Citation: Molecular Therapy — Methods & Clinical Development (2014) **1,** 5; doi:10.1038/mtm.2013.5

www.nature.com/mtm

ARTICLE

Targeted CNS delivery using human MiniPromoters and demonstrated compatibility with adeno-associated viral vectors

Charles N de Leeuw[1,2], Frank M Dyka[3], Sanford L Boye[3], Stéphanie Laprise[1], Michelle Zhou[1], Alice Y Chou[1], Lisa Borretta[1], Simone C McInerny[1], Kathleen G Banks[1], Elodie Portales-Casamar[1], Magdalena I Swanson[1], Cletus A D'Souza[4], Shannon E Boye[3], Steven JM Jones[2,4,5], Robert A Holt[2,4–6], Daniel Goldowitz[1,2], William W Hauswirth[3], Wyeth W Wasserman[1,2] and Elizabeth M Simpson[1,2,6]

Critical for human gene therapy is the availability of small promoters tools to drive gene expression in a highly specific and reproducible manner. We tackled this challenge by developing human DNA MiniPromoters (MiniPs) using computational biology and phylogenetic conservation. MiniPs were tested in mouse as single-copy knock-ins at the *Hprt* locus on the X chromosome and evaluated for lacZ reporter expression in central nervous system (CNS) and non–CNS tissue. Eighteen novel MiniPs driving expression in mouse brain were identified, 2 MiniPs for driving pan-neuronal expression and 17 MiniPs for the mouse eye. Key areas of therapeutic interest were represented in this set: the cerebral cortex, embryonic hypothalamus, spinal cord, bipolar and ganglion cells of the retina, and skeletal muscle. We also demonstrated that three retinal ganglion cell MiniPs exhibit similar cell type specificity when delivered via adeno-associated virus vectors intravitreally. We conclude that our methodology and characterization has resulted in desirable expression characteristics that are intrinsic to the MiniPromoter, not dictated by copy-number effects or genomic location, and results in constructs predisposed to success in adeno-associated virus. These MiniPs are immediately applicable for preclinical studies toward gene therapy in humans and are publicly available to facilitate basic and clinical research, and human gene therapy.

Molecular Therapy — Methods & Clinical Development (2014) **1,** 5; doi:10.1038/mtm.2013.5; published online 8 January 2014

INTRODUCTION

Several research groups have focused on genome-wide expression analyses in mouse brain.[1–4] However, these projects are limited in their ability to provide information on the location and function of specific regulatory elements that drive the expression pattern. Recently, the VISTA enhancer project has generated data regarding specific putative regulatory elements.[5] Identification of such regulatory elements facilitates the design of compact promoters that could be used in downstream clinical applications.

Gene therapies for severe brain and eye disorders hold great therapeutic promise. Adeno-associated virus (AAV) is likely to be a key delivery mechanism due to its nonpathogenic, noninsertional, and low immunogenicity characteristics.[6] However, because of its small size, the DNA payload is severely limited. To accommodate such space restrictions, compact promoters will need to be developed. Many gene therapy studies have employed ubiquitous promoters to drive expression; however, this strategy can be restricted by off-target side effects. To limit such effects, the development of region-specific or cell type–specific promoters will be crucial. In addition, physiological levels of gene expression may be most appropriate for gene-based therapeutics. The use of functional endogenous promoters, which confer physiological levels of expression, may even result in higher expression than ubiquitous promoters in particular cell types.[7] Finally, the presence of viral sequences may enhance transgene inactivation, and off-target expression may increase immunogenicity, resulting in failure of long-term therapeutic results.[8]

The Pleiades Promoter Project aims to overcome these biological challenges by generating small promoters (MiniPromoters (MiniPs)) of purely human DNA content, which exhibit highly specific expression patterns, and that could be used in space-constrained viral vectors. We previously published the first set of such Pleiades (Ple) MiniPs, each of which was 4 kilobases or less in length and able to drive regional or cell type–specific expression from the mouse genome in the mouse brain.[9] With the new work presented here, we have undertaken additional promoter development, characterize more deeply some previous designs, and, most importantly, tested a subset of MiniPs in AAV. While our primary target tissue remains adult brain for gene therapy, other central nervous system (CNS) tissues use similar transcriptional programs and are important for gene therapy. Therefore, MiniPromoter characterization has been extended to the spinal cord and retina.

Eighteen novel brain MiniPs are described herein, with 2 identified as delivering near pan-neuronal expression in the adult

[1]Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, British Columbia, Canada; [2]Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada; [3]Department of Ophthalmology, College of Medicine, University of Florida, Gainesville, Florida, USA; [4]Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada; [5]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada; [6]Department of Psychiatry, University of British Columbia, Vancouver, British Columbia, Canada.
Correspondence: EM Simpson (simpson@cmmt.ubc.ca)
Received 8 October 2013; accepted 5 November 2013

mouse brain, 13 in the spinal cord, and 6 validated for use in developmental studies. We further demonstrate 17 MiniPs with expression in the eye, mostly in the retina, but including a subset directing expression to the cornea or lens. Three MiniPs are demonstrated to retain the capacity to target expression to the ganglion cell layer when delivered to the eye in AAV vectors. These novel tools will significantly improve current methodologies of gene therapy molecular medicine through increased specificity in vector constructs for key anatomical regions of therapeutic interest.

RESULTS

33 MiniPs characterized

All MiniPs contain either a promoter fragment (Prom) or a longer Prom (LongProm) segment usually spanning a known transcription start site from a mammalian endogenous gene. In general, using computational biology and phylogenetic conservation, we designed and constructed four MiniPs for each source gene. While some MiniPs only contained the minimal promoter element, most contained additional conserved putative regulatory elements, or regulatory regions (RRs), in a configuration 5′ of the minimal promoter. All of the sequences used in generating the MiniPs are composed of entirely human sequence generated by polymerase chain reaction (PCR) from the RP11 BAC library or directly synthesized using the human genome reference sequence (hg18, March 2006 reference genome).

The 33 MiniP designs studied herein cover 75,777 base pairs (bp) of the human genome (Table 1). Most importantly, we have developed 18 novel MiniPs for expression in the brain. In addition, we have undertaken and expanded characterization of 12 previously described MiniPs from our group.[9] Finally, we characterize three previously positive MiniPs using a more sensitive reporter, lacZ.

While the primary aim of promoter design was to develop tools for adult brain gene therapy, we recognized the growing importance of gene therapy for the eye and made that an area of focused characterization for MiniPs (Table 1). As such, we generated a unique resource of 17 MiniPs for expression in the mouse eye, including the neural retina, lens, cornea, and optic nerve.

Novel brain MiniP expression ranged from ubiquitous to highly regionalized

Novel MiniPs were analyzed for brain expression (Figure 1), and detailed expression analysis can be found in the Supplementary Data. Positive MiniPs can be grouped by their primary site of expression in the mouse brain. Four MiniPs (Ple27 (*CCL27* RRs), Ple28 (*CCL27* RRs), Ple122 (*ICMT* RRs), and Ple170 (*POGZ* RRs)) expressed in nearly all brain regions. Three MiniPs, Ple15 (*C8ORF46* RRs), Ple16 (*C8ORF46* RRs), and Ple23 (*CCKBR* RRs), demonstrated expression primarily limited to the cortex. Ple240 (*UGT8* RRs) expression was restricted to the white matter. Several MiniPs expressed strongest in the diencephalon (Ple12 (*AVP* RRs), Ple29 (*CCL27* RRs), Ple97 (*GPX3* RRs), Ple112 (*HCRT* RRs), Ple146 (*NTSR1* RRs), Ple155 (*PCP2* RRs), and Ple232 (*TNNT1* RRs)). Ple67 was expressed in regions outside of the telencephalon. Four MiniPs expressed primarily in the hindbrain—Ple22 (*CCKBR* RRs), Ple201 (*SLC6A5* RRs), Ple235 (*TRH* RRs), and Ple238 (*TRH* RRs). This group of MiniPs represents a novel and diverse collection of brain promoters.

The majority of novel MiniPs showed unique expression patterns. However, we found 33% (*i.e.*, Ple12, Ple22, Ple23, Ple29, Ple155, and Ple232) of the 18 novel MiniPs delivered expression in the zona incerta of the brain, suggestive of a common regulatory property or function in this brain region despite otherwise diverse expression.

Previously, we encountered examples of MiniPs negative with the enhanced green fluorescent protein (EGFP) reporter gene, even with antibody-based detection methods. However, subsequent testing with *lacZ* found them to be positive (unpublished observation and ref. 9). This data implied that the lacZ marker is more sensitive than EGFP for evaluating MiniP expression. Thus, we set out to determine whether the use of *lacZ* with previous positive designs would broaden the expression pattern. We evaluated the expression of three of previously published MiniPs driving the EGFP reporter[9]: Ple53 (*Doublecortin* (*DCX*) RRs; expression in neurons), Ple67 (*FEV* RRs; expression in serotonergic brain regions), and Ple112 (*HCRT* RRs; expression in hypothalamus), but now driving the more sensitive *lacZ* reporter (Figure 1; and *DCX* results). All three demonstrated an overlapping but expanded expression pattern when compared with their published EGFP counterparts, suggesting that the previous EGFP reporter was expressed below antibody-detection limits in some cells.

Variability in expression was observed for three of the eight novel MiniP and new reporter germline strains (Ple12, Ple232, and Ple240; $n \geq 4$ for each strain). However, only one (Ple232) demonstrated extreme variability from nearly negative in the brain to strongly positive. Thus, we conclude that the *in vivo* expression variability of MiniPs docked at *Hprt* is generally low.

Ple26 and Ple170 MiniPs drive near pan-neuronal expression

Establishing a pan-neuronal promoter for the adult brain has proven to be challenging in the field due to the highly diverse collection of neuronal subtypes represented with numerous nonoverlapping transcriptional programs.[10] Recently, Heimer-McGinn and Young[11] published a random-insertion transgenic mouse driving cre recombinase in an apparently pan-neuronal manner. However, although a great resource for lineage tracing studies, there was no promoter fragment that can be used to create other mouse strains, or be used in viral or plasmid delivery systems. We identified two MiniPs, Ple26 (*CCL27* RRs), originally described in Portales-Casamar *et al.*,[9] and Ple170 (*POGZ* RRs; a novel MiniP), as pan-neuronal candidates. In this study, we expanded our characterization and assessed the extent to which these MiniPs label neurons. These two MiniPs were assessed for neuronal costaining using NeuN and off-target glial expression using Gfap (Figure 2).

Both MiniPs were positive throughout the cortex, the hippocampus, and the cerebellum (Figure 2a,b). The olfactory bulb was more weakly stained with Ple26 as compared with the rest of the brain and with Ple170. The general overlap with NeuN was more pronounced for Ple26 than for Ple170. This was particularly true for the cerebellum where comparatively weak labeling was observed for Ple170. Using hippocampal neurons and estimating the average coverage based on co-labeling in cornus ammonis 2–3 regions, we found a significant difference of 91.34% co-labeling with NeuN for Ple26 and 82.89% for Ple170 ($n = 3$ animals for each, and 3 sections averaged for each animal; *t*-test, $P < 0.05$). This data suggests that Ple26 would be superior to Ple170 for use in driving near pan-neuronal expression.

MiniP expression observed in spinal cord and non-CNS tissue

For 21 MiniPs, we collected adult spinal cord, heart, liver, lung, and ear-notch tissue samples to be stained for β-galactosidase (β-gal) (see Supplementary Table S1). Within this analysis, we looked at

Table 1 Summary of MiniPromoters

Ple number[a]	*Gene RRs*	*Plasmid pEMS number*[b]	*Strain category*	*Brain expression*[c]	*Eye expression*
12	*AVP*	1482	Novel MiniPromoter	This study	Negative
15	*C8ORF46*	1485	Novel MiniPromoter	This study	Negative[d]
16	*C8ORF46*	1486	Novel MiniPromoter	This study	Negative[d]
17	*C8ORF46*	1487	Expanded characterization	Portales-Casamar *et al.*[9]	Negative
22	*CCKBR*	1493	Novel MiniPromoter	This study	Negative[d]
23	*CCKBR*	1494	Novel MiniPromoter	This study	Negative[d]
24	*CCKBR*	1495	Expanded characterization	Portales-Casamar *et al.*[9]	GCL, rare INL
25	*CCKBR*	1496	Expanded characterization	Portales-Casamar *et al.*[9]	GCL, rare INL
26	*CCL27*	1497	Expanded characterization	Portales-Casamar *et al.*[9]	All layers; cornea; optic nerve
27	*CCL27*	1498	Novel MiniPromoter	This study	Negative[d]
28	*CCL27*	1499	Novel MiniPromoter	This study	All layers; lens
29	*CCL27*	1500	Novel MiniPromoter	This study	Negative[d]
32	*CLDN5*	1503	Expanded characterization	Portales-Casamar *et al.*[9]	GCL, INL
34	*CLDN5*	1505	Expanded characterization	Portales-Casamar *et al.*[9]	GCL, INL, IPL, PR
53	*DCX*	1524	Different reporter	This study	GCL, rare INL; optic nerve
55	*DCX*	1526	Expanded characterization	Portales-Casamar *et al.*[9]	GCL, rare INL; optic nerve
67	*FEV*	1538	Different reporter	This study	GCL, rare INL
88	*GFAP*	1559	Expanded characterization	Portales-Casamar *et al.*[9]	Astrocytic layer; optic nerve
97	*GPX3*	1568	Novel MiniPromoter	This study	Negative[d]
112	*HCRT*	1583	Different reporter	This study	Negative[d]
122	*ICMT*	1593	Novel MiniPromoter	This study	Negative[d]
123	*ICMT*	1594	Expanded characterization	Portales-Casamar *et al.*[9]	All layers; cornea; optic nerve
131	*MKI67*	1602	Expanded characterization	Portales-Casamar *et al.*[9]	Cornea
140	*NR2E1*	1612	Expanded characterization	Portales-Casamar *et al.*[9]	Negative
146	*NTSR1*	1617	Novel MiniPromoter	This study	GCL, rare INL
155	*PCP2*	1626	Novel MiniPromoter	This study	INL, IPL
167	*POGZ*	1091	Expanded characterization	Portales-Casamar *et al.*[9]	All layers; cornea
170	*POGZ*	1641	Novel MiniPromoter	This study	All layers; cornea
201	*SLC6A5*	1673	Novel MiniPromoter	This study	Negative[d]
232	*TNNT1*	1704	Novel MiniPromoter	This study	Negative
235	*TRH*	1707	Novel MiniPromoter	This study	Negative[d]
238	*TRH*	1710	Novel MiniPromoter	This study	Negative
240	*UGT8*	1712	Novel MiniPromoter	This study	GCL, INL, PR; cornea; optic nerve

GCL, ganglion cell layer; INL, inner nuclear layer; IPL, inner plexiform layer; PR, photoreceptors; RRs, regulatory regions.
[a]Ple MiniPromoters drive lacZ, except Ple167 which drives EGFP/cre. [b]Plasmid constructs available from AddGene (http://wwww.addgene.org/). [c]All constructs were positive in the brain. [d]These constructs were analyzed in chimeras and thus might be false negatives in the eye.

all germline strains except for Ple25 (*CCKBR* RRs), which was analyzed prior to the introduction of this protocol. Thirteen MiniPs stained positive in spinal cord (Figure 3), 12 in heart, 6 in liver, 8 in lung, and 14 in ear-notch samples. Nearly half (46%) of the 13 CNS-positive MiniPs were also expressed in all four additional tissues. Interestingly, only four (31%) were restricted to spinal cord, and hence specific to the CNS.

MiniPs that were positive in the spinal cord were often expressed in the nerve fibers exiting the spinal column and/or dorsal root ganglia. Overview images for spinal cord expression are shown in Figure 3a and summarized in Figure 3b. We observed expression patterns that consisted of only the central spinal cord ($n = 1$, Ple88 (*GFAP* RRs)), spinal cord with nerve fibers ($n = 4$; Ple17 (*C8ORF46* RRs), Ple28 (*CCL27* RRs), Ple32 (*CLDN5* RRs), and Ple170 (*POGZ* RRs)), spinal cord with dorsal root ganglia ($n = 4$; Ple24 (*CCKBR* RRs), Ple26 (*CCL27* RRs), Ple53 (*DCX* RRs), and Ple55 (*DCX* RRs)), and spinal cord with nerve fibers and dorsal root ganglia positive ($n = 4$; Ple34 (*CLDN5* RRs), Ple123 (*ICMT* RRs), Ple167 (*POGZ* RRs), and Ple240 (*UGT8* RRs)).

Embryonic analyses of six MiniPs revealed hypothalamus-specific and skeletal muscle expressing constructs

Six MiniPs were analyzed at embryonic day (E) 12.5 and assessed by X-gal staining to gain insight into their developmental expression (Figure 4). For each MiniP, at least two hemizygous males and one litter-matched wild-type were stained for X-gal. We selected three MiniPs with widespread adult brain expression data: Ple26 (*CCL27* RRs), Ple123 (*ICMT* RRs), and Ple170 (*POGZ* RRs). We selected three further MiniPs of specific interest: Ple131 (*MKI67* RRs) for the role of KI-67 in proliferation,[12] Ple140 (*NR2E1* RRs) because of its highly regionalized hypothalamic expression in adults,[9] and Ple232 (*TNNT1* RRs) for the role of TNNT1 in skeletal muscle.[13]

Ple26 (*CCL27* RRs) (Figure 4a), a pan-neuronal candidate, directed strong expression throughout the embryo, indicative of ubiquitous expression at this developmental stage. Ple123 (*ICMT* RRs) and Ple131 (*MKI67* RRs) were both predominantly expressed in the heart (Figure 4b,c), with Ple131 also marking scattered cells throughout all other regions of the embryo. Our previously published highly specific adult hypothalamus promoter, Ple140 (*NR2E1* RRs) (Figure 4d), demonstrated exclusive expression to the developing hypothalamus, consistent with the observation in the adult.[9] This early developmental specificity suggests expression in a cell type that maintains its identity throughout development. Interestingly, Ple170 (*POGZ* RRs) (Figure 4e), the second pan-neuronal candidate, is expressed more strongly throughout the limbs and tail, with a blue-green haze of expression throughout the rest of the embryo, including the nervous system. The Ple232 (*TNNT1*

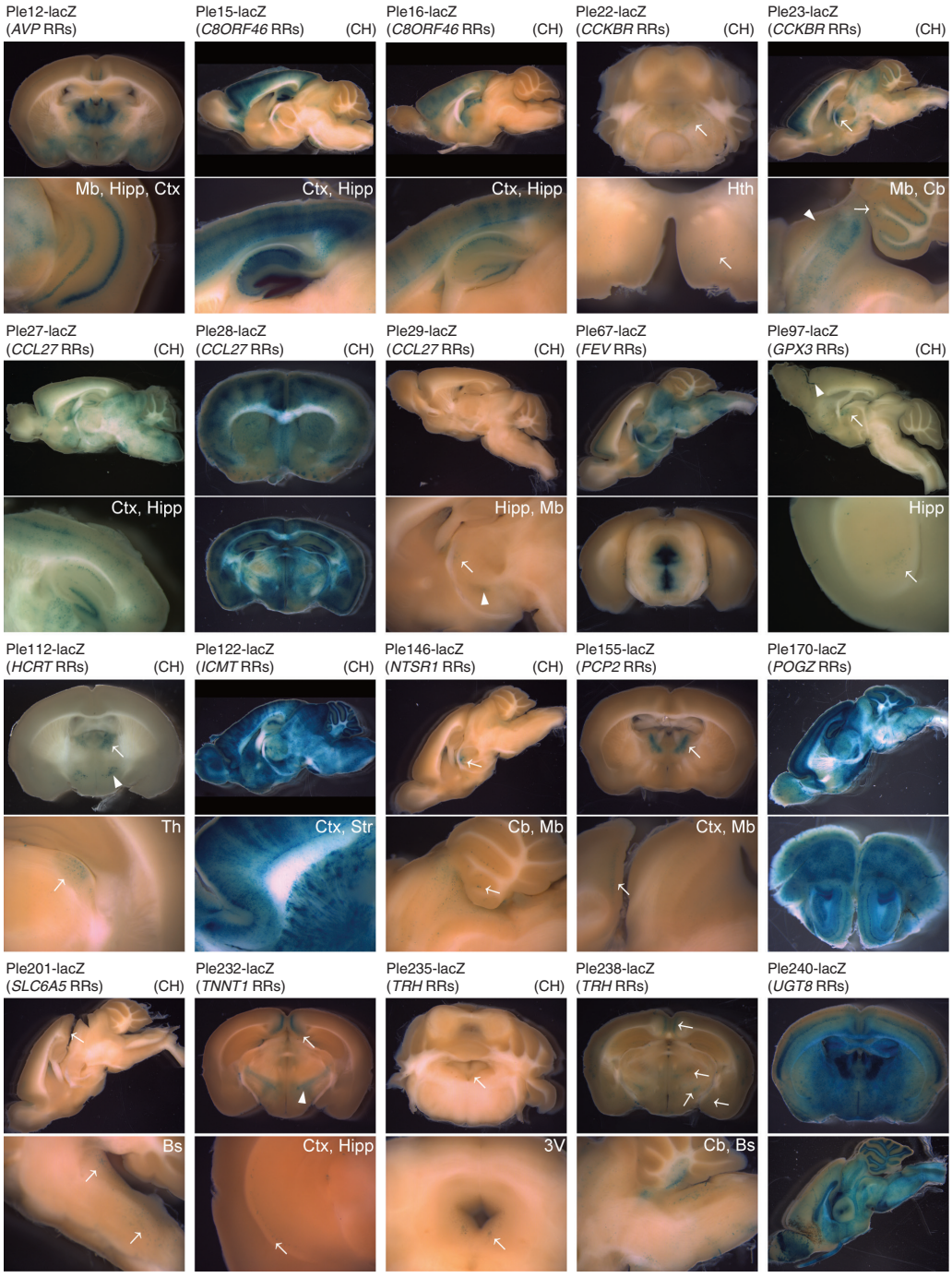


Figure 1 Brain expression patterns of 18 novel MiniPromoters and 2 additional MiniPromoter strains that were remade with the lacZ reporter. Adult mice were perfused, brains postfixed, sectioned sagittally or coronally at 1 mm, and stained overnight for lacZ using the substrate X-gal. At least two germline animals were analyzed for each established strain. Ple12 (*AVP* RRs) expressed in amygdala and thalamus (top panel), and in the ventral hippocampus and cortex layer IV (bottom panel). Ple15 (*C8ORF46* RRs) expressed throughout the cortex and throughout the hippocampal pyramidal and dentate granule neurons with strong signal in the dentate gyrus. Ple16 (*C8ORF46* RRs) showed a similar but weaker expression pattern to Ple15. Ple22 (*CCKBR* RRs) had lacZ-positive cells located in the periaqueductal gray (top panel, arrow) and in the hypothalamus (bottom panel, arrow). Ple23 displayed cortical staining throughout the brain, colliculi staining, a strong localized area at the anterior thalamus (top panel, arrow), and in the midbrain colliculi (bottom panel, arrow head), cerebellar granular and Purkinje cells (bottom panel, arrow). Ple27 and Ple28 (*CCL27* RRs) showed expression throughout most brain regions, with Ple28 being much stronger. Ple29 (*CCL27* RRs) expressed weakly in small localized areas of the thalamus and hypothalamus (bottom panel, arrow and arrowhead, respectively). Ple67 (*FEV* RRs) expressed in cortical layers II/III, cerebellar Purkinje cells, and strongly throughout most midbrain and hindbrain areas, with a definitive exclusion of striatal tissue (top panel). Raphe nuclei are clearly positive (bottom panel). Ple97 (*GPX3* RRs) stained cells localized at the anterior thalamus (top panel, arrow) and ventrolateral hippocampus (bottom panel, arrow), and also in large blood vessels (top panel, arrow head). Ple112 (*HCRT* RRs) expressed in the anteroventral thalamus (top panel, arrow) and lateral hypothalamic area (top panel, arrow head). The lateral geniculate complex was positive (bottom panel, arrow). Ple122 (*ICMT* RRs) expressed strongly throughout all major brain regions, with comparatively weaker expression found in the thalamus, and hippocampal CA2, 3, 4 and dentate gyrus (top panel). Ple146 (*NTSR1* RRs) had scattered staining in the anterior thalamus (top panel, arrow) and in Purkinje cells (bottom panel, arrow). Ple155 (*PCP2* RRs) expressed strongly in the anterodorsal, anteroventral, and

anteromedial nuclei (top panel, arrow). Additional staining was seen in the retrosplenial area (bottom panel, arrow). Ple170 (*POGZ* RRs) expressed strongly throughout all brain regions, with weaker staining in the striatal nuclei. The olfactory bulb mitral, glomerular, and granular layers are all stained (bottom panel). Ple201 (*SLC6A5* RRs) expressed weakly in the anterior and retrosplenial areas of the cortex (top panel, arrow). Staining was present in the medulla also (bottom panel, arrows). Ple232 (*TNNT1* RRs) expressed in the anterior cingulate and zona incerta (top panel, arrow and arrow head, respectively) and in the lateral hippocampus (bottom panel, arrow). Ple235 (*TRH* RRs) expressed in the pontine grey and periaqueductal grey (top and bottom panels respectively, arrows). Ple238 (*TRH* RRs) had scattered expression in many brain regions, including the cerebellar interposed and fastigial nuclei and in parts of the medulla (bottom panel). Ple240 (*UGT8* RRs) showed strong fiber-tract associated staining in all parts of the brain, with cortex layer IV strongly positive and in thalamic intralaminar and anterior nuclei groups. In addition, the olfactory bulb glomeruli and corpus callosum were clearly labeled. Bs, brainstem; CA, cornus ammonis; Cb, cerebellum; CH, chimera; Ctx, cortex; Hipp, hippocampus; Hth, hypothalamus; Mb, midbrain; RRs, regulatory regions; Str, striatum; Th, thalamus.

RRs) (Figure 4f) MiniP showed strong expression in the developing skeletal muscles of the embryo.

At E12.5, 50% of MiniPs (Ple123, Ple131, and Ple170) showed variability in expression, in contrast with the relative lack of variability in the adult. Figure 4 presents the most frequently observed staining pattern. For Ple123 (n = 7 embryos), expression levels ranged from primarily heart only (n = 5) to ubiquitous (n = 2). Ple131 (n = 5 embryos) displayed heart only (n = 1) to weak ubiquitous (n = 3) to strong ubiquitous (n = 1) expression. Ple170 (n = 3 embryos) ranged from negative (n = 2) to ubiquitous (n = 1) expression. Interestingly, in contrast to the adult brain, we did not observe variability in the embryo for Ple232 expression.

DCX-based MiniPs demonstrate expression driven by a "Prom" and a candidate regulatory element

We compared two MiniPs, both designed using the *DCX* gene, to understand differences in the RRs used (Figure 5a,b). The first design, Ple53, was based on a known promoter (LongProm) used in the literature[14] and found to be expressed as anticipated in neurogenic regions of adult brain. The second design, Ple55, featured a 5′ shortened promoter region (Prom) coupled to a highly conserved putative regulatory element from within intron 4.

We first looked at the cortical layer specificity of Ple53 and Ple55 (Figure 5c,d). Ple53 expression, as detected by X-gal and counterstained with neutral red, was limited to the deepest cortical layer VI (Figure 5c, left panel). Rare positive cells in layers II, IV, and V were observed. In contrast, Ple55 was expressed in all cortical layers, including layer I (Figure 5d, left panel). Nearly all layer II, IV, and VI neurons were labeled. Both MiniPs were expressed in the known adult neurogenic regions of the dentate subgranular zone, subventricular zone, and rostral migratory stream axis (data not shown).

To assess whether the positive cells in the cortex were primarily neurons, we compared the extent of NeuN and Gfap colabeling with X-gal in the cortex of these two strains (Figure 5c,d, center and right panels). Cells were found in both strains that colabeled (black arrows) and did not co-label (white arrows) with NeuN (Figure 5c,d). We observed some co-labeling with Gfap in the hippocampal formation, which was more extensive with Ple55 (Figure 5d, right panel). The overlapping cortical and neurogenic expression between the two designs suggests an expansion of the expression pattern with Ple55.

At E12.5 (Figure 5e,f), we noted an overall conserved staining pattern for Ple53 and Ple55, consistent with the developing nervous system. Specifically, expression was observed throughout the developing brain, eye, spinal cord, and dorsal root ganglia. Labeling extends along both the dorsal and ventral ramus. Expression in the developing kidney was found in partially cleared embryos (Figure 5e,f, white arrows). Ple55 showed additional expression in the face and limbs (black arrows). Only the nervous system expression and kidney expression have been previously documented for the mouse *Dcx* gene.[15]

In the retina (Figure 5g,h), Ple53 and Ple55 were expressed nearly exclusively in the ganglion cell layer (GCL), with rare inner nuclear layer (INL) cells positive. Both MiniPs were also expressed in the optic nerve, axonal extensions from the GCL. Studies by others in rat retina suggested horizontal cell labeling for endogenous Dcx.[16] However, only the rare INL MiniP-expressing cells (white arrows in Figure 5g,h, top middle panel) may be consistent with their localized expression in horizontal cells. Thus, we conclude that the retinal expression for Ple53 and Ple55 is ectopic and unexpected.

To further the characterization of the retinal expression, we performed co-labeling studies. Strong overlap between NeuN-positive nuclei and X-gal (Figure 5g,h, top middle panel), and with rare exceptions, a lack of overlap with the astroglial marker Gfap (Figure 5g,h, top right panel), was observed. Further analysis by immunofluorescence of β-gal using co-labeling with β-Tubulin III (Figure 5g,h, bottom left panel) and Brn3 (Figure 5g,h, bottom middle panel), both markers of retinal ganglion cells, demonstrated significant overlap, while only rare cells were positive for the cholinergic amacrine marker, Sox2 (Figure 5g,h, bottom right panel, white arrows).

To account for the expression differences observed between Ple53 and Ple55, we undertook a bioinformatics analysis (Figure 5a,b). The 5′ extension of the Prom segment found in the LongProm used in Ple53 and the additional element from intron 4 added together with the Prom segment in Ple55 were used to search for predicted transcription factor binding sites (TFBS). Ple53 was predicted to contain a MafK (a transcription factor binding site based on the *MafK* gene family) TFBS (UCSC (University of California, Santa Cruz)/ChIP (chromatin immunoprecipitation)-seq ENCODE data) and a BRACH (a transcription factor binding site based on the *brachyury* gene family) TFBS (TFBIND database), which are located in the 5′ nonoverlapping region of the LongProm and thus missing in Ple55. Furthermore, Ple55 contained two predicted serum response factor (analyzed using CONSITE with profiles from the JASPAR database) and two retinoid X receptor/retinoic acid receptor (JASPAR database) TFBSs within the conserved element 3, derived from the intron 4 region. Serum response factor is a transcription factor critical for muscle development,[17] and a TFBS for it is located in element 3 of Ple55. We propose that this serum response factor TFBS in Ple55 is responsible for the additional expression pattern in embryos and that the tissue identity is premuscle. Lastly, it appeared that there was greater co-labeling with Gfap in the hippocampal formation of adult Ple55 mouse brain, which would be consistent with a role for serum response factor in astrocyte and oligodendrogial development.[18] This unique predicted TFBS may account for the expression differences observed.

Interestingly, in contrast to all other MiniP strains, we were unable to maintain the *DCX*-based MiniP strains on the C57BL/6J background past backcross N3 due to lack of allele transmission. However, we were able to restore and maintain normal allele

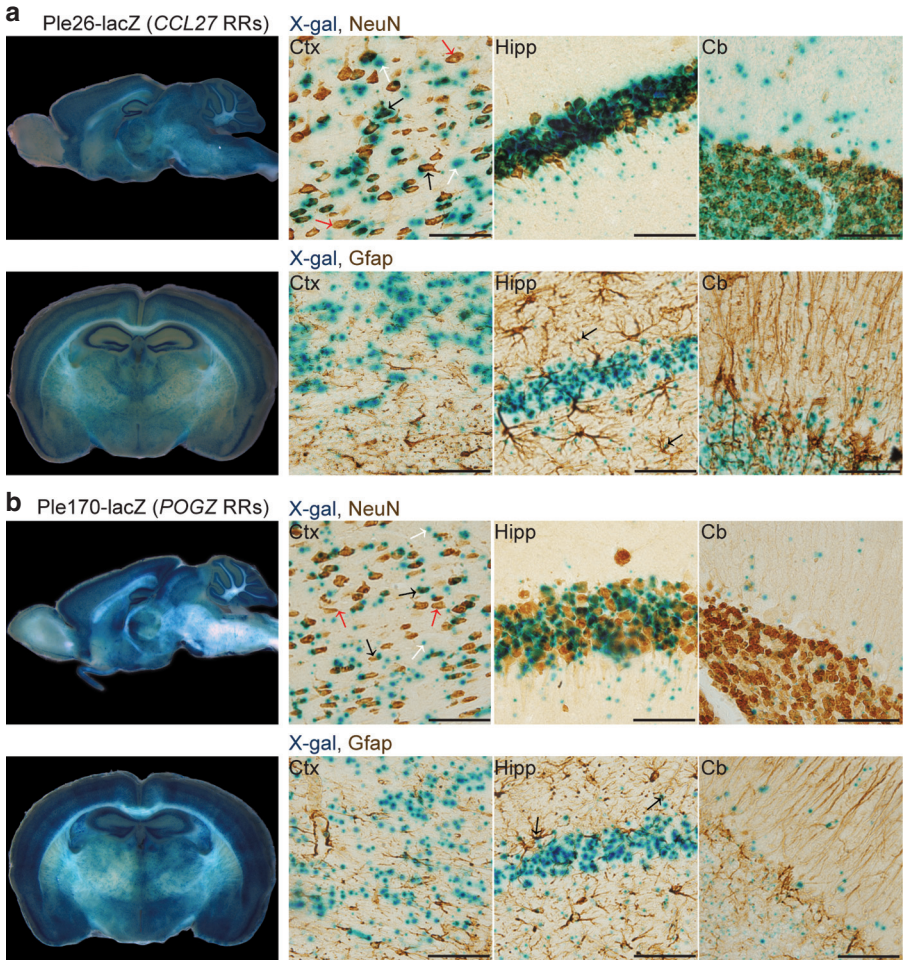


Figure 2 Two MiniPromoters, Ple26 (*CCL27* RRs) and Ple170 (*POGZ* RRs), drive near pan-neuronal expression throughout the mouse brain. For each, a representative 1-mm sagittal and coronal whole brain section is shown, as well as cryosections of NeuN (neuronal marker) and Gfap (glial marker) antibody co-labeling with X-gal in the cortex, hippocampus, and cerebellum. (**a**) Ple26 (*CCL27* RRs) showed marked expression throughout the adult brain with slightly less labeling in the olfactory bulb. In the cortex, NeuN closely co-labeled many of the X-gal positive cells (black arrows). However, there were some NeuN cells that do not co-label with X-gal (red arrows), as well as many smaller punctate X-gal nuclei that do not co-label with NeuN (white arrows). Nearly all hippocampal neurons co-label with X-gal, and extensive co-labeling was also observed in the cerebellum, with X-gal only staining also present in the molecular layer. Also, rare X-gal/Gfap-positive cells were observed (black arrows). (**b**) Ple170 (*POGZ* RRs) demonstrated a similar expression pattern to that of Ple26. The CA2 and CA3 regions of the hippocampus showed less staining than the dentate gyrus and CA1 regions. Many NeuN co-labeled cells were visible (black arrows). However, some NeuN-positive cells in the cortex were not labeled with X-gal (red arrows), as well as some punctate X-gal nuclei were NeuN-negative (white arrows). Reduced NeuN co-labeling was seen in the hippocampus and cerebellum, particularly in the cerebellar granule cell layer. Occasionally, a cell was found to be double labeled with Gfap and X-gal (black arrows). CA, cornus ammonis; Cb, cerebellum; Ctx, cortex; Hipp, hippocampus; RRs, regulatory regions. (Scale bars = 100 µm).

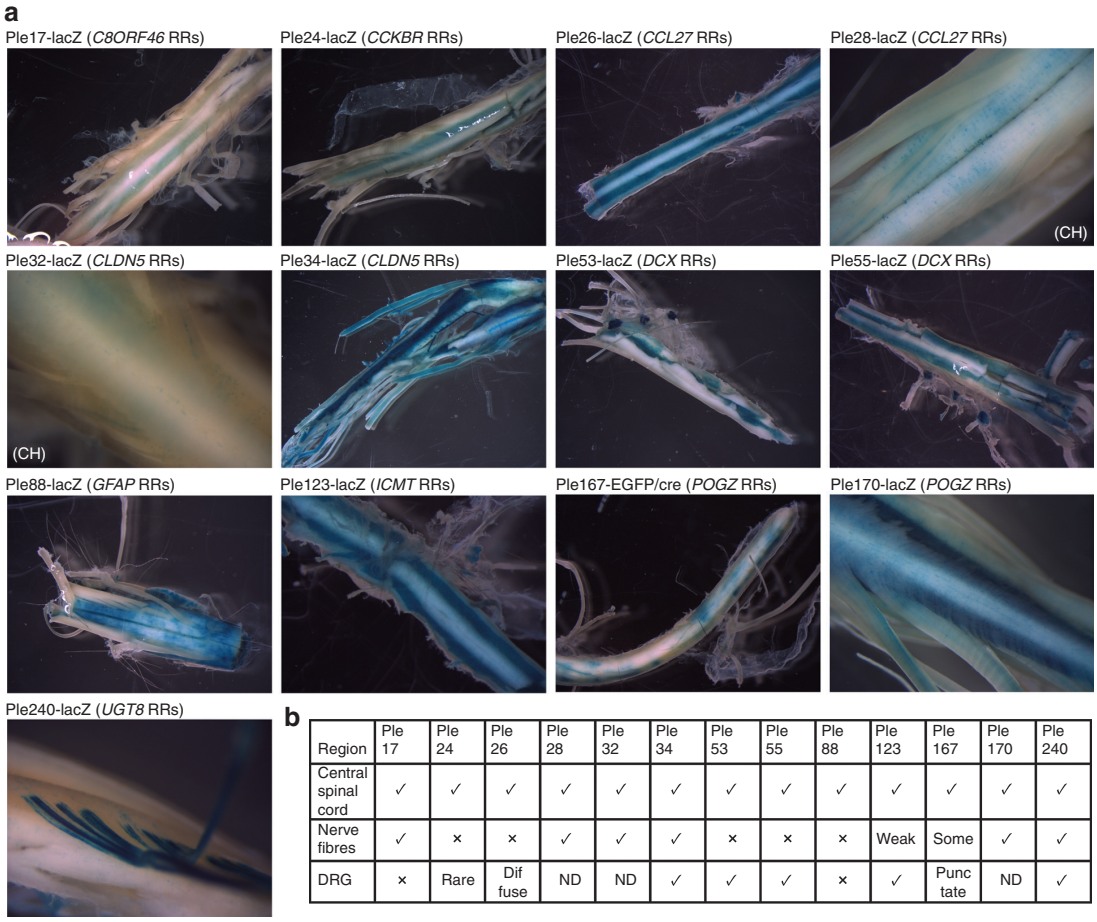
transmission after crossing the allele to the 129S1/SvImJ strain and hypothesized an unknown genotoxic effect for these alleles in C57BL/6J mice.

"Proms" typically confer the observed specificity

It has been proposed that the minimal promoter for a given gene confers the general specificity of expression but that additional regulatory elements either alter the timing or fine-tune the regional specificity.[19,20] To test this concept, we analyzed sets of MiniPs designed from the same gene in our combined dataset (this paper and ref. 9). For the purpose of this analysis, we classified MiniPs (with respect to brain only) as: (i) related, having at least some expression similar to the endogenous gene; or (ii) unrelated, having no pattern similar to the endogenous gene. Similarly, we classified sets of MiniPs as: (i) related (*i.e.,* all MiniPs are related to endogenous and similar in pattern to each other); (ii) unrelated (*i.e.,* all MiniPs do not have expression in common with the endogenous gene but are similar in pattern to each other); or (iii) mixed (*i.e.,* a set containing related and unrelated MiniPs, or alternatively, a set containing MiniPs that are all related to the endogenous expression pattern but not similar to each other). We had 16 such sets (see Supplementary Figure S1).

Nine of these sets contain MiniPs that all have an identical basal Prom (or minimal promoter) element within the set. Of these nine, six sets of MiniPs (*C8ORF46*, *CLDN5*, *HCRT*, *ICMT*, *MKI67*, and *S100β* RRs) displayed related expression, while the remaining two (*POGZ* and *RGS16* RRs) displayed unrelated expression, and one set had one related and one unrelated MiniP (*NR2E1* RRs).

An additional six sets included a combination of MiniPs that either contained the Prom element or a LongProm element (usually a known promoter in the literature and contiguous but larger than the Prom element) within the given set (*i.e.,* overlapping basal



b

Region	Ple 17	Ple 24	Ple 26	Ple 28	Ple 32	Ple 34	Ple 53	Ple 55	Ple 88	Ple 123	Ple 167	Ple 170	Ple 240
Central spinal cord	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Nerve fibres	✓	×	×	✓	✓	✓	×	×	×	Weak	Some	✓	✓
DRG	×	Rare	Dif fuse	ND	ND	✓	✓	✓	×	✓	Punc tate	ND	✓

Figure 3 Thirteen MiniPromoters expressed in the mouse spinal cord. LacZ expression was detected by X-gal immunohistochemistry in whole mount thoracic spinal cord. (**a**) Ple17 (*C8ORF46* RRs), Ple28 (*CCL27* RRs), Ple32 (*CLDN5* RRs), and Ple170 (*POGZ* RRs) showed scattered expression in the central spinal cord and in nerve bundles. Ple24 (*CCKBR* RRs) expression was seen primarily in the spinal cord with rare puncta in the dorsal root ganglia (data not shown). Ple26 (*CCL27* RRs), Ple53 (*DCX* RRs), and Ple55 (*DCX* RRs) expressed throughout the gray matter regions and also in dorsal root ganglia; the latter particularly strong for Ple53 and Ple55. Ple34 (*CLDN5* RRs) and Ple123 (*ICMT* RRs) expressed throughout all of the spinal cord, while Ple123 demonstrated weaker expression in nerve bundles but strong expression in dorsal root ganglia. Ple88 (*GFAP* RRs) was expressed only in patches along white matter regions of the central spinal cord. Ple167 (*POGZ* RRs) expressed in patches in the spinal cord with some nerve fibers and occasional dorsal root ganglia puncta also positive (data not shown). Ple240 (*UGT8* RRs) is expressed strongly and predominantly along spinal nerve cord bundles with weaker expression in the central spinal cord tissue and in dorsal root ganglia. (**b**) Summary of expression pattern for each construct in the three key regions of the spinal cord. CH, chimera; DRG, dorsal root ganglion; ND, not done; RRs, regulatory regions; √, expression present; ×, expression absent.

promoters). Five of these six sets (*CCKBR*, *DBH*, *DCX*, *GFAP*, and *PITX3* RRs) had 5′ extensions compared with the Prom, and one set (*TRH* RRs) contained a 3′ extension compared with the Prom. All six sets displayed related expression.

Lastly, the *CCL27*-based set contained MiniPs with two nonoverlapping Prom1 and Prom2 designs. Prom1 was located further upstream and annotated as a putative cortical expression promoter based on 5′ cap analysis of gene expression data. A Prom2-based MiniP, located near the known transcription start site, was still related but distinctly different from those based on Prom1. Due to this difference, we classified this as a mixed set.

Seventy-five percent (12/16) of the MiniP sets displayed related expression in all positive MiniPs. Thus, we conclude that the common Prom element is generally sufficient to confer specificity, as compared with the endogenous mouse gene, and added putative regulatory elements either enhance or reduce expression levels, or alternatively expand or narrow the expression pattern.

Seventeen MiniPs deliver expression to the adult mouse eye

The gene set for MiniP design was originally chosen for region-specific and cell type–specific adult brain expression; however, since the brain and eye are both generated from diencephalon, it is likely that many designs would also be positive in the eye. Thus, we undertook a detailed investigation of eyes in otherwise brain-positive MiniPs.

We analyzed the expression pattern of 17 MiniPs found to be positive in the adult mouse eye (Table 1). Within this set, we observed expression in nearly all major retinal cell types. In extra-retinal tissue, we identified seven MiniPs positive in the cornea, six in the optic nerve, and two in the lens. Fifteen eye positive MiniPs are presented in Figure 6, and two additional ones are presented in Figure 5, as part of the *DCX* MiniP comparison.

We found Ple24 and Ple25 *(CCKBR* RRs) were expressed in the GCL and occasionally in the inner part of the INL. Ple26 (*CCL27* RRs) was expressed in the cornea, optic nerve, and all retinal layers. Another design for this gene, Ple28, showed expression in all layers and also in the lens but no expression in the cornea or optic nerve.

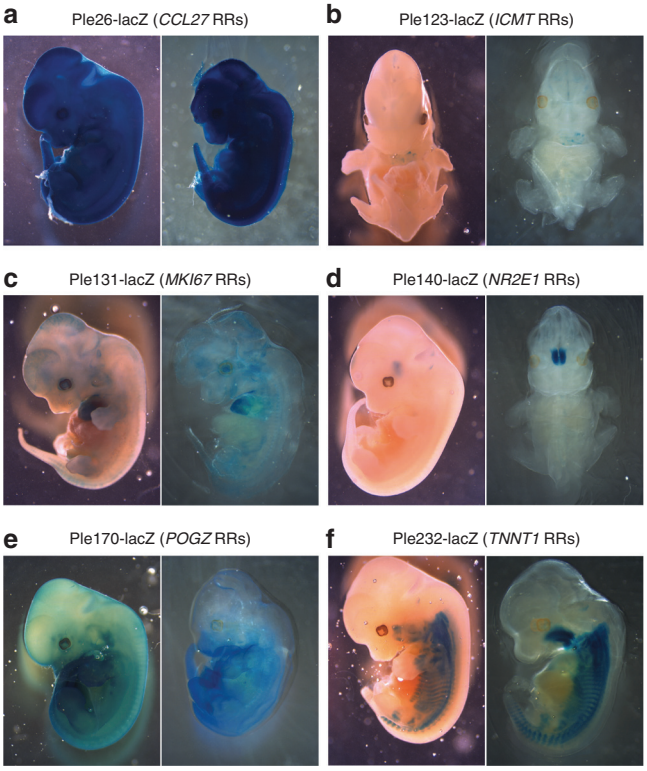


Figure 4 Expression of six MiniPromoters during development. E12.5 embryos were harvested and stained for lacZ expression using X-gal. Whole-mount (first panels) and partially cleared whole-mount embryos (second panels) were photographed. (**a**) Ple26 (*CCL27* RRs) stained throughout the entire embryo. (**b**) Ple123 (*ICMT* RRs) displayed variability, with the most common expression pattern being limited to the developing heart as shown. (**c**) Ple131 (*MKI67* RRs) was expressed throughout the embryo at varying levels, most prominently in the heart. (**d**) Ple140 (*NR2E1* RRs) showed a very strongly defined pattern in the hypothalamus of the embryo. (**e**) Ple170 (*POGZ* RRs) demonstrated staining in all parts of the embryo, but generally stronger outside of the nervous system. (**f**) Ple232 (*TNNT1* RRs) stained strongly in skeletal muscle and the tongue of the embryo but was not visible in the brain or spinal cord. RRs, regulatory regions.

Ple32 (*CLDN5* RRs) was expressed in the GCL and INL. Ple34 (*CLDN5* RRs) was positive in the INL and GCL and was expressed rarely in cells of the outer nuclear layer (photoreceptors). The morphology of cell nuclei located in the INL, with processes extending to the GCL, was suggestive of bipolar ON cells. Ple67 (*FEV* RRs) stained cells in the GCL and rare INL cells. Ple88, based on a known promoter for *GFAP*,[21] was expressed in the astrocytic layer as expected. Ple123 (*ICMT* RRs) stained all retinal layers including the cornea and optic nerve. Ple131 (*MKI67* RRs), based on proliferative maker KI-67, stained patches of the cornea only, consistent with other studies showing expression of KI-67 in basal epithelial cells in the cornea.[22] Ple146 (*NTSR1* RRs) stained rare GCL and INL cells (INL-positive cell not pictured). Ple155 *(PCP2* RRs) stained predominantly in the INL and inner plexiform layer and was strongly suggestive of bipolar ON cells, which is expected expression for this gene. Ple167 and Ple 170 (*POGZ* RRs) stained cells in all layers of the neural retina including the cornea. Ple240 (*UGT8* RRs) stained positive in the retinal inner segment (photoreceptors) and occasionally in the INL with fibers extending into the astrocytic layer, suggesting a Müller glial or bipolar cell type. The cornea and optic nerve were also positive for Ple240.

This characterization has led to a set that is unique compared with currently available eye promoters (see ref. 23 for a brief list).

Knock-in MiniPs that express in retinal ganglion cells predict expression of MiniP-GFP using AAV

To assess the feasibility of using the MiniPs for AAV-mediated delivery, we cloned three (Ple25, Ple53, and Ple67) MiniPs that were expressed primarily in the retinal ganglion cell layer, into the AAV2-(quad Y-F) vector. This vector, containing four tyrosine to phenylalanine mutations, has previously been shown to transduce cells of the retina uniformly, with a transduction profile that includes photoreceptors, Müller glia, and horizontal, bipolar, amacrine, and ganglion cells.[24] Two (Ple25 and Ple53) of the three MiniPs were 5′ trimmed by ~200–400 bp to fit into the viral vector due to limits on the carrying capacity of AAV constructs. Ple25 was trimmed from 3,740 bp (KI allele) to 3,312 bp (AAV), and Ple53 was trimmed from 3,520 bp (KI allele) to 3,310 bp (AAV), resulting in Ple25(Fmd1) and Ple53(Fmd1), respectively. Ple67 (2,202 bp) did not require trimming. Mice were injected intravitreally at postnatal day (P) 28, harvested 4 weeks later, and analyzed for green fluorescent protein (GFP) expression. The transduction profile of AAV2(quad Y-F) using the ubiquitous chicken beta-actin promoter demonstrated expression in all retinal layers (Figure 7a). Figure 7b,d,f show a side-by-side comparison of the knock-in mouse strain and the AAV vector–injected mice. All three AAV vectors supported expression of the reporter gene in the retinal ganglion cell layer. For Ple53 and Ple67, we also observed rare positive cells in the innermost aspect of the INL. Importantly, the 5′ trimming of ~200–400 bp for Ple25 and Ple53 to accommodate AAV space constraints did not alter their specificity, indicating that the deleted segments were nonessential for specificity. We conclude that for at least these three MiniPs, the expression pattern in the knock-in strain predicted that obtained using a virus-based delivery to the mouse retina. These data suggest that MiniPromoter specificity may be maintained in AAV vectors.

The retinal ganglion cell (RGC) layer in mice contains ~40% ganglion cells and 60% amacrine cells.[25] Nadal-Nicolás *et al.*[26] estimated that Brn3a+ cells label ~92.2% of RGCs, as co-labeled with the Fluorogold tracer in rat retina. Similarly, Fluorogold was estimated to label 97.8–98.4% of RGCs in rats.[27] Thus, we considered that ~90% of RGCs in the adult mouse retina should be labeled with Brn3a and could be used as a reliable marker of retinal ganglion cells. We undertook Brn3a co-labeling in virus-injected mice for Ple25, Ple53, and Ple67 to assess the extent of expression in ganglion cells, with the remainder of GFP-positive cells in the GCL assumed to be amacrine. As shown in Figure 7c,e,g, the MiniPs ranked qualitatively from most to least overlap with Brn3a: Ple53 > Ple25 > Ple67. We were also interested in how much expression was observed in amacrine (Brn3a-negative) cells in the GCL. Similarly, qualitative ranking of the same MiniPs from most to least expression in Brn3a-negative but GFP-positive cells ordered the MiniPs: Ple67 > Ple25 > Ple53. Thus, we conclude that Ple53 supports expression in the greatest proportion of retinal ganglion cells and is most specific to this cell type, while Ple67 supports expression in a subset of retinal ganglion cells in addition to numerous amacrine cells.

DISCUSSION

Interest in the identification of small promoters for capturing the expression pattern of biologically unique genes is longstanding.

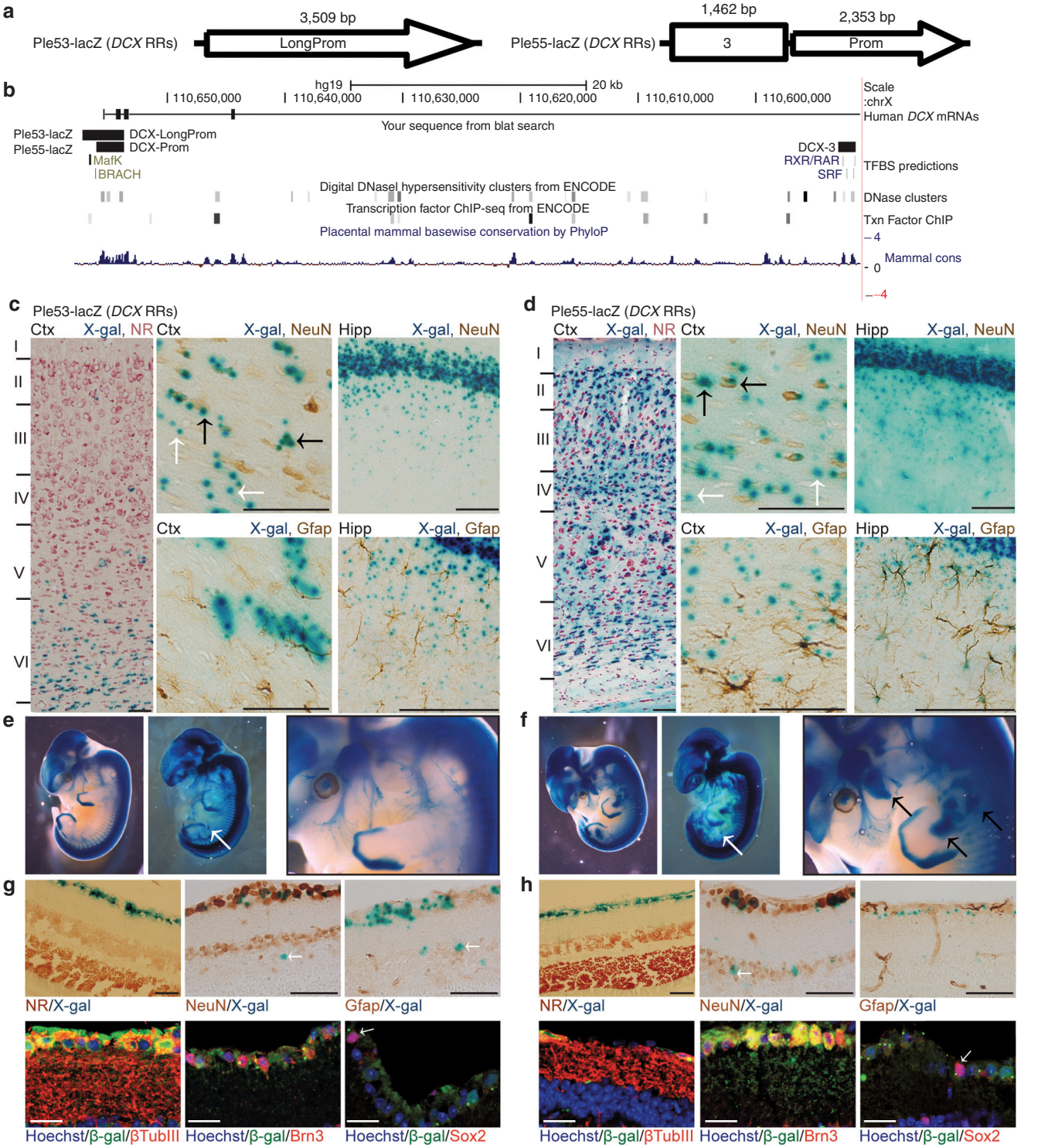


Figure 5 Comparison of two *DCX*-based MiniPromoter constructs with similar but nonidentical expression patterns identifies putative functional TFBS. (**a**) Diagram showing the differences between Ple53 and Ple55. Both MiniPromoters contained a minimal "Prom" sequence of 2,353 bp; however, Ple53 contained an additional but contiguous sequence and thus was named "LongProm", whereas Ple55 contained a noncontiguous segment 3 in addition to the Prom. (**b**) In the UCSC genome browser, predicted TFBS for Mafk and BRACH protein families were observed in the extended LongProm sequence specific to Ple53. Predicted RXR/RAR heterodimer and SRF sites were found in the segment 3 specific to Ple55. Comparison of the expression of lacZ using the substrate X-gal (blue) demonstrated that (**c, e, g**) Ple53-lacZ was expressed as a subset of the more expansive (**d, f, h**) Ple55-lacZ expression pattern. (**c, d**) The comparison demonstrated that with the Ple53 construct, the majority of cells labeled in the cortex were located in the deeper cortical layers, whereas with Ple55, they were distributed throughout the cortex. There were also more cells in Ple55 that did not co-label with NeuN. (**e, f**) Expression analysis in E12.5 X-gal stained whole-mount and partially cleared embryos for each MiniPromoter demonstrated that Ple55 contained all of the neuronal based expression observed in Ple53 plus staining in the precartilage or premuscle of the face and limb buds. Both MiniPromoters exhibited expression at the distal edge of the limb buds, and both included staining of the developing kidneys (white arrows). A close-up of the face and front limb showed the additional staining in Ple55 (black

arrows). (**g, h**) (upper panels) MiniPromoter lacZ expression as detected by X-gal reaction in the retina (left panel). Expression for the *DCX*-based designs was limited to the ganglion cell layer (GCL) for both Ple53 and Ple55, with rare cells observed in the inner nuclear layer. NeuN (marking all retinal ganglion and amacrine cells) co-labeled with X-gal (center panel), but Gfap (retinal astrocytes) did not (right panel). Rare cells did not co-label with NeuN (white arrows). (**g, h**) (lower panels) Immunofluorescence co-labeling with Brn3, βtubIII, and Sox2. Co-labeling demonstrated that most lacZ-positive cells stained with βtubIII and Brn3 markers (ganglion cells) for both MiniPromoters (left and center panels), and only rarely with Sox2-positive cholinergic amacrine cells (right panel, white arrows). Ctx, cortex; Hipp, hippocampus; NR, neutral red; RAR, retinoic acid receptor; RRs, regulatory regions; RXR, retinoid X receptor; SRF, serum response factor; TFBS, transcription factor binding sites; UCSC, University of California, Santa Cruz. (Scale bars = 100 μm, except for (**g**, bottom panel) fluorescence images 20 μm).

The Pleiades approach, with methodology described in detail previously,[9] has the ability to generate novel promoters in a high-throughput manner. Here, we present 18 additional promoters identified by this technique, and the further characterization of 15 previous designs.

Our approach focuses on physiologically relevant expression. The reasoning for this is threefold: (i) overexpression of therapeutic molecules can be deleterious[28]; (ii) overexpression of purely research-related molecules such as cre can cause cytotoxicity[19,29,30]; and (iii) introduction of artificial sequences to enhance expression can pose regulatory challenges due to their unknown specificity. Our strategy of mimicking endogenous expression levels using endogenous sequences is aimed at overcoming these challenges and thus enabling improved gene therapy designs.

However, there are some limitations to the Pleiades design pipeline. First, a small number of MiniPs displayed highly variable expression when docked at *Hprt*. We were unable to find previous reports of variable expression for constructs knocked-in at *Hprt*, which is generally a neutral docking location with consistent expression of docked constructs.[31] Expression variability was also not a general feature of our previous MiniP work.[9] Thus, we hypothesize that the variability observed is construct and sequence specific and reflects the large number of MiniPs tested. For the most part, the extent of variability we saw in embryos correlated with that observed for adults. It remains to be seen whether such construct-dependent variability is genome context specific or would be similarly reproduced in other systems (*e.g.,* virus-based delivery or knock-in at the *ROSA26* locus).

Additionally, the Pleiades data was limited by the inability to attain germline animals for a small subset of MiniPs. In these cases, chimeras were studied. Although each chimera may not display the entire cellular expression profile, the observed chimeric expression has been a reliable predictor of germline expression.[9] Lastly, we were limited in characterization of the specific cell types of expression, particularly for the brain where hundreds or thousands of different neuronal cell types are predicted to exist.[32] However, this data on regionalized expression of MiniP constructs remains a valuable scientific and therapeutic resource.

Importantly, this data represents numerous novel and biologically significant findings: first, we identified the zona incerta as a common region of expression for MiniP designs. We hypothesize that overrepresentation of the zona incerta brain region in this set is due to the diverse range of neurochemical cells located in this region, which includes neurons varying in soma size and shape (see ref. 33 for a comprehensive review on the zona incerta). In addition, extensive networking exists between cells of the zona incerta and numerous other brain regions (*e.g.,* the cerebral cortex, diencephalon, basal ganglia, brainstem, and spinal cord),[33] further suggesting a heterogeneous population in this region. The persistent finding of staining in this region suggests caution for therapeutic delivery where expression in the zona incerta could cause harm.

The brain-positive MiniPs provide novel tools for researchers and clinicians alike. Within our toolbox of MiniPs, we have now added constructs that would drive expression in the cortex and hippocampus (Ple15 and Ple16), in oligodendroglia (Ple240), in a near pan-neuronal expression (Ple26 and Ple170), or in the developing nervous system (Ple53 and Ple55).

Selected designs were also validated for their use in developmental studies. We tested six adult brain–positive MiniPs and showed expression in the brain for five of these at E12.5. The remaining MiniP, Ple232, was expressed only outside of the embryonic brain, in contrast to its expression in the adult. This finding is not entirely unexpected, since not all genes will be expressed in the brain at the time point used in this study (E12.5, mid-point of neurogenesis). Remarkably, we showed highly specific expression of Ple140 (*NR2E1* RRs) in the hypothalamus at E12.5, with no evidence of its expression in other embryonic tissue. These embryonic studies allow us to rapidly detect the functioning of MiniPs outside of the CNS, like Ple232 (*TNNT1* RRs) which could prove a valuable tool for researchers studying, and developing treatments for, skeletal muscle disorders. Overall, this embryonic dataset suggests that many other Pleiades MiniPs are likely to be positive during development.

We were surprised by the portion of MiniPs expressed in the mouse eye. This represents a novel toolset for driving retinal gene expression, and these MiniPs offer great therapeutic potential. Bipolar cells are the next step in retinal processing after photoreceptors are activated by light. In diseases where photoreceptors are nonfunctional or have undergone cell death, one strategy is to target the bipolar cells.[34] Similarly, since retinal ganglion cells form the last step in the visual processing pathway, expression of therapeutic molecules could be a viable option here also,[34] with perhaps more generic retinal disease application. Our efforts have developed at least two MiniPs with bipolar cell expression (Ple34 and Ple155) and six with ganglion cell layer expression (Ple24, Ple25, Ple53, Ple55, Ple67, and Ple146). The finding of such a large proportion (35.3%) of eye positives specific to the GCL implies that selection of candidates for MiniP design using brain expression data may be biased toward generating RGC-positive designs. We hypothesize this may be due to the similarities between RGCs and other CNS neurons with regards to (i) the ability to fire action potentials and (ii) similarities in gene expression between RGC-specific transcripts and other CNS neurons that fire action potentials.[35–39] Interestingly, the ganglion cell-marking Brn3 transcription factors are also highly expressed in the sensory neurons of the dorsal root ganglia.[40] Of the six GCL-positive MiniPs, we were able to assess dorsal root ganglion expression in three and found all three positive (Ple24, Ple53, and Ple55). Thus, our selection strategy has not only preferentially generated MiniPs positive in the GCL of the retina but also in the dorsal root ganglion (eight out of ten MiniPs tested).

Other MiniPs that expressed in the eye also demonstrate potential utility: Ple88 (*GFAP* RRs) could be used to aid in the treatment of retinal disorders with increased astrogliosis (defined as increased GFAP expression), and Ple131 (*MKI67* RRs) may be useful for corneal repair. Curiously, we found occasional discordance

between cell type expression in the brain and eye. The Ple155 (*PCP2* RRs) MiniP was expressed as expected in the mouse retina (bipolar ON cells), but with seemingly unrelated expression in the brain (MiniP expression present in cortex, thalamus, and hypothalamus but absent in Purkinje cells for which *PCP/L7* is well known).[41,42] This data further supports the notion that cell type specificity is regulatory element driven in specific organs or tissues.[43] Altogether, this represents a powerful set of tools for driving gene expression in the eye.

One concern was whether the observed expression pattern of MiniPs would be reliably reproduced when used in viral vectors. Viral vector gene delivery is the preferred therapeutic method for the brain and eye, primarily because the number of treatment options via gene delivery is vast. We showed that a subset of three MiniPs (Ple25, Ple53, and Ple67), which expressed in the retinal ganglion cell layer in knock-in mice, exhibited a similar expression pattern in the mouse retina when delivered intravitreally and expressing from recombinant AAV2. We hypothesize that the characterization by single-copy site-specific knock-ins, and use of exclusively human DNA, may have predisposed these MiniPs to success in AAV. Using this methodology, the desirable expression characteristics are more likely to be inherent to the MiniP only and not dictated by copy-number or the genomic location. In addition, human DNA should be less prone to expression variability compared with more distantly related sequence or viral promoters. Finally, critical to the success of this work was the choice of an AAV serotype with a broad transduction profile in the retina, which included the target cell types of the MiniPs. Interestingly, with this success, the MiniPs can also be used to map the specificity of novel serotypes, providing a valuable tool for researchers working on AAV capsid biology.

An interesting pattern emerged using these three MiniPs: we observed an inverse qualitative relationship between the overlap of stained retinal ganglion cells (defined here as Brn3a-positive in the adult retina) and amacrine cells (defined as Brn3a-negative) located in the GCL. This result was surprising, as we were anticipating that MiniP expression may target different subsets of ganglion or amacrine cells, with no expression in the other cell types. However, we observed overlapping expression in both cell types, and our data implies that transcriptional regulation between amacrine and RGC cells may not be distinctly separate. Indeed, in support of this idea are the results of Kunzevitzky *et al.*,[44] who evaluated the transcriptional profiles using microarrays of amacrine and retinal ganglion cells at E20, P5, and P11 in the rat. They found ~70% overlap in expression between the two cell types at all time points (see Figure 5 in the study of Kunzevitsky *et al.*). Both amacrine and ganglion cells are born between E14 and E18. Evidence for potential overlapping transcriptional regulation between these two cell types also comes from work that showed VC1.1+ progenitors can generate both VC1.1+ and VC1.1− progenitors.[45,46] The former produces amacrine and horizontal progeny, while the latter produces retinal ganglion cells and cones.[46]

Two of these three Ple MiniPs were trimmed prior to use in AAV. In the future, the process of making Pleiades MiniPs even smaller will be informed by data generated by novel technologies. Such data includes that of the ENCODE project,[47] FANTOM projects,[48–50] and VISTA Enhancer project.[5] Specifically, the ENCODE transcription factor ChIP-seq and DNase I hypersensitivity data, along with FANTOM cap analysis of gene expression,[50] and p300 enhancer protein binding[51] data will allow for anchoring of, and supportive evidence for function of, RRs. In addition, the development of ChIP-exo,[52] a

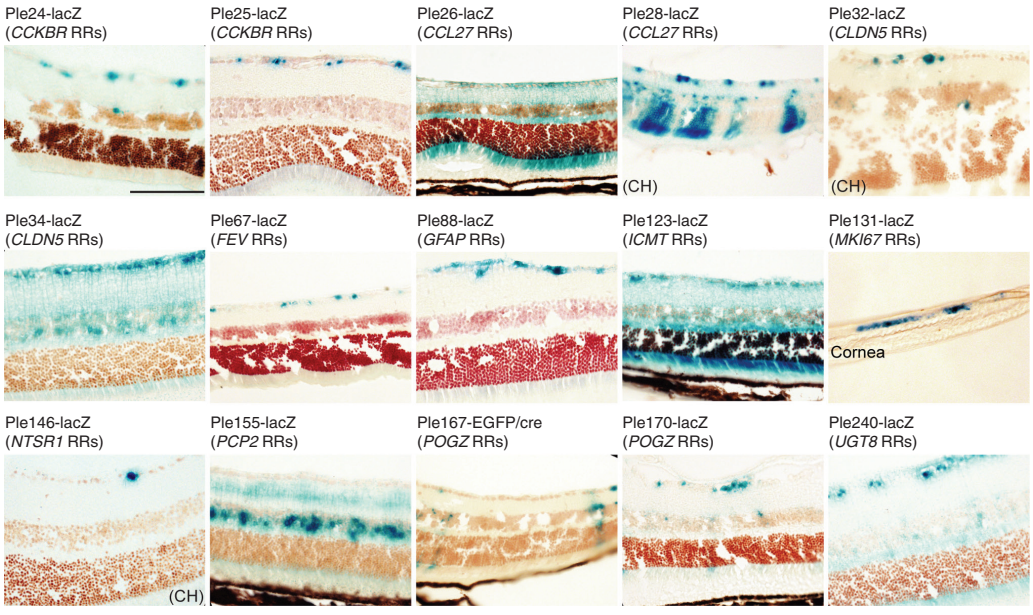


Figure 6 Fifteen MiniPromoters expressed in the mouse eye. Mice were perfused, eyes harvested, and stained for lacZ using the substrate X-gal (blue). Eyes were cryoprotected, embedded, and sectioned. Ple24 (*CCKBR* RRs), Ple25 (*CCKBR* RRs), Ple32 (*CLDN5* RRs), Ple67 (*FEV* RRs), and Ple146 (*NTSR1* RRs) expressed in a subset of GCL cells. Rarely, an amacrine cell in the innermost part of the INL was positive in Ple24, or a rare positive INL cell in Ple32. Ple26 and Ple28 (both based on *CCL27* RRs) expressed in the INL, and in photoreceptors as evidenced by the columnar staining of the ONL and photoreceptor segments. Ple26 showed extensive IPL and OPL fiber staining and Ple28 was also positive in the GCL. Ple34 (*CLDN5* RRs) and Ple155 (*PCP2* RRs) have staining characteristic of bipolar cells, with cell bodies in the INL and processes extending to the GCL (Ple34) or inner half of the IPL (Ple155). Ple88 (*GFAP* RRs) stained in the retinal astrocytic layer. Ple123 (*ICMT* RRs) expressed throughout the retina. Ple131 (*MKI67* RRs) expressed in cells of the corneal epithelium. Ple167 and Ple170 (both based on *POGZ* RRs) demonstrated similar staining patterns in the GCL, INL, OPL and photoreceptor segments. Ple240 (*UGT8* RRs) had fiber-like staining spanning the IPL and clear cell body labeling in a subset of INL cells. CH, chimera; INL, inner nuclear layer; IPL, inner plexiform layer; ONL, outer nuclear layer; OPL, outer plexiform layer; RRs, regulatory regions. (Scale bar = 100 μm).

refinement to ChIP-seq, brings single-nucleotide resolution and in-depth analysis of regulatory motifs using massively parallel reporter assays.[53,54]

In conclusion, this work confirms that MiniP design is a tractable high-throughput option for tissues and shows that it can identify novel promoters that translate across transgene delivery systems. We attribute the success of our design strategy to: (i) the use of evolutionary conserved sequence and human-only DNA limiting transgene inactivation; (ii) paring of putative regulatory elements along with their respective core basal promoters instead of unrelated or viral core promoters, the importance of which is summarized in the study by Haeussler and Joly[20]; and (iii) characterization of reproducible, single-copy knock-ins in a consistent neutral genomic environment. These MiniPs immediately enable greater specificity in gene therapy vectors, and the resource will greatly benefit both basic research and future human gene therapy.

MATERIALS AND METHODS

Ethics statement

All procedures involving animals were in accordance with the Canadian Council on Animal Care and University of British Columbia Animal Care Committee (protocols #A05-1258, A05-1748, A09-0980, and A09-0981). All virus experiments were approved by the University of Florida Institutional Animal Care and Use Committee and were conducted in accordance with the Association for Research in Vision and Opthalmology Statement for the Use of Animals in Ophthalmic and Vision Research and with National Institutes of Health regulations.

Generation of knock-in mice

Candidate genes for MiniP design were identified in previous datasets[55] as previously described.[9] MiniPs were designed and single-copy knock-in mice at the *Hprt*$^{b\text{-}m3}$ locus on the X-chromosome were generated as previously described.[9,56] Animals were housed at the Centre for Molecular Medicine and Therapeutics, in a barrier pathogen-free facility. Animals were kept on a 14-hour on light cycle and provided food and water *ad libitum*. MiniP strains were backcrossed to either C57BL/6J (JAX Stock#000664) or 129S1/SvImJ (JAX Stock#002448). Animals used for analysis were at least two male germline mice unless otherwise noted with the chimera indicator (CH)—chimeras were analyzed for MiniPs that never entered the germline despite extensive breeding of chimeras and often multiple rounds of microinjection of different embryonic stem cell clones. Each strain and embryonic stem cell line was deposited at The Jackson Laboratory as indicated in the Supplementary Table S2.

PCR analysis of genomic DNA

Vector NTI software (Invitrogen, Carlsbad, CA) was used to design PCR assays for the different constructs. MiniP-specific PCR genotyping assays are

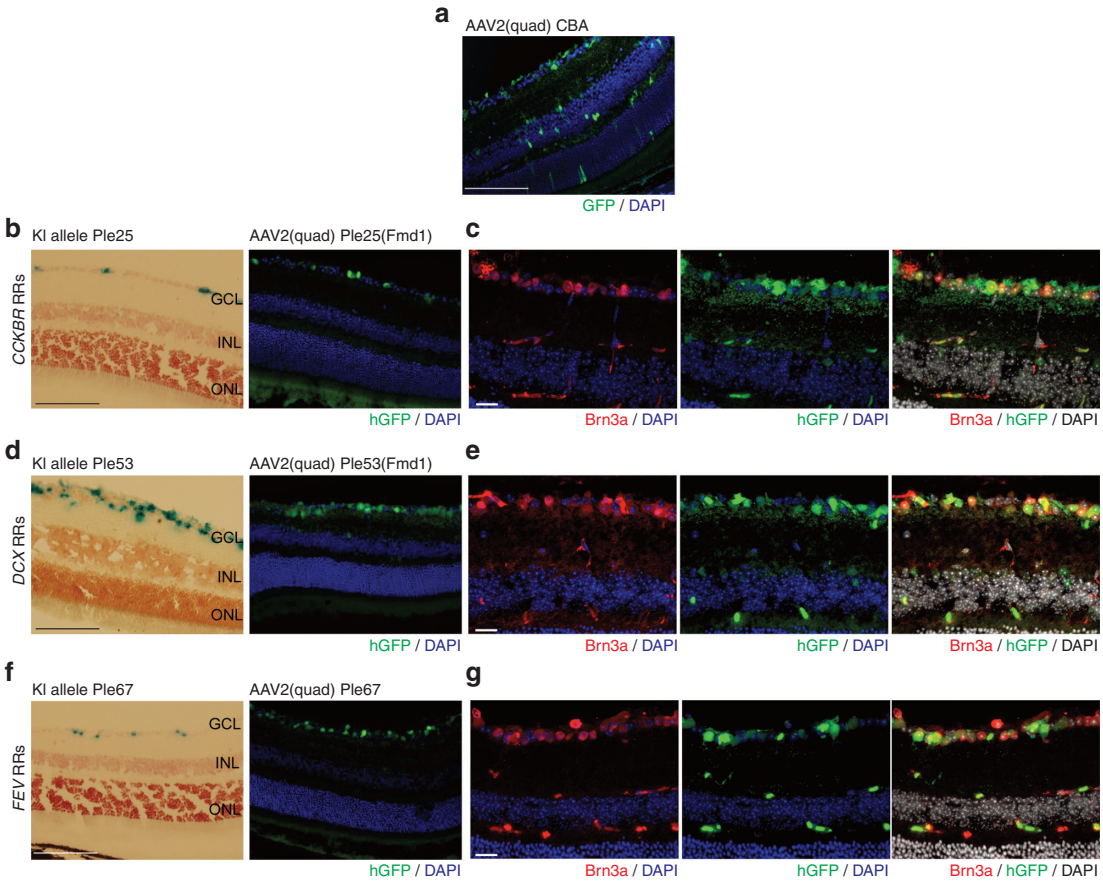


Figure 7 Three MiniPromoters that expressed in the retinal ganglion cell layer in knock-in mice maintained this restricted expression from AAV viral vectors. (**a**) Transduction profile of the AAV2(quad Y-F) vector with the CBA promoter driving GFP. Three MiniPromoter constructs (**b, c**) Ple25, (**d, e**) Ple53, and (**f, g**) Ple67, which expressed in the GCL when docked 5′ of *Hprt* driving β-galactosidase, were recloned into AAV2(quad Y-F) driving hGFP. (**b, d, f**) Knock-in mouse retinas stained for lacZ using the substrate X-gal (blue) and counterstained with neutral red (first panel) showed similar ganglion expression to intravitreal injected AAV2(quad) with the MiniPromoter driving hGFP (second panel, epifluorescence). (**c, e, g**) Virus-injected retinas were immunostained for Brn3a (a marker of ganglion cells). First panel shows Brn3a (red) with DAPI (blue), second panel hGFP (green) with DAPI (blue), and third panel the merge (yellow) of Brn3a (red), hGFP (green), with DAPI (gray). CBA, chicken beta-actin promoter; DAPI, 4′,6-diamidino-2-phenylindole; GCL, ganglion cell layer; GFP, green fluorescent protein; INL, inner nuclear layer; ONL, outer nuclear layer; RRs, regulatory regions. (Scale bars = 100 µm, except for fluorescence images 20 µm).

available on the http://www.pleiades.org Web site or by contacting the corresponding author.

Immunohistochemistry of adult brains

One millimeter sectioned tissues. The mice were anesthetized with avertin and perfused transcardially with 4% paraformaldehyde (PFA) and postfixed 2–4 hours. The brains were then removed from 4% PFA and immediately sectioned. To section, brains were placed ventral side up into the adult mouse coronal or sagittal Rodent Brain Matrix (ASI Instruments, Warren, MI). Brains were sectioned lateral through medial to lateral (for sagittal) or rostral to caudal (for coronal) using single-edge razor blades (Electron Microscopy Sciences, Hatfield, PA). All slices from one brain were placed into 1 well of a 12-well plate containing 1× phosphate-buffered saline (PBS) (Invitrogen) or 0.1% sodium azide in 1× PBS until staining. Expression of β-gal (lacZ) was detected with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) staining as previously described.[9] Briefly, the staining was performed with 3–5 ml of X-gal staining solution (25 mg/ml X-gal, 1 mol/l $MgCl_2$, 50 mmol/l potassium ferri-cyanide, 50 mmol/l potassium ferro-cyanide, and 1× PBS to volume) per well in a 24-well plate. The plate was wrapped in aluminum foil and incubated at 37 °C for 10–16 hours with gentle agitation on a rotator board. Subsequently, the sections were transferred into PBS, examined under a dissecting microscope, and photographed using a CoolSNAP-Pro_{cf} color camera (Media Cybernetics, Rockville, MD) mounted on a Leica MZ12.5 stereomicroscope (Leica Microsystems, Buffalo Grove, IL) and Image-Pro Express v.4.5.1.3 software (Media Cybernetics).

Cryosectioned tissues. The mice were anesthetized with avertin and perfused transcardially with 4% PFA in 0.1 mol/l phosphate buffer (pH 7.4) for 15 minutes. Brains were dissected and postfixed in the same solution for 2 hours and transferred into 25% sucrose-PBS overnight. Each brain was sectioned using a cryostat, and 20-µm sagittal sections were collected. X-gal staining and imaging was performed as described for 1-mm sections. Co-staining of X-gal and NeuN or Gfap was performed sequentially. Primary antibodies used were mouse anti-GFAP (1:1,000; Millipore, Billerica, MA) and mouse anti-NeuN (1:500; Chemicon, Billerica, MA). Bound primary antibodies were detected using the VectaStain Elite ABC Kit (Vector Labs, Burlingame, CA; PK-6100), as per manufacturer's instructions. Slides were gradually dehydrated in ethanol and xylene series and mounted with coverslips. Bright-field images were taken with an Olympus (Center Valley, PA) BX61 motorized microscope using the DP Controller software, Olympus.

Other tissues. Spinal cord, heart, liver, lung, and ear notch were collected from perfused mice and stained as described above.

Embryo harvesting and staining

For timed pregnancies, female heterozygous MiniP mice were crowded for 7–10 days and set-up in the afternoon with C57BL/6J or 129S1/SvImJ studs. Subsequently, females were checked for copulation plugs for 5 days in the morning. The day of plug was considered E0.5. At E12.5, embryos were harvested, yolk sac tissue taken for PCR analysis, and embryos fixed in 4% PFA/PBS for 2–6 hours. Embryos were stained in X-gal solution for 18 hours, postfixed overnight at 4 °C, and stored in 70% ethanol. Pictures were taken of uncleared embryos. The embryos were then partially cleared based on the method of Schatz *et al.*[57] In brief, embryos were incubated in step 1 solution (20% glycerol with 0.8% KOH) for 3 days at 37 °C with swirling once daily. Further clearing was done in step 2 solution (50% glycerol with 0.5% KOH) for 3 days at 37 °C. Postfixing of fragile cleared embryos was done in 4% PFA for 2 hours at room temperature. Embryos were stored until imaging in 70% ethanol, after which they were equilibrated to 100% glycerol for imaging. Imaging was performed as per 1-mm brain sections (above). All embryo expression patterns were seen in at least two animals.

Immunohistochemistry on eyes

LacZ staining was detected either directly from MiniPs driving β-gal, or by driving expression of EGFP/cre resulting in recombination of the *Gt(ROSA)26Sor*^{tm1Sor} locus. After whole body perfusion of mice, eyes were postfixed in 4% PFA. LacZ staining was performed using the X-gal substrate for 18 hours at 37 °C. Eyes were then postfixed in 4% PFA for 2 hours prior to cryoprotection in 25% sucrose-PBS overnight. Eyes were embedded in optimal cutting temperature (OCT), along with positive and negative controls, and cryosectioned at 12 µm and mounted on SuperFrost Plus slides, Fisher Scientific. Eye sections were counterstained with neutral red for 45 seconds to 1 minute. Eye expression was confirmed in at least two animals unless otherwise noted with the chimera indicator (CH). Bright-field images were taken using Olympus BX61 motorized microscope using the DP Controller software.

Immunofluorescence on eyes

Eyes were harvested and fixed in 4% PFA for 2 hours prior to cryoprotection in 25% sucrose-PBS overnight. Eyes were embedded in OCT, and cryosectioned at 12 µm and mounted on SuperFrost Plus slides prior to immunofluorescence. Standard immunofluorescence procedures were followed.[9] Primary antibodies include: chicken anti-β-gal (Abcam Cambridge, MA; ab9361), rabbit anti-β-tubulin III (Covance, Princeton, NJ; PRB-435P), rabbit anti-Brn3 (Santa Cruz Biotechnology, Santa Cruz, CA; sc28595), and rabbit anti-Sox2 (Abcam; ab97959). Secondary antibodies used were (all from Invitrogen): chicken anti-rabbit Alexa-594, goat anti-chicken Alexa-488, goat anti-chicken Alexa-594, goat anti-mouse Alexa-488, goat anti-mouse Alexa-594, goat anti-rabbit Alexa-488, and goat anti-rabbit Alexa-594. Nuclei were stained using 1 µg/ml of Hoechst 33342 (Sigma, St Louis, MO; 14533). Slides were mounted with Prolong Gold (Invitrogen) antifade reagent and imaged using an Olympus BX61 microscope and the InVivo software (Olympus), or alternatively using confocal laser scanning microscopy with a Leica SP5 II microscope (Leica Microsystems) and LAF software (Leica Microsystems).

Adeno-associated virus

Recombinant AAV vector plasmids containing a 2,202 bp Ple67 (*FEV* RRs), a 3,310 bp Ple53 (*DCX* RRs) or 3,312 bp Ple25 (*CCKBR* RRs) MiniPromoter construct, an SV40 splice donor–acceptor site followed by the cDNA for GFP were constructed. To accommodate the packaging limit of recombinant AAV, Ple25 and Ple53 were shortened relative to the knock-in MiniP allele as follows: MiniPs were amplified by PCR, and the respective restriction sites were added via the oligonucleotides. The following oligonucleotides were used: Ple25 forward *Kpn*I: GTCGGGTACCAGAATTTCAGGAGCCCATATG, Ple25 reverse *Xho*I: GTCGCTCGAGTGCCTTACTCAGCTCGACC, Ple53 forward *Kpn*I: GTCGGGTACCTTATCCCATCCATTGTTGTATC, Ple53 reverse *Xho*I: GTCGCTCGAGGGTGGAACCTCAGAGACCTG, Ple67 forward *Eco*RI: GCTGGAATTCGGAGAAGATCAAGACTCAGG, and Ple67 reverse *Xho*I: GCTGCTCGAGGAGGCAGGCTTTGCGCTC, (restriction sites underlined). The plasmids pEMS1496, pEMS1524, and pEMS1538 were used as templates. The shorter Ple25 promoter construct is referred to as Ple25(Fmd1), and the shorter Ple53 promoter construct as Ple53(Fmd1).

Identity for all promoters was confirmed by sequencing the recombinant AAV vector plasmids before packaging into AAV2-(quad Y-F), in which the VP3 capsid protein contains four surface exposed tyrosine to phenylalanine mutations (Y272F, Y444F, Y500F, and Y730F).[24] Vectors were purified and titered according to previously published methods.[58] At least 5×10^9 vector genomes (vg) per eye were injected intravitreally in 4-week-old C57BL/6 mice.

GFP expression was analyzed 4 weeks postinjection. Mice were euthanized, and the eyes were fixed in 4% PFA for 2 hours at room temperature. Eyes were then treated with a stepwise sucrose gradient (from 4 to 20% sucrose-PBS) and embedded and frozen in Tissue-Tek OCT compound (Sakura Finetek, Torrance, CA). Ten-micron-thick sections of each eye were then washed in PBS for 15 minutes, mounted in VECTASHIELD mounting medium (Vector Labs) with 4′,6-diamidino-2-phenylindole, and analyzed for GFP expression by confocal microscopy.

Image processing

Images were processed using ImageJ (http://rsbweb.nih.gov/ij/), Photoshop, and Illustrator (Adobe, San Jose, CA). Brightness, contrast, and scaling adjustments were performed. For images where brains were not properly aligned (Figure 2), they were traced with magnetic lasso and repositioned on a black background. The auto-tone feature in Photoshop was used to correct for color balance and exposure settings for images of eyes (Figure 6).

Data and resource distribution

Data on MiniPs can also be found on Pleiades Promoter Project website (http://www.pleiades.org/). All plasmids and sequence information are available from Addgene (http://www.addgene.org/). Embryonic stem cells

and mice have been deposited at, and are available from, the Mutant Mouse Regional Resource Centers (http://www.mmrrc.org/) and The Jackson Laboratory (http://www.jax.org/).

CONFLICT OF INTEREST

C.N.d.L., E.P.-C., C.A.D., S.J.M.J., R.A.H., D.G., W.W.W., E.M.S., and the University of British Columbia have filed for US patents on a subset of the MiniPs. W.W.H. and the University of Florida have a financial interest in the use of AAV therapies, and own equity in a company (AGTC) that might, in the future, commercialize some aspects of this work.

ACKNOWLEDGMENTS

We thank the entire Pleiades Promoter Project team for their pipeline work, which directly facilitated the generation of the mouse strains used in this manuscript. We thank Nazar Babyak, Shadi Khorasan-zadeh, and Tara R. Candido for their tissue culture expertise; Andrea J. Korecki and Russell J. Bonaguro for molecular biology expertise; Jacek Mis, Jing Chen, and Kristi Hatakka for microinjection; Penny Qu for assistance with imaging; Anthony Mathelier for bioinformatic discussion; Ximena Corso-Díaz for critical reading of the manuscript; and Marina Campbell for aid in manuscript preparation. This work was supported as part of the Pleiades Promoter Project by grants from Genome Canada (grant number: 20R03061); Genome British Columbia (grant number: 045PLE); GlaxoSmithKline R&D (grant number: 20R62517); BC Mental Health and Addiction Services (grant number: 20R62601); Child and Family Research Institute (grant number: 20R42459); University of British Columbia (UBC) Institute of Mental Health (grant number: 20R78630); and UBC Office of the Vice President Research (grant number: 20R78603) to S.J.M.J., R.A.H., D.G., W.W.W., and E.M.S. We also acknowledge partial support of this work from the National Institutes of Health (NIH) (grant number: EY021721) and grants from the Macula Vision Research Foundation, Foundation Fighting Blindness, Eldon Family Foundation, Overstreet Fund, and Research to Prevent Blindness to W.W.H. In addition, we acknowledge partial support from NIH (grant number: R01GM084875) to W.W.W. Finally, we acknowledge salary support from Canadian Research Chairs to D.G. and E.M.S., Canadian Institutes of Health Research New Investigator Award to W.W.W., Canadian Institutes of Health Research Canada Graduate Scholarships to C.N.d.L., and Michael Smith Foundation for Health Research Awards to C.N.d.L., S.J.M.J., R.A.H., and W.W.W.

REFERENCES

1 Ng, L, Bernard, A, Lau, C, Overly, CC, Dong, H-W, Kuan, C, *et al*. (2009). An anatomic gene expression atlas of the adult mouse brain. *Nat Neurosci* **12**: 356–362.
2 Gong, S, Zheng, C, Doughty, ML, Losos, K, Didkovsky, N, Schambra, UB *et al*. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**: 917–925.
3 Visel, A, Thaller, C and Eichele, G (2004). GenePaint.org: an atlas of gene expression patterns in the mouse embryo. *Nucleic Acids Res* **32** (Database issue): D552–D556.
4 Magdaleno, S, Jensen, P, Brumwell, CL, Seal, A, Lehman, K, Asbury, A *et al*. (2006). BGEM: an in situ hybridization database of gene expression in the embryonic and adult mouse nervous system. *PLoS Biol* **4**: e86.
5 Visel, A, Minovitsky, S, Dubchak, I and Pennacchio, LA (2007). VISTA Enhancer Browser–a database of tissue-specific human enhancers. *Nucleic Acids Res* **35** (Database issue): D88–D92.
6 Snyder, RO (1999). Adeno-associated virus-mediated gene delivery. *J Gene Med* **1**: 166–175.
7 Xu, R, Janson, CG, Mastakov, M, Lawlor, P, Young, D, Mouravlev, A *et al*. (2001). Quantitative comparison of expression with adeno-associated virus (AAV-2) brain-specific gene cassettes. *Gene Ther* **8**: 1323–1332.
8 Cordier, L, Gao, GP, Hack, AA, McNally, EM, Wilson, JM, Chirmule, N *et al*. (2001). Muscle-specific promoters may be necessary for adeno-associated virus-mediated gene transfer in the treatment of muscular dystrophies. *Hum Gene Ther* **12**: 205–215.
9 Portales-Casamar, E, Swanson, DJ, Liu, L, de Leeuw, CN, Banks, KG, Ho Sui, SJ *et al*. (2010). A regulatory toolbox of MiniPromoters to drive selective expression in the brain. *Proc Natl Acad Sci USA* **107**: 16589–16594.
10 Molyneaux, BJ, Arlotta, P, Menezes, JR and Macklis, JD (2007). Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* **8**: 427–437.
11 Heimer-McGinn, V and Young, P (2011). Efficient inducible Pan-neuronal cre-mediated recombination in SLICK-H transgenic mice. *Genesis* **49**: 942–949.
12 Gerdes, J, Schwab, U, Lemke, H and Stein, H (1983). Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int J Cancer* **31**: 13–20.
13 Novelli, G, Gennarelli, M, Zelano, G, Sangiuolo, F, Lo Cicero, S, Samson, F *et al*. (1992). Polymerase chain reaction in the detection of mRNA transcripts from the slow skeletal troponin T (TNNT1) gene in myotonic dystrophy and normal muscle. *Cell Biochem Funct* **10**: 251–256.
14 Couillard-Despres, S, Winner, B, Karl, C, Lindemann, G, Schmid, P, Aigner, R *et al*. (2006). Targeted transgene expression in neuronal precursors: watching young neurons in the old brain. *Eur J Neurosci* **24**: 1535–1545.
15 Reiner, O, Coquelle, FM, Peter, B, Levy, T, Kaplan, A, Sapir, T *et al*. (2006). The evolving doublecortin (DCX) superfamily. *BMC Genomics* **7**: 188.
16 Wakabayashi, T, Kosaka, J, Mori, T, Takamori, Y and Yamada, H (2008). Doublecortin expression continues into adulthood in horizontal cells in the rat retina. *Neurosci Lett* **442**: 249–252.
17 Li, S, Czubryt, MP, McAnally, J, Bassel-Duby, R, Richardson, JA, Wiebel, FF *et al*. (2005). Requirement for serum response factor for skeletal muscle growth and maturation revealed by tissue-specific gene deletion in mice. *Proc Natl Acad Sci USA* **102**: 1082–1087.
18 Lu, PP and Ramanan, N (2012). A critical cell-intrinsic role for serum response factor in glial specification in the CNS. *J Neurosci* **32**: 8012–8023.
19 Boulaire, J, Balani, P and Wang, S (2009). Transcriptional targeting to brain cells: engineering cell type-specific promoter containing cassettes for enhanced transgene expression. *Adv Drug Deliv Rev* **61**: 589–602.
20 Haeussler, M and Joly, JS (2011). When needles look like hay: how to find tissue-specific enhancers in model organism genomes. *Dev Biol* **350**: 239–254.
21 Brenner, M, Kisseberth, WC, Su, Y, Besnard, F and Messing, A (1994). GFAP promoter directs astrocyte-specific expression in transgenic mice. *J Neurosci* **14** (3 Pt 1): 1030–1037.
22 Joyce, NC, Meklir, B, Joyce, SJ and Zieske, JD (1996). Cell cycle protein expression and proliferative status in human corneal cells. *Invest Ophthalmol Vis Sci* **37**: 645–655.
23 Hashimoto, T (2008). Development of viral vectors with optimal transgene expression for ocular gene therapies. *Adv Exp Med Biol* **613**: 113–119.
24 Petrs-Silva, H, Dinculescu, A, Li, Q, Deng, WT, Pang, JJ, Min, SH *et al*. (2011). Novel properties of tyrosine-mutant AAV2 vectors in the mouse retina. *Mol Ther* **19**: 293–301.
25 Jeon, CJ, Strettoi, E and Masland, RH (1998). The major cell populations of the mouse retina. *J Neurosci* **18**: 8936–8946.
26 Nadal-Nicolás, FM, Jiménez-López, M, Sobrado-Calvo, P, Nieto-López, L, Cánovas-Martínez, I, Salinas-Navarro, M *et al*. (2009). Brn3a as a marker of retinal ganglion cells: qualitative and quantitative time course studies in naive and optic nerve-injured retinas. *Invest Ophthalmol Vis Sci* **50**: 3860–3868.
27 Salinas-Navarro, M, Mayor-Torroglosa, S, Jiménez-López, M, Avilés-Trigueros, M, Holmes, TM, Lund, RD *et al*. (2009). A computerized analysis of the entire retinal ganglion cell population and its spatial distribution in adult rats. *Vision Res* **49**: 115–126.
28 Schedl, A, Ross, A, Lee, M, Engelkamp, D, Rashbass, P, van Heyningen, V *et al*. (1996). Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* **86**: 71–82.
29 Buerger, A, Rozhitskaya, O, Sherwood, MC, Dorfman, AL, Bisping, E, Abel, ED *et al*. (2006). Dilated cardiomyopathy resulting from high-level myocardial expression of Cre-recombinase. *J Card Fail* **12**: 392–398.
30 Forni, PE, Scuoppo, C, Imayoshi, I, Taulli, R, Dastrù, W, Sala, V *et al*. (2006). High levels of Cre expression in neuronal progenitors cause defects in brain development leading to microencephaly and hydrocephaly. *J Neurosci* **26**: 9593–9602.
31 Farhadi, HF, Lepage, P, Forghani, R, Friedman, HC, Orfali, W, Jasmin, L *et al*. (2003). A combinatorial network of evolutionarily conserved myelin basic protein regulatory sequences confers distinct glial-specific phenotypes. *J Neurosci* **23**: 10214–10223.
32 Masland, RH (2004). Neuronal cell types. *Curr Biol* **14**: R497–R500.
33 Mitrofanis, J (2005). Some certainty for the "zone of uncertainty"? Exploring the function of the zona incerta. *Neuroscience* **130**: 1–15.
34 Cepko, CL (2012). Emerging gene therapies for retinal degenerations. *J Neurosci* **32**: 6415–6420.
35 Trimarchi, JM, Stadler, MB, Roska, B, Billings, N, Sun, B, Bartch, B *et al*. (2007). Molecular heterogeneity of developing retinal ganglion and amacrine cells revealed through single cell gene expression profiling. *J Comp Neurol* **502**: 1047–1065.
36 Kong, WC and Cho, EY (1999). Antibodies against neurofilament subunits label retinal ganglion cells but not displaced amacrine cells of hamsters. *Life Sci* **64**: 1773–1778.
37 Julien, JP, Tretjakoff, I, Beaudet, L and Peterson, A (1987). Expression and assembly of a human neurofilament protein in transgenic mice provide a novel neuronal marking system. *Genes Dev* **1**: 1085–1095.
38 Reh, TA, Tetzlaff, W, Ertlmaier, A and Zwiers, H (1993). Developmental study of the expression of B50/GAP-43 in rat retina. *J Neurobiol* **24**: 949–958.
39 Capone, GT, Bendotti, C, Oster-Granite, ML and Coyle, JT (1991). Developmental expression of the gene encoding growth-associated protein 43 (Gap43) in the brains of normal and aneuploid mice. *J Neurosci Res* **29**: 449–460.
40 Badea, TC, Williams, J, Smallwood, P, Shi, M, Motajo, O and Nathans, J (2012). Combinatorial expression of Brn3 transcription factors in somatosensory neurons: genetic and morphologic analysis. *J Neurosci* **32**: 995–1007.
41 Berrebi, AS, Oberdick, J, Sangameswaran, L, Christakos, S, Morgan, JI and Mugnaini, E (1991). Cerebellar Purkinje cell markers are expressed in retinal bipolar neurons. *J Comp Neurol* **308**: 630–649.

42 Oberdick, J, Smeyne, RJ, Mann, JR, Zackson, S and Morgan, JI (1990). A promoter that drives transgene expression in cerebellar Purkinje and retinal bipolar neurons. *Science* **248**: 223–226.
43 Visel, A, Bristow, J and Pennacchio, LA (2007). Enhancer identification through comparative genomics. *Semin Cell Dev Biol* **18**: 140–152.
44 Kunzevitzky, NJ, Almeida, MV and Goldberg, JL (2010). Amacrine cell gene expression and survival signaling: differences from neighboring retinal ganglion cells. *Invest Ophthalmol Vis Sci* **51**: 3800–3812.
45 Alexiades, MR and Cepko, CL (1997). Subsets of retinal progenitors display temporally regulated and distinct biases in the fates of their progeny. *Development* **124**: 1119–1131.
46 Cepko, CL, Austin, CP, Yang, X, Alexiades, M and Ezzeddine, D (1996). Cell fate determination in the vertebrate retina. *Proc Natl Acad Sci USA* **93**: 589–595.
47 ENCODE Project Consortium, Bernstein, BE, Birney, E, Dunham, I, Green, ED, Gunter, C, Snyder, M. (2013). An integrated encyclopedia of DNA elements in the human genome. *Nature* **488**: 57–74.
48 The FANTOM Consortium (2005). The transcriptional landscape of the mammalian genome. *Science* **309**: 1559–1563.
49 Ravasi, T, Suzuki, H, Cannistraci, CV, Katayama, S, Bajic, VB, Tan, K *et al*. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* **140**: 744–752.
50 Plessy, C, Bertin, N, Takahashi, H, Simone, R, Salimullah, M, Lassmann, T *et al*. (2010). Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat Methods* **7**: 528–534.
51 Visel, A, Blow, MJ, Li, Z, Zhang, T, Akiyama, JA, Holt, A *et al*. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**: 854–858.
52 Rhee, HS and Pugh, BF (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**: 1408–1419.
53 Kheradpour, P, Ernst, J, Melnikov, A, Rogov, P, Wang, L, Zhang, X *et al*. (2013). Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. *Genome Res* **23**: 800–811.
54 Patwardhan, RP, Hiatt, JB, Witten, DM, Kim, MJ, Smith, RP, May, D *et al*. (2012). Massively parallel functional dissection of mammalian enhancers in vivo. *Nat Biotechnol* **30**: 265–270.
55 D'Souza, CA, Chopra, V, Varhol, R, Xie, YY, Bohacec, S, Zhao, Y *et al*. (2008). Identification of a set of genes showing regionally enriched expression in the mouse brain. *BMC Neurosci* **9**: 66.
56 Yang, GS, Banks, KG, Bonaguro, RJ, Wilson, G, Dreolini, L, de Leeuw, CN *et al*. (2009). Next generation tools for high-throughput promoter and expression analysis employing single-copy knock-ins at the Hprt1 locus. *Genomics* **93**: 196–204.
57 Schatz, O, Golenser, E and Ben-Arie, N (2005). Clearing and photography of whole mount X-gal stained mouse embryos. *Biotechniques* **39**: 650, 652, 654 passim.
58 Jacobson, SG, Acland, GM, Aguirre, GD, Aleman, TS, Schwartz, SB, Cideciyan, AV *et al*. (2006). Safety of recombinant adeno-associated virus type 2-RPE65 vector delivered by ocular subretinal injection. *Mol Ther* **13**: 1074–1084.

Supplementary Information accompanies this paper on the *Molecular Therapy—Methods & Clinical Development* website (http://www.nature.com/mtm)