20, 2011). doi:[10.5772/17361](https://doi.org/10.5772/17361).
29. Han, J., Tam, K., Ma, F., Tam, C., Aleshe, B., Wang, X., Quintos, J. P., Morselli, M., Pellegrini, M., Hollis, R. P., *et al.*  $\beta$ -Globin Lentiviral Vectors Have Reduced Titers due to Incomplete Vector RNA Genomes and Lowered Virion Production. *Stem Cell Reports* **16**, 198–211. doi:[10.1016/j.stemcr.2020.10.007](https://doi.org/10.1016/j.stemcr.2020.10.007) (Jan. 12, 2021).
30. Poletti, V. & Mavilio, F. Designing Lentiviral Vectors for Gene Therapy of Genetic Diseases. *Viruses* **13**. Number: 8, 1526. doi:[10.3390/v13081526](https://doi.org/10.3390/v13081526) (Aug. 2021).
31. De Ravin, S. S., Liu, S., Sweeney, C. L., Brault, J., Whiting-Theobald, N., Ma, M., Liu, T., Choi, U., Lee, J., O’Brien, S. A., *et al.* Lentivector cryptic splicing mediates increase in CD34+ clones expressing truncated HMGA2 in human X-linked severe combined immunodeficiency. *Nature Communications* **13**, 3710. doi:[10.1038/s41467-022-31344-x](https://doi.org/10.1038/s41467-022-31344-x) (June 28, 2022).
32. May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K. M. L., Luzzatto, L. & Sadelain, M. Therapeutic haemoglobin synthesis in  $\beta$ -thalassaemic mice expressing lentivirus-encoded human  $\beta$ -globin. *Nature* **406**, 82–86. doi:[10.1038/35017565](https://doi.org/10.1038/35017565) (July 2000).

33. Cooper, A. R., Lill, G. R., Gschweng, E. H. & Kohn, D. B. Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter. *Nucleic Acids Research* **43**, 682–690. doi:[10.1093/nar/gku1312](https://doi.org/10.1093/nar/gku1312) (Jan. 9, 2015).
34. Moiani, A., Paleari, Y., Sartori, D., Mezzadra, R., Miccio, A., Cattoglio, C., Cocchiarella, F., Lidonnici, M. R., Ferrari, G. & Mavilio, F. Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. *The Journal of Clinical Investigation* **122**, 1653–1666. doi:[10.1172/JCI61852](https://doi.org/10.1172/JCI61852) (May 1, 2012).
35. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., *et al.* Transfusion independence and HMGA2 activation after gene therapy of human  $\beta$ -thalassaemia. *Nature* **467**, 318–322. doi:[10.1038/nature09328](https://doi.org/10.1038/nature09328) (Sept. 2010).
36. Curtin, J. A., Dane, A. P., Swanson, A., Alexander, I. E. & Ginn, S. L. Bidirectional promoter interference between two widely used internal heterologous promoters in a late-generation lentiviral construct. *Gene Therapy* **15**, 384–390. doi:[10.1038/sj.gt.3303105](https://doi.org/10.1038/sj.gt.3303105) (Mar. 2008).
37. Reichenbach, P., Giordano Attianese, G. M. P., Ouchen, K., Cribioli, E., Triboulet, M., Ash, S., Saillard, M., Vuillefroy de Silly, R., Coukos, G. & Irving, M. A lentiviral vector for the production of T cells with an inducible transgene and a constitutively expressed tumour-targeting receptor. *Nature Biomedical Engineering* **7**, 1063–1080. doi:[10.1038/s41551-023-01013-5](https://doi.org/10.1038/s41551-023-01013-5) (Sept. 2023).
38. Maetzig, T., Galla, M., Brugman, M. H., Loew, R., Baum, C. & Schambach, A. Mechanisms controlling titer and expression of bidirectional lentiviral and gammaretroviral vectors. *Gene Therapy* **17**, 400–411. doi:[10.1038/gt.2009.129](https://doi.org/10.1038/gt.2009.129) (Mar. 2010).
39. Lenasi, T., Contreras, X. & Peterlin, B. M. Transcriptional Interference Antagonizes Proviral Gene Expression to Promote HIV Latency. *Cell host & microbe* **4**, 123–133. doi:[10.1016/j.chom.2008.05.016](https://doi.org/10.1016/j.chom.2008.05.016) (Aug. 14, 2008).
40. Johnstone, C. P., Love, K. S., Kabaria, S. R., Jones, R. D., Blanch-Asensio, A., Ploessl, D. S., Peterman, E. L., Lee, R., Yun, J., Oakes, C. G., *et al.* Gene syntax defines supercoiling-mediated transcriptional feedback. *Science* **392**, eadw1925. doi:[10.1126/science.adw1925](https://doi.org/10.1126/science.adw1925) (Apr. 30, 2026).
41. Lensch, S., Herschl, M. H., Ludwig, C. H., Sinha, J., Hinks, M. M., Mukund, A., Fujimori, T. & Bintu, L. Dynamic spreading of chromatin-mediated gene silencing and reactivation between neighboring genes in single cells. *eLife*. doi:[10.7554/eLife.75115](https://doi.org/10.7554/eLife.75115) (June 9, 2022).
42. Osti, D., Marras, E., Ceriani, I., Grassini, G., Rubino, T., Viganò, D., Parolaro, D. & Perletti, G. Comparative analysis of molecular strategies attenuating positional effects in lentiviral vectors carrying multiple genes. *Journal of Virological Methods* **136**, 93–101. doi:[10.1016/j.jviromet.2006.04.003](https://doi.org/10.1016/j.jviromet.2006.04.003) (Sept. 1, 2006).
43. Tian, J. & Andreadis, S. T. Independent and high-level dual-gene expression in adult stem-progenitor cells from a single lentiviral vector. *Gene therapy* **16**, 874–884. doi:[10.1038/gt.2009.46](https://doi.org/10.1038/gt.2009.46) (July 2009).
44. West, S. & Proudfoot, N. J. Transcriptional Termination Enhances Protein Expression in Human Cells. *Molecular Cell* **33**, 354–364. doi:[10.1016/j.molcel.2009.01.008](https://doi.org/10.1016/j.molcel.2009.01.008) (Feb. 13, 2009).

45. Cabrera, A., Edelstein, H. I., Glykofrydis, F., Love, K. S., Palacios, S., Tycko, J., Zhang, M., Lensch, S., Shields, C. E., Livingston, M., *et al.* The sound of silence: Transgene silencing in mammalian cell engineering. *Cell Systems* **13**, 950–973. doi:[10.1016/j.cels.2022.11.005](https://doi.org/10.1016/j.cels.2022.11.005) (Dec. 21, 2022).
46. Arana, S., Du, P. P., Vaughan-Jackson, A., Enright, N., Spees, K., Valbuena, R., Garcia, C. A., Nguyen, T., Venida, A., Seczynska, M., *et al.* Reduced Cas9 transgene silencing by incorporation of intron sequences. *Nature Communications* **16**, 10656. doi:[10.1038/s41467-025-65669-0](https://doi.org/10.1038/s41467-025-65669-0) (Nov. 27, 2025).
47. Uenaka, T., Napole, A., Saha, A. D., Sun, D., Singavarapu, A., Calzada, E., Chen, J., Erlebach, L., McQuade, A., Ramos, D. M., *et al.* Prevention of transgene silencing during human pluripotent stem cell differentiation. *Cell Stem Cell* **33**, 517–530.e8. doi:[10.1016/j.stem.2026.01.007](https://doi.org/10.1016/j.stem.2026.01.007) (Mar. 5, 2026).
48. 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.* High-resolution profiling reveals coupled transcriptional and translational regulation of transgenes. *Nucleic Acids Research* **53**, gkaf528. doi:[10.1093/nar/gkaf528](https://doi.org/10.1093/nar/gkaf528) (June 24, 2025).
49. Nóbrega, C., Mendonça, L. & Matos, C. A. in *A Handbook of Gene and Cell Therapy* (eds Nóbrega, C., Mendonça, L. & Matos, C. A.) 39–90 (Springer International Publishing, Cham, 2020). doi:[10.1007/978-3-030-41333-0\\_3](https://doi.org/10.1007/978-3-030-41333-0_3).
50. Zufferey, R., Donello, J. E., Trono, D. & Hope, T. J. Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *Journal of Virology* **73**, 2886–2892. doi:[10.1128/jvi.73.4.2886-2892.1999](https://doi.org/10.1128/jvi.73.4.2886-2892.1999) (Apr. 1999).
51. Schambach, A., Bohne, J., Baum, C., Hermann, F. G., Egerer, L., von Laer, D. & Giroglou, T. Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression. *Gene Therapy* **13**, 641–645. doi:[10.1038/sj.gt.3302698](https://doi.org/10.1038/sj.gt.3302698) (Apr. 2006).
52. Duran, A. G., Schweska, M., Nazari-Shafti, T. Z., Neuber, S., Stamm, C. & Gossen, M. Limiting Transactivator Amounts Contribute to Transgene Mosaicism in Tet-On All-in-One Systems. *ACS Synthetic Biology* **11**, 2623–2635. doi:[10.1021/acssynbio.2c00036](https://doi.org/10.1021/acssynbio.2c00036) (Aug. 19, 2022).
53. Love, K. S., Johnstone, C. P., Peterman, E. L., Gaglione, S., Birnbaum, M. E. & Galloway, K. E. Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells. *Cell Systems*, 101269. doi:[10.1016/j.cels.2025.101269](https://doi.org/10.1016/j.cels.2025.101269) (Apr. 28, 2025).
54. Mustafina, K., Fukunaga, K. & Yokobayashi, Y. Design of Mammalian ON-Riboswitches Based on Tandemly Fused Aptamer and Ribozyme. *ACS Synthetic Biology* **9**, 19–25. doi:[10.1021/acssynbio.9b00371](https://doi.org/10.1021/acssynbio.9b00371) (Jan. 17, 2020).
55. Babos, K. N., Galloway, K. E., Kisler, K., Zitting, M., Li, Y., Shi, Y., Quintino, B., Chow, R. H., Zlokovic, B. V. & Ichida, J. K. Mitigating Antagonism between Transcription and Proliferation Allows Near-Deterministic Cellular Reprogramming. *Cell Stem Cell* **25**, 486–500.e9. doi:[10.1016/j.stem.2019.08.005](https://doi.org/10.1016/j.stem.2019.08.005) (Oct. 3, 2019).
56. Wang, N. B., Lende-Dorn, B. A., Beitz, A. M., Han, P., Adewumi, H. O., O’Shea, T. M. & Galloway, K. E. Proliferation history and transcription factor levels drive direct conversion to motor neurons. *Cell Systems* **16**, 101205. doi:[10.1016/j.cels.2025.101205](https://doi.org/10.1016/j.cels.2025.101205) (Apr. 16, 2025).

57. Wang, N. B., Adewumi, H. O., Lende-Dorn, B. A., Beitz, A. M., O'Shea, T. M. & Galloway, K. E. Compact transcription factor cassettes generate functional, engraftable motor neurons by direct conversion. *Cell Systems* **16**, 101206. doi:[10.1016/j.cels.2025.101206](https://doi.org/10.1016/j.cels.2025.101206) (Apr. 16, 2025).
58. Lende-Dorn, B. A., Atkinson, J. C., Bae, Y. & Galloway, K. E. Chemogenetic tuning reveals optimal MAPK signaling for cell-fate programming. *Cell Reports* **44**, 116226. doi:[10.1016/j.celrep.2025.116226](https://doi.org/10.1016/j.celrep.2025.116226) (Sept. 23, 2025).
59. Kabaria, S. R., Bae, Y., Ehmann, M. E., Lende-Dorn, B. A., Beitz, A. M., Peterman, E. L., Love, K. S., Ploessl, D. S. & Galloway, K. E. Programmable promoter editing for precise control of transgene expression. *Nature Biotechnology*, 1–16. doi:[10.1038/s41587-025-02854-y](https://doi.org/10.1038/s41587-025-02854-y) (Oct. 13, 2025).
60. Askou, A. L., Aagaard, L., Kostic, C., Arsenijevic, Y., Hollensen, A. K., Bek, T., Jensen, T. G., Mikkelsen, J. G. & Corydon, T. J. Multigenic lentiviral vectors for combined and tissue-specific expression of miRNA- and protein-based antiangiogenic factors. *Molecular Therapy Methods & Clinical Development* **2**. doi:[10.1038/mtm.2014.64](https://doi.org/10.1038/mtm.2014.64) (Jan. 1, 2015).
61. Tycko, J., DelRosso, N., Hess, G. T., Aradhana, Banerjee, A., Mukund, A., Van, M. V., Ego, B. K., Yao, D., Spees, K., *et al.* High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell* **183**, 2020–2035.e16. doi:[10.1016/j.cell.2020.11.024](https://doi.org/10.1016/j.cell.2020.11.024) (Dec. 23, 2020).
62. Tycko, J., Van, M. V., Aradhana, DelRosso, N., Ye, H., Yao, D., Valbuena, R., Vaughan-Jackson, A., Xu, X., Ludwig, C., *et al.* Development of compact transcriptional effectors using high-throughput measurements in diverse contexts. *Nature Biotechnology* **43**, 1525–1538. doi:[10.1038/s41587-024-02442-6](https://doi.org/10.1038/s41587-024-02442-6) (Sept. 2025).
63. Li, H.-S., Israni, D. V., Gagnon, K. A., Gan, K. A., Raymond, M. H., Sander, J. D., Roybal, K. T., Joung, J. K., Wong, W. W. & Khalil, A. S. Multidimensional control of therapeutic human cell function with synthetic gene circuits. *Science* **378**, 1227–1234. doi:[10.1126/science.ade0156](https://doi.org/10.1126/science.ade0156) (Dec. 15, 2022).
64. Hooijberg, E., Bakker, A. Q., Ruizendaal, J. J. & Spits, H. NFAT-controlled expression of GFP permits visualization and isolation of antigen-stimulated primary human T cells. *Blood* **96**, 459–466. doi:[10.1182/blood.V96.2.459](https://doi.org/10.1182/blood.V96.2.459) (July 15, 2000).
65. Zhang, L., Kerkar, S. P., Yu, Z., Zheng, Z., Yang, S., Restifo, N. P., Rosenberg, S. A. & Morgan, R. A. Improving Adoptive T Cell Therapy by Targeting and Controlling IL-12 Expression to the Tumor Environment. *Molecular Therapy* **19**, 751–759. doi:[10.1038/mt.2010.313](https://doi.org/10.1038/mt.2010.313) (Apr. 1, 2011).
66. Bonfá, G., Martino, G., Sellitto, A., Rinaldi, A., Tedeschi, F., Caliendo, F., Melchiorri, L., Perna, D., Starkey, F., Nikolados, E., *et al.* Design of novel synthetic promoters to tune gene expression in T cells Jan. 9, 2025. doi:[10.1101/2025.01.09.632034](https://doi.org/10.1101/2025.01.09.632034).
67. Wu, M. R., Nissim, L., Stupp, D., Pery, E., Binder-Nissim, A., Weisinger, K., Enghuus, C., Palacios, S. R., Humphrey, M., Zhang, Z., *et al.* A high-throughput screening and computation platform for identifying synthetic promoters with enhanced cell-state specificity (SPECS). *Nature Communications* **10**, 1–10. doi:[10.1038/s41467-019-10912-8](https://doi.org/10.1038/s41467-019-10912-8) (Dec. 2019).

## Supplement

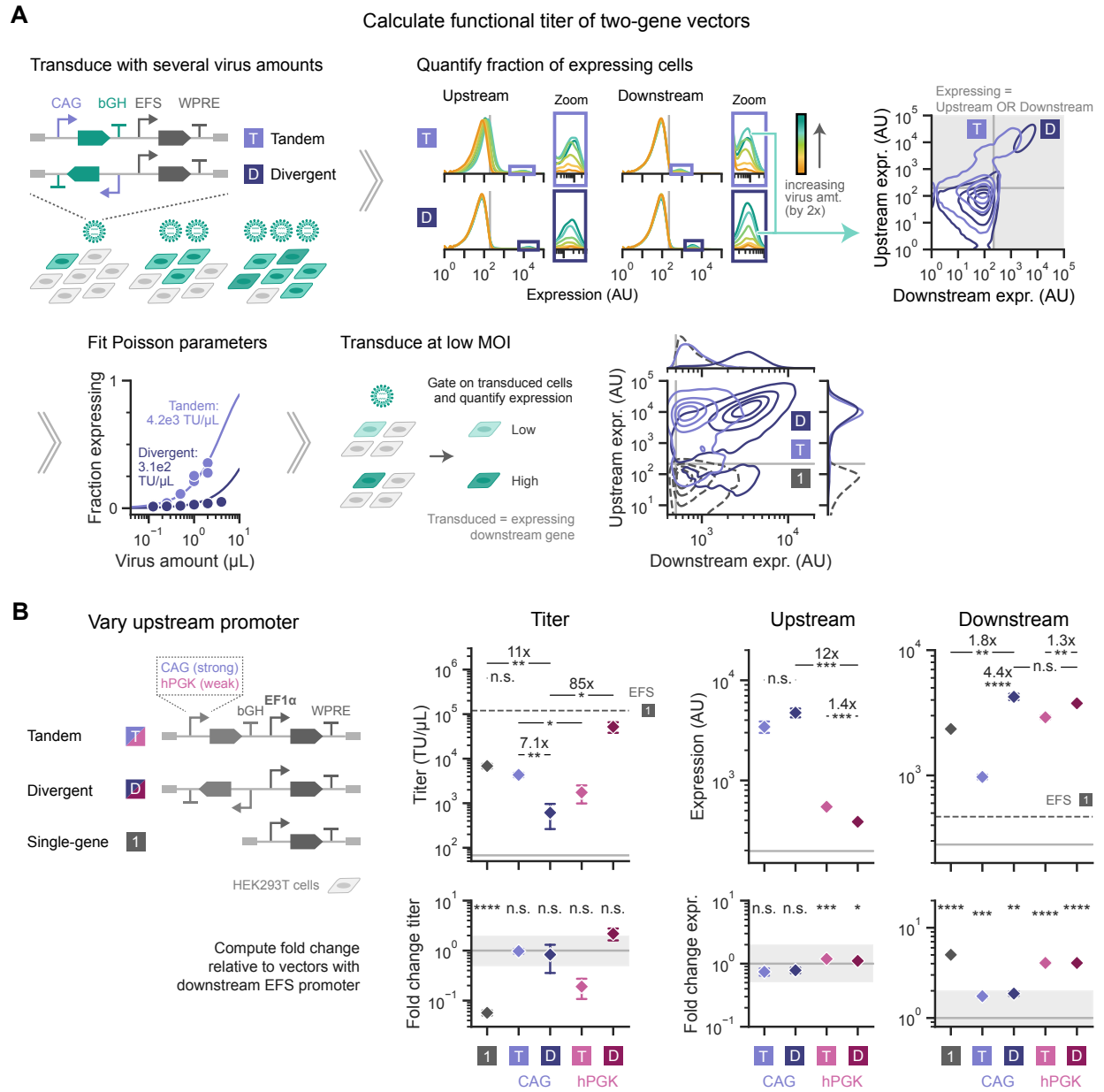
- [Tables S1](#) and [S2](#)
- [Figs. S1–S7](#)

Primer name	Sequence (5' to 3')
mCherry splicing switch forward	TGCAGGACGGCGAGTTCATC
mCherry splicing switch reverse	CGCGTTCGTACTGTTCCACGA

**Table S1. Primers used for PCRs of genomic DNA from cells transduced with the splicing switch.** Primers bind in the mCherry coding sequence, flanking the site where the splicing switch is inserted. See [Fig. S6E](#) for more information.

Gene	Fluorophore	Laser (nm)	Filter (nm)
Downstream gene (splicing switch vectors)	TagBFP	405	440 / 50
Downstream gene (other vectors)	mGreenLantern	488	530 / 30
iMN transgenic reporter	EGFP		510 / 10
Upstream gene (splicing switch vectors)	mCherry	561	615 / 25
Upstream gene (other vectors)	mRuby2		

**Table S2. Fluorophores and cytometer settings for flow cytometry experiments.** Fluorophores are listed for genes shown in plots. Lasers (wavelength in nm) and bandpass filters (center / width in nm) represent relevant settings on the Attune NxT Flow Cytometer.



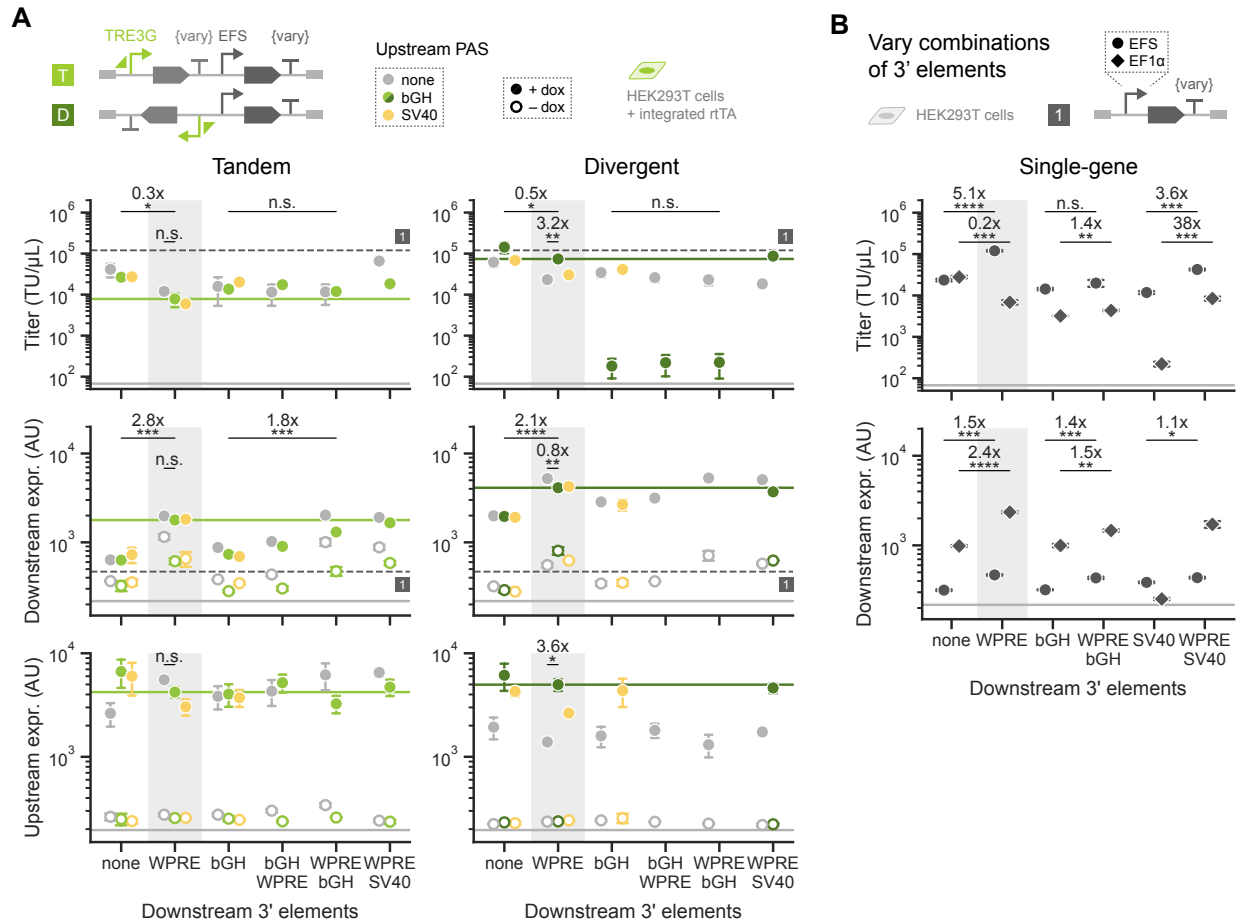
**Figure S1. Functional titer quantifies production efficiency for vectors with diverse genetic parts.**

**A.** Analogous to Fig. 1C, workflow for measuring functional titer and expression for two-gene vectors, shown for vectors with tandem (T) and divergent (D) syntax containing an upstream gene driven by the strong, synthetic CAG promoter and a second fluorescent protein driven by the EFS promoter downstream. **Top:** HEK293T cells were transduced with several different amounts of each virus. Expression of upstream and downstream genes were measured for each virus condition via fluorescence in arbitrary units (AU, log scale) from a flow cytometer. Light gray vertical line depicts the expression gate, used to calculate the fraction of expressing cells in each condition. Colored boxes indicate zoomed regions of the distributions. 2D density distributions depict expression (expr.) for conditions with the second-highest virus amount. Gray shading indicates cells considered expressing, i.e., above the expression gate for either gene. Plots show one representative batch of virus. **Bottom:** For the same vectors, fraction expressing as a function of virus volume was fit to a Poisson distribution to obtain titer in transducing units (TU) per  $\mu\text{L}$  virus. Points represent technical replicates for a representative batch of virus, and lines depict the Poisson curve fits. The computed titer is annotated on the plot, and virus volume is plotted on a log scale. To quantify cargo expression, fresh HEK293T cells were transduced at a low multiplicity of infection (MOI). 2D density distributions depict expression (AU, log scale) of cells expressing the downstream gene. A representative biological replicate is shown for the same two-gene vectors (solid lines) and for the EFS-driven single-gene vector from Fig. 1C (dashed lines).

(continued)

**Figure S1.** (continued)

**B.** Performance of single-gene and two-gene vectors as in Figs. 2C and 2D, except with EF1 $\alpha$  expressing the downstream gene. **Top:** Titer is in transducing units (TU) per  $\mu$ L (log scale), and expression is the geometric mean fluorescence in arbitrary units (AU, log scale). Dark gray dashed lines show values for the single-gene EFS vector in Fig. 2C for reference. **Bottom:** Values above normalized to the analogous vectors with downstream EFS in Figs. 2C and 2D, plotted on a log scale. Solid gray line indicates a fold change of 1, and gray shading spans a two-fold change in either direction (0.5 to 2) for reference. Statistics compare unnormalized values. Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points.

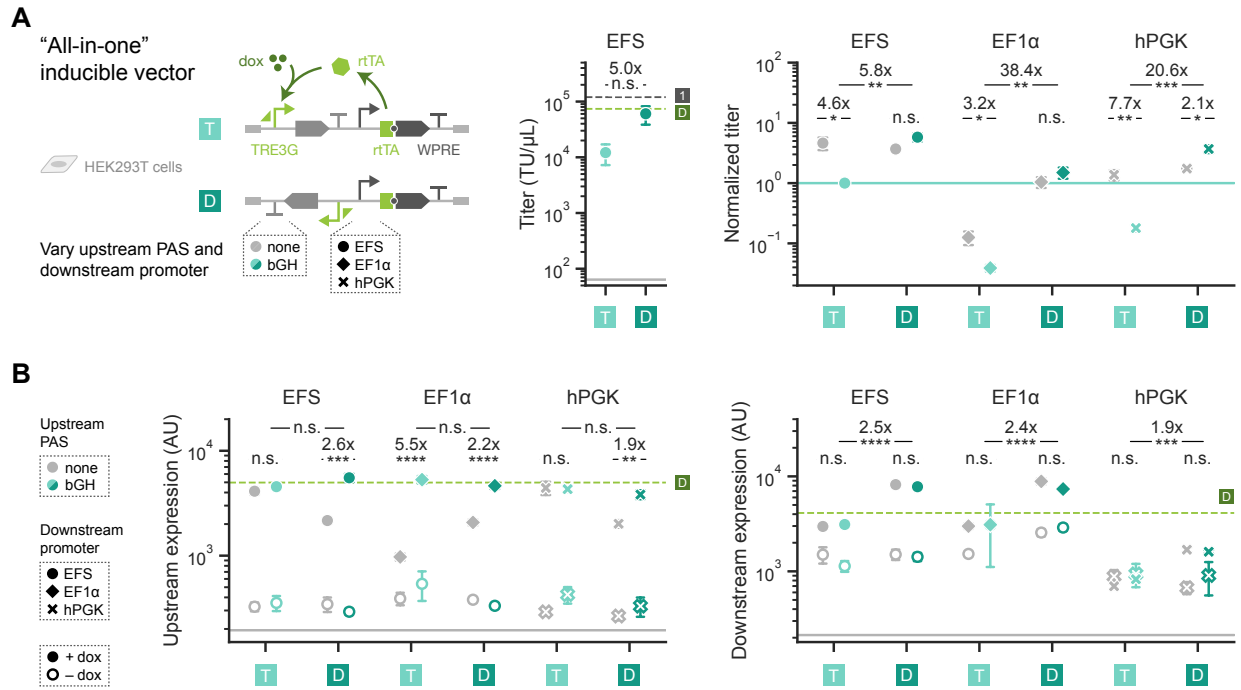


**Figure S2. Varying downstream 3' elements and upstream PAS impacts vector titer and expression.**

**A.** Performance of two-gene vectors with an inducible upstream gene as in Fig. 3, varying the upstream PAS and downstream 3' elements. Elements include combinations of the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE), bovine growth hormone (bGH) polyadenylation signal (PAS), simian virus 40 (SV40) PAS, or none. Vectors were transduced into HEK293T cells containing a separately integrated rTA cassette. Cells were treated with 1  $\mu$ g/mL dox (filled points) or untreated (open points). Solid green lines depict values for the vectors in Fig. 3, and dark gray dashed lines show values for the single-gene vector in Fig. 2 for reference. Annotations show the fold change between indicated points for conditions treated with dox.

**B.** Performance of single-gene vectors with the EFS promoter (circles) or EF1 $\alpha$  promoter (diamonds) and varying 3' elements. Vectors were transduced into HEK293T cells. Annotations show the fold change between indicated points.

Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Titer is in transducing units (TU) per  $\mu$ L (log scale). Expression is the geometric mean fluorescence in arbitrary units (AU, log scale). Solid light gray lines represent titers calculated for untransduced cells or expression gates. Gray shading highlights vectors with WPRE as the downstream 3' element, found in Figs. 2 and 3. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ .

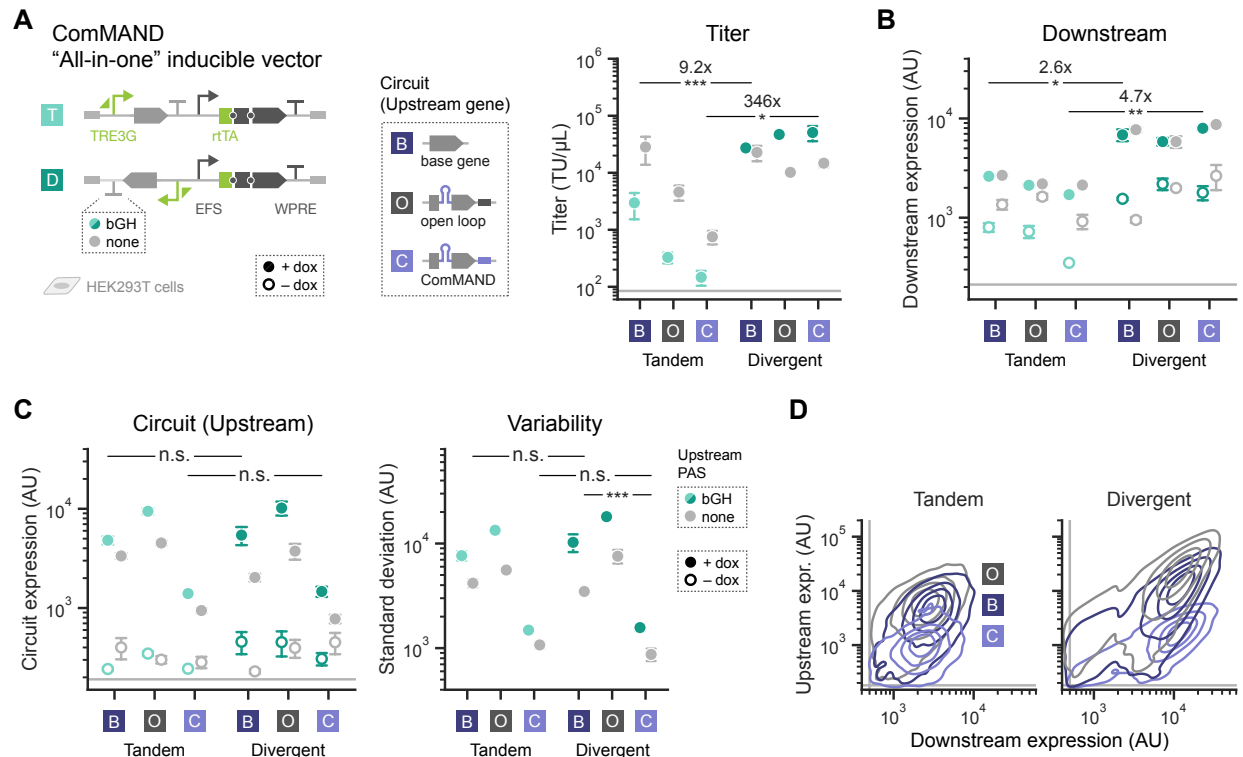


**Figure S3. “All-in-one” inducible vectors perform similarly with various genetic parts.**

**A.** “All-in-one” inducible vectors as in Fig. 4 were constructed varying the upstream PAS (none or bGH) and downstream promoter (EF1α, EFS, or hPGK). **Left:** Unnormalized viral titer for vectors in Fig. 4 in transducing units (TU) per μL (log scale). Dashed dark gray line depicts the value for the single-gene vector with the EFS promoter from Fig. 1C, and dashed green line shows the value for the inducible vector with divergent syntax from Fig. 3 for reference. Solid light gray line represents titer calculated for untransduced cells. **Right:** As in Fig. 4, viral titer is shown normalized to the vector with tandem syntax, upstream bGH PAS, and downstream EFS promoter (also indicated by the solid teal line) for each batch of virus (log scale).

**B.** Expression of the upstream and downstream genes for vectors in A transduced into HEK293T cells treated with 1 μg/mL dox (filled points) or untreated (open points). Solid light gray lines represent the expression gates. Dashed green lines depict values for the inducible vector with divergent syntax treated with dox from Fig. 3 for reference. Expression is the geometric mean fluorescence in arbitrary units (AU, log scale).

Points represent means ± standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student’s t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points for conditions treated with dox.



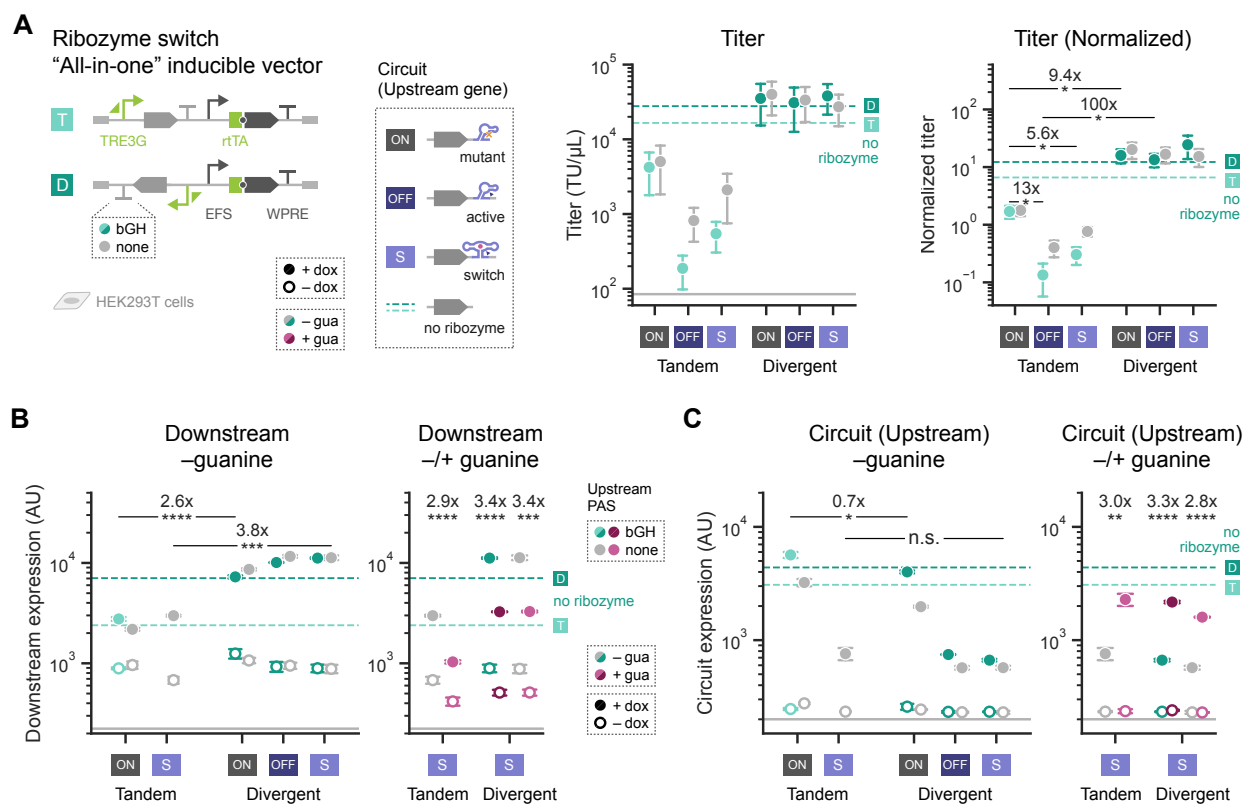
**Figure S4. Additional "all-in-one" inducible vectors expressing ComMAND match trends.**

**A.** "All-in-one" inducible vectors were constructed as in Fig. 5A varying the upstream PAS (bGH or none). The upstream gene expresses ComMAND (C), an open-loop circuit (O) containing the intronic microRNA and a mismatched target site as a control, or the output coding sequence alone (base gene, B). Here and in Fig. 5A, the downstream gene consists of rtTA, a fluorescent protein, and a puromycin resistance gene separated by 2A "self-cleaving" peptides. Titer is in transducing units (TU) per  $\mu\text{L}$  (log scale).

**B–C.** Expression of the downstream gene and the circuit (upstream gene) for HEK293T cells transduced with the same vectors and treated with  $1\ \mu\text{g}/\text{mL}$  dox (filled points) or untreated (open points). Expression is the geometric mean fluorescence in arbitrary units (AU, log scale) of populations gated on cells expressing both genes, and standard deviation (AU, log scale) is shown for expression of the upstream gene.

**D.** 2D density distributions depict upstream and downstream expression (AU, log scale) for a representative biological replicate of conditions in B–C, where tandem vectors lack a PAS on the upstream gene and divergent vectors include a bGH PAS as in Fig. 5A. Plots show cells treated with  $1\ \mu\text{g}/\text{mL}$  dox and gated on populations expressing both genes.

Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Solid light gray lines depict titer for untransduced cells or expression gates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points for conditions treated with dox.



**Figure S5. Additional “all-in-one” inducible vectors expressing the ribozyme switch match trends.**

**A.** “All-in-one” inducible vectors were constructed as in Fig. 5B varying the upstream PAS (bGH or none). The 3' UTR of the upstream gene contains either a mutant, minimally cleaving ribozyme (expression ON); a constitutively active ribozyme (expression OFF); a guanine (gua)-responsive ribozyme switch (S); or no ribozyme (dashed teal lines). **Left:** Titer is in transducing units (TU) per  $\mu\text{L}$  (log scale). **Right:** Titer is normalized to a vector lacking a ribozyme (not shown) for each batch of virus (log scale).

**B–C.** Expression of the downstream gene and the circuit (upstream gene) for HEK293T cells transduced with the same vectors and treated with 1  $\mu\text{g}/\text{mL}$  dox (filled points) or untreated (open points). Vectors containing the ribozyme switch include additional conditions where cells were also treated with 100  $\mu\text{M}$  guanine (pink), with or without dox. Expression is the geometric mean fluorescence in arbitrary units (AU, log scale).

Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Solid light gray lines depict titer for untransduced cells or expression gates. Statistics are two-sided Student's t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points for conditions treated with dox.

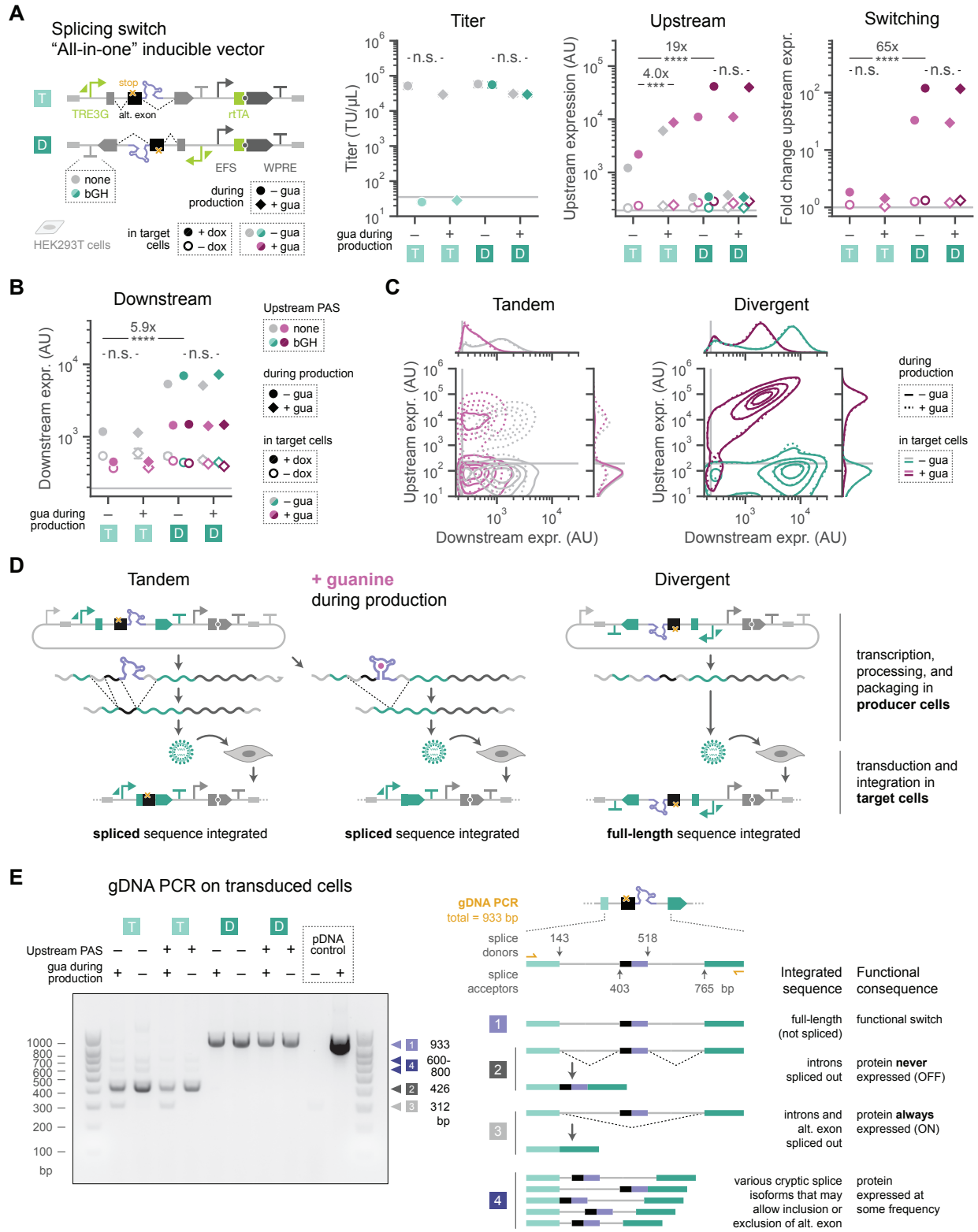


Figure S6. Additional "all-in-one" inducible vector designs expressing a splicing switch.

(caption on next page)

**Figure S6.** (continued)

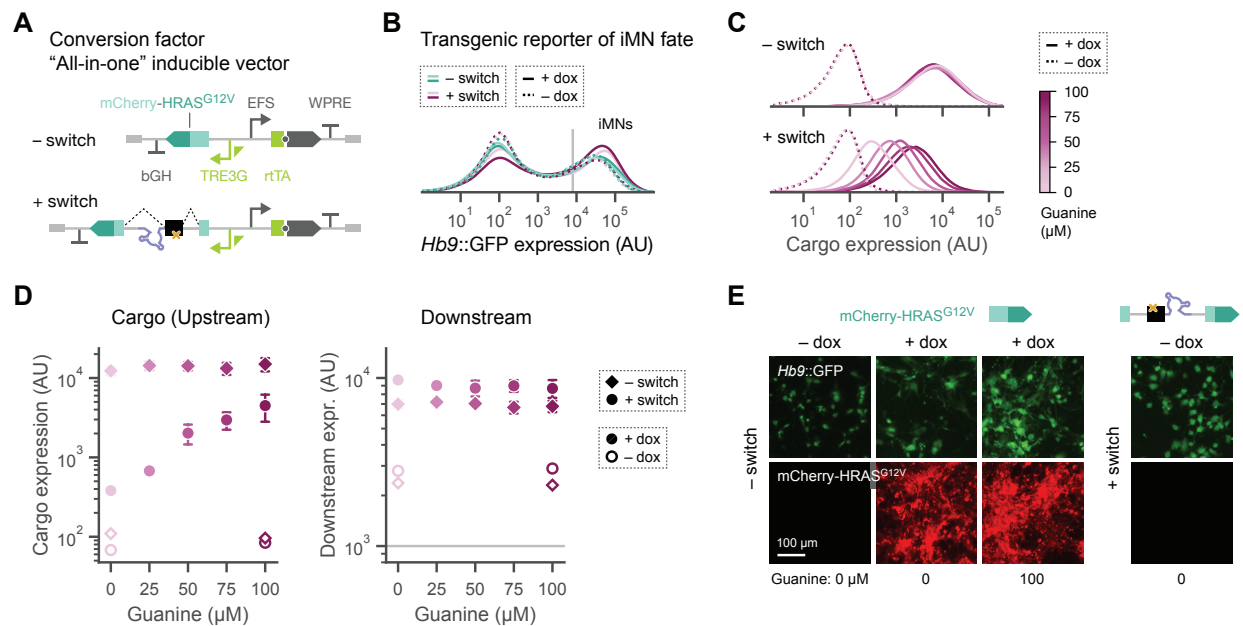
**A–B.** “All-in-one” inducible vectors as in Fig. 6A encode the splicing switch within the upstream gene, with divergent or tandem syntax and with no PAS or the bGH PAS on the upstream gene. Vectors were produced in HEK293T cells cultured without (circles) or with (diamonds) 100  $\mu$ M guanine (gua). Titer is in transducing units (TU) per  $\mu$ L (log scale). Expression of upstream and downstream genes is the geometric mean fluorescence in arbitrary units (AU, log scale) for HEK293T cells treated with (filled points) or without (open points) 1  $\mu$ g/mL dox and with (pink) or without 100  $\mu$ M guanine. Fold change in upstream expression shows conditions treated with guanine relative to those without.

**C.** 2D density distributions depict upstream and downstream expression for the conditions in A–B.

**D.** The splicing switch can interfere with virus production. When encoded on the sense strand of the viral genome, the switch may retain activity in producer cells, leading to packaging and integration of spliced sequences. Producer cells cultured with or without guanine can generate different splice products for tandem vectors, resulting in integration of different sequences lacking the splicing switch in target cells (undesired). In contrast, for the divergent vector, the viral transcript is not spliced in producer cells, leading to integration of the full-length sequence in target cells (desired).

**E. Left:** The gel from Fig. 6B is shown uncropped. The gel depicts PCR products amplified from genomic DNA (gDNA) of cells transduced with the vectors in A. Vectors have the bGH PAS (+) or no PAS (–) on the upstream gene and were produced in the presence (+) or absence (–) of 100  $\mu$ M guanine. As controls, plasmid DNA for one of the vectors was included in a separate reaction (pDNA +), or no template was added to the reaction (pDNA –). Lengths of the ladder bands are labeled in base pairs (bp), and the most prominent bands in the conditions are annotated with their predicted lengths and sequence. **Right:** The amplicon spans the site where the splicing switch was added, and primers (orange) bind in the mCherry coding sequence. Splice donor and acceptor sites are annotated (arrows) with their location on the amplicon (in bp). Predicted splice products and integrated sequences are shown for the most prominent bands on the gel, noting their functional consequence for protein expression in target cells.

Solid light gray lines depict titer for untransduced cells or expression gates. Points represent means  $\pm$  standard error for  $n \geq 3$  biological replicates. Statistics are two-sided Student’s t-tests, n.s.  $p \geq 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . Annotations show the fold change between indicated points for conditions treated with dox. Alt. exon, alternative exon.



**Figure S7. A splicing switch can regulate expression of a conversion factor during reprogramming.**

- A.** The splicing switch regulates expression of the oncogenic HRAS mutant HRAS<sup>G12V</sup>, which is fused to the fluorescent protein mCherry and delivered via the the best-performing divergent “all-in-one” inducible vector. Full vectors with or without the splicing switch used in Figs. 6C and 6D are shown.
- B.** Primary mouse embryonic fibroblasts were directly converted to induced motor neurons (iMNs) by virally delivering conversion factors including the vectors in A containing (pink) or lacking (teal) the splicing switch. Cells were treated with 1 μg/mL dox (solid line) or untreated (dashed line) for the duration of the conversion process. The transgenic reporter *Hb9::GFP* is an indicator of iMN fate. At 14 days post-transduction, GFP expression was quantified via fluorescence in arbitrary units (AU, log scale) to identify iMNs. Solid light gray line indicates the expression gate.
- C.** Expression of the mCherry-HRAS<sup>G12V</sup> cargo (upstream gene) is measured via fluorescence (AU, log scale) for iMNs identified as in B. Plots show distributions for cells transduced with vectors containing or lacking the splicing switch, treated with 1 μg/mL dox (solid lines) or without dox (dashed lines), and treated with 0–100 μM guanine.
- D.** Cargo and downstream expression are shown for conditions in C as a function of guanine concentration. Expression is geometric mean fluorescence (AU, log scale). Solid light gray line shows the expression gate. Points represent means ± standard error for  $n = 5$  biological replicates.
- E.** As in Fig. 6D, images show fluorescence of *Hb9::GFP* (top) and mCherry-HRAS<sup>G12V</sup> (bottom) at 14 days post-transduction for cells transduced with vectors lacking or containing the splicing switch, treated with or without dox, and treated with varying concentrations of guanine. Scale bar represents 100 μm.