



Multiplexed genome regulation in vivo with hyper-efficient Cas12a

Lucie Y. Guo^{1,2,5} , Jing Bian^{1,5}, Alexander E. Davis², Pingting Liu² , Hannah R. Kempton¹, Xiaowei Zhang¹ , Augustine Chemparathy¹, Baokun Gu¹, Xueqiu Lin¹, Draven A. Rane¹ , Xiaoshu Xu¹, Ryan M. Jamiolkowski¹, Yang Hu² , Sui Wang² and Lei S. Qi^{1,3,4}

Multiplexed modulation of endogenous genes is crucial for sophisticated gene therapy and cell engineering. CRISPR–Cas12a systems enable versatile multiple-genomic-loci targeting by processing numerous CRISPR RNAs (crRNAs) from a single transcript; however, their low efficiency has hindered in vivo applications. Through structure-guided protein engineering, we developed a hyper-efficient *Lachnospiraceae* bacterium Cas12a variant, termed hyperCas12a, with its catalytically dead version hyperdCas12a showing significantly enhanced efficacy for gene activation, particularly at low concentrations of crRNA. We demonstrate that hyperdCas12a has comparable off-target effects compared with the wild-type system and exhibits enhanced activity for gene editing and repression. Delivery of the hyperdCas12a activator and a single crRNA array simultaneously activating the endogenous *Oct4*, *Sox2* and *Klf4* genes in the retina of post-natal mice alters the differentiation of retinal progenitor cells. The hyperCas12a system offers a versatile in vivo tool for a broad range of gene-modulation and gene-therapy applications.

Clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR associated protein (Cas) nucleases and their nuclease-deactivated Cas (dCas) variants have revolutionized the field of genome editing and regulation^{1,2}. The discovery of alternative CRISPR systems beyond the widely used type II *Streptococcus pyogenes* Cas9 has expanded the toolkit for genetic manipulation^{3,4}. One set of nucleases of great interest is the type V-A Cas12a nucleases, which include *Acidaminococcus* Cas12a (AsCas12a) and *Lachnospiraceae* bacterium Cas12a (LbCas12a)^{3,5}. Unlike Cas9, which requires a separate RNase III protein for maturation of its guide RNAs, Cas12a possesses intrinsic RNase activity and can thus process multiple functional CRISPR RNAs (crRNAs) from a single long transcript^{6,7}. This unique characteristic of Cas12a enables facile multiplexed targeting, making it a powerful approach for versatile gene modulation with applications in cell reprogramming and combinatorial genetic screening⁸.

Despite these advantages, Cas12a has not been adopted in vivo as readily as its Cas9 counterpart, in part due to its less ideal editing and regulation efficiency, and more variable target-dependent indel efficiencies compared with Cas9 (refs. ^{4,9,10}). As the wild-type (WT) Cas12a exhibits a more restricted protospacer-adjacent-motif (PAM) requirements (TTTV) compared with Cas9 orthologues³ (for example, *S. pyogenes* Cas9, NGG)¹¹, efforts have been made to engineer Cas12a variants^{6,8,10,12–18} with expanded applications in combinatorial screening⁸ and ex vivo cell therapy¹⁴.

However, utility based on the unique feature of Cas12a for multiplexed epigenetic and transcriptional modulation has not been explored for in vivo applications. The inferior in vivo performance of Cas systems is often linked to the low copy numbers of Cas and crRNA molecules^{19,20}. We speculate that in vivo utilization of Cas12a requires optimization of its performance under restricting molecular concentrations. To address this key challenge, here we

applied synthetic biology to engineer Cas12a variants with a higher performance for diverse genome engineering applications and demonstrate its utility for in vivo multiplexed gene modulation in mammalian retinas.

Results

Development of hyperdCas12a for enhanced CRISPR activation.

We focused on LbdCas12a, given that previous studies showed that the WT LbdCas12a-fused transcription activator (LbdCas12a–VPR) achieved an approximately fivefold higher performance than the WT AsdCas12a–VPR for single-gene activation^{5,12}. To enhance the activity of the native LbdCas12a, we adopted structure-guided protein engineering. We focused on negatively charged (aspartate and glutamate) residues of LbdCas12a that reside within 10 Å of the target DNA (Protein Data Bank accession number 5XUS)²¹ and systematically mutated a subset of these residues to positively charged arginine (Fig. 1a). We hypothesized that this might increase the binding affinity of Cas12a to its negatively charged target DNA. Importantly, unlike previous studies that focused on residues adjacent to the PAM^{5,22} to increase the targeting range, we tested both proximal and distant residues to the PAM (Fig. 1a). To systematically characterize the library of variants for their performance, we engineered a HEK293T cell line with a genomically integrated TRE3G–green fluorescent protein (GFP) reporter²³ (Fig. 1b). Co-transfection of plasmids encoding the mutated dCas12a variants fused to a miniaturized VPR (miniVPR)²⁴ and the TRE3G promoter-targeting crRNA into reporter cells allowed for quantitative comparison of the performance of variants through flow cytometry (Extended Data Fig. 1a–d).

Although many single mutations exhibited decreased dCas12a-mediated gene activation, a few mutants (D122R, E125R, D156R, E159R, D235R, E257R, E292R, D350R, E894R, D952R

¹Department of Bioengineering, Stanford University, Stanford, CA, USA. ²Department of Ophthalmology, Stanford University, Stanford, CA, USA.

³Department of Chemical and Systems Biology, Stanford University, Stanford, CA, USA. ⁴Stanford ChEM-H Institute, Stanford University, Stanford, CA, USA. ⁵These authors contributed equally: Lucie Y. Guo, Jing Bian. ✉e-mail: lucieguo@stanford.edu; suiwang@stanford.edu; stanley.qi@stanford.edu

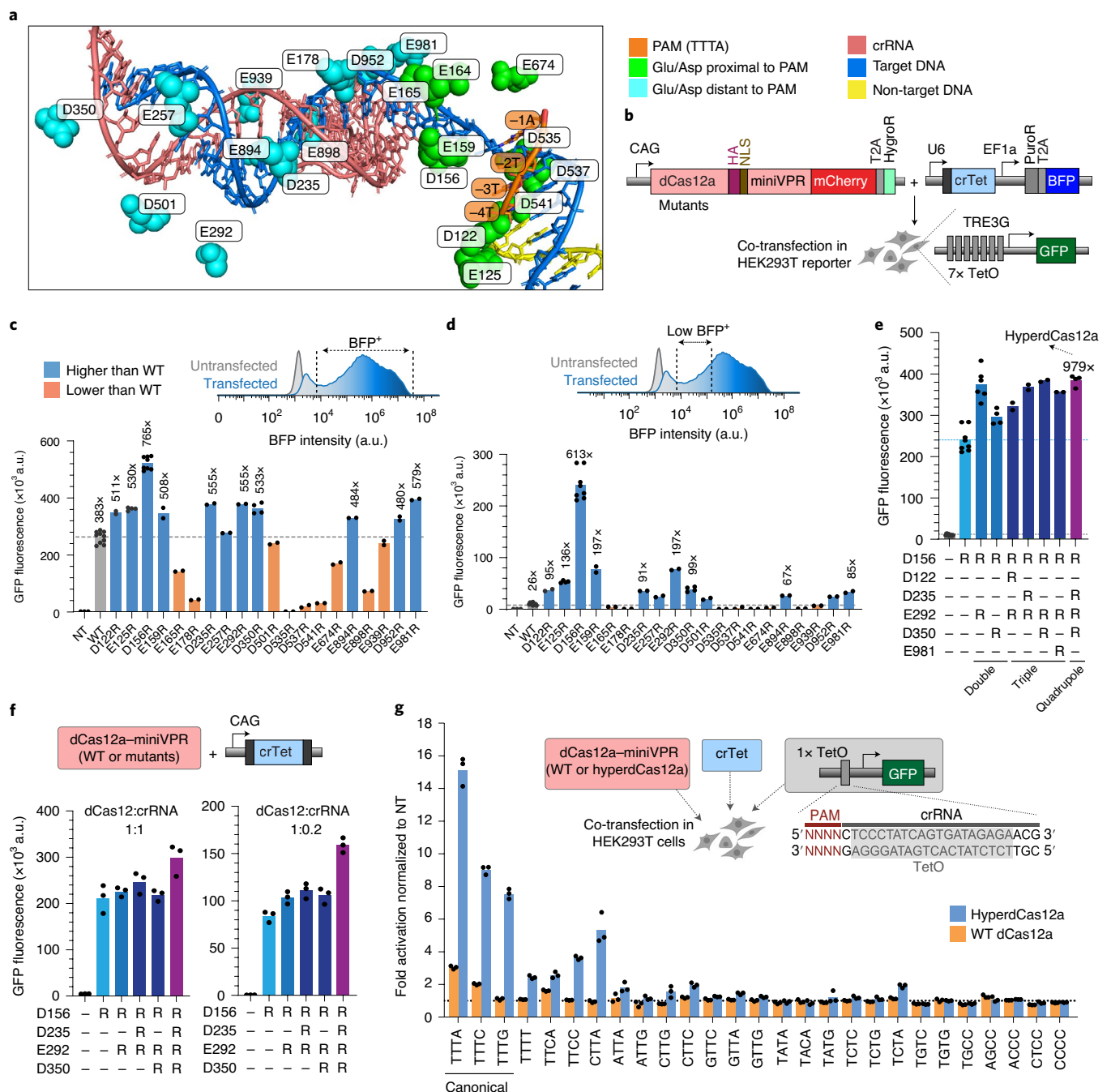


Fig. 1 | Development of combinatorial dCas12a mutants with superior activity under conditions of low crRNA levels. **a**, Structure of LbCas12a (Protein Data Bank accession number 5XUS) highlighting glutamate and aspartate residues within 10 Å of the target DNA. **b**, Constructs used for co-transfection to test CRISPR activation. HEK293T cells stably expressing GFP driven by the inducible TRE3G promoter were transfected with Tet crRNA (crTet), driven by a U6 promoter, as well as dCas12a mutants and analysed through flow cytometry 2 d after transfection. **c**, Representative flow cytometry histogram of the BFP intensity in untransfected and transfected cells (top). Level of GFP fluorescence in reporter cells for WT dCas12a and its mutants (bottom). **d**, Representative flow cytometry histogram of BFP intensity comparing untransfected and transfected cells, showing a subset of cells expressing low levels of BFP (top). Level of GFP fluorescence in reporter cells for WT dCas12a and its mutants (bottom). **e**, GFP fluorescence in the low-BFP cells transfected with WT dCas12a, single mutants and combinatorial variants with several of the most-potent single mutations from **c**. The quadruple mutant (D156R + D235R + E292R + E350R) is referred to as hyperdCas12a. **c–e**, Fold changes in GFP fluorescence were calculated relative to non-targeting (NT) crLacZ and shown above each bar. Grey dotted lines represent the levels of the WT, and the blue dotted line (**e**) represents the level of the single D156R mutant. Each data point represents the mean GFP intensity of an independent experiment, with each bar representing the average of two or more independent experiments. **f**, Level of GFP fluorescence of WT dCas12a and the dCas12a mutants following co-transfection into HEK293T reporter cells at a crRNA/dCas12a ratio of 1:1 (bottom left) and 0.2:1 (bottom right) as per the schematic (top). **g**, Parental HEK293T cells were co-transfected with hyperdCas12a or WT dCas12a and crTet as well as a third plasmid containing a truncated TRE3G promoter with a single TetO element preceded by 27 various PAMs as per the schematic (inset). The cells were gated for mCherry and low BFP expression. Fold changes were calculated relative to non-targeting crLacZ (bottom). The dotted line represents the level of the non-targeting crRNA. **f,g**, Each data point represents the mean GFP intensity of an independent experiment, with each bar showing the average of three independent experiments. a.u., arbitrary units.

and E981R) exhibited enhanced activation compared with the WT dCas12a (Fig. 1c and Extended Data Fig. 1e,f). Whereas WT dCas12a produced a fold change of approximately 383× over the non-targeting control, the single D156R mutation enabled >700-fold activation and several other single mutations enabled 400–600-fold activation. We next examined the performance of these variants after gating for a cell population expressing blue fluorescent protein (BFP) at low levels ('low BFP'), which serves as a proxy for cells with lower crRNA concentrations, which is particularly relevant for in vivo applications where there would be less of the crRNA–Cas12a complex compared with in vitro settings (Fig. 1d). Among the low-BFP cell population, WT dCas12a exhibited a decrease in activity, only enabling a 26-fold activation of GFP over the non-targeting control. Notably, several mutants performed substantially better than the WT protein in this condition: the single D156R mutation enabled >600-fold activation and several others enabled 90–200-fold activation (Fig. 1d).

We next chose a few of the best enhancing single mutations (D122R, D156R, D235R, E292R, D350R and E981R), combined them into double, triple and quadruple mutants, and tested their activation in low-BFP cells using the GFP reporter assay (as per Fig. 1d). We observed further enhancement, in comparison with the best single mutation, using several combinatorial mutations, including a quadruple mutant harbouring four mutations (D156R, D235R, E292R and D350R; Fig. 1e and Extended Data Fig. 1g). With an optimized nuclear localization signal (2×Myc NLS; Extended Data Fig. 2a–d), the quadruple mutant still outperformed the WT protein, especially in the low-crRNA population (Extended Data Fig. 2e–g).

Note that D156 of LbCas12a is homologous to E174 of AsCas12a (Extended Data Fig. 2h,i), which is one of the mutated residues in the previously published enhanced AsCas12a (enAsCas12a)⁵, whereas the other three mutations have not been reported to date. We also tested combinations of homologous mutations based on enAsdCas12a (E174R, S542R and K548R)⁵. Interestingly, these combinations did not increase activation over the single LbdCas12a (D156R; Extended Data Fig. 2h–j), which was outperformed by the quadruple mutant (Fig. 1e).

Using dCas12a for multiplex genome-regulation applications would require that the protein maintains its RNase ability to process a functional crRNA from a longer poly-crRNA transcript. To easily test this using the same GFP reporter system, we compared the performance of the dCas12a mutants with the WT protein

using crRNA expressed by an RNA polymerase II promoter (CAG promoter in this case) so that dCas12a would be required to process the crRNA before activation of the target gene. GFP activation using WT dCas12a was reduced using the CAG promoter-driven crRNA compared with the U6 promoter-driven crRNA (compare the GFP fluorescence of the WT in Fig. 1c and 1f), but the single and combinatorial mutants greatly enhanced the level of activation. Notably, the quadruple mutant (D156R, D235R, E292R and D350R) achieved the highest level of activation—approximately 60-fold above the level achieved by the WT protein (Fig. 1f, left). We then tested the mutants under restricted crRNA conditions (crRNA/dCas12a ratio of 0.2:1). The quadruple mutant outperformed all other mutants, with >300-fold activation above the level achieved by the WT protein (Fig. 1f, right). In addition, it outperformed enAsdCas12a at low crRNA concentrations for both the CAG-driven crRNA and dual crRNA array (Extended Data Fig. 2k–m). We henceforth refer to this quadruple mutant as 'hyper-efficient dCas12a' (hyperdCas12a) for further characterization and in vivo gene targeting.

Although our mutagenesis focused on increasing efficiency (instead of broadening the targeting range as in previous studies^{5,22}), we tested the PAM preferences of this mutant specifically for gene activation. We used a truncated TRE3G promoter containing a single TetO preceded by a PAM and found that hyperdCas12a outperformed WT dCas12a for all three canonical PAMs (TTTA, TTTC and TTG) as well as several of the non-canonical PAMs (TTTT, CTTA, TTCA and TTCC; Fig. 1g). Given that of the four mutated residues of hyperdCas12a, only the D156R mutation is proximal to the PAM motif (Fig. 1a), it is logical that several of these PAMs are also accessible by the homologous E174R single mutant of AsdCas12a⁵ and that the PAM range of hyperdCas12a may be stricter than that of enAsdCas12a⁵.

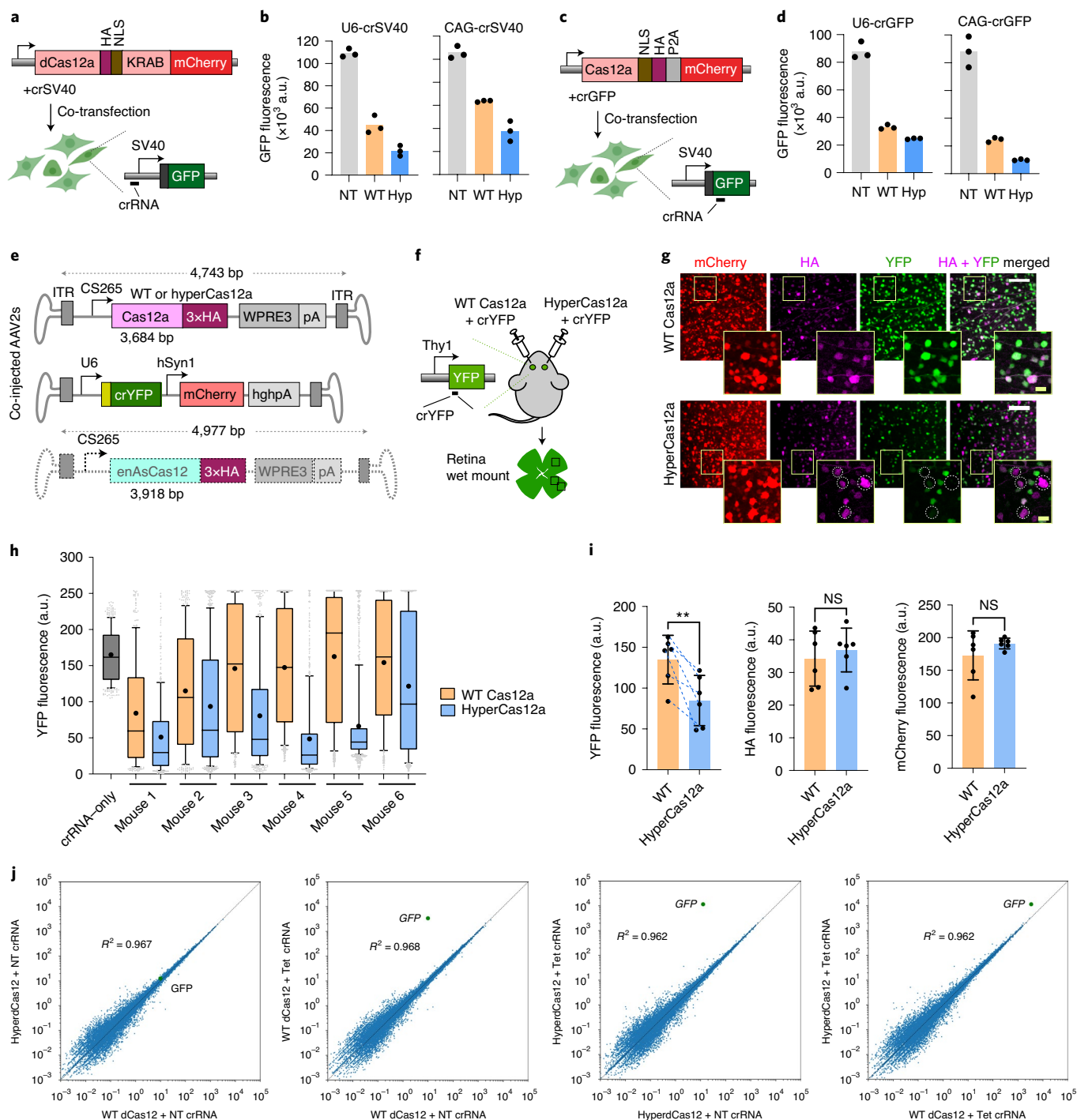
HyperdCas12a improves gene repression and in vivo gene editing. We next tested whether hyperdCas12a also improved other Cas12-based applications. Using a HEK293T reporter cell line with a genomically integrated constitutive simian vacuolating virus 40 (SV40)–GFP cassette¹, fusion of hyperdCas12a to a KRAB repressor domain with crRNA targeting the SV40 promoter enabled improved repression over the non-targeting control compared with its WT equivalent (Fig. 2a,b). This suggests that hyperdCas12a can be modularly coupled to effectors and exhibit enhanced effects for transcriptional and epigenetic modulation.

Fig. 2 | HyperCas12a outperforms WT dCas12 in CRISPR repression and in vivo gene editing. **a**, Constructs to test CRISPR repression. HEK293T cells stably expressing SV40–GFP were co-transfected with WT dCas12a or hyperdCas12a fused to the transcriptional repressor KRAB as well as crRNA (driven by either U6 or CAG promoters) targeting the SV40 promoter and analysed 5 d after transfection. **b**, Levels of GFP fluorescence in the assay described in **a**. **c**, Constructs to test gene editing. HEK293T cells stably expressing SV40–GFP were transfected with nuclease-active WT Cas12a or hyperCas12a as well as crGFP (driven by either U6 or CAG promoters) targeting a coding region of GFP and analysed 5 d after transfection. **d**, Levels of GFP fluorescence in the gene editing assay described in **c**. **b,d**, Each data point represents the mean GFP intensity of an experiment, with each bar showing the average of three independent experiments. NT, non-targeting crRNA; WT, WT dCas12a; and hyp, hyperdCas12a. **e**, AAV constructs for in vivo gene editing. **f**, Schematic of the intravitreal injection experiment, where AAV-hyperCas12a + AAV-crYFP was delivered into one eye and AAV-WT Cas12a + AAV-crYFP was delivered to the other eye. **g**, Immunohistochemistry images of retinal wet mounts. The dotted circles indicate mCherry⁺HA⁺ cells without YFP expression. Scale bars, 100 μm (main images) and 20 μm (insets, magnified views of the regions in the yellow squares). **h**, Levels of YFP fluorescence in mCherry⁺ cells for each mouse, determined by automated segmentation analyses. Data for all six mice—that is, six independent biological replicates—are displayed; $n = 250$ –800 cells were analysed for each mouse (the exact n values are provided in the source data). As additional control, one mouse was injected with AAV-crYFP only, without Cas12a (grey). In the box-and-whisker plots, the boxes show the 25th to 75th percentiles (with the bar indicating the median and the dot indicating the mean) and the whiskers extend to the 10th and 90th percentiles, with individual data points shown for the lowest and highest 10% of each dataset. **i**, Mean YFP (left), HA (middle) and mCherry (right) fluorescence values for each mouse, measured using automated segmentation analysis. The mean $\pm$ s.d. as well as individual data points are shown for $n = 6$ independent animals. The P values were calculated using a paired two-tailed Student's t -test; ** $P = 0.0078$ and NS, not significant. For the YFP graph (left), blue dotted lines were drawn to connect values for each mouse to facilitate comparison of this paired dataset. **j**, HEK293T cells stably expressing TRE3G–GFP (as per Fig. 1b) were co-transfected with plasmids with dCas12a–miniVPR (WT or hyper) and crRNA, and collected for genome-scale RNA-seq 2 d after transfection. The GFP gene is labelled in green and the coefficient of multiple correlation (R^2) is provided. The plots represent representative results from two independent RNA-seq experiments (for reproducibility plots, refer to Extended Data Fig. 4b). a.u., arbitrary units.

We further introduced the four activity-enhancing mutations into the nuclease-active form of Cas12a and observed that the resulting hyperCas12a endonuclease enabled more effective gene knockout (Fig. 2c,d). Analysis of the indel formation frequency at each nucleotide position showed that editing by hyperdCas12a peaked around positions 18–23 base pairs (bp) relative to the PAM, similar to the WT protein (Extended Data Fig. 3).

To test gene editing in vivo, we packaged hyperCas12a in an adenovirus-associated virus (AAV) serotype 2 with a retinal ganglion cell-specific promoter further miniaturized from a previous study²⁵ (265 bp), a truncated Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) (245 bp)²⁶ and a small

synthetic poly-A tail (49 bp; Fig. 2e). We co-delivered AAV–hyper-Cas12a by intravitreal injection along with AAV–crRNA (targeting yellow fluorescent protein, YFP) into one eye of transgenic mice expressing Thy1–YFP²⁷, and we injected its WT counterpart in the contralateral eye as a side-by-side control (Fig. 2f). For all mice tested, hyperCas12a showed improved YFP knockout compared with WT Cas12a (Fig. 2g–i). Despite using minimal versions of all regulatory elements, the AAV containing hyperdCas12a (4,743 bp) teetered on the AAV packaging limit (approximately 4.7 kb); by being 234 bp larger, enAsdCas12a exceeded this limit (Fig. 2e). This highlights the utility of hyperCas12a for enhanced AAV-based in vivo gene editing.



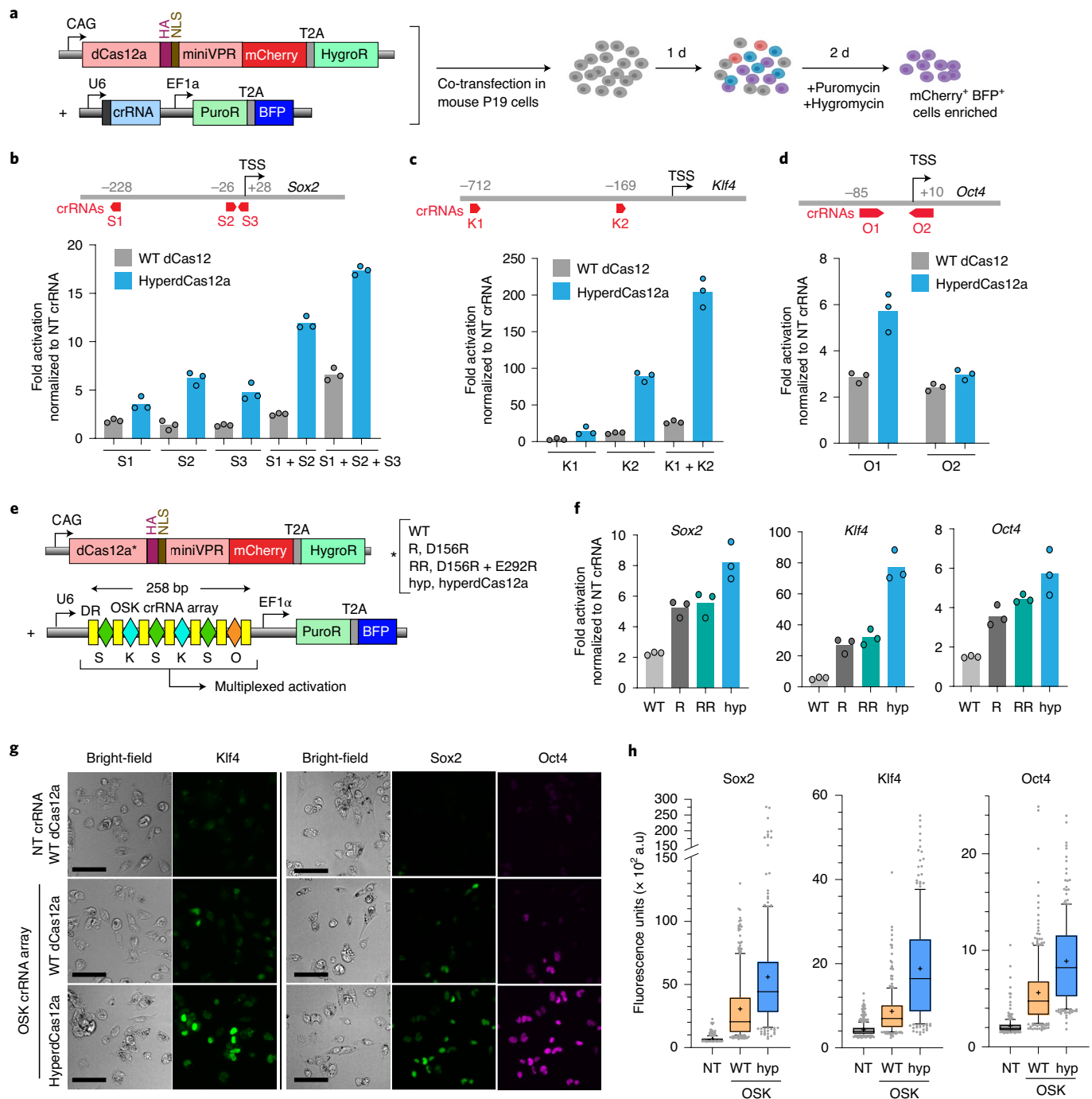


Fig. 3 | HyperdCas12a enables multiplex activation of endogenous genes. **a**, Mouse P19 cells were co-transfected with the indicated plasmids (left) and then selected with puromycin and hygromycin 24 h after transfection to enrich for mCherry⁺BFP⁺ cells (for flow cytometry data, refer to Extended Data Fig. 5). The cells were collected for analysis 72 h after transfection. **b–d**, Schematics (top) of crRNAs targeting the promoters of *Sox2* (**b**), *Klf4* (**c**) and *Oct4* (**d**) as well as the levels of transcriptional activation of each target gene, determined using quantitative PCR with reverse transcription (RT-qPCR; bottom), by WT dCas12 and hyperdCas12a relative to non-targeting (NT) crRNA. TSS, transcriptional start site. The genomic positions of the first T in PAM (relative to the TSS, which is at '0') are shown for each crRNA targeting the corresponding promoter. **e**, Constructs used for multiplex activation, with a poly-crRNA array consisting of six crRNAs targeting *Oct4*, *Sox2*, *Klf4* (OSK) separated by direct repeats (DR). **f**, Comparison of multiplex transcriptional activation levels of each target gene in the different dCas12a variants relative to non-targeting crRNA (determined by RT-qPCR). **g**, Immunostaining using antibodies targeting endogenous *Sox2*, *Klf4* and *Oct4*. Representative images of one experiment from **e,f**, with multiple fields of view, are shown. Scale bars, 100 μ m. **h**, Levels of *Sox2* (left), *Klf4* (middle) and *Oct4* (right) expression in mCherry⁺ cells from the immunofluorescence images in **g**. For each condition, 200–400 cells were quantified. In the box-and-whisker plots, the boxes show the 25th to 75th percentiles (with the bar indicating the median and the cross indicating the mean) and the whiskers extend to the 10th and 90th percentiles, with individual data points shown for the lowest and highest 10% of each dataset; a.u., arbitrary units. **b,c,f**, The bars represent the mean values of three independent biological replicates.

CRISPR activation by hyperdCas12a is specific. Cas12a is known to be highly precise in human cells²⁸, with a specificity that is likely to be higher in comparison to Cas9 (refs. ^{9,12}). To evaluate the specificity of CRISPR activation by hyperdCas12a compared with WT Cas12a on a genome-wide scale, we carried out whole-transcriptome RNA sequencing (RNA-seq) of HEK293T cells with the TRE3G–GFP reporter (Fig. 1b) transfected with either WT dCas12a or hyperdCas12a combined with the TRE3G-targeting crRNA. We also included a non-targeting crRNA as a negative control for each case (Fig. 2j). As expected, with the targeting crRNA, the GFP transcript exhibited an increase in abundance, consistent with flow cytometry data showing stronger transcriptional activation by hyperdCas12a compared with the WT dCas12a in Fig. 1c (Fig. 2j). In a comparison of targeting versus non-targeting crRNAs, both WT dCas12a and hyperdCas12a showed similar specificity (Fig. 2j and Supplementary Table 4). Two biological replicates were analysed separately and showed similar results (Extended Data Fig. 4a,b).

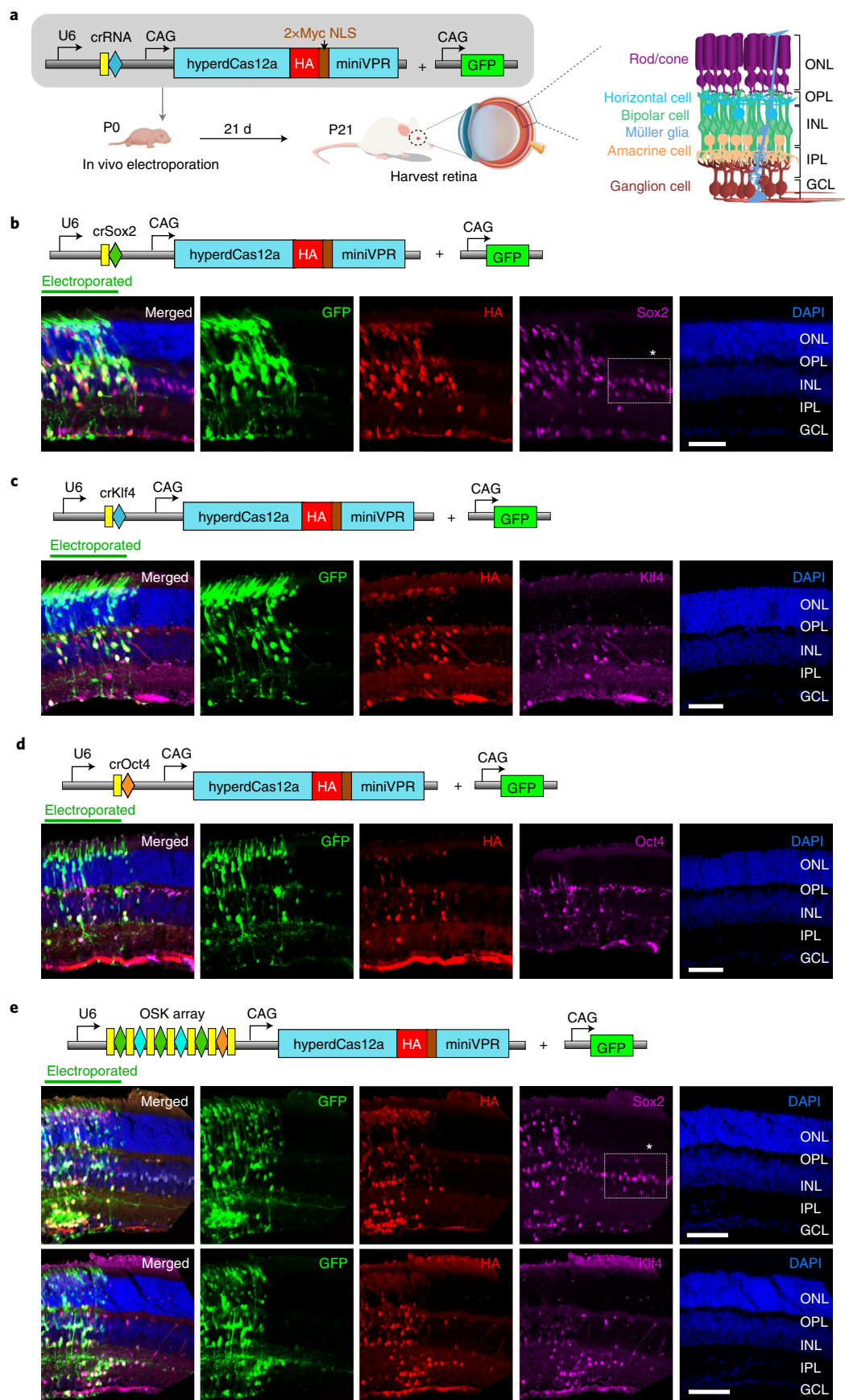
HyperdCas12a effectively activates endogenous genes. We moved beyond the HEK293T GFP reporter cell lines to evaluate the ability of hyperdCas12a to activate endogenous genes. We used P19 cells derived from embryonic teratocarcinoma in mice, in which we achieved a moderate (approximately 21%) co-transfection efficiency of two plasmids (Extended Data Fig. 5a–c). To enrich for transfected cells, we used a dual-selection approach by treating the cells with both puromycin and hygromycin for 48 h (Fig. 3a), which resulted in approximately 90% of the cells containing both the crRNA and dCas12a plasmids to allow for facile comparisons between different crRNAs and dCas12a variants (Extended Data Fig. 5a–c).

We tested crRNAs targeting promoters of the endogenous transcription factor genes *Oct4* (also known as *Pou5f1*), *Sox2* and *Klf4* due to their known synergistic regenerative role in multiple contexts^{29,30}. We designed Cas12a crRNAs targeting the promoter of each gene (Extended Data Figs. 6–8), encompassing regions previously targeted by dCas9–SunTag–VP64 in mouse embryonic stem cells³¹. We performed immunostaining to visualize the expression of the target proteins in cells and identified crRNAs that enabled effective transcriptional activation of *Sox2* (Extended Data Fig. 6), *Klf4* (Extended Data Fig. 7) and *Oct4* (Extended Data Fig. 8). Furthermore, for *Sox2* and *Klf4*, we achieved synergistic activation using paired crRNAs (despite the target sequences for the *Klf4* crRNAs being >500 bp apart) and further synergy in *Sox2* activation using a ‘triplet’ of three separate *Sox2* crRNAs (Extended Data Figs. 6,7). In contrast, we did not achieve synergy with paired crRNAs for *Oct4*, possibly due to higher basal levels of *Oct4* expression in P19 cells (Extended Data Fig. 8). Using a subset of the validated crRNAs, we compared the level of endogenous gene activation by WT dCas12a with hyperdCas12a. All of the tested crRNAs, including paired and triplet crRNAs, exhibited enhanced activation using hyperdCas12a compared with WT dCas12a (Fig. 3b–d).

HyperdCas12a drives enhanced multiplex activation. Given that *Oct4*, *Sox2* and *Klf4* are known to work synergistically, there is a strong rationale for their multiplexed activation³². To simultaneously activate these three targets, we generated a single crRNA array driven by the U6 promoter encoding six crRNAs (Fig. 3e) based on our crRNA screening in P19 cells. The crRNA combination that achieved the highest activation in P19 cells was chosen for each target: three crRNAs for *Sox2* (S1 + S2 + S3), two for *Klf4* (K1 + K2) and one for *Oct4* (O1; Extended Data Figs. 6–8). With dCas12a (D156R) and a double mutant (D156R + E292R) we achieved enhanced activation compared with WT dCas12a and further enhancement with hyperdCas12a, which achieved increased activation of approximately five-, eight- and 70-fold for *Oct4*, *Sox2* and *Klf4*, respectively (Fig. 3f–h), and also outperformed enAsdCas12a (Extended Data Fig. 9). Interestingly, although previous studies using WT dCas12a demonstrated decreased expression of crRNAs at and beyond the fourth position in the crRNA array^{4,12,33}, with hyperdCas12a we observed compelling *Oct4* activation in P19 cells despite its location as the sixth crRNA. We observed decreased activation of each target gene compared with the level achieved by single crRNAs (compare Fig. 3f with Fig. 3b–d), consistent with previous observations^{12,34}, possibly due to decreased copies of the longer pre-crRNA array expressed by the U6 promoter compared with shorter individual crRNAs. Nevertheless, hyperdCas12a performed robustly when a single CRISPR array was used to activate multiple endogenous targets. In addition, the enhanced performance of hyperdCas12a over the single D156R mutant and the double D156R + E292R mutant in this assay highlights the additive power of the combinatorial mutations and points to hyperdCas12a as a logical protein of choice for multiplex genome engineering in mammalian cells.

Multiplexed gene activation by hyperdCas12a in mouse retina. We targeted the retina for in vivo applications given the high interest in using genome engineering for eye diseases, its relative immune privilege and accessibility, and the global burden of degenerative retinal diseases. We used the well-validated in vivo electroporation technique^{35–37}, which has advantages over other methods of gene transfer, such as more lenient size limitation of the transgene. Transgenes persist up to a few months in retina cells in vivo³⁵. We constructed single plasmids consisting of hyperdCas12a with an optimized nuclear-targeting sequence (NLS; Extended Data Fig. 2a–d) and single crRNAs to *Sox2*, *Klf4* or *Oct4*, and validated its ability to drive in vivo CRISPR activation of single gene targets 21 d after in vivo electroporation at post-natal day 0 (P0; Fig. 4a–d and Extended Data Fig. 10a–d); non-targeting crRNA did not show substantial activation (Extended Data Fig. 10e). The CAG–GFP plasmid was co-electroporated to serve as the electroporation efficiency control. We observed numerous haemagglutinin (HA)⁺ cells within the electroporated GFP⁺ patches in the retina, indicating successful delivery and expression of hyperdCas12a (Fig. 4).

Fig. 4 | In vivo gene activation by hyperdCas12a with single crRNA and poly-crRNA arrays. **a**, Constructs and experimental schematic of in vivo plasmid electroporation in the retinas of post-natal mice. CAG–GFP was used to mark the electroporated patch. Wild-type CD-1 pups were electroporated on the day of birth (P0) and euthanized after 21 d (P21) to assess the retinal histology. **b–d**, Representative retinal slices of mouse retinas electroporated with a plasmid containing a single crRNA and hyperdCas12a to activate the expression of the endogenous genes *Sox2* (**b**), *Klf4* (**c**) and *Oct4* (**d**; bottom). Note that the GFP signals mark the boundary of the electroporated patch; thus, the area that did not receive electroporated plasmids serves as an internal control that aids the interpretation of the specificity of immunostaining. Cells that received the plasmid with hyperdCas12a and crRNA express HA. Immunostaining was performed with antibodies to *Sox2* (**b**), *Klf4* (**c**) or *Oct4* (**d**) to identify cells that achieved CRISPR endogenous gene activation by single crRNA. Schematics of the plasmids used are provided (top). **e**, Representative retinal slices of mouse retinas electroporated with a single plasmid containing a poly-crRNA array and hyperdCas12a driving activation of *Sox2* (middle) and *Klf4* expression (bottom). Basal expression of *Sox2* in the inner nuclear layer outside the electroporation boundary can be seen in the box labelled with an asterisk. Images are representative slides of $n = 4$ or 5 independent biological replicates (and are quantitated in Fig. 5j,k). A schematic of the plasmid used is provided (top). **b–d**, Images are representative slices from $n = 3$ independent biological replicates (and are quantitated in Extended Data Fig. 10d). INL, inner nuclear layer; ONL, outer nuclear layer; OPL, outer plexiform layer; and DAPI, 4,6-diamidino-2-phenylindole. Scale bars, 50 μ m.



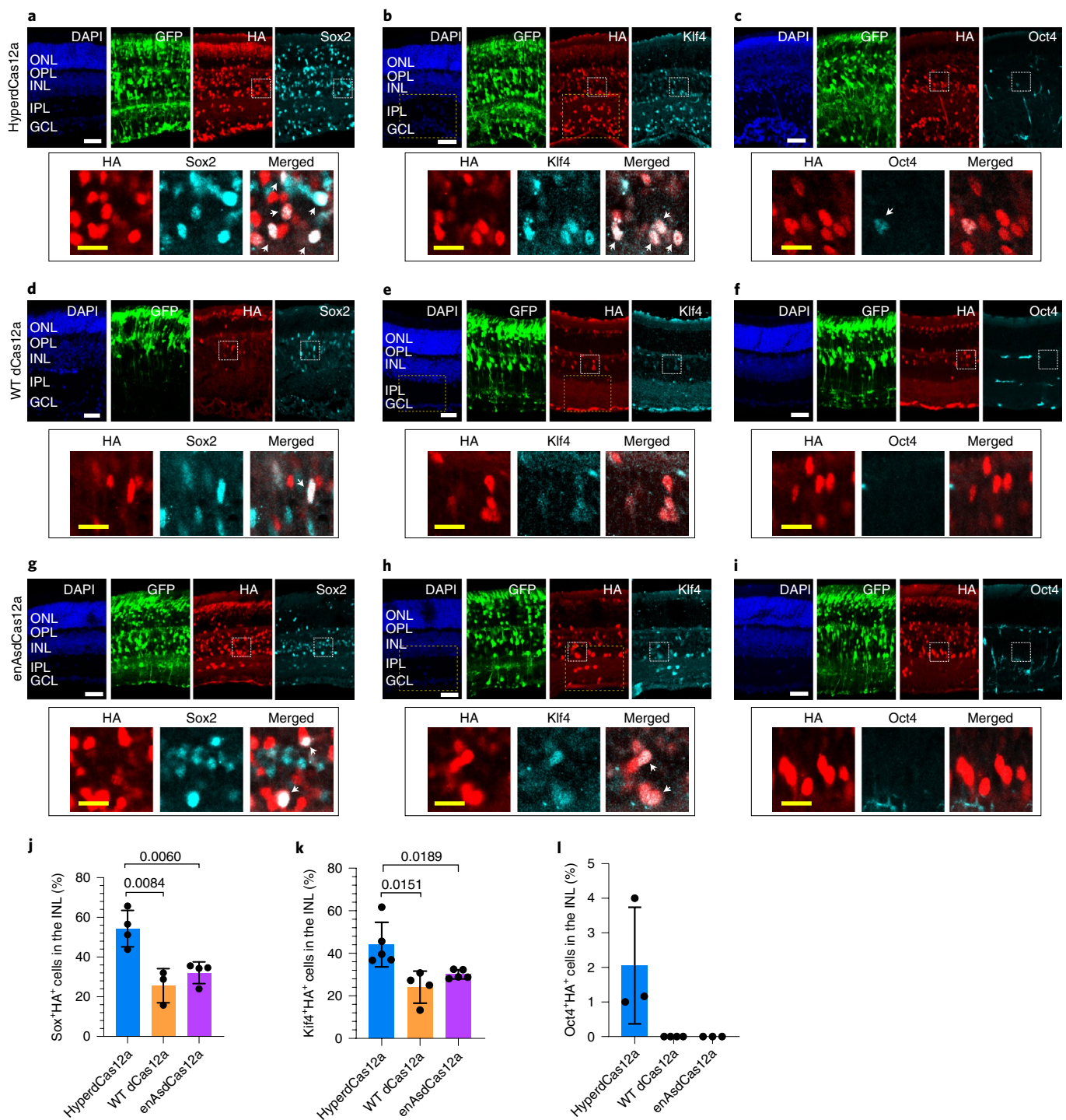


Fig. 5 | In vivo multiplex gene activation by hyperdCas12a compared with dCas12a alternatives. a–i, Representative retinal slices after in vivo electroporation with crRNA array and hyperdCas12a (**a–c**), WT LbdCas12a (**d–f**) or enAsdCas12a (**g–i**) to activate endogenous *Sox2*, *Klf4* and *Oct4* expression (top). Magnified views of the regions in the white boxes highlighting HA⁺ cells in the inner nuclear layer (bottom). Scale bars, 50 μ m (white) and 20 μ m (yellow). Arrows highlight HA⁺ cells that strongly express *Sox2*, *Oct4* or *Klf4*. **b,e,h,** Magnified views of the regions in the yellow boxes are shown in Fig. 6a. **j–l,** Quantitative comparison of the percentage of Sox2⁺ (**j**), Klf4⁺ (**k**) and Oct4⁺ (**l**) cells in the HA⁺ cell population in the INL layer of mouse retinas electroporated with plasmids containing the crRNA array and hyperdCas12a, WT dCas12a or enAsdCas12a. The mean $\pm$ s.d. and individual data points are shown for $n = 3$ –5 independent biological replicates. **j,k,** P values were calculated using an unpaired two-tailed Student's t -test and are indicated on the graphs. ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; and GCL, ganglion cell layer.

Individual overexpression of *Sox2*, *Oct4* and *Klf4* has been shown to direct the differentiation of retinal progenitor cells (RPCs)^{38–40} towards specific fates, and the synergistic co-activation of these ectopic transcription factors can induce the formation of induced

pluripotent stem cells in vitro^{31,41} and rejuvenate mature retinal ganglion cells in vivo³⁰. We tested whether our hyperdCas12a system can synergistically activate endogenous *Sox2*, *Klf4* and *Oct4* in post-natal RPCs in vivo and whether this manipulation affects the

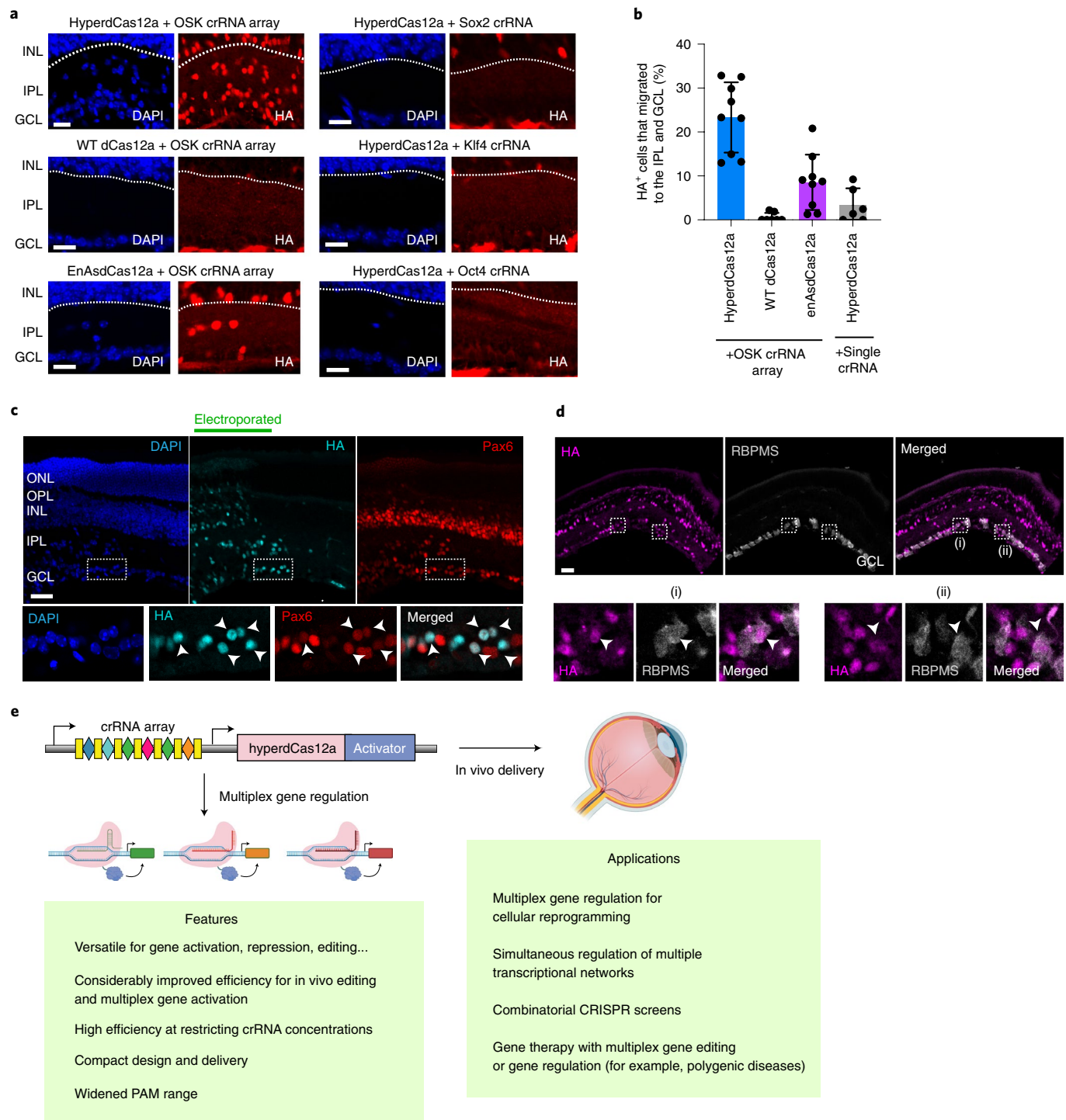


Fig. 6 | Multiplex gene activation by hyperdCas12a induces the migration of retina progenitor cells and altered differentiation in vivo. **a**, Representative images showing the presence of HA⁺ cells in the IPL and GCL layers, induced by hyperdCas12a or the dCas12 alternatives, for single crRNA and with the poly-crRNA array. Scale bars, 20 μ m. White lines illustrate the boundary between inner nuclear layer (INL) and inner plexiform layer (IPL). **b**, Quantitation of the experiment in **a** for $n=6-9$ independent biological replicates. The individual data points with the mean $\pm$ s.d. are shown. **c**, HyperdCas12a-mediated activation of endogenous *Oct4*, *Sox2* and *Klf4* in PO RPCs induced the formation of Pax6⁺ cells. Insets: magnified views of the region in the white boxes showing co-localization of Pax6, HA and DAPI staining. **d**, HyperdCas12a activation of endogenous *Oct4*, *Sox2* and *Klf4* induces the formation of RBPMS⁺ cells. Insets: magnified views of the regions in the two white boxes in the main image on the right. **c,d**, Scale bars, 50 μ m. **e**, Illustration of features and potential applications of hyperdCas12a. Created with [Biorender.com](https://www.biorender.com). ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; and GCL, ganglion cell layer.

differentiation capacity of RPCs. We delivered a single plasmid of hyperdCas12a with a poly-crRNA array targeting *Sox2*, *Klf4* and *Oct4*, and observed strong expression of *Sox2* (Figs. 4e and 5a) and *Klf4* (Figs. 4e and 5b), and mild activation of *Oct4* in HA⁺ cells (Fig. 5c). The level of in vivo activation of all three gene targets was stronger with hyperdCas12a than with WT dCas12a (Fig. 5d–f,j–l) and enAsdCas12a (Fig. 5g–i), which is consistent with the in vitro results (Extended Data Fig. 9).

Multiplex gene regulation alters cell differentiation. We examined the fates of HA⁺ cells that had received the hyperdCas12a and poly-crRNA array plasmid. The in vivo electroporation technique delivers DNA mainly to mitotic cells and at P0, mitotic RPCs give rise to rod photoreceptors, Müller glia as well as bipolar and amacrine neurons⁴², which migrate to and reside in the outer or inner nuclear layers but not in the ganglion cell layer (GCL). We noted that activation by hyperdCas12a–miniVPR with our poly-crRNA array resulted in a strong enriched population of HA⁺ cells in the GCL and inner plexiform layer (IPL). This effect was less pronounced with enAsdCas12a and almost absent with WT dCas12a (Fig. 6a,b). Furthermore, the enrichment of HA⁺ cells in the GCL and IPL was absent when hyperdCas12a was used to individually activate *Sox2*, *Klf4* or *Oct4*, suggesting the importance of simultaneous multiplexed gene activation for driving phenotypic changes (Fig. 6a,b and Extended Data Fig. 10a–d).

We observed the expression of Pax6 (marker for displaced amacrine and ganglion cells in the GCL) in most of the HA⁺ cells that migrated into the GCL (Fig. 6c). A minority of GCL HA⁺ cells expressed RBPMS (Fig. 6d). These data suggest that activation of endogenous *Sox2* and *Klf4* (and weakly, *Oct4*) can direct P0 RPCs to differentiate into displaced amacrine- and ganglion-like cells, drive their migration into the GCL and IPL layers and support the conclusion that hyperdCas12a can activate multiple endogenous genes to alter phenotypes in vivo.

Discussion

In this work we developed an optimized LbCas12a variant, which we termed hyperdCas12a, for in vivo CRISPR multiplexed genome modulation. HyperdCas12a enabled the simultaneous activation of endogenous targets in post-natal retina and altered the differentiation of RPCs. HyperdCas12a outperformed WT dCas12a, particularly under restricting concentrations of crRNAs and dCas12a, a condition that is especially relevant for in vivo applications. Despite its enhanced activity, hyperdCas12a showed comparable specificity to the WT dCas12a protein. HyperdCas12a also improved CRISPR-mediated gene repression and its nuclease-active version (hyperCas12a) achieved more effective gene editing in vivo in retinal ganglion cells. In addition, hyperdCas12a-mediated endogenous gene activation outperformed WT dCas12a and enAsdCas12a in vivo.

The enhanced dCas12a system is broadly useful for both in vitro and in vivo applications that require multiple genetic manipulations, which is limited with the existing tools. Our system enables simultaneous modulation at multiple genomic loci, thus paving the way for CRISPR-based regulation of multiple pathways or synergistic targets, as is required in the case of non-monogenic diseases, which represent a large proportion of human diseases. Moreover, the system could be useful as a platform for regenerative biology and therapy. There is high interest in the direct reprogramming of lineage-determined cells from one fate to another as a therapeutic strategy to compensate for the loss of certain cell populations.

The exact level of activation required for the desired phenotypic changes would depend on the context. For example, CRISPRa has been useful as a tool to rescue haplo-insufficiency by rescuing the expression of the endogenous functional allele, in which an increase of only twofold is sufficient to reverse the disease phenotype⁴³.

In other instances, especially for directing cell differentiation, higher levels of activation may be needed to induce a stronger phenotype—we previously showed that CRISPR activation of endogenous *Oct4* or *Sox2* to >15-fold in mouse embryonic fibroblasts was sufficient for reprogramming to pluripotency³¹. Our hyperCas12a system can overcome the technical barrier of simultaneous activation of multiple transcription factors and may facilitate in vivo reprogramming applications via the improved multigene modulation efficiency (Fig. 6g). Although the Cas12a system has other intrinsic limitations (for example, uneven array processing^{6,44}) that are not fully addressed by our present advances in protein engineering, we anticipate that this enhanced protein would be compatible with published crRNA array optimizations^{6,10,18,45} to achieve even greater potential of the Cas12a system.

CRISPR-mediated transcriptional activation of endogenous genes at their native chromatin loci would minimize the risks of supraphysiologic expression associated with forced overexpression of transgenes, which may be particularly important for multiplexed gene regulation to balance their relative expression in single cells. In addition, targeted endogenous gene regulation may also remodel chromatin status and epigenetic modifications at endogenous loci, and lead to persistent gene regulation effects compared with overexpression of corresponding complementary DNA^{46–48}. Furthermore, cDNA-based approaches require a priori decisions about which isoform of the gene to express, whereas CRISPR activation can target regulatory elements such as a promoter to potentially activate all isoforms. CRISPR activators can also be useful in cases of large cDNAs^{49,50} (or in the future, for large intergenic noncoding RNAs) that may not fit due to the limited cargo capacity of viral delivery vehicles. Our system enables the simultaneously manipulation of the endogenous expression of combinations of fate-determining transcription factors, which will open new avenues for in vivo genetics research, regenerative biology and gene therapy applications.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41556-022-00870-7>.

Received: 20 January 2021; Accepted: 15 February 2022;

Published online: 12 April 2022

References

- Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
- Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR–Cas9. *Science* **346**, 1258096 (2014).
- Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR–Cas system. *Cell* **163**, 759–771 (2015).
- Zetsche, B. et al. Multiplex gene editing by CRISPR–Cpf1 using a single crRNA array. *Nat. Biotechnol.* **35**, 31–34 (2017).
- Kleinstiver, B. P. et al. Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282 (2019).
- Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods* **16**, 887–893 (2019).
- Fonfara, I., Richter, H., Bratovi, M., Le Rhun, A. & Charpentier, E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517–521 (2016).
- Gier, R. A. et al. High-performance CRISPR–Cas12a genome editing for combinatorial genetic screening. *Nat. Commun.* **11**, 3455 (2020).
- Kim, D. et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **34**, 863–868 (2016).
- Bin Moon, S. et al. Highly efficient genome editing by CRISPR–Cpf1 using CRISPR RNA with a uridylylate-rich 3′-overhang. *Nat. Commun.* **9**, 3651 (2018).

11. Li, F. et al. Comparison of CRISPR/Cas endonucleases for in vivo retinal gene editing. *Front. Cell. Neurosci.* **14**, 570917 (2020).
12. Tak, Y. E. et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat. Methods* **14**, 1163–1166 (2017).
13. Ling, X. et al. Improving the efficiency of CRISPR-Cas12a-based genome editing with site-specific covalent Cas12a-crRNA conjugates. *Mol. Cell* **81**, 4747–4756 (2021).
14. Zhang, L. et al. AsCas12a ultra nuclease facilitates the rapid generation of therapeutic cell medicines. *Nat. Commun.* **12**, 3908 (2021).
15. Jones, S. K. et al. Massively parallel kinetic profiling of natural and engineered CRISPR nucleases. *Nat. Biotechnol.* **39**, 84–93 (2021).
16. Liu, P. et al. Enhanced Cas12a editing in mammalian cells and zebrafish. *Nucleic Acids Res.* **47**, 4169–4180 (2019).
17. Kocak, D. D. et al. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* **37**, 657–666 (2019).
18. Nguyen, L. T., Smith, B. M. & Jain, P. K. Enhancement of *trans*-cleavage activity of Cas12a with engineered crRNA enables amplified nucleic acid detection. *Nat. Commun.* **11**, 4906 (2020).
19. Wang, D., Zhang, F. & Gao, G. CRISPR-based therapeutic genome editing: strategies and in vivo delivery by AAV vectors. *Cell* **181**, 136–150 (2020).
20. Zhang, Y. et al. Enhanced CRISPR–Cas9 correction of Duchenne muscular dystrophy in mice by a self-complementary AAV delivery system. *Sci. Adv.* **6**, eaay6812 (2020).
21. Yamano, T. et al. Structural basis for the canonical and non-canonical PAM recognition by CRISPR–Cpf1. *Mol. Cell* **67**, 633–645 (2017).
22. Gao, L. et al. Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35**, 789–792 (2017).
23. Kempton, H. R. et al. Multiple input sensing and signal integration using a split Cas12a system. *Mol. Cell* **78**, 184–191 (2020).
24. Vora, S. et al. Rational design of a compact CRISPR–Cas9 activator for AAV-mediated delivery. Preprint at *bioRxiv* <https://doi.org/10.1101/298620> (2018).
25. Wang, Q. et al. Mouse γ -synuclein promoter-mediated gene expression and editing in mammalian retinal ganglion cells. *J. Neurosci.* **40**, 3896–3914 (2020).
26. Levy, J. M. et al. Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. *Nat. Biomed. Eng.* **4**, 97–110 (2020).
27. Feng, G. et al. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000).
28. Kleinstiver, B. P. et al. Genome-wide specificities of CRISPR–Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874 (2016).
29. Liu, Y. et al. CRISPR activation screens systematically identify factors that drive neuronal fate and reprogramming. *Cell Stem Cell* **23**, 758–771 (2018).
30. Lu, Y. et al. Reprogramming to recover youthful epigenetic information and restore vision. *Nature* **588**, 124–129 (2020).
31. Liu, P., Chen, M., Liu, Y., Qi, L. S. & Ding, S. CRISPR-based chromatin remodeling of the endogenous Oct4 or Sox2 locus enables reprogramming to pluripotency. *Cell Stem Cell* **22**, 252–261 (2018).
32. Lu, Y. et al. Reprogramming to recover youthful epigenetic information and restore vision. *Nature* **588**, 124–129 (2020).
33. Gonatopoulos-Pournatzis, T. et al. Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9–Cas12a platform. *Nat. Biotechnol.* **38**, 638–648 (2020).
34. Breinig, M. et al. Multiplexed orthogonal genome editing and transcriptional activation by Cas12a. *Nat. Methods* **16**, 51–54 (2019).
35. Matsuda, T. & Cepko, C. L. Controlled expression of transgenes introduced by in vivo electroporation. *Proc. Natl. Acad. Sci. USA* **104**, 1027–1032 (2007).
36. Wang, S., Sengel, C., Emerson, M. M. & Cepko, C. L. A gene regulatory network controls the binary fate decision of rod and bipolar cells in the vertebrate retina. *Dev. Cell* **30**, 513–527 (2014).
37. Chan, C. S. Y. et al. Cell type- and stage-specific expression of Otx2 is regulated by multiple transcription factors and cis-regulatory modules in the retina. *Development* **147**, dev187922 (2020).
38. Rocha-Martins, M. et al. De novo genesis of retinal ganglion cells by targeted expression of Klf4 in vivo. *Development* **146**, dev176586 (2019).
39. Sharma, P. et al. Oct4 mediates Müller glia reprogramming and cell cycle exit during retina regeneration in zebrafish. *Life Sci. Alliance* **2**, e201900548 (2019).
40. Lin, Y. P., Ouchi, Y., Satoh, S. & Watanabe, S. Sox2 plays a role in the induction of amacrine and Müller glial cells in mouse retinal progenitor cells. *Investig. Ophthalmol. Vis. Sci.* **50**, 68–74 (2009).
41. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
42. Venkatesh, A., Ma, S., Langellotto, F., Gao, G. & Punzo, C. Retinal gene delivery by rAAV and DNA electroporation. *Curr. Protoc. Microbiol.* <https://doi.org/10.1002/9780471729259.mc14d04s28> (2013).
43. Matharu, N. et al. CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **363**, eaau0629 (2019).
44. Liao, C. et al. Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. *Nat. Commun.* **10**, 2948 (2019).
45. Magnusson, J. P. et al. Enhanced Cas12a multi-gene regulation using a CRISPR array separator. *eLife* **10**:e66406 (2021).
46. Black, J. B. et al. Targeted epigenetic remodeling of endogenous loci by CRISPR/Cas9-based transcriptional activators directly converts fibroblasts to neuronal cells. *Cell Stem Cell* **19**, 406–414 (2016).
47. Nuñez, J. K. et al. Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**, 2503–2519 (2021).
48. Nakamura, M., Ivec, A. E., Gao, Y. & Qi, L. S. Durable CRISPR-based epigenetic silencing. *BioDesign Res.* **2021**, 981582 (2021).
49. Kemaladewi, D. U. et al. A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. *Nature* **572**, 125–130 (2019).
50. Xu, X. et al. Engineered miniature CRISPR–Cas system for mammalian genome regulation and editing. *Mol. Cell* **81**, 4333–4345 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Cell culture. HEK293T cells (American Type Culture Collection) and the reporter cell line²³ with stable expression of TRE3G–GFP were cultured in DMEM + GlutaMAX (Thermo Fisher) supplemented with 10% fetal bovine serum (FBS; Alstern) and 100 U ml⁻¹ penicillin–streptomycin (Life Technologies). P19 cells (American Type Culture Collection) were cultured in alpha-MEM medium with nucleosides (Invitrogen) with same FBS and penicillin–streptomycin concentrations as for the HEK293T cells. The cells were maintained at 37°C and 5% CO₂, and passaged using standard cell culture techniques. For transient transfection of HEK293T cells, the cells were seeded the day before transfection at 1 × 10⁵ cells ml⁻¹. Transient transfections were performed using 3 ml of TransIT-LT1 transfection reagent (Mirus) per mg of plasmid. The HEK293T cells were analysed 2–5 d post transfection, as specified in each experiment. For transient transfection of P19 cells, the cells were seeded the day before transfection at a density of 2 × 10⁵ cells ml⁻¹. Transient transfections were performed using 3 µl of Mirus X2 transfection reagent per µg of plasmid. For double selection, the cells were treated with 500 µg ml⁻¹ hygromycin and 2 µg ml⁻¹ puromycin. The P19 cells were analysed 3 d post transfection, as indicated.

Plasmid cloning. Standard molecular cloning techniques were used to assemble constructs in this study. Nuclease-dead dCas12a from *Lachnospiraceae* bacterium and its crRNA backbone were modified from a previous publication²³. The step-by-step protocol for generating crRNA arrays is provided at *Protocol Exchange*⁵¹.

Flow cytometry. Cells were dissociated using 0.05% trypsin–EDTA (Life Technologies), resuspended in PBS + 10% FBS and analysed for fluorescence using a CytoFLEX S flow cytometer (Beckman Coulter). Cells (10,000) from the population of interest (for most experiments, mCherry⁺ and BFP⁺ gated based on the untransfected control) were collected from each sample and analysed using FlowJo (v10).

Targeted deep sequencing and data analysis. Targeted deep sequencing primers were designed to generate an amplicon of 223 bp encompassing the crYFP target with generic adaptors (YFP-R, 5′-GACTGGAGTTCAGACGTGTGCTCTCCG ATCTCGGTGGTGCAGATGAACCTCAGG-3′; and YFP-F, 5′-ACACTCTTTCC CTACACGACGCTCTTCCGATCTGCTCATGGTGAAGGCGC-3′). DNA was extracted after PCR amplification and gel purification using a Qiagen gel extraction kit. Equal quantities of the samples were pooled and a mixed barcoded library was sequenced by the GENEWIZ Amplicon-EZ sequencing service. More than 350,000 reads were generated for each sample using an Illumina platform. Data analysis was performed using CRISPResso2 (ref. ⁵²).

RT-qPCR. RNA was isolated from transfected cells using a Qiagen RNeasy plus kit, followed by reverse transcription of 100 ng RNA to cDNA using an iScript kit (BioRad). A quantitative real-time PCR reaction was performed using SYBR master mix (BioRad) according to the manufacturer's protocol. Quantification of RNA expression was normalized based on the expression levels of *glyceraldehyde 3-phosphate dehydrogenase* and calculated using the $\Delta\Delta C_t$ method. The list of qPCR primers used is provided in Supplementary Table 3.

Immunostaining and quantitation. P19 cells were seeded onto black flat-bottomed 96-well plates 48 h after transfection (continuing in dual-selection media) and fixed with 1×DPBS containing 4% formaldehyde 24 h after seeding. Each well was permeabilized with 1×DPBS with 0.25% Triton X-100, blocked with 1×DPBS containing 5% donkey serum and then incubated overnight at 4°C with the following primary antibodies diluted in 1×DPBS containing 5% donkey serum: mouse anti-Oct4 (1:200; BD Bioscience, 611203), rabbit anti-Sox2 (1:200; Cell Signaling Technology, 14962) and goat anti-Klf4 (1:200; R&D Systems, AF3158). Each well was washed three times with 1×DPBS and then incubated for 1 h with Alexa Fluor 488- or 647-conjugated donkey secondary antibodies (Life Tech) diluted 1:500 in the same buffer used for the primary antibodies. Each well was then washed three times with 1×PBS and immersed in 1×PBS. No nuclear dye was used. Imaging was performed using a Leica DMI8 inverted microscope with a ×20 objective and Leica DFC9000 CT camera. The cell fluorescence intensities were quantitated using a semi-automatic image analysis pipeline based on MATLAB (R2019a) available at <https://github.com/QilabGitHub/dCas12a-microscopy>. Threshold-based segmentation was performed based on the mCherry channel representing dCas12. Morphological operations were then applied to remove noise and yield masks for single cells. Based on the masks, the mean fluorescent intensities of all corresponding channels were collected for every cell for further statistical analysis. For display, one representative image was used per condition, with roughly equal numbers of cells between conditions to facilitate qualitative comparison.

RNA-seq. HEK293T reporter cells stably expressing TRE3G–GFP were seeded in a six-well plate at a density of 2 × 10⁵ cells ml⁻¹ and co-transfected the next day with Tet crRNA or LacZ non-target crRNA with dCas12aWT or hyperdCas12a, in duplicate. One day after transfection, the transfected cells were placed in antibiotic selection medium (containing 500 µg ml⁻¹ hygromycin and 2 µg ml⁻¹ puromycin) for 2 d before harvest. Total RNA was isolated using an RNeasy plus mini kit

(Qiagen). Library preparation and next-generation sequencing were performed by Novogene as described previously²⁹. The Spliced Transcripts Alignment to a Reference (STAR) software⁵³ was used to index the hg19 genome and GFP sequence, and then to map paired-end reads to the genome. HTSeq-Count⁵⁴ was used to quantify gene-level expression. The gene-level fragments per kilobase of transcript per million mapped reads (FPKM) values were calculated using a custom Python script available at <https://github.com/QilabGitHub/FPKMcalculation>. Pearson's correlation coefficients were obtained using QR decomposition and regression.

Animals. For the in vivo electroporation experiments, WT neonatal mice were obtained from timed pregnant CD-1 mice (Charles River Laboratories). For the AAV experiments, Thy1-YFP-17 transgenic mice were originally generated by G. Feng and J. Sanes²⁷, and were acquired from Z. He; 6–8-week-old male mice were used. All animal studies were approved by the Institutional Animal Care and Use Committee at Stanford School of Medicine.

AAV production and intravitreal injection. AAV serotype 2 virions were produced by AAVnGene using previously described approaches²⁵. The AAV titres were determined by real-time PCR. AAV–Cas12a and AAV–crYFP were mixed at a ratio of 2:1. AAV–Cas12a was diluted to 4.5 × 10¹² vector genome (vg) ml⁻¹ and AAV–crYFP was diluted to 2.25 × 10¹² vg ml⁻¹. For intravitreal injection, mice were anaesthetized with xylazine and ketamine based on their body weight (0.01 mg g⁻¹ xylazine + 0.08 mg g⁻¹ ketamine). A pulled and polished microcapillary needle was inserted into the peripheral retina just behind the ora serrata. Approximately 2 µl of the vitreous was removed to allow injection of 2 µl AAV into the vitreous chamber to achieve 9 × 10⁹ vg per retina of Cas12a and 4.5 × 10⁹ vg per retina of crYFP. The mice were killed 10 weeks after the AAV injection. Transcardiac perfusion was performed as described²⁵. For the retina whole mounts, the retinas were dissected out and washed extensively in PBS before blocking in staining buffer (10% normal goat serum and 2% Triton X-100 in PBS) for 1 h. RBPMs guinea pig antibody was made at ProSci according to previous publications⁵⁵ and used at 1:4,000, and rat HA (clone 3F10, 1:200; Roche) was diluted in the same staining buffer. Floating retinas were incubated with primary antibodies at 4°C overnight and washed three times with PBS (30 min each wash). Secondary antibodies (Cy2-, Cy3- or Cy5-conjugated) were then applied (1:200; Jackson ImmunoResearch) and incubated for 1 h at room temperature. The retinas were again washed three times with PBS (30 min each wash) before a coverslip was attached with Fluoromount-G (SouthernBiotech). Quantitation of the fluorescence of individual cells utilized a custom semi-automatic image analysis pipeline based on MATLAB (version R2019a) available at <https://github.com/QilabGitHub/dCas12a-microscopy>. For the analysis of mouse-retina wet mounts, threshold-based segmentation was performed based on the fluorescent channel representing crRNA, which had the highest signal-to-noise ratio and distributes evenly throughout the cytoplasm. Morphological operations were then applied to remove noise and thus yield masks for individual cells. Based on the masks, the mean fluorescent intensities of all corresponding channels for every cell were collected for further statistical analyses.

In vivo plasmid electroporation. Plasmid DNA was injected into the sub-retinal space of neonatal mice and electrical pulses were applied with tweezer-style electrodes as described^{35,36}. More details are provided at *Protocol Exchange*⁵¹. Plasmid with WT dCas12a was mixed with the CAG-GFP construct at a 5:1 ratio and electroporated at a concentration of up to 2 µg µl⁻¹ total plasmid at P0. Five pulses of 80 V (50 ms each) at intervals of 950 ms were applied to neonatal mouse pups. Dissected mouse eyeballs were processed as described³⁷. The eyeballs were fixed in 4% paraformaldehyde in 1×PBS (pH 7.4) for 2 h at room temperature. The retinas were dissected and equilibrated at room temperature in a series of sucrose solutions (5% sucrose in 1×PBS, 5 min; 15% sucrose in 1×PBS, 15 min; 30% sucrose in 1×PBS, 1 h; and 1:1 mixed solution of OCT compound and 30% sucrose in PBS, overnight at 4°C), frozen and stored at –80°C. A Leica CM3050S cryostat (Leica Microsystems) was used to prepare 20-µm cryosections. Retinal cryosections were briefly washed in 1×PBS, incubated in 0.2% Triton X-100, followed by 1×PBS for 20 min and blocked for 30 min in a blocking solution of 0.1% Triton X-100, 1% bovine serum albumin and 10% donkey serum (Jackson ImmunoResearch Laboratories) in 1×PBS. The slides were incubated with primary antibodies diluted in blocking solution in a humidified chamber at room temperature at 4°C overnight. After three washes in 0.1% Triton X-100 in 1×PBS, the slides were incubated with secondary antibodies and DAPI (Sigma-Aldrich; D9542) for 1–2 h, washed three times with 0.1% Triton X-100 in 1×PBS and mounted in Fluoromount-G (Southern Biotechnology Associates). The primary antibodies to Oct4, Sox2 and Klf4 described in the 'Immunostaining and quantitation' section were used. Additional primary antibodies used were: rat anti-HA (Roche, 3F10), guinea pig anti-RBPMs (PhosphoSolutions, 1832) and rabbit anti-Pax6 (Thermo, 42-6600). The retinal slices were imaged using an LSM710 confocal inverted laser scanning microscope, with a Plan Apochromat objective ×40/1.4 oil (FWD = 0.13 mm) with 405, 488, 561 and 633 nm lasers. Quantitation was performed as described previously³⁶ using the Fiji software.

Statistics and reproducibility. Most in vitro data are represented as bar graphs showing the mean (without error bars) and the individual data points. For the in vivo experiments, data are presented as bar graphs of the mean $\pm$ s.d. or box-and-whisker plots. The statistical analysis was based on sample size (n), indicating the number of biologically independent experiments or animals, as described in the respective figure legends. No statistical method was used to pre-determine the sample size. Each experiment was performed a minimum of three times to make sure that similar results were reproducible, except for the initial single mutant screen (Fig. 1c–e) that was performed with two independent biological replicates, the crRNA screens (Extended Data Figs. 6–8) and the whole-transcriptome RNA-seq (two independent biological replicates; Extended Data Fig. 4). Unless otherwise indicated, micrographs are representative images from at least three independent experiments. No samples or animals were excluded from the analyses. Data collection and analysis were not blinded. Data were compared between groups of animals using a two-tailed Student's t -test. Statistical tests were performed using Prism 9.1.0 (GraphPad Software). For RNA-seq analysis, Pearson's correlation coefficients were obtained using QR decomposition and regression.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Whole-transcriptome sequencing data can be accessed in Gene Expression Omnibus under the accession code [GSE166817](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE166817). Key constructs and plasmids will be available on Addgene (https://www.addgene.org/Stamley_Qi/). All other data supporting the findings of this study are available from the corresponding authors on reasonable request. Source data are provided with this paper.

Code availability

The gene-level FPKM values were calculated using a custom Python script available at <https://github.com/QilabGitHub/FPKMcalculation>. The semi-fluorescence intensities of individual cells (microscopy) were quantitated with a semi-automated image analysis pipeline based on MATLAB (version R2019a) available at <https://github.com/QilabGitHub/dCas12a-microscopy>.

References

51. Guo, L. Y. et al. Multiplex CRISPR genome regulation in mouse retina with hyper-efficient Cas12a. *Protoc. Exch.* <https://doi.org/10.21203/rs.3.pex-1811/v1> (2022).
52. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
53. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
54. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
55. Kwong, J. M. K., Caprioli, J. & Piri, N. RNA binding protein with multiple splicing: a new marker for retinal ganglion cells. *Investig. Ophthalmol. Vis. Sci.* **51**, 1052–1058 (2010).

Acknowledgements

L.Y.G. acknowledges support from a Stanford's National Eye Institute (NEI) T32 Vision Training Grant (grant no. 5T32EY020485), the Knights Templar Eye Foundation (KTEF) and the VitreoRetinal Surgery Foundation (VRSF). L.S.Q. acknowledges support from the Li Ka Shing Foundation, National Science Foundation CAREER award (award no. 2046650), National Institutes of Health (grant no. 1U01DK127405), Chan Zuckerberg Initiative Neurodegeneration Challenge Network, Stanford Maternal and Child Health Research Institute (MCHRI) through the Uytensu–Hamilton 22q11 Neuropsychiatry Research Award Program, Stanford–Coulter Translational Research Grant, and California Institute for Regenerative Medicine (CIRM; grant no. DISC-12669). S.W. is supported by funding from the American Diabetes Association (grant no. 1-16-INI-16), NIH grant no. 1R01EY03258501, NIH grant no. R01NS109990, NIH-NEI grant no. P30-EY026877 and Research to Prevent Blindness, Inc. [BioRender.com](https://www.biorender.com) was used to generate illustrations in Fig. 1b, 1g, 2a, 2c, 2f, 3a, 4a and 6e. The members of the Qi, Wang and Hu laboratories provided valuable discussion and support for this project.

Author contributions

L.Y.G. and L.S.Q. conceived the idea for this study. L.Y.G., J.B., S.W. and L.S.Q. designed the experiments. L.Y.G., J.B., X.Z., H.R.K., B.G., D.A.R. and R.M.J. performed the ex vivo experiments. L.Y.G., J.B., A.E.D. and P.L. performed the in vivo experiments with guidance from S.W. and Y.H. L.Y.G., J.B., S.W. and L.S.Q. analysed the experimental data. A.C. and X.L. performed the computational analyses of the sequencing data, and X.Z. and R.M.J. analysed the imaging data. X.X. provided reagents. L.Y.G. and L.S.Q. wrote the manuscript with input from all authors.

Competing interests

The authors have filed a provisional patent via Stanford University related to the work (US patent no. 63/148,652). L.S.Q. is a founder and scientific advisory board member of Epicrispr Biotechnologies.

Additional information

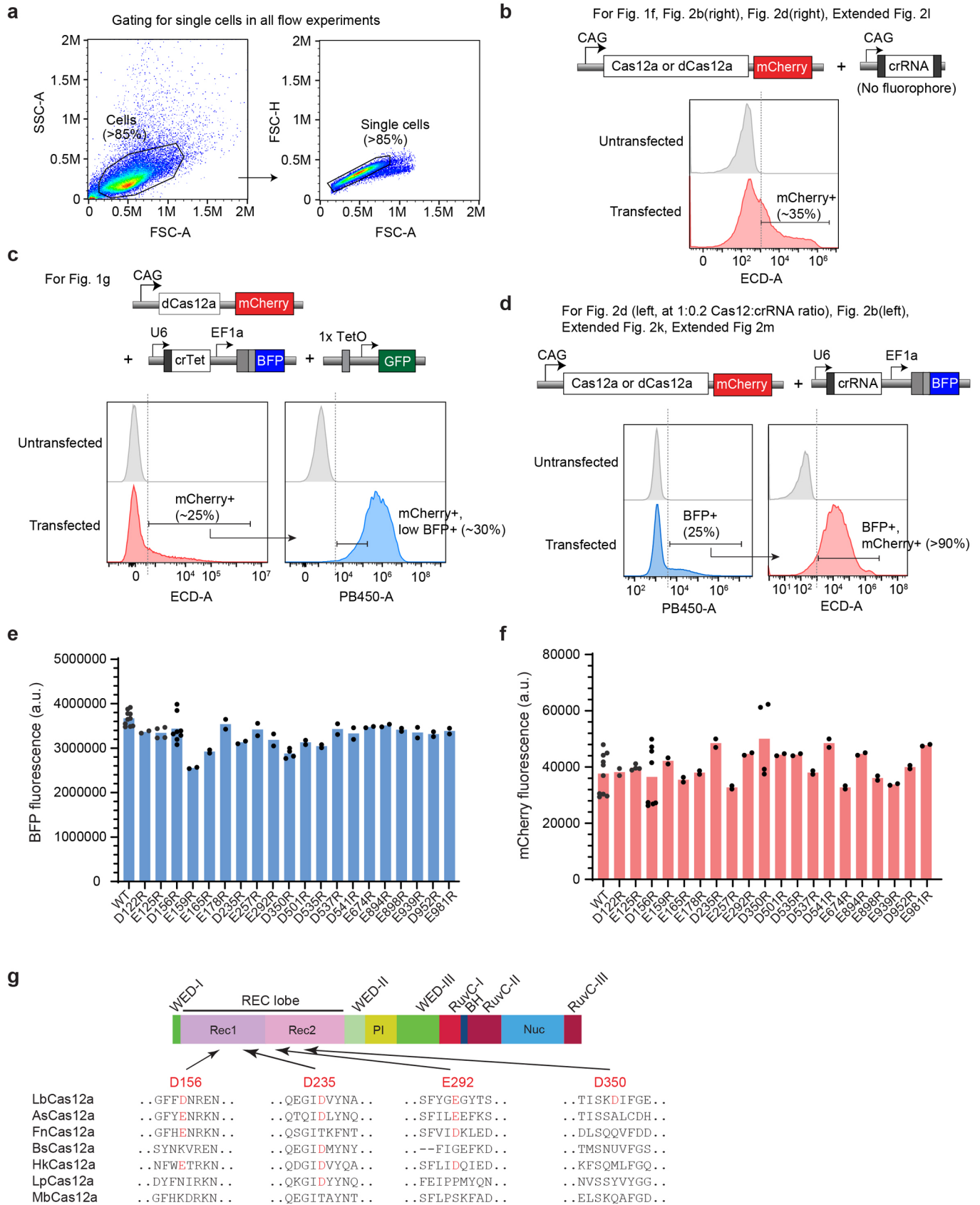
Extended data is available for this paper at <https://doi.org/10.1038/s41556-022-00870-7>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41556-022-00870-7>.

Correspondence and requests for materials should be addressed to Lucie Y. Guo, Sui Wang or Lei S. Qi.

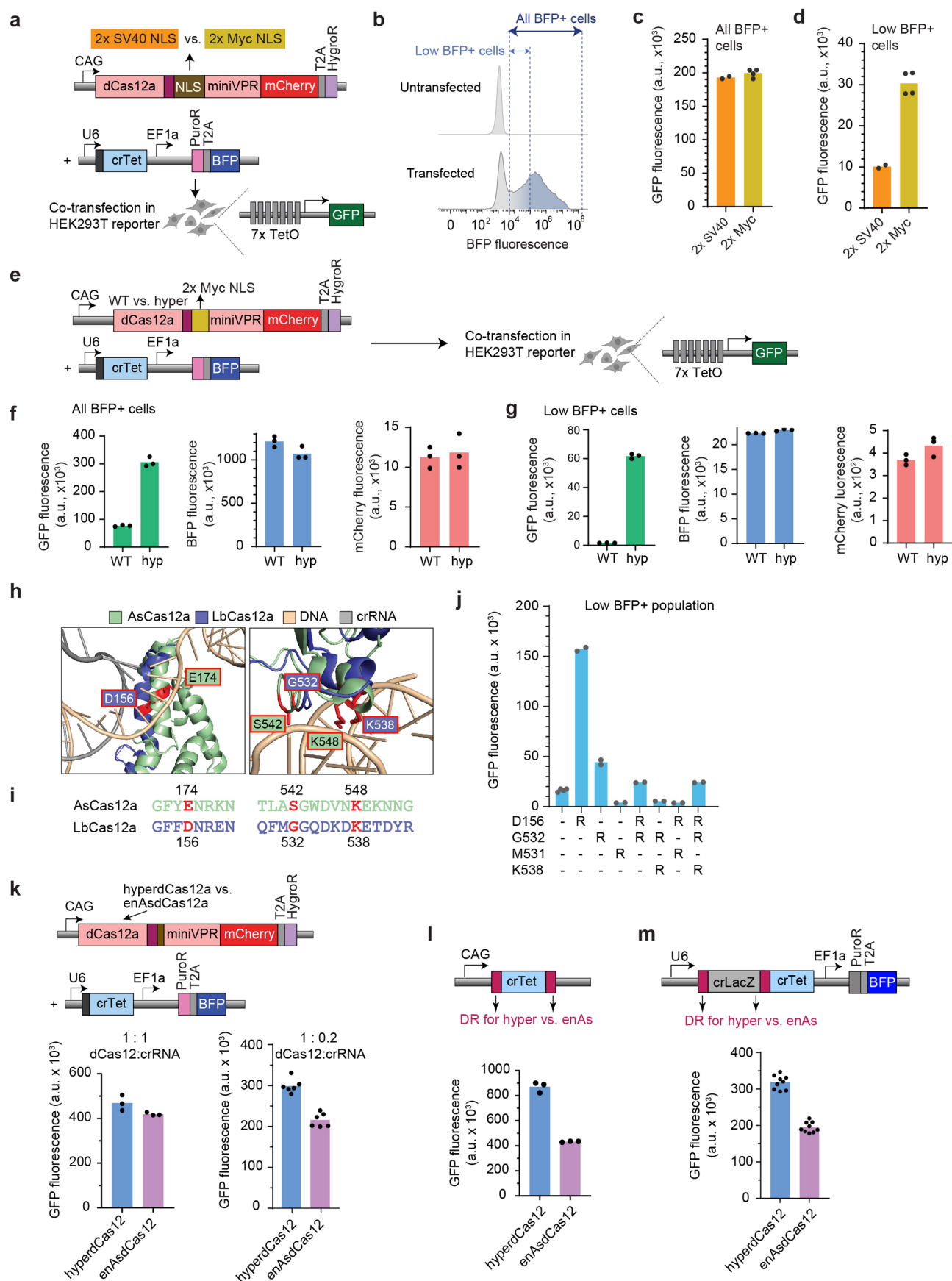
Peer review information *Nature Cell Biology* thanks Stephen H. Tsang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.



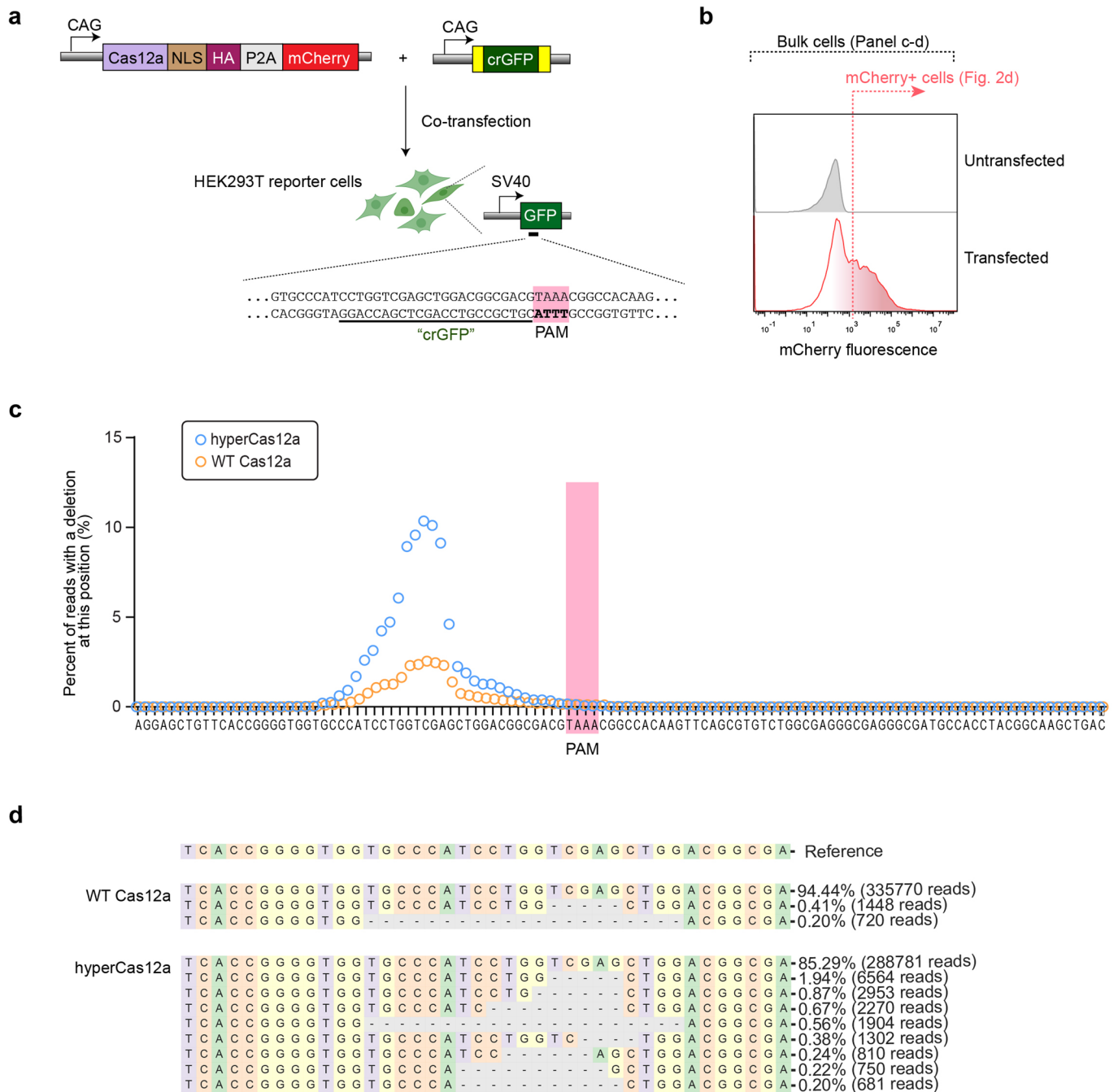
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Gating strategies for flow cytometry. **a**, Standard strategy for gating cells based on forward scatter (FSC) and side scatter (SSA) (left), then further gating for singlets based on FSC-height (FSC-H) and FSC-area (FSC-A) (right). To analyse transfected cells, further gating is applied to the singlet population based on fluorescence intensity. Please note that for Fig. 1c-e, and Extended Data Fig. 2a-j, that gating strategy is included within the figure. **b**, Gating strategy for experiments with Cas12a-mCherry plasmid and a CAG-crRNA plasmid (without fluorophore), thus mCherry+ cells are used for analysis. **c**, Gating strategy for Fig. 1g, in which 3 plasmids are co-transfected. **d**, Gating strategy for some experiments with Cas12a-mCherry plasmid and U6-crRNA (with BFP). **e**, Mean BFP fluorescence across the mutants tested in Fig. 1c. **f**, Mean mCherry fluorescence among mutants tested in Fig. 1c. In e-f, each data point represents the mean GFP intensity of an independent experiment, with each bar representing the average of 2 or more independent experiments. **g**, Schematic of the LbCas12a protein domains and location of four of the most potent point mutants, with alignment across various Cas12 species. The relevant Asp (D) or Glu(E) residues are highlighted in red.

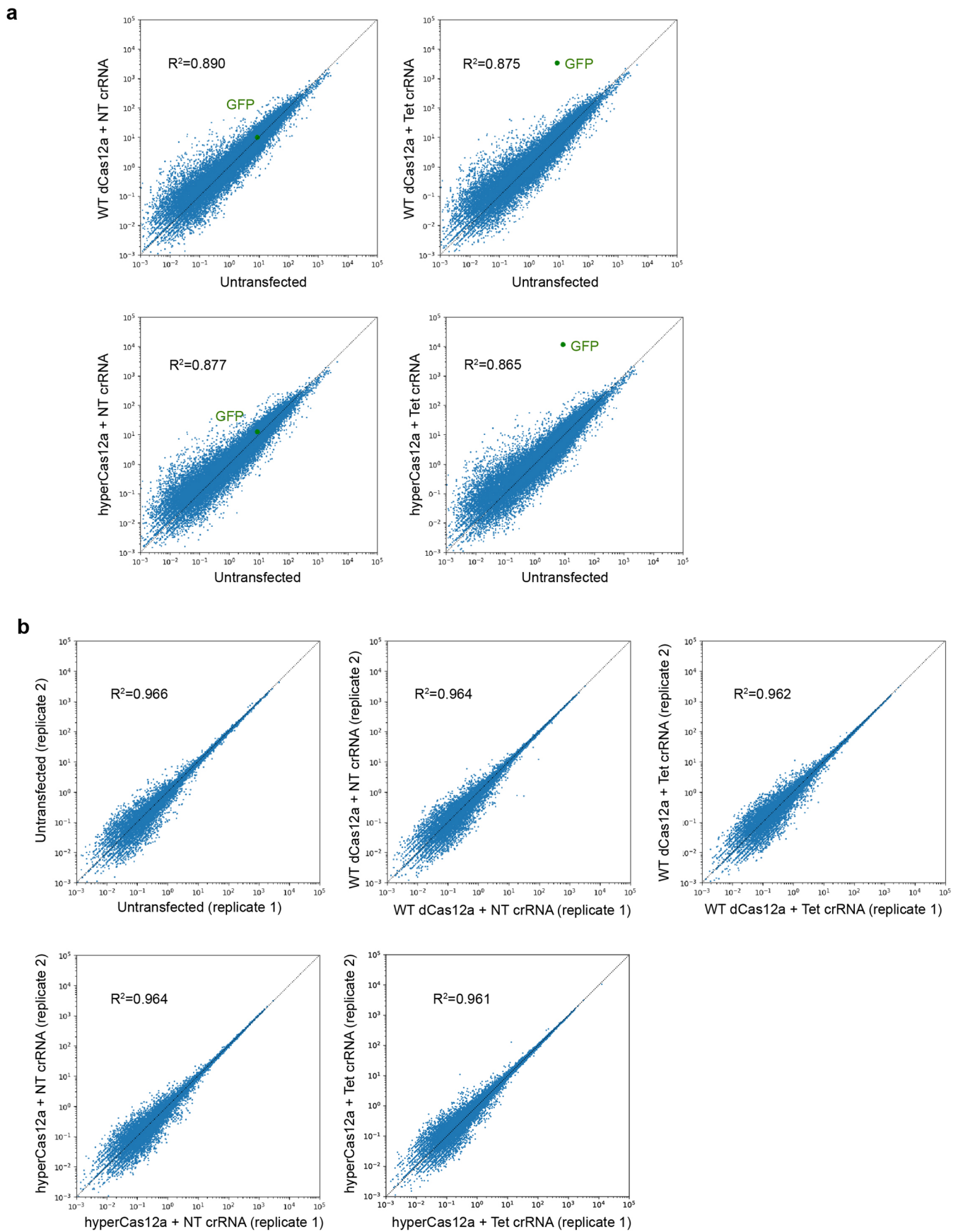


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Optimizing the nuclear localizing signal and comparing to enAsdCas12a. **a**, Schematic to test two different nuclear localization signals. Constructs containing either 2xSV40 or 2xMyc NLS fused with WT dCas12a are co-transfected with Tet crRNA in TRE3-GFP HEK293T reporter cells. **b**, Representative flow cytometry histogram of BFP intensity, showing threshold for BFP+ cells, and subset of 'low BFP+' cells (similar to Fig. 1d). **c-d**, GFP fluorescence in BFP+ (c) or 'low BFP+' cells (d). **e-g**, GFP fluorescence in BFP+ (f) or low BFP+ cells (g) to compare WT versus hyperdCas12a (hyp) with 2xMyc NLS, as well as BFP and mCherry average fluorescence in each gated BFP group. **h-i**, Alignment of the structure of LbCas12a versus AsCas12a proteins (h) and alignment of peptide sequences (i) encompassing mutations harboured by enAsCas12a, a reported enhanced variant of Cas12a from *Acidaminococcus* with the E174R/S542R/K548R mutations⁵ corresponding to homologous residues (D156R/G532R/K538R) mutations in LbCas12a. **j**, Comparison of variants containing mutations of homologous residues in LbCas12a in 'low BFP+' cells. Interestingly, D156R combined with G532R and/or K538R did not achieve activation higher than the single D156R mutant, in contrast to results with homologous residues in AsCas12a⁵. **k**, Comparison of hyperdCas12a versus enAsdCas12a with a single crRNA driven by U6 promoter in 1:1 versus 1:0.2 ratio of dCas12a:crRNA, in TRE3G-GFP HEK293T cells. **l**, Comparison of hyperdCas12a versus enAsdCas12a with single crRNA driven by CAG promoter flanked by direct repeats (DR) specific to LbCas12a versus AsCas12a. **m**, Comparison of hyperdCas12a versus enAsdCas12a with dual crRNAs containing crTet on the second position and non-targeting crLacZ on the first position flanked by As or Lb direct repeats (DR). All transfections in this figure were carried out in TRE3G-GFP HEK293T reporter cells. Bar graph in f, g and k-m shows the mean of $n \geq 3$ independent experiments; bar graph in j shows the mean of $n \geq 2$ independent experiments; each data point represents value of an independent experiment.

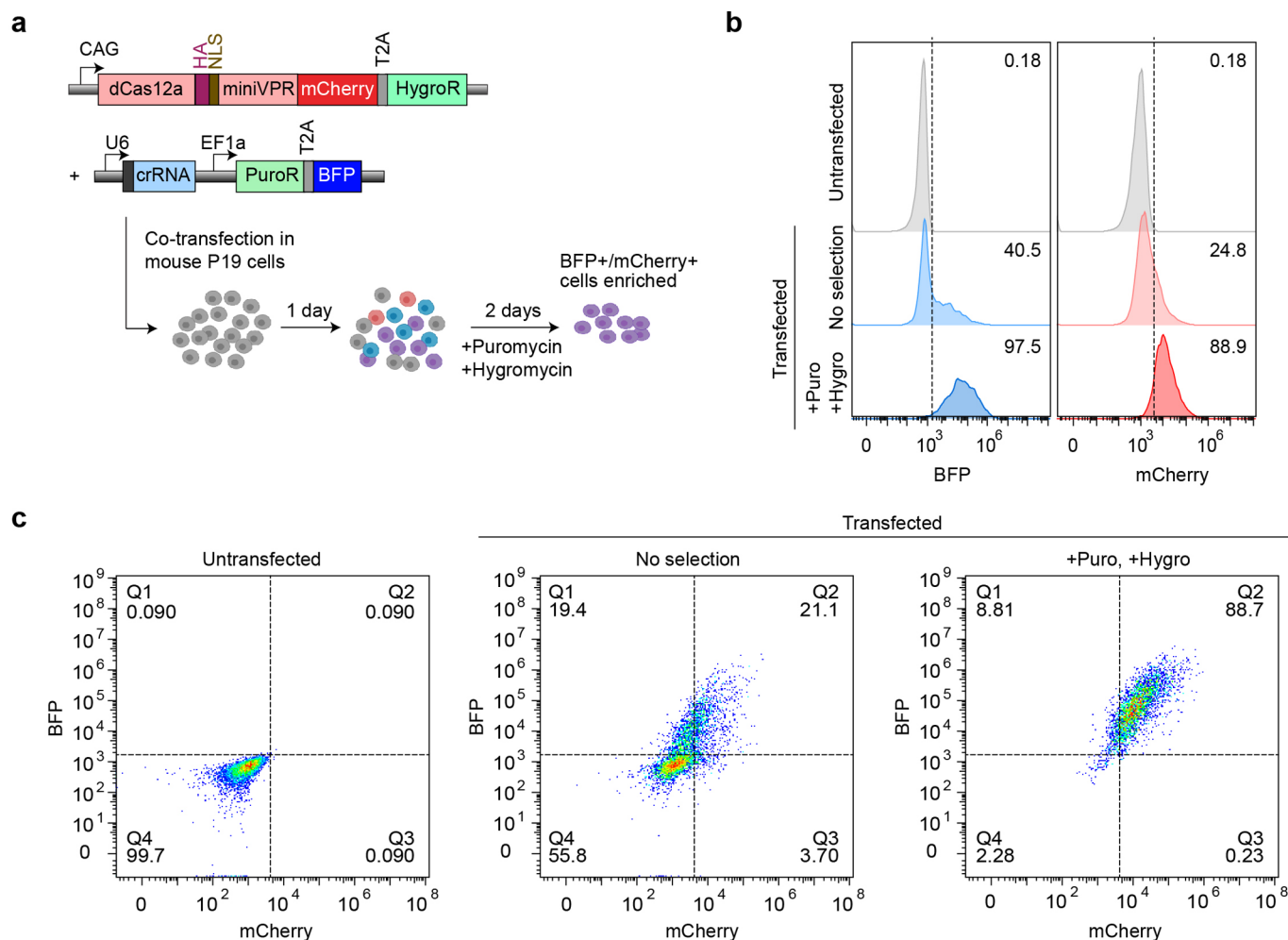


Extended Data Fig. 3 | Improved gene editing by hyperCas12a. **a**, Nuclease-active WT Cas12a versus hyperCas12a were co-transfected with crGFP into HEK293T cells stably expressing GFP driven by SV40 promoter. **b**, Representative flow cytometry histogram showing threshold for mCherry+ cells. Analysis of mCherry+ cells are shown in Fig. 2d, while bulk cells (without sorting) were used for indel analysis (panels c-d). **c**, Indel activity at each nucleotide position, shown as percentage of total reads with a deletion at the position. The PAM is highlighted in pink. **d**, Indel patterns and corresponding ratios in total reads detected by deep sequencing as analysed by CRISPResso2⁵².

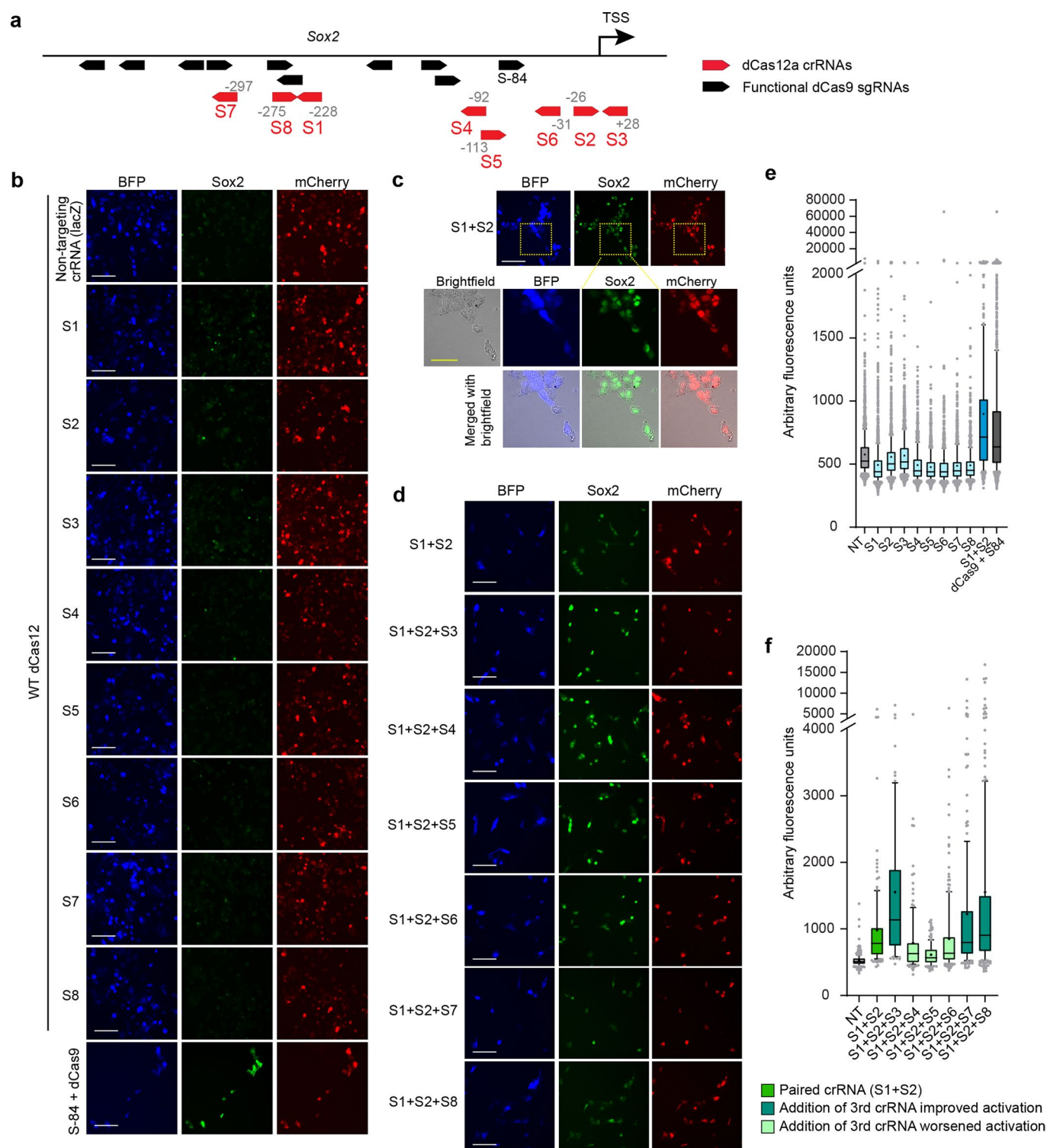


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Characterization of off-target effects of hyperdCas12a. **a.** RNA-seq FPKMs (fragments per kilobase million fragments mapped) are plotted as transfected versus non-transfected cells. The transfected samples are TRE3G-GFP HEK293T reporter cells co-transfected with WT dCas12a or hyperdCas12a, and with non-targeting crRNA or crRNA targeting TRE3G. **b.** RNA-seq plots showing FKPM (Fragments Per Kilobase Million) between two biological duplicates for each condition. The calculated Pearson correlation coefficient for each condition is shown on the graph.

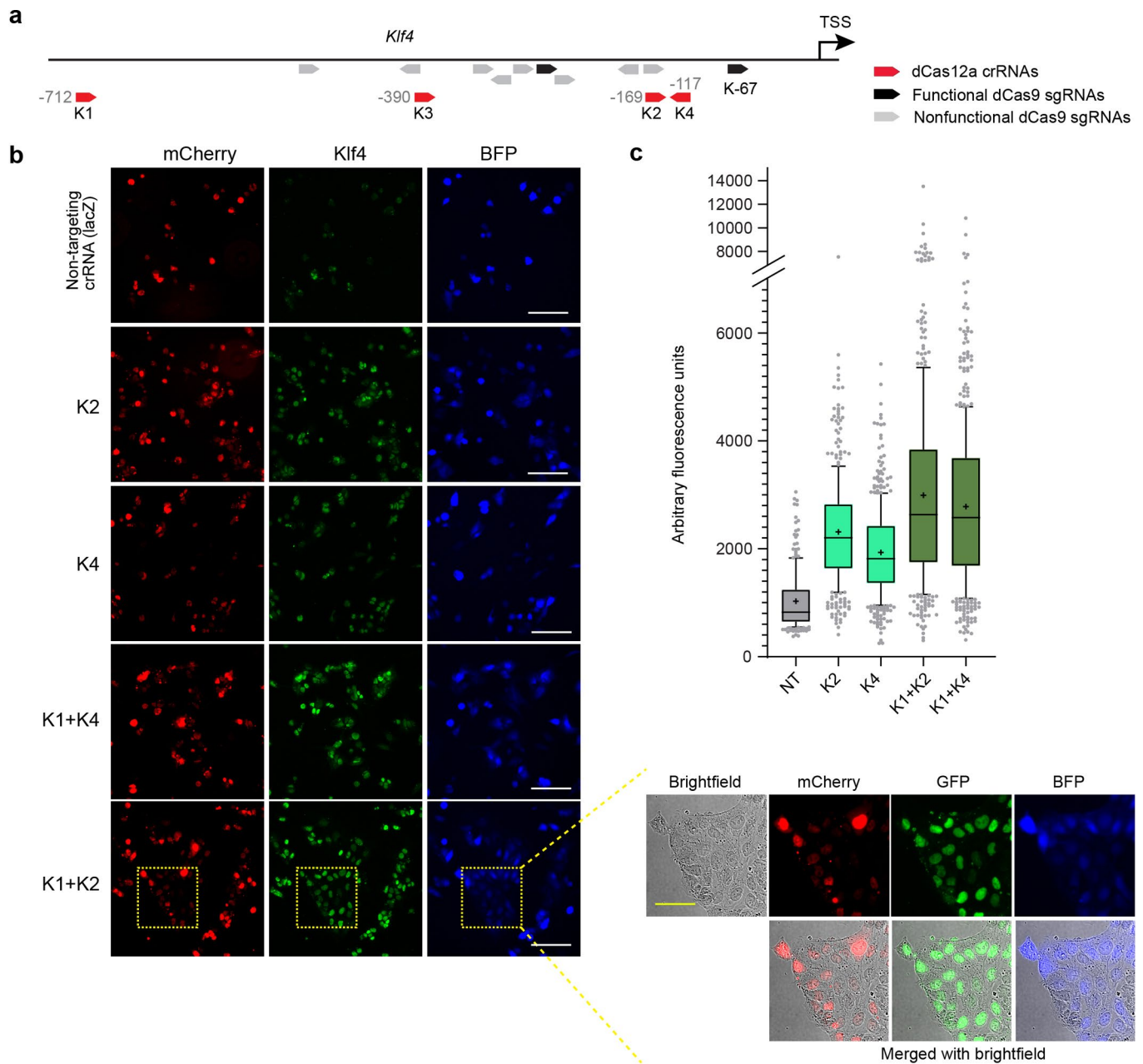


Extended Data Fig. 5 | Dual antibiotic selection for co-transfected mouse P19 cells. **a**, Mouse P19 cells were co-transfected with constructs expressing puromycin resistance (PuroR) and hygromycin resistance (HygroR), then selected with puromycin and hygromycin at 24 hr after transfection. Cells were collected for analysis 72 hr after transfection. **b**, Histograms showing percentage of BFP + (crRNA) and mCherry + (dCas12a) cells for non-transfected, non-selected, and Puro/Hygro-selected cells. **c**, Flow cytometry plots. Data in panels b-c are representative plots of $n = 3$ independent experiments.

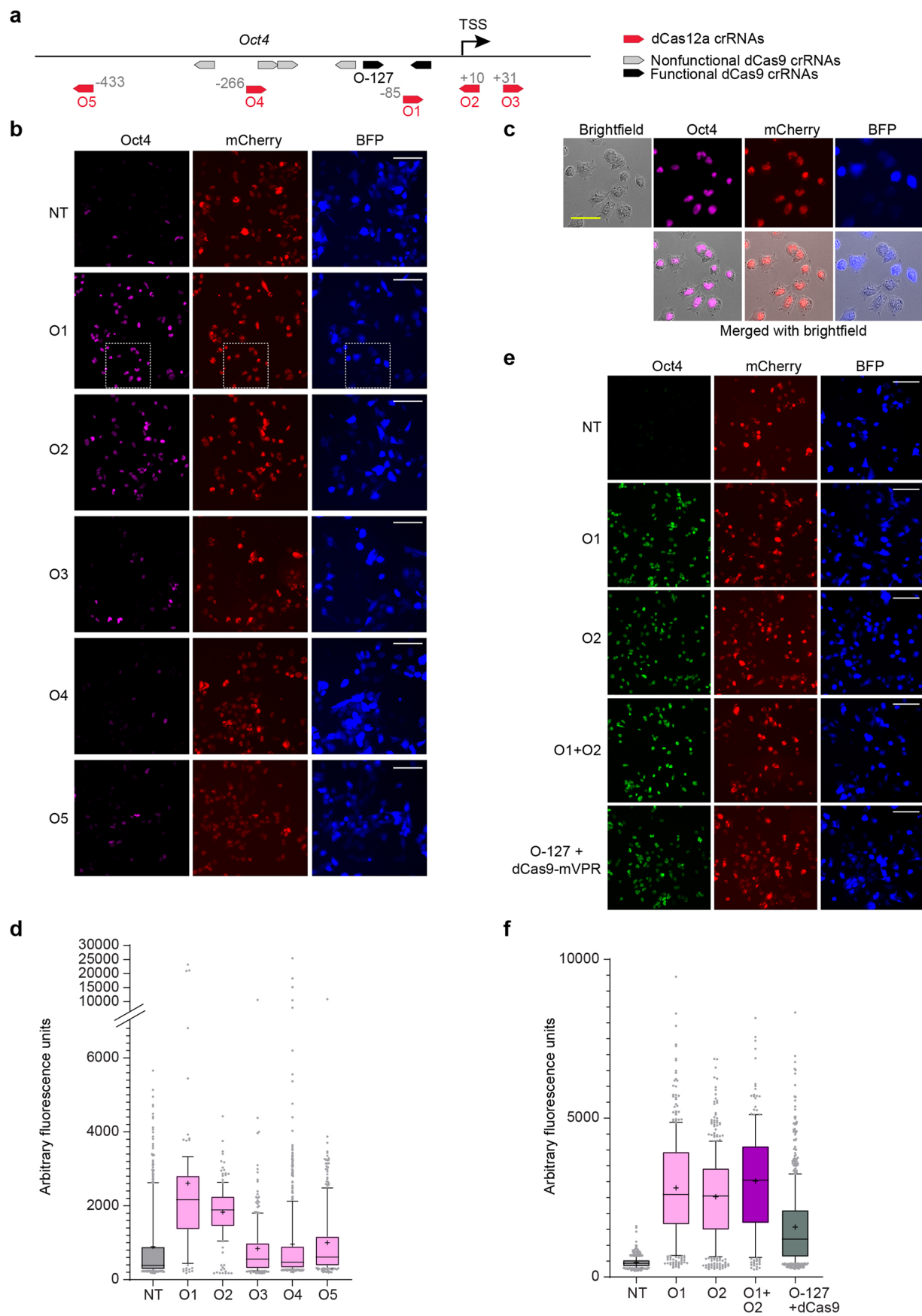


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Screening dCas12a crRNAs for activating endogenous Sox2. **a**, Schematic of dCas12a crRNAs (red) targeting promoter of Sox2, and their relative positions to validated dCas9 sgRNAs³¹ that are functional (black) for activating Sox2. Arrows indicate sense or antisense binding of crRNAs/sgRNAs to target DNA. The genomic position of the first 'T' in PAM (relative to TSS, which is '0') are shown for each crRNA targeting the Sox2 promoter. **b**, Immunostaining of Sox2 expression from activation by WT dCas12a-miniVPR with various Sox2 single crRNAs, compared to activation by dCas9-miniVPR (using a validated sgRNA, S84)³¹. Scale bar, 100 μ m. **c-d**, Immunostaining of Sox2 expression and co-localization with BFP and mCherry for a pair of crRNAs (c) and a panel of 'triplet' crRNAs (d), demonstrating additive or synergistic effect when multiple crRNAs are used in tandem. Interestingly, addition of a third crRNA targeting a region between the paired crRNAs S1 and S2 decreases the level of activation. Inset (c) shows brightfield image to demonstrate nuclear localization of mCherry (hyperdCas12a) and target (Sox2), since BFP on crRNA plasmid precludes the use of an additional nuclear dye. White scale bar, 100 μ m; yellow scale bar (within inset), 50 μ m. **e-f**, Automated quantitation of images in panels b-d. In panel **e**, 350-2000 cells for each condition were quantitated for one screening experiment with multiple fields of view. The exact number of cells for each condition is listed in the Source Data for Extended Data Fig. 6. NT, non-targeting crRNA. In panel **f**, 70-250 cells for each condition were quantitated for one screening experiment with multiple fields of view. For box-and-whisker plots, the box shows 25-75% (with bar at median, dot at mean), and whiskers encompass 10-90%, with individual data points shown for the lowest and highest 10% of each dataset. The exact number of cells for each condition are listed in the Source Data for Extended Data Fig. 6. NT, non-targeting crRNA.

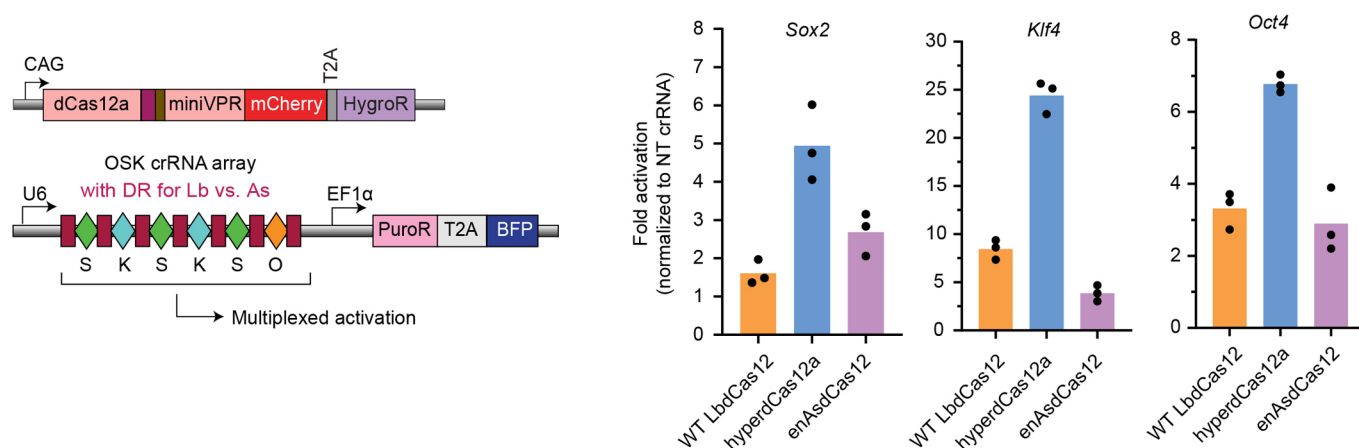


Extended Data Fig. 7 | Screening dCas12a crRNAs for activating endogenous *Klf4*. **a**, Schematic of dCas12a crRNAs (red) targeting promoter of *Klf4* and their relative positions to known dCas9 sgRNAs³¹ that are functional (black) or non-functional (grey) for activating *Klf4*. Arrows indicate sense or antisense binding of crRNAs/sgRNAs to the target DNA. The genomic position of the first 'T' in PAM (relative to TSS, which is 'O') are shown for each crRNA targeting to the *Klf4* promoter. **b**, Immunostaining of *Klf4*. Inset shows brightfield image to demonstrate nuclear localization of mCherry (hyperdCas12a) and target (*Klf4*), since BFP on crRNA plasmid precludes the use of an additional nuclear dye. White scale bar, 100 μ m; yellow scale bar (within inset), 50 μ m. **c**, Automated quantitation of images in panel b, where 200–600 cells for each condition were quantitated for one screening experiment with multiple fields of view. The exact number of cells for each condition is listed in the Source Data for Extended Data Fig. 7. NT, non-targeting crRNA.

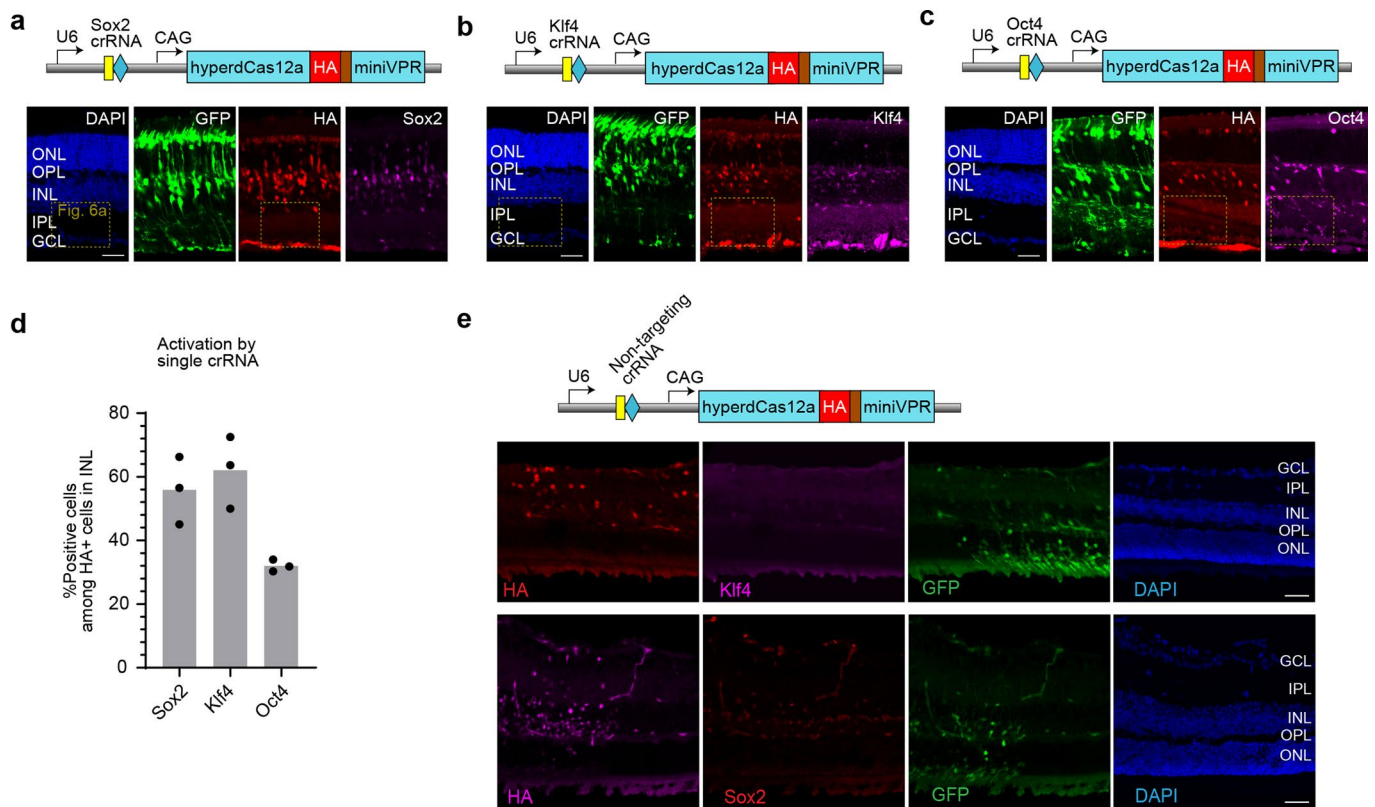


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Screening dCas12a crRNAs for activating endogenous *Oct4*. **a.** Schematic of dCas12a crRNAs (red) targeting promoters of *Oct4* and their relative positions to known dCas9 sgRNAs³¹ that are functional (black) or non-functional (grey) for activating *Oct4*. Arrows indicate sense or antisense binding of crRNAs/sgRNAs to the target DNA. The genomic position of the first 'T' in PAM (relative to TSS, which is '0') are shown for each crRNA targeting to the *Oct4* promoter. **b.** Immunostaining of *Oct4*. White scale bar, 100 μ m. **c.** Inset shows merge with brightfield to demonstrate nuclear localization of mCherry (hyperdCas12a) and target (*Oct4*), since BFP on crRNA plasmid precludes the use of an additional nuclear dye. Yellow scale bar, 50 μ m. **d.** Quantification of panel b, where 100–600 cells for each condition were quantitated over one screening experiment with multiple fields of view. The exact number of cells for each condition are listed in the Source Data for Extended Data Fig. 8. **e.** Immunostaining of *Oct4* after activation by paired crRNA consisting of the two most potent crRNAs (O1 + O2), which shows lack of additive effect. White scale bar, 100 μ m. **f.** Quantification in panels e, where 200–700 cells for each condition were quantitated over one screening experiment with multiple fields of view. For box-and-whisker plots, the box shows 25–75% (with bar at median, dot at mean), and whiskers encompass 10–90%, with individual data points shown for the lowest and highest 10% of each dataset. The exact number of cells for each condition is listed in the Source Data for Extended Data Fig. 8.



Extended Data Fig. 9 | Enhanced multiplex activation by hyperdCas12a. Multiplex endogenous gene activation by hyperdCas12a versus enAsdCas12a and 6-crRNA array targeting *Oct4*, *Sox2* and *Klf4* (OSK) in mouse P19 cells as measured by qPCR, in similar experiment as described in Fig. 3a. Each data point shows one independent measurement, and each bar shows the average of n=3 independent experiments. DR, direct repeat; NT, non-targeting.



Extended Data Fig. 10 | In vivo single crRNA activation by hyperdCas12a. **a-c**, Constructs containing hyperdCas12a and single crRNA to Sox2 (**a**), *Klf4* (**b**) or Oct4 (**c**) for *in vivo* electroporation in post-natal mouse retina and representative immunofluorescence images. CAG-GFP is used to mark the electroporated patch. Scale bar, 50 μ m. Magnified views of the regions in the yellow boxes are shown in Fig. 6a. **d**, Quantification of percentages of Oct4+, Sox2+ or Klf4+ cells among HA+ cells in INL. Bar graph shows the mean of 3 independent experiments, and each data point represents value of an independent experiment. **e**, Immunofluorescence images of *in vivo* electroporation in mouse retina with hyperdCas12a with non-targeting LacZ crRNA. Scale bar, 50 μ m.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Associated bundled software for acquisition of data using flow cytometry and live cell microscopy. For RNAseq, RNA was extracted with standard kit and library preparation and next-generation sequencing were performed by Novogene. For gene editing experiments, DNA was extracted with standard kit and sequenced by GENEWIZ Amplicon-EZ sequencing service.

Data analysis CRISPResso2 was used for analysis of next generation sequencing data. FlowJo (v10) was used for analysis of flow cytometry data. Prism (GraphPad) was used for statistical analysis. FIJI and custom Matlab script for image analysis. Spliced Transcripts Alignment to a Reference (STAR) software4 was used to index hg19 genome and GFP sequence, and then to map paired end reads to the genome. HTSeq-Count5 was used to quantify gene-level expression. Gene-level fragments per kilobase of transcript per million mapped reads (FPKM) were calculated using a custom Python script. The script is available at <https://github.com/QilabGitHub/FPKMcalculation>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

We have included a "data availability" statement in the manuscript. Whole-transcriptome sequencing data can be accessed in Gene Expression Omnibus

under the accession code GSE166817. Key constructs and plasmids will be available on Addgene (https://www.addgene.org/Stamley_Qi/). Source data are provided with this study. All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes (3-6 for ex vivo experiments and in vivo experiments) were chosen to be consistent with previously published literature in genome engineering experiments (PMID: 32027839, 31937940)
Data exclusions	No data was excluded.
Replication	All attempts at reproducibility succeeded, as defined by (at a minimum) two or three positive results.
Randomization	Randomization was not used in this project. For intravitreal injection experiments, internal control was used for each mouse to minimize covariates: right eye was injected with WT Cas12a, and left eye was injected with hyperCas12a. For subretinal delivery experiments, experimental and control mice were age-matched, and each litter (or several litters) was allocated into different treatment groups; no covariates were controlled; most images were taken at the boundary of the electroporated area (see Fig 4) so that the tissue without GFP (marker for electroporation) can serve as an internal non-treated control.
Blinding	Quantitation of fluorescence intensities in microscopy images with semi-automated analysis pipeline was carried out by a blinded operator, with limited involvement of the experimenter.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Pax6 (Thermo Fisher, Rabbit polyclonal, 42-6600) 1:500 (in vivo), RBPMS (PhosphoSolutions, Guinea Pig polyclonal, 1832-RBPMS) 1:4000 (in vivo), HA (Roche, Rat HA clone 3F10, 11867423001) 1:100 (in vivo) 1:200 (in vitro), Oct4 (BD bioscience, mouse clone 40/Oct-3, 611203) 1:100 (in vivo), 1:200 (in vitro), Sox2 (Cell signaling, rabbit polyclonal, 14962) 1:100 (in vivo), 1:200 (in vitro), Klf4 (R&D, goat polyclonal, AF3158) 1:100 (in vivo), 1:200 (in vitro), Donkey anti-rat Cy3 (Jackson ImmunoResearch, 711-166-152) 1:1000 (in vivo), Donkey anti-goat Cy5 (Jackson ImmunoResearch, 705-175-147) 1:1000 (in vivo), Donkey anti-rabbit Cy5 (Fisher Scientific, NC0254454) 1:1000 (in vivo), Donkey anti-mouse Cy5 (Jackson ImmunoResearch, 715-605-151) 1:1000 (in vivo), Donkey anti-mouse Alexa Fluor 488 (Thermo Fisher, R37114) 1:500 (in vitro), Donkey anti-rabbit Alexa Fluor 647 (Thermo Fisher, A-31573) 1:500 (in vitro), Donkey anti-goat Alexa Fluor 647 (Thermo Fisher, A32849) 1:500 (in vitro)
Validation	The antibodies undergo standard QC testing by the manufacturers. Validation per manufacturer website: Oct4: https://www.bdbiosciences.com/eu/applications/research/stem-cell-research/cancer-research/human/purified-mouse-anti-oct-34-40oct-3/p/611203 Sox2: https://www.cellsignal.com/products/primary-antibodies/sox2-d1c7j-xp-rabbit-mab/14962 Klf4: https://www.rndsystems.com/products/mouse-klf4-antibody_af3158 HA: https://www.sigmaaldrich.com/catalog/product/roche/roahaha?lang=en&region=US

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HEK293T (ATCC) and derivative cells, mouse P19 cells (ATCC)
Authentication	Cell lines were not authenticated
Mycoplasma contamination	Cell lines were not tested for mycoplasma
Commonly misidentified lines (See ICLAC register)	None

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Wild-type neonatal mice were obtained from timed pregnant CD1 mice (Charles River Laboratories, #022). Thy1-YFP-17 transgenic mice were obtained from collaborators as described per Methods. Only male mice were used, at age 7-9 weeks. All mice were housed in standard cages on a 12-hour light-dark cycle.
Wild animals	We did not use any wild animals in this study.
Field-collected samples	We did not use any field-collected samples in this study.
Ethics oversight	All animal studies were approved by the Institutional Animal Care and Use Committee at Stanford School of Medicine.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For flow cytometry, cells (HEK293T or mouse P19 cells) were dissociated using 0.05% Trypsin-EDTA (Life Technologies), resuspended in PBS+10% FBS.
Instrument	CytoFLEX S flow cytometer (Beckman Coulter).
Software	FlowJo software (BD Biosciences) was used for analysis
Cell population abundance	10,000 cells from the population of interest (for most experiments, mCherry+ and BFP+ gated based on non-transfected control) were collected for each sample
Gating strategy	FSC / SSC determined by bulk population and comparison to other cells. Singlet gate, followed by appropriate fluorescent gates (detection of plasmid presence, if available). Please refer to Extended Figure 1 for gating strategy.
☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.	