Molecular Therapy
Methods & Clinical Development
Original Article

5′ Transgenes drive leaky expression of 3′ transgenes in Cre-inducible bi-cistronic vectors

Yasuyuki Osanai,[1,2] Yao Lulu Xing,[1,9] Shinya Mochizuki,[3] Kenta Kobayashi,[4,5] Jihane Homman-Ludiye,[1] Amali Cooray,[1] Jasmine Poh,[1] Ayumu Inutsuka,[6] Nobuhiko Ohno,[2,7] and Tobias D. Merson[1,8]

[1]Australian Regenerative Medicine Institute, Monash University, 15 Innovation Walk, Clayton, VIC 3800, Australia; [2]Department of Anatomy, Division of Histology and Cell Biology, School of Medicine, Jichi Medical University, Shimotsuke, Tochigi 329-0431, Japan; [3]Department of Anatomy, Bioimaging and Neuro-cell Science, Jichi Medical University, Shimotsuke, Tochigi 329-0431, Japan; [4]Section of Viral Vector Development, National Institute for Physiological Sciences, Myodaiji, Okazaki 444-8585, Japan; [5]The Graduate University for Advanced Studies (SOKENDAI), Shonan Village, Hayama, Kanagawa 240-0193, Japan; [6]Division of Brain and Neurophysiology, Department of Physiology, Jichi Medical University, Shimotsuke, Tochigi 329-0431, Japan; [7]Division of Ultrastructure Research, National Institute for Physiological Sciences, Myodaiji, Okazaki 444-8585, Japan; [8]Oligodendroglial Interactions Group, Systems Neurodevelopment Laboratory, National Institute of Mental Health, Bethesda, MD 20892, USA

Molecular cloning techniques enabling contemporaneous expression of two or more protein-coding sequences provide an invaluable tool for understanding the molecular regulation of cellular functions. The Cre-lox system is used for inducing the expression of recombinant proteins encoded within a bi-/poly-cistronic cassette. However, leak expression of transgenes is often observed in the absence of Cre recombinase activity, compromising the utility of this approach. To investigate the mechanism of leak expression, we generated Cre-inducible bi-cistronic vectors to monitor the expression of transgenes positioned either 5′ or 3′ of a 2A peptide or internal ribosomal entry site (IRES) sequence. Cells transfected with these bi-cistronic vectors exhibited Cre-independent leak expression specifically of transgenes positioned 3′ of the 2A peptide or IRES sequence. Similarly, AAV-*FLEX* vectors encoding bi-cistronic cassettes or fusion proteins revealed the selective Cre-independent leak expression of transgenes positioned at the 3′ end of the open reading frame. Our data demonstrate that 5′ transgenes confer promoter-like activity that drives the expression of 3′ transgenes. An additional *lox*-STOP-*lox* cassette between the 2A sequence and 3′ transgene dramatically decreased Cre-independent transgene expression. Our findings highlight the need for appropriate experimental controls when using Cre-inducible bi-/poly-cistronic constructs and inform improved design of vectors for more tightly regulated inducible transgene expression.

INTRODUCTION

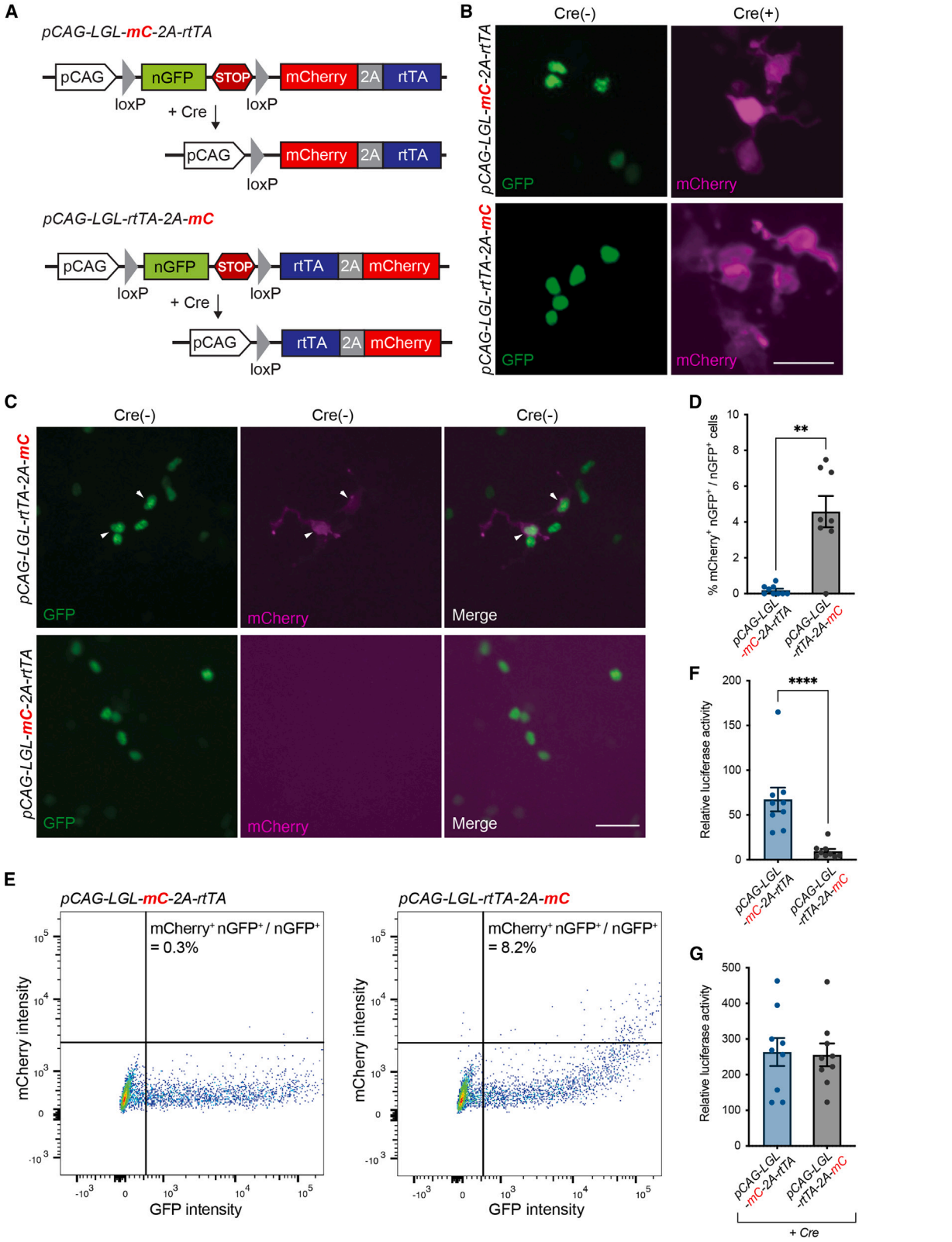
Cell type-specific bi-/poly-cistronic expression systems are powerful tools to investigate cell morphology and physiology.[1,2] To express recombinant proteins in a cell type-specific manner, a transgene regulated by a cell type-specific promoter or a combination of cell type-specific Cre transgenic mice with *lox*-STOP-*lox* or flip-excision (FLEx) vectors is widely used. The *lox*-STOP-*lox* cassette comprises two identical unidirectional *lox* sequences that flank an STOP sequence (poly(A) signal) designed to terminate transcription.[3,4] Inserting a fluorescent reporter gene, such as tdTomato immediately before the STOP sequence and inserting another reporter, such as GFP after the *lox*-STOP-*lox* cassette, enables one to distinguish between cells that have exhibited Cre activity (tdTomato$^-$/GFP$^+$) vs. those that have not been exposed to Cre (tdTomato$^+$/GFP$^-$).[5,6] An alternative strategy for Cre-inducible expression of transgenes is to use a FLEx switch, including a double-floxed inverted open reading frame (DIO). FLEx switches typically use two pairs of non-compatible *lox* sequence variants, such as a pair of *loxP* sites and a pair of either *lox2272*, *loxN* or *lox511*, to flip the orientation of a protein-coding sequence (CDS) from antisense to sense orientation upon Cre-mediated flip/excision.[7,8] Both *lox*-STOP-*lox* sequences and FLEx switches have been used to drive the overexpression of transgenes in cells that express Cre recombinase in a cell type-specific manner.[9] The smaller size of the FLEx switch (<300 bp) compared to most *lox*-STOP-*lox* cassettes (approximately 1,000 bp) make it particularly attractive for use in AAV vectors where cargo size is limited to <5 kb, including promoter, CDS, and poly(A) sequence.[10]

In many instances, the expression of multiple transgenes under the control of a single promoter is desired. One approach is to generate a fusion protein by concatenating the CDSs of two or more proteins into a single CDS. Fusion proteins are useful for identifying the subcellular localization of a target protein, for example, by expressing a protein of interest with an expression tag such as GFP. However, this strategy cannot be used if the expression of proteins with different

Received 3 November 2023; accepted 21 June 2024;
https://doi.org/10.1016/j.omtm.2024.101288.

[9]Present address: Department of Neurosurgery, Stanford University School of Medicine, CA, USA

Correspondence: Yasuyuki Osanai, Department of Anatomy, Division of Histology and Cell Biology, School of Medicine, Jichi Medical University, Shimotsuke, Tochigi 329-0431, Japan.
E-mail: yasuyuki.osanai@jichi.ac.jp
Correspondence: Tobias D. Merson, Oligodendroglial Interactions Group, Systems Neurodevelopment Laboratory, National Institute of Mental Health, Bethesda, MD 20892, USA.
E-mail: toby.merson@nih.gov



(legend on next page)

subcellular localizations. In this case, the CDS encoding multiple proteins can be separated by an internal ribosomal entry site (IRES) or self-cleaving 2A peptide. The use of 2A peptides has become the preferred method to express multiple genes under the control of a single promoter because of the smaller size and higher transgene expression.[11]

For maximal utility, an inducible expression system should allow for stringent control of transgene expression. Yet leaky expression of transgenes in the absence of an induction signal is frequently observed.[12–14] For inducible transgene expression systems that employ Cre recombinase as the induction signal, low levels of Cre-independent expression may be tolerable, depending upon the proteins being expressed. Low-level GFP expression, for instance, may be acceptable in an animal injected with an adeno-associated virus (AAV) designed to express the light-sensitive neuronal silencer archaerhodopsin and GFP (AAV-*FLEX-ArchT-GFP*). However, in some cases, even very low levels of leaky transgene expression can have a significant impact on cell function and on the interpretation of experimental data, as illustrated by the following three examples. First, leaky expression of diphtheria toxin fragment A (DTA), a transgene widely used as a suicide gene to ablate cells of interest,[15,16] would be unacceptable since a single molecule of DTA is sufficient to kill a cell.[17] Second, the loss of astrocyte-specific expression of mCherry in AAV-*GFAP-NeuroD1-T2A-mCherry* transfected cells, as observed by Wang and colleagues,[18] could lead to the incorrect conclusion that glial cells can transdifferentiate into neurons. Third, attempts to identify synaptically connected cells by infecting cells with an EnvA-pseudotyped recombinant rabies virus can be hampered by leaky expression of avian leukosis and sarcoma virus subgroup A receptor (TVA) in Cre-negative cells due to the strength of the EnvA-TVA interaction.[12,19–21]

Understanding the cause of leaky transgene expression is key to improving the stringency of inducible expression systems. It has been demonstrated that Cre-independent expression of transgenes encoded by bi-cistronic FLEx vectors can be reduced by repositioning the ATG start codon to a region 5′ of the *loxP* site just downstream of the promoter.[14,21] However, approaches to eliminate Cre-independent leaky expression are yet to be fully explored.[12,18]

In this study, we demonstrate that transgenes positioned 3′ of a 2A or IRES sequence in Cre-inducible bi-/poly-cistronic expression constructs are prone to Cre-independent expression. We reveal that leak expression is not derived from a typical promoter, but arises from promoter-like activity of the transgene that is positioned 5′ of the 2A sequence. Next, we show that introducing a second *lox*-STOP-*lox* sequence, using non-compatible *lox* sites, between the 2A sequence and 3′ CDS, eliminates leaky expression of the 3′ transgene. Finally, we demonstrate that *in vivo* delivery of AAV-*FLEX* vectors encoding bi-cistronic cassettes or fusion proteins results in the selective Cre-independent leak expression of transgenes positioned at the 3′ end of the open reading frame. Our findings have significant implications for transgenesis and gene therapy studies where the optimal design of constructs to precisely control the expression of multiple CDSs under the regulatory control of a single promoter is desired.

RESULTS

Leaky expression of the 3′ transgene in Cre-inducible bi-cistronic expression vectors

To investigate factors that influence Cre-independent gene expression of Cre-regulated gene expression vectors, we generated two variants of a *lox*-STOP-*lox* vector designed to express nuclear-targeted H2B-GFP (nGFP) in the absence of Cre activity and to express both the reverse tetracycline transactivator (rtTA) and membrane-targeted mCherry following Cre-mediated recombination. The two vectors differed only in the position of the rtTA and mCherry coding sequences relative to an intervening P2A sequence separating the two CDSs. In the first variant, *pCAG-loxP-nGFP-STOP-loxP-mCherry-P2A-rtTA*, hereafter denoted *pCAG-LGL-mC-2A-rtTA*, mCherry was positioned 5′ of the P2A sequence (Figure 1A). In the second variant, *pCAG-loxP-nGFP-STOP-loxP-rtTA-P2A-mCherry*, hereafter denoted *pCAG-LGL-rtTA-2A-mC*, mCherry was positioned 3′ of the P2A sequence. Each construct was transfected alone into HEK293T cells or in the presence of a Cre recombinase expression vector, *pCAG-Cre*. As expected, observation by fluorescence microscopy revealed that cells transfected with *pCAG-LGL-mC-2A-rtTA* alone expressed nGFP, whereas cells co-transfected with both *pCAG-LGL-mC-2A-rtTA* and *pCAG-Cre* expressed mCherry (Figure 1B). While we also observed nGFP expression in cells transfected with *pCAG-LGL-rtTA-2A-mC* alone, we noted that some cells co-expressed mCherry in the absence of *pCAG-Cre*, indicating leaky expression of mCherry (Figure 1C). The percentage of $nGFP^+$ cells that expressed mCherry was significantly higher in cells transfected with *pCAG-LGL-rtTA-2A-mC* compared to *pCAG-LGL-mC-2A-rtTA* (4.58 ± 0.87% vs. 0.19 ± 0.09%, $p = 0.002$, n = 8–9) (Figure 1D)

Figure 1. Transgenes located in the 3′ region of bi-cistronic vectors are leaky

(A) Schematic diagram of *pCAG-LGL-mC-2A-rtTA* and *pCAG-LGL-rtTA-2A-mC* vectors before and after Cre-mediated recombination. The two vectors differ in the position of rtTA and mCherry CDS. (B) HEK293T cells transfected with *pCAG-LGL-mC-2A-rtTA* (top) or *pCAG-LGL-rtTA-2A-mC* (bottom) expressing nucleus-targeted GFP in Cre (−) condition or membrane-targeted mCherry in the Cre (+) condition. Scale bar, 50 μm. (C) *pCAG-LGL-rtTA-2A-mC*-transfected HEK293T cells expressing leaky mCherry in the Cre (−) condition (top), while that transfected with *pCAG-LGL-mC-2A-rtTA* do not express mCherry (bottom). Scale bar, 50 μm. (D) Comparison of the percentage of $nGFP^+$ cells that exhibit leaky mCherry expression in the Cre (−) condition following transfection with either *pCAG-LGL-mC-2A-rtTA* or *pCAG-LGL-rtTA-2A-mC* ($n = 9$ wells for three independent experiment). (E) Flow cytometric analysis of mCherry and GFP expression among HEK293T cells transfected with *pCAG-LGL-mC-2A-rtTA* (left) or *pCAG-LGL-rtTA-2A-mC* (right). (F) Relative luciferase activity of HEK293T cells transfected with *pTRE-Luciferase* and either *pCAG-LGL-mC-2A-rtTA* or *pCAG-LGL-rtTA-2A-mC* compared to control cells transfected with *pTRE-Luciferase* alone, in the Cre (−) condition ($n = 9$ wells from three independent experiment). (G) Relative luciferase activity of HEK293T cells co-transfected with *pCAG-Cre* plus *pTRE-Luciferase* and either *pCAG-LGL-mC-2A-rtTA* or *pCAG-LGL-rtTA-2A-mC* as compared to control cells transfected with *pTRE-Luciferase* alone ($n = 9$ wells from three independent experiment). $^{**}p < 0.01$, $^{****}p < 0.0001$ using two-tailed Mann-Whitney *U* test. Data represent mean ± SEM.

and was corroborated by flow cytometric analysis (Figure 1E). We hypothesized that the mCherry CDS could be prone to leaky expression specifically when positioned 3′ of the P2A sequence because it is located further from the STOP cassette. To test whether the rtTA CDS is also subject to similar position-dependent leaky expression, we analyzed rtTA expression by transfecting HEK293T cells with either the *pCAG-LGL-rtTA-2A-mC* or *pCAG-LGL-mC-2A-rtTA* vector in combination with the *pTRE-Luciferase* plasmid, which expresses luciferase in response to rtTA expression. Luciferase assays revealed that, in the absence of *pCAG-Cre*, the expression of rtTA was significantly higher in cells transfected with *pCAG-LGL-mC-2A-rtTA* as compared to *CAG-LGL-rtTA-2A-mC* (67.3 ± 13.3 vs. 9.43 ± 2.67 RLU, $p < 0.0001$) (Figure 1F). In the presence of *pCAG-Cre*, luciferase activity was similar for both expression vectors (*pCAG-LGL-rtTA-2A-mC*: 255.6 ± 31.9 RLU, *pCAG-LGL-mC-2A-rtTA*: 263.7 ± 39.1 RLU, $p = 0.80$) (Figure 1G), indicating that the expression of rtTA following Cre-mediated recombination is similar, irrespective of its position relative to P2A. Collectively, these results indicate that transgenes positioned 3′ of the P2A sequence in the bi-cistronic expression cassette are prone to Cre-independent leak expression.

Changing the promoter does not prevent leaky expression

The CAG promoter is known to drive high-level gene expression in a constitutive manner. To establish whether the use of an alternative cell-type specific promoter prevents the leaky expression of the CDS positioned 3′ of the P2A sequence, we replaced the CAG promoter with the mouse myelin basic protein (MBP) proximal promoter, which is specifically active in oligodendrocytes, a type of glial cell in the central nervous system of vertebrates.[22] *pMBP-LGL-rtTA-2A-mC* and *pMBP-LGL-mC-2A-rtTA* were transfected into the oligodendroglial cell line, CG4.[23] In fluorescence microscope observations, the percentage of nGFP$^+$ cells that expressed mCherry was significantly higher in cells transfected with *pMBP-LGL-rtTA-2A-mC* as compared to *pMBP-LGL-mC-2A-rtTA* (2.79 ± 0.78% vs. 0.14 ± 0.14%, $p = 0.0013$) (Figure 2A) in the Cre (−) condition. Leaky mCherry expression in *pMBP-LGL-rtTA-2A-mC*-transfected cells was further confirmed by flow cytometry (Figure 2B). Similarly, Cre-independent leak expression of rtTA was significantly higher in *pTRE-Luciferase*-transfected cells that were co-transfected with *pMBP-LGL-mC-2A-rtTA* as compared to *pMBP-LGL-rtTA-2A-mC* (6.71 ± 1.29 vs. 1.19 ± 0.05 RLU, $p < 0.0001$) (Figure 2C). By contrast, when CG4 cells were transfected with both *pCAG-Cre* and *pTRE-Luciferase*, the expression level of rtTA was similar when co-transfected with either *pMBP-LGL-rtTA-2A-mC* or *pMBP-LGL-mC-2A-rtTA* (164.3 ± 28.7 vs. 129.2 ± 21.4 RLU, $p = 0.2982$) (Figure 2D). This result demonstrated that, in the recombined state, the expression of rtTA is not subject to position-dependent effects. Thus, changing the promoter or the target cell type did not alter the propensity for transgenes positioned 3′ of the P2A sequence to exhibit leaky expression.

Having demonstrated that both mCherry and rtTA exhibit leaky expression when positioned 3′ of the P2A sequence in our bi-cistronic vectors, we next sought to determine whether different sets of transgenes exhibit similar position-dependent leak expression in Cre-inducible poly-cistronic vectors. We found that placing tdTomato in the 3′ position of a poly-cistronic vector designed to express three different transgenes separated by P2A sequences led to a high degree of leak expression (16% of GFP$^+$ cells) in the absence of Cre recombinase (Figure 3). By contrast, leaky expression of the mCherry transgene was not observed when it was positioned in the 5′ position of other vectors (*pMBP-LGL-mC-P2A-DTR* or *pMBP-LGL-mC-P2A-FLAG/H2A-P2A-rtTA*) (Figure 3). Similarly, position-dependent leak expression of rtTA was observed in poly-cistronic vectors containing rtTA (Figure S1). We also found that switching the P2A sequence for an IRES sequence did not prevent Cre-independent leak expression of a tdTomato transgene positioned 3′ of the IRES sequence in the vector *pMBP-loxP-nGFP-STOP-loxP-DTR-IRES-tdT*, hereafter denoted *pMBP-LGL-DTR-IRES-tdT* (Figure 3). Following transfection of *pMBP-LGL-DTR-IRES-tdT* into CG4 cells, we observed that 8.9% of nGFP$^+$ cells also expressed tdTomato in the absence of Cre recombinase. Taken together, our data provide evidence that leaky expression of 3′ transgenes in bi-cistronic and poly-cistronic Cre-inducible vectors can occur in the absence of Cre recombinase, irrespective of promoter choice, selection of an IRES vs. P2A sequence, and in different cell types.

Leaky expression in cells transfected with promoterless constructs

Theoretically, the floxed STOP cassette in our Cre-dependent bi-cistronic expression vectors should be sufficient to block the transcription of downstream coding sequences. Our results thus far demonstrated that the floxed STOP cassette was effective at suppressing the expression of the immediate downstream CDS positioned 5′ of the P2A sequence, but was not sufficient to block expression of the CDS positioned 3′ of P2A. One possible explanation for this phenomenon is that the 5′ CDS serves to drive the expression of the 3′ transgene independent of the upstream promoter. To test this hypothesis, we removed the upstream promoter and STOP cassette from the pMBP expression vectors by enzymatic digestion and purified the *rtTA-2A-mC* and *mC-2A-rtTA* fragments by gel extraction (Figure 4A). These *rtTA-2A-mC* and *mC-2A-rtTA* segments were transfected into CG4 cells, and the expression of mCherry or rtTA was analyzed. Interestingly, some *rtTA-2A-mC*-transfected cells expressed mCherry, despite the removal of the MBP promoter (Figure 4B). The number of cells expressing mCherry was significantly higher for CG4 cells transfected with *rtTA-2A-mC* as compared to *mC-2A-rtTA* (4.15 ± 1.07 vs. 0.25 ± 0.12 cells/well, $p = 0.0009$) (Figure 4C). Notably, mCherry-expressing cells did not co-express nGFP (Figure 4B), consistent with removal of the nGFP CDS by enzymatic digestion. Although there was no statistically significant difference, the average luciferase activity of *mC-2A-rtTA* was higher than that of *rtTA-2A-mC* (7.18 ± 2.46 vs. 2.55 ± 0.49 RLU, $p = 0.0951$) (Figure 4D). Together, these results demonstrate that leak expression of the CDS positioned 3′ of the P2A sequence persists in the absence of a *bona fide* promoter. Thus, bi-cistronic Cre-dependent expression cassettes comprising two or more concatenated transgenes can exhibit leaky expression of the 3′ transgene in the absence of Cre activity and independent of a *bona fide* promoter.

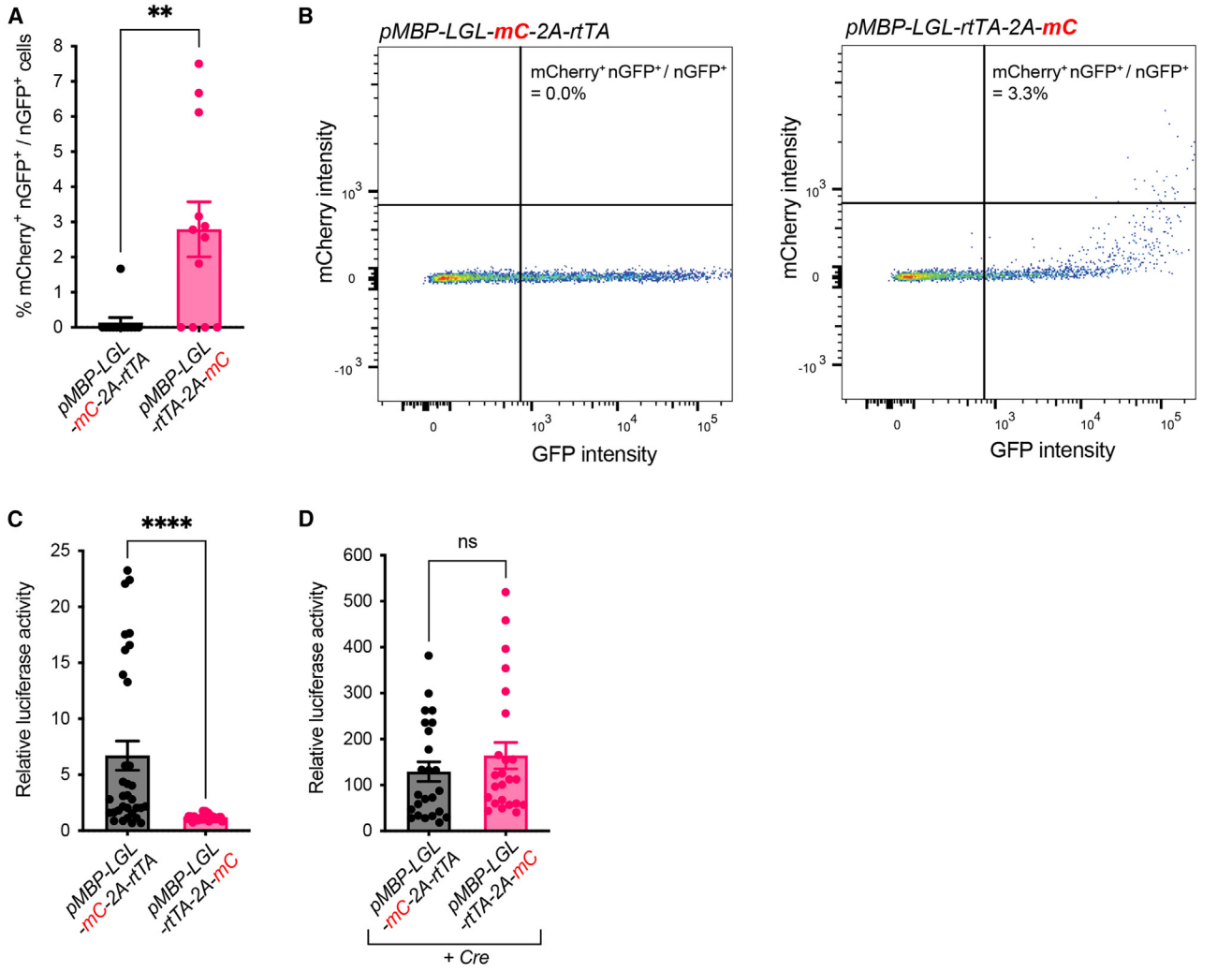


Figure 2. Inducible bi-cistronic vectors employing a cell type-specific promoter exhibit leaky 3′ transgene expression

(A) Comparison of the percentage of transfected CG4 cells that exhibit leaky mCherry expression following transfection with either *pMBP-LGL-mC-2A-rtTA* or *pMBP-LGL-rtTA-2A-mC* in the Cre (–) condition (n = 12 wells from four independent experiments). (B) Flow cytometric analysis of mCherry and GFP expression among CG4 cells transfected with either *pMBP-LGL-mC-2A-rtTA* or *pMBP-LGL-rtTA-2A-mC*. (C) Relative luciferase activity of CG4 cells transfected with *pTRE-Luciferase* and either *pMBP-LGL-mC-2A-rtTA* or *pMBP-LGL-rtTA-2A-mC* compared to control cells transfected with *pTRE-Luciferase* alone, in the Cre (–) condition (n = 15 wells from 5 independent experiments). (D) Relative luciferase activity of CG4 cells co-transfected with *pCAG-Cre* and *pTRE-Luciferase* and either *pMBP-LGL-mC-2A-rtTA* or *pMBP-LGL-rtTA-2A-mC* compared to control cells transfected with *pTRE-Luciferase* alone (n = 12 wells from four independent experiments). $^{**}p < 0.01$, $^{****}p < 0.0001$ using two-tailed Mann-Whitney U tests. Data represent mean ± SEM.

Next, to investigate whether the mCherry or P2A sequences were responsible for leak expression of the 3′ transgene, we separately PCR-amplified the *rtTA*, *P2A-rtTA*, and *mC-P2A-rtTA* segments, using the *pMBP-LGL-mC-2A-rtTA* vector as template (Figure 4E). Each segment was co-transfected with *pTRE-Luciferase* into CG4 cells and luciferase activity was assayed. Cells transfected with either the *rtTA* segment or *P2A-rtTA* segment did not exhibit luciferase activity above control levels whereas cells transfected with the *mC-P2A-rtTA* segment exhibited a significantly higher luciferase activity compared to blank control wells (2.21 ± 0.38 vs. 1.00 ± 0.07 RLU, p = 0.0005) (Figure 4F). Taken together, these data demonstrate that the 5′ CDS (mCherry or rtTA), but not the P2A sequence, was necessary and sufficient to drive leaky expression of 3′ transgene.

Inserting transgenes into the genome reduces leaky transgene expression

Although our results indicated that transient transfection of inducible bi-cistronic vectors led to a significant amount of leaky expression of 3′ transgenes, some studies using transgenic mice containing Cre-inducible bi-cistronic cassettes have indicated that leakage is rarely observed in mouse tissues *in vivo*.[24,25] To determine the extent of leak expression following genomic integration of Cre-inducible

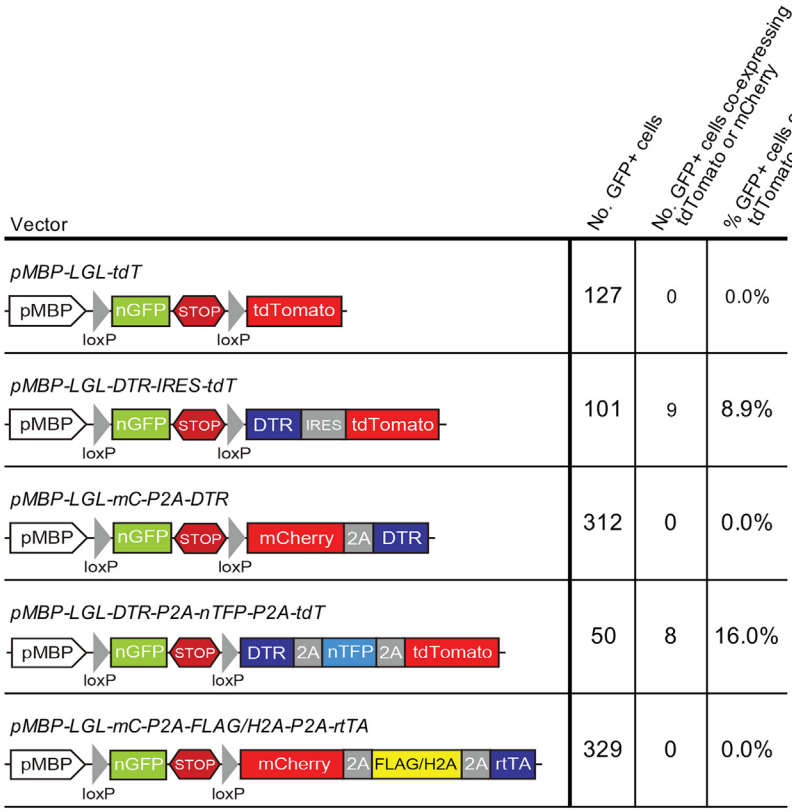


Figure 3. Position of tdTomato or mCherry CDS within inducible bi-cistronic and poly-cistronic expression vectors influences the degree of leaky expression

Vectors were transfected into HEK293T cells and the number of cells expressing GFP was quantified. Additionally, the absolute number and percentage of GFP$^+$ cells that co-expressed tdTomato was recorded. ns, not significant.

bi-cistronic expression vectors we generated stable cell lines HEK-*pCAG-LGL-rtTA-2A-mC* and HEK-*pCAG-LGL-mC-2A-rtTA*. We observed a non-significant trend for increased leak expression of 3′ transgenes in stably transfected HEK293T cells. Specifically, we observed low-level expression of mCherry in a subset of nGFP$^+$ cells stably transfected with either *pCAG-LGL-rtTA-2A-mC* or *pCAG-LGL-mC-2A-rtTA* (0.30 ± 0.14 vs. 0.05 ± 0.03%, $p = 0.17$) (Figures 5A and 5B). Luciferase assays in the Cre (−) condition also revealed a non-significant trend for higher rtTA expression in HEK-*pCAG-LGL-mC-2A-rtTA* cells as compared to HEK-*pCAG-LGL-rtTA-2A-mC* (3.11 ± 0.78 vs. 1.67 ± 0.30 RLU, $p = 0.12$) (Figure 5C). Importantly, mCherry and rtTA leakage levels in stably transfected cells were on average approximately 25-fold and approximately 22-fold lower, respectively, than those observed following transient transfection (% of mCherry$^+$ nGFP$^+$/nGFP$^+$ cells in stable vs. transient transfectants for *pCAG-LGL-rtTA-2A-mC* was 0.30 ± 0.14% vs. 7.49 ± 3.01%, $p = 0.0018$ [Figure 5D]; relative luciferase activity in stable vs. transient transfectants of *pCAG-LGL-mC-2A-rtTA* was 3.11 ± 0.78 vs. 67.3 ± 13.3 RLU, $p < 0.0001$ [Figure 5E]). In the presence of Cre, the stable cell line HEK-*pCAG-LGL-rtTA-2A-mC* exhibited abundant mCherry expression (Figure 5F). Stable transfectants also exhibited lower relative luciferase activity in the Cre (+) condition as compared to transiently transfected cells (stable HEK-*pCAG-LGL-mC-2A-rtTA* vs. HEK-*pCAG-LGL-rtTA-2A-mC*, 125.8 ± 48.5 *vs.* 161.5 ± 45.3 RLU; transient *pCAG-LGL-mC-2A-rtTA* vs. *pCAG-LGL-rtTA-2A-mC* 263.7 ± 39.1 *vs.* 255.6 ± 31.9 RLU; two-way ANOVA comparing mode of transfection: $p = 0.009$) (Figure 5G). Interestingly, whereas rtTA leak expression for *pCAG-LGL-mC-2A-rtTA* in the Cre (−) condition was approximately 22-fold lower in stable vs. transient transfectants (Figure 5E), in the Cre (+) condition, rtTA expression was just approximately 2-fold lower in stable vs. transient transfectants (stable HEK-*pCAG-LGL-mC-2A-rtTA*: 125.8 ± 48.5 RLU, transient *pCAG-LGL-mC-2A-rtTA*: 263.7 ± 39.1 RLU) (Figure 5G). These data demonstrate that, even though rtTA expression in the presence of Cre was higher in transiently transfected cells, stable transfectants exhibited a greater fold change in rtTA expression following Cre recombination. Collectively, these data suggest that leak expression of the 3′ transgene was more pronounced in transient expression experiments than in stable expression experiments.

STOP cassette insertion into the 3′ region of P2A prevents leak expression of 3′ transgene

We reasoned that inserting another *lox*-STOP-*lox* cassette between P2A and the 3′ transgene would serve to prevent transcription of the 3′ transgene initiated by promoter-like activity of the 5′ CDS within the bi-cistronic expression cassette. To test this idea, we inserted another *lox*-STOP-*lox* site after the P2A sequence to generate the vector *pMBP-lox2272-nG-STOP-lox2272-mC-2A-loxP-STOP-loxP-rtTA* (hereafter denoted *pMBP-LGL-mC-2A-LSL-rtTA*) (Figure 6A). This construct expressed nGFP but not mCherry in the Cre (−) condition, and it expressed both mCherry and rtTA in Cre (+) condition (Figure 6B) because both *lox2272-nG*-STOP-*lox2272* and *loxP*-STOP-*loxP* sites are independently excised via exclusive recombination events that occur between respective pairs of *lox2272* and *loxP* sites. Further, we inserted an M3 promoter domain into the MBP promoter, which is known to increase the MBP promoter activity and specificity.[22,26,27] Cells transfected with the *pMBP-LGL-mC-2A-LSL-rtTA* did not express leaky mCherry (none of 2,130 nGFP$^+$ cells were found to co-express mCherry by fluorescence microscopic observation). Using flow cytometric analysis, the minimal leakage of mCherry in the Cre (−) condition and high-level expression of mCherry in Cre (+) condition were corroborated by comparing flow cytometry profiles of cells transfected with

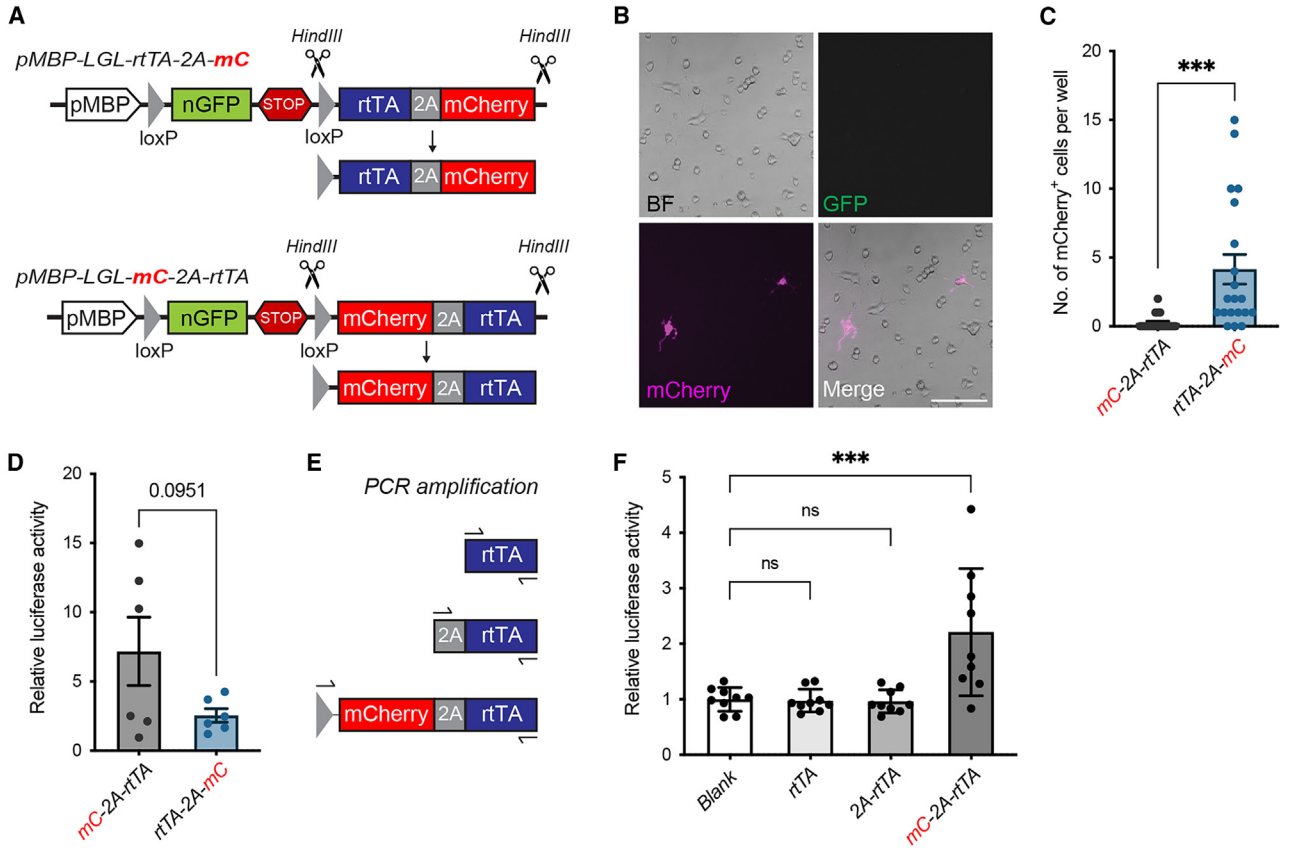


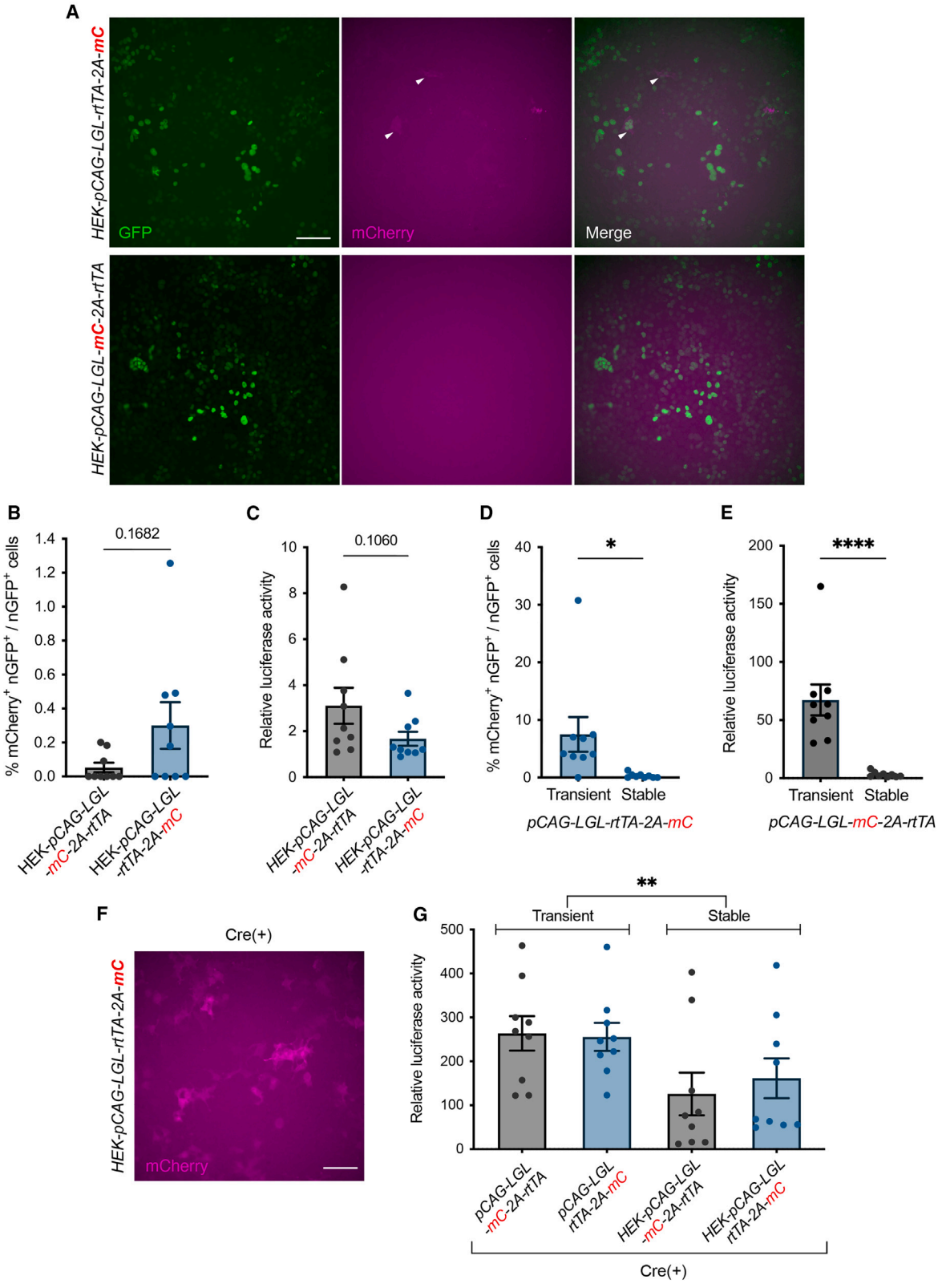
Figure 4. Leaky expression was observed without a general promoter

(A) Schematic diagram of the preparation of *rtTA-2A-mC* and *mC-2A-rtTA* segments. (B) Photomicrograph of *rtTA-2A-mC* segment-transfected CG4 cells expressing mCherry. Scale bar, 100 μm. (C) Comparison of the percentage of cells exhibiting leaky mCherry expression for *rtTA-2A-mC* segment- or *mC-2A-rtTA* segment-transfected CG4 cells in the Cre (−) condition (n = 20 wells from 6 independent experiments). (D) Relative luciferase activity of CG4 cells transfected with *pTRE-Luciferase* and either the *rtTA-2A-mC* or *mC-2A-rtTA* segment, as compared to control cells transfected with *pTRE-Luciferase* alone (data represent means from 6 independent experiments, n = 3 wells per experiment). (E) Schematic diagram of PCR-amplified *rtTA* segments, *2A-rtTA* segments, and *mC-2A-rtTA* segments. (F) Relative luciferase activity of CG4 cells transfected with *pTRE-Luciferase* and either PCR-amplified *rtTA*, *2A-rtTA*, or *mC-2A-rtTA*, as compared to control cells transfected with *pTRE-Luciferase* alone (n = 9 wells from 3 independent experiments). *$p < 0.05$, ***$p < 0.001$, two-tailed Mann-Whitney U test (C), unpaired t test, (D) or one-way ANOVA with Dunnett's test (F). Data represent mean ± SEM. ns, not significant.

pMBP-LGL-rtTA-2A-mC, *pMBP-LGL-mC-2A-rtTA* or *pMBP-LGL-mC-LSL-rtTA* (Figure 6C). As expected, the level of leak expression of rtTA in cells transfected with *pMBP-LGL-mC-2A-LSL-rtTA* was significantly lower than that of *pMBP-LGL-mC-2A-rtTA*-transfected cells in the Cre (−) condition (*pMBP-LGL-rtTA-2A-mC*: 1.17 ± 0.07%, *pMBP-LGL-mC-2A-rtTA*: 9.38% ± 1.94%, *pMBP-LGL-mC-2A-LSL-rtTA*: 1.43 ± 0.12%; *pMBP-LGL-rtTA-2A-mC* vs. *pMBP-LGL-mC-2A-rtTA*, $p < 0.0001$; *pMBP-LGL-mC-2A-rtTA* vs. *pMBP-LGL-mC-2A-LSL-rtTA*; $p < 0.0099$) (Figure 6D). The expression of rtTA in *pMBP-LGL-mC-2A-LSL-rtTA* transfected cells in the Cre (+) condition was approximately 30% of the level observed with the original constructs (*pMBP-LGL-rtTA-2A-mC*: 156.3 ± 40.8 RLU, *pMBP-LGL-mC-2A-rtTA*: 147.6 ± 27.4 RLU, *pMBP-LGL-mC-2A-LSL-rtTA*: 47.1 ± 8.0 RLU; *pMBP-LGL-rtTA-2A-mC* vs. *pMBP-LGL-mC-2A-LSL-rtTA*, $p = 0.0003$; *pMBP-LGL-mC-2A-rtTA* vs. *pMBP-LGL-mC-2A-LSL-rtTA*, $p = 0.0136$) (Figure 6E). These results support our findings that the 5′ CDS drives 3′ transgene expression, suggesting that the expression could be prevented via the insertion of an additional *lox-STOP-lox* cassette preceding the 3′ transgene.

Leak expression of the 3′ transgene in widely used AAV-*FLEX* constructs

Finally, we analyzed whether widely used Cre-inducible expression vectors designed to enable inducible expression of concatenated CDSs exhibit 3′ transgene leakage *in vivo*. We prepared AAV from three AAV-*FLEX* constructs which contained the GFP CDS in either a 5′ position within a bi-cistronic cassette (AAV-*EF1a-FLEX-GFP-2A-TVA*) or in the 3′ position of a CDS encoding a GFP fusion protein (AAV-*βactin-FLEX-DTR-GFP* and AAV-*EF1a-FLEX-ArchT-GFP*, Figure 7A). AAV2 particles generated from each AAV-*FLEX* construct were injected into the mouse eye together with AAV-*CMV-mCherry* or AAV-*CAG-Cre-mCherry* (each AAV 1×10^9 vg,



(legend on next page)

total volume of AAV solution 1 μL) (Figure 7B), and the mRNA expression level of 5′ and 3′ transgenes relative to *mCherry* expression was measured using droplet digital PCR (ddPCR) (Figures 7C–7E). In the absence of Cre recombinase, we found that leak expression of the 3′ transgene encoded by two of the three AAV-*FLEX* vectors was significantly higher than the 5′ transgene for eyes co-injected with AAV-*CMV-mCherry* (AAV-*EF1a-FLEX-GFP-2A-TVA*: *GFP* 32.9 ± 2.1%, *TVA* 88.6 ± 7.4%, $p < 0.0001$ [Figure 7C]AAV-*βactin-FLEX-DTR-GFP*: *DTR* 84.7 ± 24.34%, *GFP* 174.2 ± 55.3%, $p = 0.024$ [Figure 7D]). There was no statistically significant difference between GFP and ArchT expression in retinae transduced with AAV-*EF1a-FLEX-ArchT-GFP* and AAV-*CMV-mCherry* (*ArchT* 24.0% ± 4.0%, *GFP* 27.3% ± 4.0%, $p = 0.153$) (Figure 7E). However, we found no significant difference in the level of expression of 5′ vs. 3′ transgenes when eyes were co-injected with AAV particles generated from an AAV-*FLEX* construct together with AAV-*CAG-Cre-mCherry* (AAV-*EF1a-FLEX-GFP-2A-TVA*: *GFP* 169.5 ± 54.0%, *TVA* 229.2 ± 60.9%, $p = 0.22$ [Figure 7C]; AAV-*βactin-FLEX-DTR-GFP*: *DTR* 201.1 ± 113.5%, *GFP* 282.1 ± 139.9%, $p = 0.154$ [Figure 7D]; AAV-*EF1a-FLEX-ArchT-GFP*: *ArchT* 1475 ± 685.4%, *GFP* 1479 ± 690.9%, $p = 0.933$ [Figure 6E]). These results indicate that 3′ transgenes were expressed independent of Cre activity and at higher levels than 5′ transgenes.

Next, we stereotaxically injected the brains of anesthetized mice with AAV particles generated from either one of each AAV-*FLEX* construct together with AAV-*CMV-mCherry* or AAV-*CAG-Cre-mCherry* and counted GFP/mCherry double-positive cells in brain sections 14 days later (Figures 7F–7H). As predicted based on our previous results, in the absence of Cre recombinase, the percentage of mCherry$^+$ cells that co-expressed GFP was more than 8.8-fold higher for AAV-*FLEX* constructs with GFP in the 3′ position as compared to the condition where GFP was in the 5′ position (AAV-*EF1a-FLEX-GFP-2A-TVA*: 2.55 ± 0.85%, AAV-*βactin-FLEX-DTR-GFP*: 11.7 ± 3.16%, AAV-*EF1a-FLEX-ArchT-GFP*: 22.6 ± 1.65, *GFP-2A-TVA* vs. *DTR-GFP*: $p = 0.0178$, *GFP-2A-TVA* vs. *ArchT-GFP*: $p < 0.0001$) (Figure 7H). When we switched to co-injection of AAV-*CAG-Cre-mCherry* instead of AAV-*CMV-mCherry*, the percentage of mCherry$^+$ cells co-expressing GFP increased for all AAV-*FLEX* constructs (AAV-*EF1a-FLEX-GFP-2A-TVA*: 46.4 ± 6.48%, AAV-*βactin-FLEX-DTR-GFP*: 20.9 ± 3.23%, AAV-*EF1a-FLEX-ArchT-GFP*: 39.6 ± 2.45, *GFP-2A-TVA* vs. *DTR-GFP*: $p = 0.007$) (Figure 7I). Strikingly however, the increase in the percentage of mCherry+ cells co-expressing GFP was most marked for the AAV-*FLEX* vector in which GFP occupied the 5′ position (>18-fold increase) compared to those where GFP was in the 3′ position (<1.8-fold increase).

These results indicated that the 3′ transgene was leaky in bi-cistronic AAV vector-infected cells *in vivo*. Importantly, the fact that we observed leaky GFP expression in the absence of Cre recombinase when using the AAV-*βactin-FLEX-DTR-GFP* and AAV-*EF1a-FLEX-ArchT-GFP* vectors indicates that the 3′ CDS of fusion proteins are also subject to leaky expression in addition to bi-cistronic cassettes where the CDS is separated by 2A or IRES sequences. Taken together, our data demonstrate that 3′ transgenes in inducible bi-cistronic or fusion-protein vectors are prone to leaky expression both *in vitro* and *in vivo*.

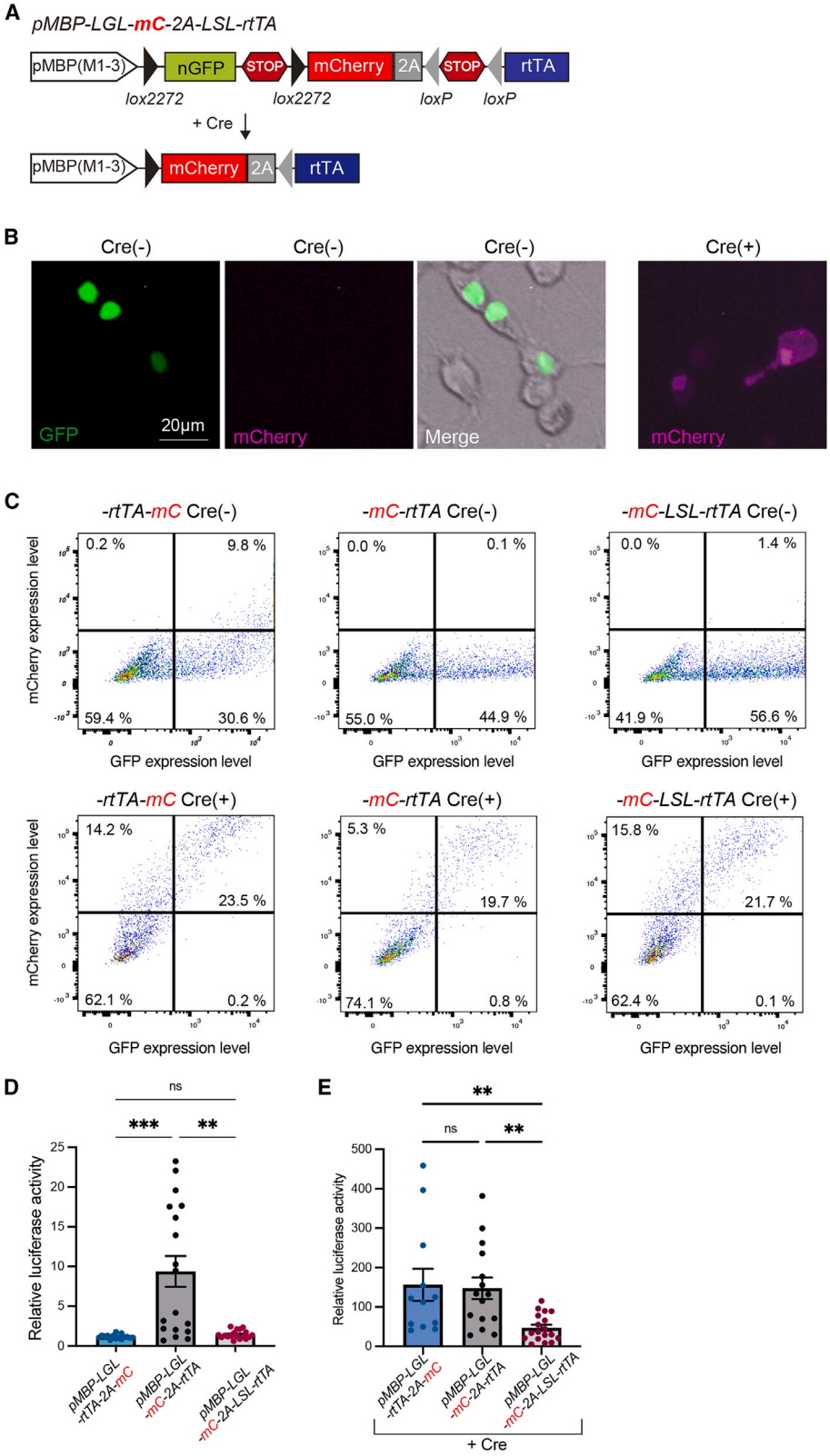
DISCUSSION

In this study, we observed that CDSs located in the 3′ regions of poly-cistronic Cre-inducible expression vectors are prone to Cre-independent leak expression. Our data demonstrate that this leaky expression depends on the promoter-like activity of 5′ CDSs rather than the activity of a *bona fide* promoter whose activity is incompletely blocked by a downstream floxed STOP cassette. The broad implications of our findings are that tight regulation of the expression of 3′ coding sequences within fusion proteins or poly-cistronic expression cassettes, where coding sequences are separated by IRES or 2A peptides, cannot be guaranteed by using inducible regulatory elements such as *lox*-STOP-*lox* cassettes, FLEx/DIO cassettes, or cell type-specific promoters, especially in transient expression experiments.

Although there have been previous reports of leaky transgene expression by researchers utilizing inducible gene expression strategies,[12,18,28,29] the mechanisms underlying leak expression have not been thoroughly explored, and the implications of leak expression have been largely dismissed. Low-level Cre-independent leak expression is often disregarded if it is assumed to have no impact on the

Figure 5. Significant reduction in 3′ transgene leakage in stable cell lines with inducible bi-cistronic expression

(A) Representative photomicrographs of the HEK-*pCAG-LGL-rtTA-2A-mC* cell line (top) or HEK-*pCAG-LGL-mC-2A-rtTA* cell line (bottom). Both cell lines correctly expressed nGFP in the Cre (−) condition; however the HEK-*pCAG-LGL-rtTA-2A-mC* cell line expressed leaky mCherry in some cells (arrowheads). Scale bar, 100 μm. (B) Comparison of the percentage of nGFP$^+$ cells exhibiting leaky mCherry expression for the HEK-*pCAG-LGL-mC-2A-rtTA* and HEK-*pCAG-LGL-rtTA-2A-mC* cell lines in the Cre (−) condition ($n = 9$ wells from 3 independent experiments). (C) Relative luciferase activity of the HEK-*pCAG-LGL-mC-2A-rtTA* and HEK-*pCAG-LGL-rtTA-2A-mC* cell lines transiently transfected with *pTRE-Luciferase* as compared to control HEK293T cells transiently transfected with *pTRE-Luciferase* alone ($n = 9$ wells from 3 independent experiments). (D) Comparison of the percentage of nGFP$^+$ HEK293T cells exhibiting leaky mCherry expression following transient vs. stable transfection with *pCAG-LGL-rtTA-2A-mC* in the Cre (−) condition. Datasets were obtained from Figure 1D or Figure 4B, respectively ($n = 9$ wells from three independent experiments). (E) Relative luciferase activity of *pCAG-LGL-mC-2A-rtTA*-transfected HEK293T cells and HEK-*pCAG-LGL-mC-2A-rtTA* stably-transfected cells following transient transfection with *pTRE-Luciferase* as compared to control HEK293T cells transiently transfected with *pTRE-Luciferase* alone. Datasets were obtained from Figure 1F or Figure 4C, respectively ($n = 9$ wells from 3 independent experiments). (F) Photomicrograph of mCherry expression by HEK-*pCAG-LGL-rtTA-2A-mC* stably-transfected cells after transient transfection with *pCAG-Cre.* Scale bar, 100 μm. (G) Comparison of relative luciferase activity of HEK293T cells transiently transfected with *pCAG-Cre* and *pTRE-Luciferase* and either *CAG-LGL-rtTA-mC* or *CAG-LGL-mC-2A-rtTA*, vs. *pCAG-Cre* and *pTRE-Luciferase* co-transfected HEK-*CAG-LGL-rtTA-2A-mC* and HEK-*CAG-LGL-mC-rtTA* stable cell lines ($n = 9$ wells from 3 independent experiments). Datasets for transient expression were obtained from Figure 1G. $^*p < 0.05$, $^{****}p < 0.001$ or actual p values are indicated, two-tailed Mann-Whitney *U* test (B-E), or two-way ANOVA with Tukey test (G). Data represent mean ± SEM.



(legend on next page)

experiment's outcomes. For example, if the level of leaky expression of GFP or any other fluorescent protein is sufficiently low, this background signal may be set below the detection threshold for real expression and data analysis may not be significantly affected. However, we observed that more than 7% of cells transfected with *pCAG-LGL-rtTA-2A-mC* alone expressed mCherry, in the absence of Cre recombinase, at levels that could be readily detected by fluorescence microscopy or flow cytometric analysis without signal amplification or the use of anti-mCherry antibodies. Moreover, positioning a tdTomato CDS at the 3′ end of a poly-cistronic expression cassette comprising three CDS separated by 2A sequences resulted in leaky tdTomato expression by 16% of transfected cells. By contrast, the expression of mCherry was minimal in the absence of Cre recombinase when the CDS was positioned at the 5′ end of poly-cistronic expression cassettes. We observed similar position-dependent effects on the expression of the rtTA. We conclude from these analyses, that when it comes to inducible expression vectors, (1) the expression level of 3′ transgenes can be readily observed in a subset of transfected cells without the transcriptional inducer and (2) each transgene encoded within a poly-cistronic expression vector cannot be inferred by assaying the expression level of a single CDS, especially in the absence of the transcriptional inducer. This finding is particularly problematic when the intention is to regulate the expression of proteins for which just a few molecules are sufficient to execute their functions. For example, DTA is a widely used suicide gene,[15,16] and one molecule of DTA is sufficient to kill a cell.[17] Similarly, the expression of DNA recombinases such as Cre, Dre and Flp can induce recombination of site-specific recombinase (SSR)-dependent reporter alleles even when expressed transiently and at a low level. Based on our observations, inducible expression of suicide genes (e.g., DTA) or SSRs may only be possible if they are placed in the 5′ region of a bi-cistronic vector since the 5′ CDS gene expression is much more tightly regulated compared to 3′ CDS transgene expression.

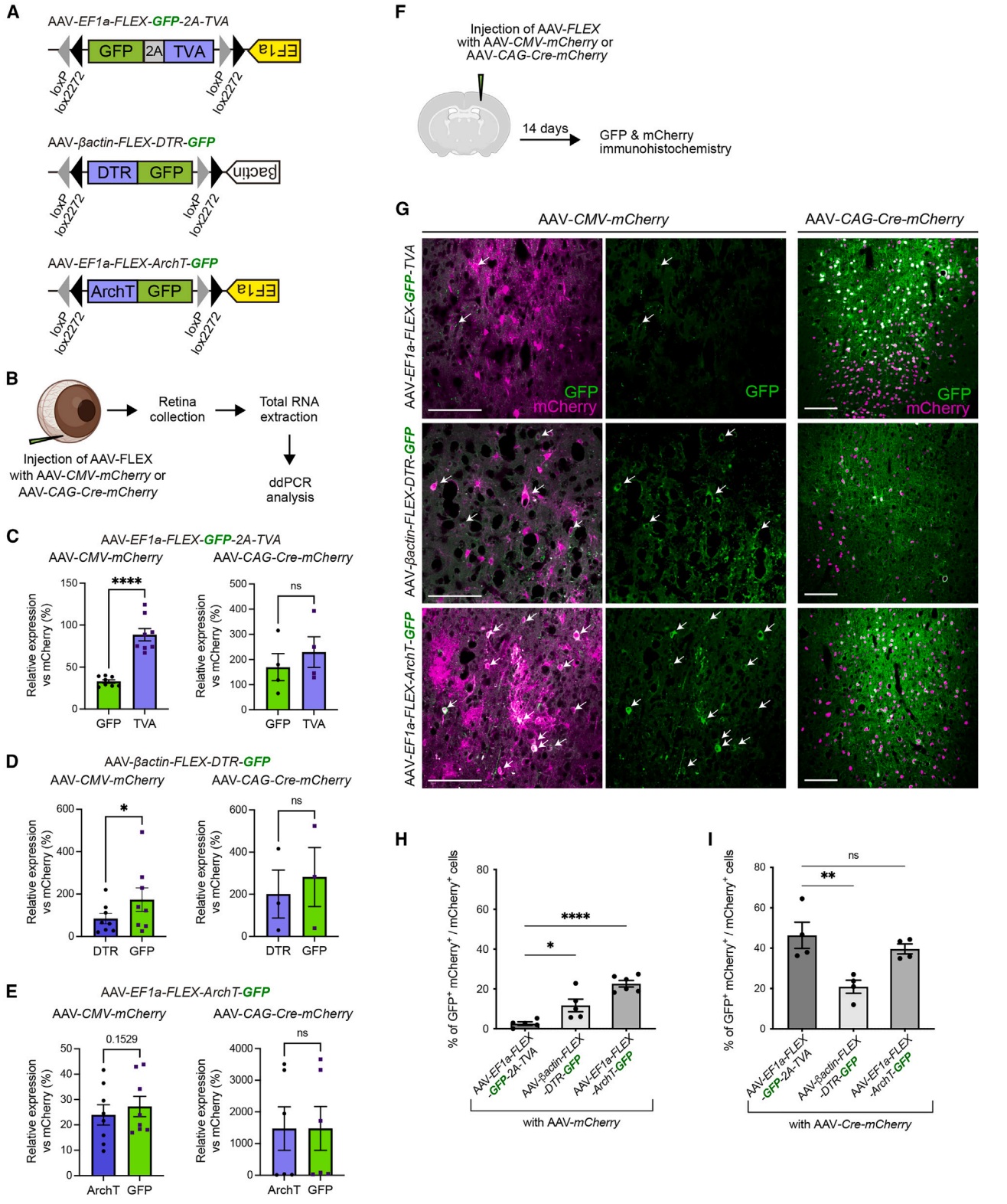
We found that all tested CDSs, including mCherry, rtTA, GFP, TFP, DTR, and ArchT, when positioned in the 5′ position of a bi-cistronic vector, were able to drive leaky expression of the 3′ CDS. Moreover, bi-cistronic Cre-inducible vectors incorporating either a 2A or IRES sequence were prone to leaky expression (Figure 3), as were vectors containing FLEx/DIO cassettes where the CDSs were positioned in reverse orientation downstream of the promoter (Figure S1). Thus, it is likely that all bi-/poly-cistronic gene expression vectors with an inducible gene expression system have a risk of leaky 3′ CDS expression.

An additional implication of our study is that Cre-independent leak expression of 3′ transgenes is not restricted to transient transfection of cultured cells using plasmid DNA. We also observed Cre-independent leak expression of 3′ transgenes following *in vivo* viral transduction of the mouse retina or brain using AAVs incorporating FLEx cassettes encoding a bi-cistronic CDS or fusion protein. On the one hand, ddPCR analysis revealed that leakage of the 3′ CDS was significantly higher compared to that of 5′ CDS in retinae injected with AAV-*EF1a-FLEX-GFP-2A-TVA* or AAV-*βactin-FLEX-DTR-GFP*, whereas there was no significant difference in the relative expression of *ArchT* and *GFP* in retinae injected with *EF1a-FLEX-ArchT-GFP* (Figures 7C–7E). On the other hand, Cre-independent leak GFP expression was greatest in the brains of mice injected with *EF1a-FLEX-ArchT-GFP* (Figure 7H). These observations suggest that, as in conventional promoters, the promoter-like activity of the 5′ CDSs may vary depending on the cell types and organs. We also noted a large variation in the level of expression of the 3′ transgene in tissue injected with both AAV-*Cre* and the AAV-*FLEX* viruses (Figures 7C–7E). The greater variability in 3′ transgene expression level compared to tissue infected with the AAV-*FLEX* virus alone can be attributed to varying degrees of viral load of target cells with the AAV-*Cre* and AAV-*FLEX* viruses. In the presence of Cre recombinase, high-level expression of both 5′ and 3′ transgenes reflects the activity of the *bone fide* promoter which becomes oriented immediately upstream of the CDS following the Cre-mediated FLEx recombination event. Even when both viruses are injected into the tissue, some cells are likely to be transduced with AAV-*FLEX* alone and thus the expression level of the 3′ transgene would reflect levels observed in the Cre-negative state. Accordingly, the decreased variability in the level of expression of the 3′ transgene that we observed when tissue is transduced with the AAV-*FLEX* virus alone can be explained by the fact that, in the Cre-negative condition, leak expression of the 3′ transgene reflects activity of the 5′ CDS rather than the *bone fide* promoter which, in the absence of Cre recombinase remains in the opposite orientation to the transgenes.

Results of segment transfection experiments with *rtTA*, *P2A-rtTA*, and *mC-P2A-rtTA* segments clearly showed that the 5′ region CDS was responsible for leaky rtTA expression. It is still not clear how the 5′ region CDS drove the expression of 3′ region CDS. It is widely known that RNA polymerase II should bind to the promoter region to initiate CDS transcription.[30] Our data demonstrate that 5′ transgenes can drive 3′ transgene leakage, suggesting that 5′ transgenes possess cryptic promoter activity. The molecular mechanisms underlying cryptic promoter activity in this context remain to be explored. A

Figure 6. Insertion of a *lox-STOP-lox* cassette between the P2A peptide and 3′ transgene reduced the leaky expression of the 3′ transgene

(A) Schematic diagram of *pMBP-LGL-mC-2A-LSL-rtTA* expression vector before and after Cre-mediated recombination. (B) CG4 cells transfected with the *pMBP-LGL-mC-2A-LSL-rtTA* vector express nGFP in the Cre (−) condition or mCherry in Cre (+) condition. Scale bar, 50 μm. (C) Flow cytometric analysis of mCherry and GFP expression among CG4 cells transfected with either *pMBP-LGL-mC-2A-rtTA*, *pMBP-LGL-rtTA-2A-mC* or *pMBP-LGL-mC-LSL-2A-rtTA* in the Cre (+) or Cre (−) condition. (D) Relative luciferase activity of *pMBP-LGL-rtTA-2AmC-*, *pMBP-LGL-mC-2A-rtTA-* or *pMBP-LGL-mC-2A-LSL-rtTA*-transfected cells in Cre (−) condition, as compared to control CG4 cells transfected with *pTRE-Luciferase* alone (n = 15, 18, and 18, respectively). (E) Relative luciferase activity of CG4 cells co-transfected with *pCAG-Cre* and either *pMBP-LGL-rtTA-2AmC*, *pMBP-LGL-mC-2A-rtTA* or *pMBP-LGL-mC-2A-LSL-rtTA*, as compared to control CG4 cells transfected with *pTRE-Luciferase* alone (n = 9, 12, and 18, respectively). $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$, Kruskal-Wallis test with Dunn's correction (D and E). Data represent mean ± SEM.



(legend on next page)

potential mechanism that could be examined is whether DNA sequences located within the 5′ transgenes of concatenated ORFs can function as transcription factor (TF)-binding sites. TF-binding activity within the cDNA sequence of the 5′ transgene could help to recruit the pre-initiation complex to initiate transcription at ATG start codons that are in-frame with the CDS of 3′ transgenes. As demonstrated by the Callaway group,[14] removing the ATG start codon in the 5′ CDS may decrease the leakage; however, it is not feasible to remove all ATGs within a CDS since methionine is only encoded by the ATG codon and removal of all ATGs in an attempt to prevent any downstream expression, particularly in-frame ATGs, may cause loss of protein function.

We successfully decreased leaky gene expression by integrating the bi-cistronic construct into the genome or by inserting an additional *lox*-STOP-*lox* cassette between the P2A sequence and 3′ CDS region (Figure 5). In *pMBP-LGL-mC-P2A-LSL-rtTA*-transfected cells, mCherry and rtTA expression was tightly regulated by Cre-mediated recombination. However, we found that the rtTA expression level following Cre recombination was lower than that of normal constructs. The lower-level expression of rtTA in *pCAG-Cre* + *pMBP-LGL-mC-P2A-LSL-rtTA*-transfected cells could reflect several possibilities. First, it could result from reduced recombination efficiency of the construct which requires two independent recombination events at both *lox2272-nG-STOP-lox2272* and *loxP-STOP-loxP* sites. It is conceivable that, in cells with low-level expression of Cre recombinase, only one of the *lox-STOP-lox* cassettes undergoes recombination, leaving the *lox-STOP-lox* immediately upstream of the 3′ transgene intact, thereby blocking expression. Second, the presence of additional amino acids at the N-terminal domain of the rtTA protein encoded by the single *loxP* sequence that is preserved following Cre-mediated recombination may alter rtTA activity. Although the inclusion of an additional *lox*-STOP-*lox* cassette preceding the 3′ transgene reduced transgene expression in recombined cells, the construct design is preferable for situations where transgene expression is strictly Cre-dependent and lower levels of transgene expression following recombination are acceptable. For instance, it would be particularly important for Cre-dependent regulation of TVA or DTR expression, where Cre-independent leak expression would represent a significant experimental confound.

An alternative approach for regulating the expression of two transgenes under the control of a specific promoter is to use a construct that contains two copies of the same promoter in reverse orientation, enabling bidirectional regulation of transgene expression. Bidirectional promoter constructs such as MXS_BidirectionalCAG are available from Addgene (#78175).[31] Although we have not tested the MXS_BidirectionalCAG bidirectional vector in our own laboratory, it may be a better alternative to use of a bi-cistronic vector with an inducible gene expression system.

We showed that Cre-independent leak expression of the 3′ transgene is significantly reduced by genomic insertion, as demonstrated when generating stably transfected HEK293T cells (Figure 5). Based on these results, one might predict that lentivirus-mediated integration of Cre-dependent bi-cistronic transgenes would also reduce Cre-independent leak expression of the 3′ transgene. Although we have not performed experiments to investigate this possibility, several published reports characterizing transgene expression in genetically modified mice suggest that genomic insertion alone does not insulate against leak expression. For instance, it has been reported that the knock-in *somatostatin-IRES-Cre* mouse (RRID:IMSR_JAX:013044) expresses leaky Cre recombinase in some non-somatostatin-expressing neurons.[32] Even low-level leaky expression of Cre recombinase could be sufficient to drive recombination of a Cre-dependent reporter line such as the Ai14 reporter.[32] The *somatostatin-IRES-Cre* line was generated by inserting an IRES, Cre recombinase CDS, polyA sequence, and an frt-flanked *neo* cassette into the 3′ untranslated region of the somatostatin locus, after the translational termination site.[33] It is plausible, and warrants evaluation, as to whether the leaky Cre expression observed in the *somatostatin-IRES-Cre* line could be driven by promoter-like activity of the somatostatin CDS.

Our data have implications for understanding the observations of leak expression documented by other research groups. For example, Botterill and colleagues (2021)[29] observed that stereotaxic delivery of AAV-*DIO-hSyn-hM3Dq-mCherry* into the mouse cerebral cortex resulted in leak expression of mCherry in the absence of Cre recombinase, which was particularly evident after antibody-mediated amplification. Since mCherry-positive cells did not exhibit c-fos expression following administration of the hM3Dq agonist, C21, the authors concluded that leak expression does not yield notable functional

Figure 7. Bi-cistronic AAV-*FLEX* vectors express 3′ transgene leakage *in vivo*

(A) Schematic diagram of AAV-*EF1a-FLEX-GFP-2A-TVA*, AAV-*βactin-FLEX-DTR-GFP*, and AAV-*EF1a-FLEX-ArchT-GFP* vectors. (B) Experimental design of AAV injection of the eye followed by ddPCR analysis of transgene expression in the virally transduced retina. (C) Relative expression of GFP and TVA in the retina following co-injection of AAV-*EF1a-FLEX-GFP-2A-TVA* and either AAV-*CMV-mCherry* (left) or AAV-*CAG-Cre-mCherry* (right). (D) Relative expression of DTR and GFP in the retina following co-injection of AAV-*βactin-FLEX-DTR-GFP* and either AAV-*CMV-mCherry* (left) or AAV-*CAG-Cre-mCherry* (right). (E) Relative expression of ArchT and GFP in the retina following co-injection of AAV-*EF1a-FLEX-ArchT-GFP* and either AAV-*CMV-mCherry* (left) or AAV-*CAG-Cre-mCherry* (right). (F) Schematic diagram of AAV injection into the mouse brain and downstream analysis. (G) Representative images of cerebral cortex of mouse brains injected with AAV-*EF1a-FLEX-GFP-2A-TVA*, AAV-*βactin-FLEX-DTR-GFP*, and AAV-*EF1a-FLEX-ArchT-GFP* and co-injected with AAV-*CMV-mCherry* or AAV-*CAG-Cre-mCherry*. Arrows in Cre (−) condition indicate cells co-expressing mCherry and leaky GFP. Scale bar, 100 μm. (H) Comparison of the percentage of mCherry$^+$ cells that exhibit leaky GFP expression for cortical hemispheres co-injected with AAV-*CMV-mCherry* and either AAV-*EF1a-FLEX-GFP-2A-TVA*, AAV-*βactin-FLEX-DTR-GFP* or AAV-*EF1a-FLEX-ArchT-GFP* (n = 5 or 6 hemispheres per vector). (I) Comparison of the percentage of mCherry$^+$ cells that co-express GFP in cortical hemispheres injected with AAV-*CAG-Cre-mCherry* and either AAV-*EF1a-FLEX-GFP-2A-TVA*, AAV-*βactin-FLEX-DTR-GFP* or AAV-*EF1a-FLEX-ArchT-GFP* (n = 4 hemispheres per vector). $^*p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$, or actual p values are indicated, paired two-tailed Student's t-test (B–D), or one-way ANOVA with Dunnett's test (H and I). Data represent mean ± SEM. Illustrations were created with BioRender.com. ns, not significant.

effects in their experimental system. They also excluded behavioral repercussions of potential leak hM3Dq expression. Our data provide a plausible explanation for their observations and further illustrate that leak expression of transgenes can result in functional effects, depending on the position of the transgenes in Cre-dependent constructs containing concatenated open reading frames. For the AAV-*DIO-hSyn-hM3Dq-mCherry* construct employed by Botterill and colleagues, the hM3Dq coding sequence is in the 5′ position of the hM3Dq-mCherry fusion protein. As a result, we would expect hM3Dq to exhibit minimal leak expression compared to mCherry, which occupies the 3′ position and retains an intact ATG start codon and Kozak sequence. Based on our observations of Cre-independent positional effects of CDSs within fusion proteins, if the position of mCherry and hM3Dq in the AAV construct were reversed, it is possible that leak expression of hM3Dq may be more evident and potentially represent an experimental confound. Similarly, we are aware of at least one other mouse line that exhibits Cre-independent expression of the 3′ transgene positioned downstream of a *lox*-STOP-*lox* cassette.

We contacted the inventor of the *RphiGT* mouse line (RRID:IMSR_JAX:024708) in which *pCAG-lox-STOP-lox-RABVgp4-IRES-TVA* was introduced into the *Gt(ROSA)26Sor* locus to enable Cre-mediated expression of recombinant rabies glycoprotein G and TVA (chicken) for studies employing monosynaptic neuron tracing.[34] The author kindly informed us that there were signs of leaky TVA expression (3′ of the IRES), whereas RABVgp4 (positioned 5′ of the IRES) was correctly regulated by Cre recombination (J. Takatoh, personal communication). Despite these limitations, the mouse line can still be useful to target the peripheral nervous system.[35]

This study has revealed some limitations to the use of bi- and polycistronic expression vectors as part of an inducible gene expression strategy. To avert undesirable leaky transgene expression in inducible expression systems, we recommend using two separate vectors with an identical promoter to induce the expression of two transgenes or using a vector containing a bi-directional promoter. For studies in which the generation of a bi- or poly-cistronic vector is unavoidable, it is preferable to place genes for which leaky expression would be particularly deleterious in the 5′ position of the expression cassette. Leaky expression of 3′ transgenes can be reduced by inserting additional sets of *lox-STOP-lox* sequences, each possessing distinct Cre recognition sites, immediately 5′ of each CDS within a poly-cistronic expression cassette. Adopting these approaches will aid in the generation of inducible expression constructs that offer more precise control over the expression of multiple transgenes for basic studies and gene therapy approaches.

MATERIALS AND METHODS

Molecular cloning

Plasmid vectors were constructed using restriction cloning, PCR cloning, and Gibson cloning.[36] Some DNA segments were synthesized using gBlock gene fragments (Integrated DNA Technologies). DNA segments obtained by restriction digestion or PCR amplification were purified using the QIAquick spin column (28115; QIAGEN). The IRES sequence of the encephalomyocarditis virus (EMCV) was used in the construction of the vector *pMBP-LGL-DTR-IRES-tdT* and was obtained from the plasmid MSCV-IRES-Tomato MEF2C (Addgene #89715).[37] The full sequence of the IRES is provided in the supplemental information. All plasmid constructs were verified by Sanger sequencing at Micromon (Monash University).

Generation of plasmid constructs

pCAG-LGL-rtTA-2A-mC *(GenBank: OR372645)*

Isothermal assembly technique was used for making this construct. The *pCAG-Cre* plasmid (Addgene #13775)[38] was digested with *EcoRI* and *NotI* and the *pCAG* backbone was purified by gel extraction. The *loxP-nG-STOP-loxP-rtTA* fragment was obtained by PCR amplification using *pMBP-loxP-nG-STOP-loxP-rtTA-P2A-TFP* plasmid as template (see supplemental information). *TPC1* forward primer (5′-GCTGTCTCATCATTTTGGCAAAGCGAATAATTCTAGGGTCGAC-3′) and *TPC1* reverse primer (5′-TCCGCTTCCTCCTGGTAACATGT-3′), *P2A* fragment was amplified by PCR using synthetic *P2A-mtdTomato* plasmid (GeneArt, Invitrogen) as a template and *TPC2* forward primer (5′-ACATGTTACCAGGAGGAAGCGGAGGAAGCGGAGCTACTAACTTCAG-3′) and *TPC2* reverse (5′-GGTCTTGGAGAAACAGCAACCCAT-3′) primer, *mC* fragment was PCR amplified using *pMBP-loxP-nG-STOP-loxP-mC-P2A-DTR* plasmid as a template (see supplemental information) and *TPC3* forward primer (5′-ATGGGTTGCTGTTTCTCCAAGACC-3′) and *TPC3* reverse primer (5′-GGCAGCCTGCACCTGAGGAGTGCTTATCCGCTTCCCATATGCTTGTA-3′). This *pCAG* fragment, *loxP-nG-STOP-loxP-rtTA* fragment, *P2A* fragment, and *mC* fragment were purified by gel extraction and assembled in an isothermal Gibson assembly reaction (Gibson Assembly, New England Biolabs), followed by transformation into NEB 5-alpha competent *E. coli* (New England Biolabs). The DNA sequence of the *pCAG-LGL-rtTA-2A-mC* plasmid was confirmed by Sanger sequencing.

pCAG-LGL-mC-2A-rtTA *(GenBank: OR372646)*

This construct was generated by an isothermal Gibson assembly reaction of *pCAG* fragment, *loxP-nG-STOP-loxP-mC-P2A* fragment and *rtTA* fragment. The *pCAG* fragment was obtained by *EcoRI-NotI* digestion of *pCAG-Cre* as described above. The *loxP-nG-STOP-loxP-mC-P2A* fragment was generated through PCR amplification using *pMBP-loxP-nG-STOP-loxP-mC-P2A-FLAG-H2A-P2A-rtTA-WPRE* plasmid (see supplemental information) as a template and *CPT1* forward primer (5′-GCTGTCTCATCATTTTGGCAAAGCGAATAATTCTAGGGTCGAC-3′) and *CPT1* reverse primer (5′-GTCTTTGTAGTCGGTACCAGG-3′), *rtTA* segment was generated using *CMV-Tet3G* (Clontech) as a template and *CPT2* forward primer (5′-CCTGGTACCGACTACAAAGACTCTAGACTGGACAAGAGCAAAG-3′) and *CPT2* reverse primer (5′-ggcagcctgcacctgaggagtgcatgtctggatccttacttag-3′). These *pCAG* fragment, *loxP-nG-STOP-loxP-mC-P2A* fragment and *rtTA* segment were gel extracted and assembled as described above.

pMBP-LGL-mC-P2A-DTR *(GenBank: OR372642)*

EcoRV-Membrane mCherry (hereafter denoted mC) *-P2A-KpnI-DTR-KpnI-XbaI* fragment was synthesized (gBlocks Gene Fragment,

IDT) and PCR amplified. The fragment was digested by *EcoRV* and *XbaI* and inserted into *EcoRV-XbaI* site of *pMBP-loxP-nG-STOP-loxP plasmid*.

pMBP-LGL-mC-P2A-FLAG/H2A-P2A-rtTA *(GenBank: OR372643)*

The *pMBP-LGL-mC-2A-DTR* construct was digested with *KpnI* and the *pMBP-LGL-mC-P2A* fragment was purified by gel extraction. A *KpnI-FLAG-H2A-P2A-rtTA-WPRE-KpnI* fragment was synthesized (gBlocks Gene Fragment, IDT), PCR amplified, and subsequently digested with *KpnI*. The *pMBP-LGL-mC-P2A* fragment and *KpnI-FLAG-H2A-P2A-rtTA-WPRE-KpnI* fragment were ligated at the *KpnI* site.

pMBP-LGL-rtTA-2A-mC *(GenBank: OR372647)*

The *pCAG-LGL-rtTA-P2A-mC* plasmid was digested by *ScaI-EcoRI* and *nG-STOP-loxP-rtTA-P2A-mC* segment was collected by gel extraction. *pMBP-LGL-mC-P2A-DTR* was digested by *ScaI-EcoRI* followed by phenol-chloroform extraction and ethanol precipitation. The DNA solution containing *ScaI-EcoRI* digested fragments (2,651 bp, 2,528 bp, and 4,189 bp) was further digested by *SphI* to make gel extraction easier (2,651 bp, 1751 bp, 777 bp, and 4,189 bp), and the *pMBP-loxP* fragment (2,651 bp) was collected using gel extraction. The *pMBP-loxP* fragment and *nG-STOP-loxP-rtTA-P2A-mC* fragment were ligated with *ScaI-EcoRI* site.

pMBP-LGL-mC-2A-rtTA *(GenBank: OR372648)*

pMBP-loxP fragment (2,651 bp) was obtained as described above. *nG-STOP-loxP-mC-P2A-rtTA* fragment was collected by *ScaI-EcoRI* digestion of *pCAG-loxP-nG-STOP-loxP-mC-P2A-rtTA* and gel extraction. *pMBP-loxP* fragment and *nG-STOP-loxP-mC-P2A-rtTA* fragment were ligated with *ScaI-EcoRI* site.

pMBP-LGL-mC-2A-LGL-rtTA *(GenBank: OR372653)*

The *pMBP(M321)-lox2272-nG-STOP-lox2272-mC* fragment was PCR amplified using *pMBP(M321)-lox2272-nG-STOP-lox2272-mC-pMBP(M321)-SnaBI* plasmid (see supplemental information) and forward primer (5′- Atccagctcgaccaagcttg TCCACAGAATCAGGGGATA-3′) and reverse primer (5′-tcggcgcgttcgtactgttc-3′). P2A fragment was PCR amplified using *pCAG-LGL-mC-2A-rtTA* as a template and forward primer (5′-gaacagtacgaacgcgccga-3′) and reverse primer (5′-ATTTCCtttgtagtcggtaccaggtc-3′). DNA fragment containing *loxP-3xpolyA-loxP* was PCR amplified using *Ai65(RCFL-tdT)* targeting vector[39] (Addgene# 61577) as a template and forward primer (5′-CTAGGGAAGAAGAGAGACCCAGGaaatataacttcgtataat-3′) and reverse primer (5′-GTCCTTGTATTTCCGAAGACAgcaggtcgagggacctaata-3′). This *loxP-3xpolyA-loxP* sequence was PCR amplified using forward primer (5′-gtaccgactacaaaGGAAATataacttcgtat-3′) and reverse primer (5′-tccagtctagaAATGACgggacctaataacttcgtat-3′) to add DNA sequence compatible to the other segments, and named *loxP-STOP-loxP*. *rtTA-SV40polyA* fragment was PCR amplified using *pCMV-Tet3G* as a template and forward primer (5′-CCCGTCATTtctagactggacaagagcaa-3′) and reverse primer (5′-caagcttggtcgagctggat-3′). These *pMBP(M321)-lox2272-nG-STOP-lox2272-mC* fragment, *P2A* fragment, *loxP-STOP-loxP* fragment and *rtTA-SV40polyA* fragment were assembled using isothermal Gibson assembly reaction to generate *pMBP(M321)-lox2272-nG-STOP-lox2272-mC-loxP-STOP-loxP-rtTA* plasmid.

PB-CAG-LGL-mC-P2A-rtTA *(GenBank: OR372649) and* PB-CAG-LGL-rtTA-P2A-mC *(GenBank: OR392287)*

A DNA fragment containing the *SpeI* site, piggyBAC terminal repeat sequences and *BglII* site was PCR amplified using *PBCAG-mRFP* (Addene #40996)[40] as template with forward primer (5′-agctggacatcacc tcccacaacg-3′) and reverse primer (5′-tagcactagtctcgatatacagatcgata a-3′). The fragment was digested by *SpeI* and *BglII*. *pCAG-LGL-rtTA-2A-mC* construct or *pCAG-LGL-mC-2A-rtTA* construct were digested with *SpeI* and *BglII* to generate *SpeI-CAG-LGL-rtTA-2A-mC-BglII* segment or *SpeI-CAG-LGL-mC-2A-rtTA-BglII* segments, respectively. The segments and the *SpeI*-piggyBAC terminal repeat sequences-*BglII* fragment were ligated to generate the *PB-CAG-LGL-mC-P2A-rtTA* and *PB-CAG-LGL-rtTA-P2A-mC* constructs.

Gel extraction

PCR solution or enzymatically digested plasmid solution was purified via agarose gel extraction. Agarose gel (1%) was prepared in 1 × TAE and SYBR safe solution (Thermo Fisher Scientific). DNA solution was added to the 6× loading buffer and a total of 120–180 μL of solution was loaded into a combined well, where three wells of a gel maker were combined to one well (AD216, Clonetech). Gels were run at 130 V for 45 min in an electrophoresis apparatus (Horizon11.14, Thermo Fisher Scientific). Bands of correct size were excised and purified using QIAquick gel extraction kit (Qiagen). The DNA segments used for cell transfection experiments were further purified by phenol/chloroform extraction and ethanol purification.

Cell culture

HEK293T cells were cultured in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (FBS; Corning) and 1% penicillin-streptomycin-glutamine (Gibco). CG4 cells[23] were cultured in SATO medium supplemented with N-acetyl-cysteine (60 μg/mL; Sigma), biotin (10 ng/mL; Sigma), forskolin (5 μM; Sigma), insulin (5 μg/mL; Sigma), recombinant human NT3 (5 ng/mL; Peprotech), and PDGF-AA (10 ng/mL; Peprotech). SATO medium was prepared in high-glucose DMEM (Gibco, Cat. 31966-021) using 100× SATO stock solution, which was prepared as described in Cold Spring Harbor Protocols (https://doi.org/10.1101/pdb.rec077073).

Transfection

Transfections were performed one day after seeding cells into 96-well plates (1.0×10^4 HEK293T cells/well; 0.9×10^4 CG4 cells/well) or into 24-well plates (1.0×10^5 HEK293T cells/well; 5.0×10^4 CG4 cells/well) using Lipofectamine LTX reagent with PLUS reagent (Thermo Fisher Scientific) in medium without penicillin-streptomycin according to the manufacturer's protocols.

For HEK293T cells, plasmid DNA (0.5 μg per construct) and PLUS reagent (0.5 μL per construct) were diluted in 50 μL Opti-MEM

(Invitrogen). This solution was mixed with Opti-MEM containing Lipofectamine LTX (1 μL Lipofectamine LTX in 50 μL Opti-MEM, incubated for 5 min). After 20 min of incubation at room temperature, the transfection solution was added to each well (100 μL/well for 24-well plate; 20 μL/well for 96-well plate). The medium was replaced with 10% FBS-DMEM medium containing penicillin-streptomycin 4 h after transfection.

For CG4 cells, the cells were incubated in SATO medium containing PDGF-AA and NT3 during transfection. Constructs (each 0.5 μg) and PLUS reagent (0.5 μL for one construct) were diluted in 25 μL Opti-MEM (Invitrogen). The Opti-MEM containing constructs and PLUS reagents were mixed with the Opti-MEM/Lipofectamine LTX (2 μL Lipofectamine LTX in 25 μL Opti-MEM pre-incubated for 5 min). After 20 min of incubation at room temperature, the mixed solution was added to each well (50 μL/well for 24-well plates; 10 μL/well for 96-well plates). The medium was replaced with SATO medium containing T3 (40 ng/mL, Sigma) and penicillin-streptomycin 4 h after transfection.

Luciferase assays

The pTRE-Luciferase (Clontech) and constructs containing rtTA transgene were co-transfected into HEK 293T or CG4 cells using Lipofectamine LTX. Four hours after lipofectamine transfections, the medium was replaced with medium containing doxycycline (1 μg/mL) and penicillin-streptomycin. T3 (40 ng/mL) was also added to the medium to induce differentiation of CG4 cells into oligodendrocytes. Two days post transfection, luciferase assays were performed using a luciferase assay kit (Pierce) according to manufacturer's protocol. The average chemiluminescence intensity for wells transfected with pTRE-Luciferase alone was defined as having a relative luciferase activity of 1.

Flow cytometry

Flow cytometry was performed on a BD Fortessa X-20 cell analyzer with BD FACSDiva software (BD Biosciences). The voltage, compensation, scale and gating were determined by first adjusting them for negative and single fluorophore-only controls. Fluorophore Minus One controls (DAPI + mCherry, DAPI + GFP, and GFP + mCherry samples) were also used for initial adjustment of compensation. For each sample, 10,000 events were acquired. Data were first gated to include only intact cells (FSC-A vs. SSC-A) and then single cells (FSC-A vs. FSC-H), and DAPI-negative cells were gated as live cells and dead cells were removed from the analysis (FSC-A vs. BV450-A) and fluorophore expression was observed in live cells. Once the parameters were established, the same settings and gate parameters were used to acquire and analyze all other samples. The acquired data were analyzed using FlowJo version 10 (BD Biosciences), and the same compensation, scale, and gating parameters were used for each experiment. For each experiment, a plot comparing the proportion of GFP-expressing cells vs. the proportion of mCherry-expressing cells was generated, and these plots are presented in the results.

Generation of stable cell lines

HEK-*CAG-LGL-rtTA-P2A-mC* or HEK-*CAG-LGL-mC-P2A-rtTA* stable cell lines were generated using the PiggyBac transposon-based expression system.[41] Briefly, the *PB-CAG-LGL-rtTA-P2A-mC* or *PB-CAG-LGL-mC-P2A-rtTA* construct was co-transfected with *pCAGP-Base* (Addgene #40972) into HEK293T cells. The transfected cells were passaged every 2–4 days for 20 days. Following this, FACS sorting was performed to collect the top 10% of nGFP-expressing cells. The collected nGFP (+) cells were expanded and used as HEK-*CAG-LGL-rtTA-P2A-mC* or HEK-*CAG-LGL-mC-P2A-rtTA* stable cell lines.

Animals

Wild-type 8-week-old *C57BL/6J* female mice were purchased from Clea Japan or Japan SLC. Mice were housed in a 12-h light/dark cycle with free access to water and food. Animal experiments were conducted according to the guidelines of Jichi Medical University.

AAV preparation and injection

AAV2-*EF1a-FLEX-GFP-T2A-TVA*, AAV2-*βactin-FLEX-DTR-GFP*, and AAV2-*EF1a-FLEX-ArchT-GFP* were generated from pAAV-*EF1a-FLEX-GT* (Addgene #26198),[21] pAAV-*FLEX-DTR-GFP* (Addgene #124364), and pAAV-*EF1a-FLEX-ArchT-GFP* (Addgene #58851),[42] respectively, as previously described.[43,44] AAV2-*CMV-mCherry* and AAV1-*CAG-Cre-mCherry* were purchased from SignaGen Laboratories (Cat. SL101405 and SL101117, respectively). AAV-*EF1a-GFP* (AAV2-*EF1a-EGFP-WPRE*) was generated as previously described.[45] Either AAV2-*CMV-mCherry* or AAV1-*CAG-Cre-mCherry* and one of either AAV2-*EF1a-FLEX-GFP-T2A-TVA*, AAV2-*βactin-FLEX-DTR-GFP*, and AAV2-*EF1a-FLEX-ArchT-GFP* were mixed (1.0×10^{12} vg/mL for each AAV) for stereotaxic injection.

For intravitreal injection of AAV vectors, the mice were anesthetized with 2%–3% isoflurane, and a mixture of one of the AAV-*FLEX* vectors and either AAV2-*CMV-mCherry* or AAV1-*CAG-Cre-mCherry* (1×10^{12} vg/mL for each AAV, 1 μL) was injected into the vitreous cavity using a syringe with a 33G needle as described previously.[46,47] The mice were sacrificed by cervical dislocation and retinae were collected as described previously for RNA extraction.[48,49]

For transcranial AAV injections, mice were anesthetized with a cocktail of medetomidine (0.3 mg/kg), midazolam (4.0 mg/kg) and butorphanol (5.0 mg/kg) as described previously.[50] Mice were placed in a stereotaxic frame (Narishige) with a mouse adaptor. After drilling a hole into the skull to expose the injection site, 1 μL (1.0×10^{9} vg) AAV solution was stereotaxically injected into the somatosensory cortex (1.0 mm posterior and 1.5 mm lateral to bregma, at a depth of 0.3 mm) using NanojectII as described previously.[44] Mice were sacrificed 14 days after AAV injection.

RT-PCR and ddPCR gene expression assay

Retinae were collected from the AAV-injected mouse eyes under a stereomicroscope immediately after cervical dislocation. One to two

retinae were placed in 350 μL RLT buffer (Qiagen) and homogenized by passing it 7 times through a 26G needle fitted to an RNase-free syringe. The lysate was added to 300 μL of RLT buffer and inverted and mixed. The lysate was centrifuged at 10,000×*g* for 3 min, the supernatant was collected, placed in a gDNA Eliminator spin column and centrifuged at 10,000×*g* for 30 s. The flow-through was collected and total RNA was extracted using RNeasy Mini kit (Qiagen) according to the manufacturer's protocol. The quality of total RNA was checked by electrophoresis on a denaturing agarose gel. DNase treatment and DNA contamination assessment were not performed.

cDNA of *GFP-T2A-TVA*, *DTR-GFP*, *ArchT-GFP*, *or mCherry* was synthesized from 100 ng total RNA obtained from AAV-injected eyes using PrimeScript RT reagent kit (Takara) with the following primers (for *GFP-T2A-TVA* 5′- ACTAGCGGAGAACAAGTCTG-3′, for *DTR-GFP* and *ArchT-GFP* 5′-TCGTCCATGCCGAGAGTG ATC-3′, for *mCherry* 5′-TCGGCGCGTTCGTACTGTTC-3′). Briefly, 5× PrimeScript Buffer 2 μL, PrimeScript RT Enzyme Mix I 0.5 μL, a primer for *GFP-T2A-TVA*, *DTR-GFP*, *or ArchT-GFP* 0.5 μL, primer for *mCherry* 0.5 μL, total RNA 100 ng, and RNase-DNase-free water in an amount to make the total volume of the reaction solution 10 μL were added and mixed. Reverse transcription reactions were performed under the following conditions; 37°C for 15 min, 85°C for 5 s, and 4°C. The cDNA solution was stored at −20°C for up to 1 month until used for ddPCR.

ddPCR was conducted on a QX200 ddPCR system (Bio-Rad) using standard methods for measuring mRNA expression level of 5′ and 3′ transgenes in eyes injected with bi-cistronic one of AAV-*FLEX* vectors, AAV-*EF1a-FLEX-GFP-2A-TVA* (Addgene #26198), AAV-*βactin-FLEX-DTR-GFP* (Addgene #124364), *AAV-EF1a-FLEX-ArchT-GFP* (Addgene# 58851), and AAV-*CMV-mCherry* or AAV-*CAG-Cre-mCherry* (SignaGen Cat. SL101405 and SL101117, respectively). The mRNA expression level of mCherry was also measured and used as a standard for mRNA expression level.

For ddPCR analysis, the 2×ddPCR supermix 10 μL (Bio-Rad), 100 μM FW primer 0.2 μL, 100 μM RV primer 0.2 μL, 10.2 μM TaqMan MGB probe (Thermo Fisher Scientific) 0.5 μL, MilliQ 8.1 μL, and cDNA synthesized from 10 ng total RNA (1 μL from the cDNA solution synthesized by PrimeScript RT reagent kit) were mixed and total 20 μL PCR reaction solution was prepared for one sample. The 20 μL PCR reaction solution and 70 μL Droplet Generator oil (Bio-Rad) were mixed and droplets were generated by QX200 droplet generator. PCR reactions were performed on the generated droplets with the following parameters: STEP1 95°C 10 min, STEP2 94°C 30 s, STEP3 52°C 1 min, STEP2–STEP3 39 repetitions, STEP4 98°C 10 min, STEP5 4°C. Following is a list of TaqMan MGB probe (Thermo Fisher Scientific) with primer sets (for *GFP* Probe 5′-VIC-CCCAACGAGAAGCG-MGB-3′, FW primer 5′ -AGTCCGC CCTGAGCAAAGA-3′, RV primer 5′-TCCAGCAGGACCATGTGA TC-3′, amplicon length 55 bp; for TVA probe 5′-FAM-AGTGCT CCTGTGCTGC-MGB-3′, FW primer 5′-CGCATGTGGATGCTGA TCA-3′, RV primer 5′-CGATACCACCCACAGCTACCA-3′, amplicon length 61 bp; for *DTR* probe 5′-FAM-CACTTTATCCTCCAAG CCA-MGB-3′, FW primer 5′-GGCAGATCTGGACCTTTTGAGA-3′, RV primer 5′-CTAGTCCCTTGCCTTTCTTCTTTCT-3′, amplicon length 43 bp; for ArchT probe 5′-FAM-CTGGCTACCAGCCT GA-MGB-3′, FW primer 5′-TTCAGCACCATCTGCATGATC-3′, RV primer 5′-TCTGGTCCCCTCTCCTTAGCA-3′, amplicon length 80 bp; for mCherry Probe 5′-VIC-CAAGTTGGACATCACCTC-MGB-3′, FW primer 5′-AGACCACCTACAAGGCCAAGAA-3′, RV primer 5′-CGATGGTGTAGTCCTCGTTGTG-3′, amplicon length 99 bp). All primers and probes were diluted in TE buffer. *In silico* specificity screening using BLAST confirmed that genes from mice were not amplified by these primers.

The amplification signals were read by QX200 Droplet Reader and analyzed by Quanta-Soft software (Bio-Rad). cDNA containing at least 100 copies of *mCherry* cDNA was used for gene expression analysis of AAV-*FLEX* vectors. The raw data for each sample minus non-template control were described as the data for the sample. Repeatability of the ddPCR was evaluated in two replicates. The 5′ or 3′ transgene expression relative to *mCherry* expression (%) was measured and described in Figure 7.

Tissue processing and immunohistochemistry

Mice were perfused transcardially with 4% paraformaldehyde (PFA) in 0.1 M sodium phosphate buffer (pH 7.4) under deep anesthesia using 3%–5% isoflurane or by intraperitoneal injection of an anesthetic/analgesic drug cocktail (0.3 mg/kg medetomidine, 4.0 mg/kg midazolam, and 5.0 mg/kg butorphanol). The brains were collected from each animal and postfixed with 4% PFA overnight at 4°C. Postfixed brains were cryoprotected overnight in 0.1 M sodium phosphate buffer (pH 7.4) containing 30% sucrose, embedded in optimal cutting temperature compound (Sakura Finetek), and cut into 10-μm sections using a cryostat (CM3050; Leica). The sections were blocked in 10% normal goat serum in D-PBS with 0.3% Triton X-100 (PBST) then incubated with rat anti-GFP antibody (1:1,000; GF090R; Nacalai Tesque) in PBST at 4°C overnight. After washing with PBST, the sections were incubated with secondary antibodies (1:500, Alexa Fluor 488-conjugated goat anti-rabbit IgG; Molecular Probes) for 1 h at room temperature. Slides were rinsed again and counterstained with Hoechst 33342 (1 μg/mL; Invitrogen), and coverslipped as described previously.[51]

Microscopic analysis

The number of GFP (+) and/or mCherry (+) cells in the wells of 96-well plate was counted manually using the 10× objective of an Olympus IX81 inverted fluorescence microscope (Olympus). All cells in the well were counted.

Sections from AAV-injected brains were examined under 20× objective on an FV1000 confocal microscope (Olympus). Cell counts of confocal images were performed using Fiji software. Images of Hoechst, GFP, and mCherry were changed to binary images through Fiji threshold function. The binary images of Hoechst, GFP, and mCherry were merged, and co-localization of signals was analyzed

manually. Fluorescent signals exhibiting clear cellular morphology were included in the cell quantified.

Statistical analysis

Prism 7 (GraphPad) and Excel (Microsoft) were used for statistical analysis. Cell counting experiments and luciferase assays were repeated in at least three independent experiments. The statistical significance level was set at $p < 0.05$. Data that were normally distributed (determined using the D'Agostino and Pearson test) were analyzed using the Student's *t* test or one-way ANOVA with Dunnett's test. Non-parametric data were analyzed using the Mann-Whitney *U* test. All data are described as mean ± SEM.

DATA AND CODE AVAILABILITY

The primary sequence information of plasmid DNAs described in this publication are available online. *pMBP-loxP-nG-STOP-loxP* (GenBank: OR372640), *pMBP-loxP-nG-STOP-loxP-rtTA-P2A-TFP* (GenBank: OR372641), *pMBP-loxP-nG-STOP-loxP-mC-P2A-DTR* (GenBank: OR372642), *pMBP-loxP-nG-STOP-loxP-mC-P2A-FLAG-H2A-P2A-rtTa-WPRE* (GenBank: OR372643), *pMBP-loxP-nG-STOP-lox2272-DTR-loxP-lox2272* (GenBank: OR372644), *pCAG-loxP-nG-STOP-loxP-rtTA-P2A-mC* (GenBank: OR372645), *pCAG-loxP-nG-STOP-loxP-mC-P2A-rtTA* (GenBank: OR372646), *pMBP-loxP-nG-STOP-loxP-rtTA-P2A-mC* (GenBank: OR372647), *pMBP-loxP-nG-STOP-loxP-mC-P2A-rtTA* (GenBank: OR372648), *PB-CAG-loxP-nG-loxP-mC-P2A-rtTA* (GenBank: OR372649, *PB-CAG-loxP-nG-loxP-rtTA-P2A-mC* (GenBank: OR392287), *pMBP(M321)-loxP-nG-STOP-lox2272-DTR-loxP-lox2272* (GenBank: OR372650), *pMBP(M321)-loxP-nG-STOP-lox2272-DTR-loxP-lox2272-pMBP(M321)-SnaBI* (GenBank: OR372651), *pMBP(M321)-lox2272-nG-STOP-lox2272-mC-pMBP(M321)-SnaBI* (GenBank: OR372652) and *pMBP(M321)-lox2272-nG-STOP-lox2272-mC-loxP-STOP-loxP-rtTA* (GenBank: OR372653).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2024.101288.

ACKNOWLEDGMENTS

This work was supported by The Naito Foundation Subsidy for Inter-Institute Researches (Y.O.), KAKENHI Grants from the Japan Society for the Promotion of Science (20K22691 and 21K15197, Y.O.; 21H05241, 21H04786, 20K21506, and 20KK0170, N.O.), a research grant from the National Center of Neurology and Psychiatry (No. 30-5, 3–5; N.O.), a Future Fellowship from the Australian Research Council (FT150100207, T.D.M.), generous philanthropic support provided by Metal Manufactures Ltd. (T.D.M.) and was supported (in part) by the Intramural Research Program of the NIMH (ZIA MH002985-02, T.D.M.). The Australian Regenerative Medicine Institute at Monash University is supported by grants from the State Government of Victoria and the Australian Government. We gratefully acknowledge support provided by research platforms at Monash University, including FlowCore for flow cytometry and fluorescence-activated cell sorting (FACS) and Micromon for Sanger sequencing and Monash Micro Imaging for fluorescence microscopy. We thank the following researchers for plasmids deposited at Addgene that we have used in this study: Geoff Wahl for *H2B-GFP* (Addgene #11680), Liqun Luo for *pCA-mTmG* (Addgene #26123), Connie Cepko for *pCAG-Cre* (Addgene #13775), Alex Kentsis for *MSCV-IRES-Tomato MEF2C* (Addgene #89715), Joseph Loturco for *PBCAG-mRFP* and *pCAGPBase* (Addene #40996, #40972), Michael Davidson for *mTFP1-H2A-10* (Addgene #55488), Hongkui Zeng for the *Ai65(RCFL-tdT)* targeting vector (Addgene #61577), Edward Callaway for *pAAV-EF1a-FLEX-GT* (Addgene #26198), Eiman Azim and Thomas Jessell for *pAAV-FLEX-DTR-GFP* (Addgene #124364), and Edward Boyden for *pAAV-EF1a-FLEX-ArchT-GFP* (Addgene #58851). We also thank Jun Takatoh for providing important information about *RphiGT* mouse line and Mirana Ramialison (ARMI) for advice on approaches to identify transcription factor binding motifs and Yoshihide Sehara and Fumihiro Niwa (Jichi Medical University) for advice on western blot analysis and valuable discussion.

AUTHOR CONTRIBUTIONS

Y.O. and T.D.M. conceptualized the study and performed formal analysis of the data. Y.O., N.O., and T.D.M. acquired funding. Y.O., Y.L.X., S.M., A.C., and J.P. acquired data. A.I and J.H-L. assisted with methodologies. K.K., N.O., and T.D.M. provided resources. T.D.M. and N.O. supervised the project. Y.O. wrote the original draft. Y.O., Y.L.X., and T.D.M. conducted manuscript review and editing.

DECLARATION OF INTERESTS

The authors declare no competing financial interests.

REFERENCES

1. Wang, X. (2009). Cre Transgenic Mouse Lines. In Transgenesis Techniques, 3rd edition, E.J. Cartwright, ed. (Humana Press), pp. 265–274. https://doi.org/10.1007/978-1-60327-019-9.

2. Yamamoto, M., Shook, N.A., Kanisicak, O., Yamamoto, S., Wosczyna, M.N., Camp, J.R., and Goldhamer, D.J. (2009). A multifunctional reporter mouse line for Cre- and FLP-dependent lineage analysis. Genesis *47*, 107–114. https://doi.org/10.1002/dvg.20474.

3. Lakso, M., Sauer, B., Mosinger, B., Jr., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L., and Westphal, H. (1992). Targeted oncogene activation by site-specific recombination in transgenic mice. Proc. Natl. Acad. Sci. USA *89*, 6232–6236. https://doi.org/10.1073/pnas.89.14.6232.

4. Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. *13*, 133–140. https://doi.org/10.1038/nn.2467.

5. Werdien, D., Peiler, G., and Ryffel, G.U. (2001). FLP and Cre recombinase function in Xenopus embryos. Nucleic Acids Res. *29*, E53. https://doi.org/10.1093/nar/29.11.e53.

6. Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L., and Luo, L. (2007). A global double-fluorescent Cre reporter mouse. Genesis *45*, 593–605. https://doi.org/10.1002/dvg.20335.

7. Lee, G., and Saito, I. (1998). Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. Gene *216*, 55–65. https://doi.org/10.1016/s0378-1119(98)00325-4.

8. Richier, B., and Salecker, I. (2015). Versatile genetic paintbrushes: Brainbow technologies. Wiley Interdiscip. Rev. Dev. Biol. *4*, 161–180. https://doi.org/10.1002/wdev.166.

9. Kim, H., Kim, M., Im, S.K., and Fang, S. (2018). Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. Lab. Anim. Res. *34*, 147–159. https://doi.org/10.5625/lar.2018.34.4.147.

10. Wu, Z., Yang, H., and Colosi, P. (2010). Effect of genome size on AAV vector packaging. Mol. Ther. *18*, 80–86. https://doi.org/10.1038/mt.2009.255.

11. Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E., and Hayakawa, T. (2000). IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. Mol. Ther. *1*, 376–382. https://doi.org/10.1006/mthe.2000.0050.

12. Lavin, T.K., Jin, L., Lea, N.E., and Wickersham, I.R. (2020). Monosynaptic Tracing Success Depends Critically on Helper Virus Concentrations. Front. Synaptic Neurosci. *12*, 6. https://doi.org/10.3389/fnsyn.2020.00006.

13. Kallunki, T., Barisic, M., Jaattela, M., and Liu, B. (2019). How to Choose the Right Inducible Gene Expression System for Mammalian Studies? Cells *8*, 796. https://doi.org/10.3390/cells8080796.

14. Fischer, K.B., Collins, H.K., and Callaway, E.M. (2019). Sources of off-target expression from recombinase-dependent AAV vectors and mitigation with cross-over insensitive ATG-out vectors. Proc. Natl. Acad. Sci. USA *116*, 27001–27010. https://doi.org/10.1073/pnas.1915974116.

15. Lee, P., Morley, G., Huang, Q., Fischer, A., Seiler, S., Horner, J.W., Factor, S., Vaidya, D., Jalife, J., and Fishman, G.I. (1998). Conditional lineage ablation to model human diseases. Proc. Natl. Acad. Sci. USA *95*, 11371–11376. https://doi.org/10.1073/pnas.95.19.11371.

16. Stanger, B.Z., Tanaka, A.J., and Melton, D.A. (2007). Organ size is limited by the number of embryonic progenitor cells in the pancreas but not the liver. Nature *445*, 886–891. https://doi.org/10.1038/nature05537.

17. Yamaizumi, M., Mekada, E., Uchida, T., and Okada, Y. (1978). One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell. Cell *15*, 245–250. https://doi.org/10.1016/0092-8674(78)90099-5.

18. Wang, L.L., Serrano, C., Zhong, X., Ma, S., Zou, Y., and Zhang, C.L. (2021). Revisiting astrocyte to neuron conversion with lineage tracing in vivo. Cell *184*, 5465–5481.e16. https://doi.org/10.1016/j.cell.2021.09.005.

19. Federspiel, M.J., Bates, P., Young, J.A., Varmus, H.E., and Hughes, S.H. (1994). A system for tissue-specific gene targeting: transgenic mice susceptible to subgroup A avian leukosis virus-based retroviral vectors. Proc. Natl. Acad. Sci. USA *91*, 11241–11245. https://doi.org/10.1073/pnas.91.23.11241.

20. Seidler, B., Schmidt, A., Mayr, U., Nakhai, H., Schmid, R.M., Schneider, G., and Saur, D. (2008). A Cre-loxP-based mouse model for conditional somatic gene expression and knockdown in vivo by using avian retroviral vectors. Proc. Natl. Acad. Sci. USA *105*, 10137–10142. https://doi.org/10.1073/pnas.0800487105.

21. Wall, N.R., Wickersham, I.R., Cetin, A., De La Parra, M., and Callaway, E.M. (2010). Monosynaptic circuit tracing in vivo through Cre-dependent targeting and complementation of modified rabies virus. Proc. Natl. Acad. Sci. USA *107*, 21848–21853. https://doi.org/10.1073/pnas.1011756107.

22. Farhadi, H.F., Lepage, P., Forghani, R., Friedman, H.C.H., Orfali, W., Jasmin, L., Miller, W., Hudson, T.J., and Peterson, A.C. (2003). A combinatorial network of evolutionarily conserved myelin basic protein regulatory sequences confers distinct glial-specific phenotypes. J. Neurosci. *23*, 10214–10223. https://doi.org/10.1523/jneurosci.23-32-10214.2003.

23. Louis, J.C., Magal, E., Muir, D., Manthorpe, M., and Varon, S. (1992). CG-4, a new bipotential glial cell line from rat brain, is capable of differentiating in vitro into either mature oligodendrocytes or type-2 astrocytes. J. Neurosci. Res. *31*, 193–204. https://doi.org/10.1002/jnr.490310125.

24. Madisen, L., Mao, T., Koch, H., Zhuo, J.M., Berenyi, A., Fujisawa, S., Hsu, Y.W.A., Garcia, A.J., 3rd, Gu, X., Zanella, S., et al. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nat. Neurosci. *15*, 793–802. https://doi.org/10.1038/nn.3078.

25. Sciolino, N.R., Plummer, N.W., Chen, Y.W., Alexander, G.M., Robertson, S.D., Dudek, S.M., McElligott, Z.A., and Jensen, P. (2016). Recombinase-Dependent Mouse Lines for Chemogenetic Activation of Genetically Defined Cell Types. Cell Rep. *15*, 2563–2573. https://doi.org/10.1016/j.celrep.2016.05.034.

26. Dionne, N., Dib, S., Finsen, B., Denarier, E., Kuhlmann, T., Drouin, R., Kokoeva, M., Hudson, T.J., Siminovitch, K., Friedman, H.C., and Peterson, A.C. (2016). Functional organization of an Mbp enhancer exposes striking transcriptional regulatory diversity within myelinating glia. Glia *64*, 175–194. https://doi.org/10.1002/glia.22923.

27. Dib, S., Denarier, E., Dionne, N., Beaudoin, M., Friedman, H.H., and Peterson, A.C. (2011). Regulatory modules function in a non-autonomous manner to control transcription of the mbp gene. Nucleic Acids Res. *39*, 2548–2558. https://doi.org/10.1093/nar/gkq1160.

28. Lindeberg, J., and Ebendal, T. (1999). Use of an internal ribosome entry site for bicistronic expression of Cre recombinase or rtTA transactivator. Nucleic Acids Res. *27*, 1552–1554. https://doi.org/10.1093/nar/27.6.1552.

29. Botterill, J.J., Khlaifia, A., Walters, B.J., Brimble, M.A., Scharfman, H.E., and Arruda-Carvalho, M. (2021). Off-Target Expression of Cre-Dependent Adeno-Associated Viruses in Wild-Type C57BL/6J Mice. eNeuro *8*, ENEURO.0363-21.2021. https://doi.org/10.1523/ENEURO.0363-21.2021.

30. Cramer, P. (2019). Organization and regulation of gene transcription. Nature *573*, 45–54. https://doi.org/10.1038/s41586-019-1517-4.

31. Sladitschek, H.L., and Neveu, P.A. (2016). Bidirectional Promoter Engineering for Single Cell MicroRNA Sensors in Embryonic Stem Cells. PLoS One *11*, e0155177. https://doi.org/10.1371/journal.pone.0155177.

32. Hu, H., Cavendish, J.Z., and Agmon, A. (2013). Not all that glitters is gold: off-target recombination in the somatostatin-IRES-Cre mouse line labels a subset of fast-spiking interneurons. Front. Neural Circuits *7*, 195. https://doi.org/10.3389/fncir.2013.00195.

33. Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsiani, D., Fu, Y., Lu, J., Lin, Y., et al. (2011). A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron *71*, 995–1013. https://doi.org/10.1016/j.neuron.2011.07.026.

34. Takatoh, J., Nelson, A., Zhou, X., Bolton, M.M., Ehlers, M.D., Arenkiel, B.R., Mooney, R., and Wang, F. (2013). New modules are added to vibrissal premotor circuitry with the emergence of exploratory whisking. Neuron *77*, 346–360. https://doi.org/10.1016/j.neuron.2012.11.010.

35. Zhang, Y., Zhao, S., Rodriguez, E., Takatoh, J., Han, B.X., Zhou, X., and Wang, F. (2015). Identifying local and descending inputs for primary sensory neurons. J. Clin. Invest. *125*, 3782–3794. https://doi.org/10.1172/JCI81156.

36. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345. https://doi.org/10.1038/nmeth.1318.

37. Brown, F.C., Still, E., Koche, R.P., Yim, C.Y., Takao, S., Cifani, P., Reed, C., Gunasekera, S., Ficarro, S.B., Romanienko, P., et al. (2018). MEF2C Phosphorylation Is Required for Chemotherapy Resistance in Acute Myeloid Leukemia. Cancer Discov. *8*, 478–497. https://doi.org/10.1158/2159-8290.CD-17-1271.

38. Matsuda, T., and Cepko, C.L. (2007). Controlled expression of transgenes introduced by in vivo electroporation. Proc. Natl. Acad. Sci. USA *104*, 1027–1032. https://doi.org/10.1073/pnas.0610155104.

39. Madisen, L., Garner, A.R., Shimaoka, D., Chuong, A.S., Klapoetke, N.C., Li, L., van der Bourg, A., Niino, Y., Egolf, L., Monetti, C., et al. (2015). Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron *85*, 942–958. https://doi.org/10.1016/j.neuron.2015.02.022.

40. Chen, F., and LoTurco, J. (2012). A method for stable transgenesis of radial glia lineage in rat neocortex by piggyBac mediated transposition. J. Neurosci. Methods *207*, 172–180. https://doi.org/10.1016/j.jneumeth.2012.03.016.

41. Li, Z., Michael, I.P., Zhou, D., Nagy, A., and Rini, J.M. (2013). Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. Proc. Natl. Acad. Sci. USA *110*, 5004–5009. https://doi.org/10.1073/pnas.1218620110.

42. Han, X., Chow, B.Y., Zhou, H., Klapoetke, N.C., Chuong, A., Rajimehr, R., Yang, A., Baratta, M.V., Winkle, J., Desimone, R., and Boyden, E.S. (2011). A high-light sensitivity optical neural silencer: development and application to optogenetic control of non-human primate cortex. Front. Syst. Neurosci. *5*, 18. https://doi.org/10.3389/fnsys.2011.00018.

43. Watakabe, A., Ohtsuka, M., Kinoshita, M., Takaji, M., Isa, K., Mizukami, H., Ozawa, K., Isa, T., and Yamamori, T. (2015). Comparative analyses of adeno-associated viral vector serotypes 1, 2, 5, 8 and 9 in marmoset, mouse and macaque cerebral cortex. Neurosci. Res. *93*, 144–157. https://doi.org/10.1016/j.neures.2014.09.002.

44. Osanai, Y., Shimizu, T., Mori, T., Yoshimura, Y., Hatanaka, N., Nambu, A., Kimori, Y., Koyama, S., Kobayashi, K., and Ikenaka, K. (2017). Rabies virus-mediated oligodendrocyte labeling reveals a single oligodendrocyte myelinates axons from distinct brain regions. Glia *65*, 93–105. https://doi.org/10.1002/glia.23076.

45. Inutsuka, A., Maejima, S., Mizoguchi, H., Kaneko, R., Nomura, R., Takanami, K., Sakamoto, H., and Onaka, T. (2022). Nanobody-based RFP-dependent Cre recombinase for selective anterograde tracing in RFP-expressing transgenic animals. Commun. Biol. *5*, 979. https://doi.org/10.1038/s42003-022-03944-2.

46. Cameron, E.G., Xia, X., Galvao, J., Ashouri, M., Kapiloff, M.S., and Goldberg, J.L. (2020). Optic Nerve Crush in Mice to Study Retinal Ganglion Cell Survival and Regeneration. Bio. Protoc. *10*, e3559. https://doi.org/10.21769/BioProtoc.3559.

47. Osanai, Y., Battulga, B., Yamazaki, R., Kouki, T., Yatabe, M., Mizukami, H., Kobayashi, K., Shinohara, Y., Yoshimura, Y., and Ohno, N. (2022). Dark Rearing in the Visual Critical Period Causes Structural Changes in Myelinated Axons in the Adult Mouse Visual Pathway. Neurochem. Res. *47*, 2815–2825. https://doi.org/10.1007/s11064-022-03689-8.

48. Skeie, J.M., Tsang, S.H., and Mahajan, V.B. (2011). Evisceration of mouse vitreous and retina for proteomic analyses. J. Vis. Exp. *xx*, 2795. https://doi.org/10.3791/2795.

49. Kunisawa, K., Shimizu, T., Kushima, I., Aleksic, B., Mori, D., Osanai, Y., Kobayashi, K., Taylor, A.M., Bhat, M.A., Hayashi, A., et al. (2018). Dysregulation of schizophrenia-related aquaporin 3 through disruption of paranode influences neuronal viability. J. Neurochem. *147*, 395–408. https://doi.org/10.1111/jnc.14553.

50. Yamazaki, R., Osanai, Y., Kouki, T., Shinohara, Y., Huang, J.K., and Ohno, N. (2021). Macroscopic detection of demyelinated lesions in mouse PNS with neutral red dye. Sci. Rep. *11*, 16906. https://doi.org/10.1038/s41598-021-96395-4.

51. Oluich, L.J., Stratton, J.A.S., Xing, Y.L., Ng, S.W., Cate, H.S., Sah, P., Windels, F., Kilpatrick, T.J., and Merson, T.D. (2012). Targeted ablation of oligodendrocytes induces axonal pathology independent of overt demyelination. J. Neurosci. *32*, 8317–8330. https://doi.org/10.1523/JNEUROSCI.1053-12.2012.