HHS Public Access

Author manuscript

Nat Cell Biol. Author manuscript; available in PMC 2022 October 12.

Published in final edited form as:

Nat Cell Biol. 2022 April ; 24(4): 590–600. doi:10.1038/s41556-022-00870-7.

Multiplexed genome regulation in vivo with hyper-efficient Cas12a

Lucie Y. Guo[1,2,*,†], **Jing Bian**[1,*], **Alexander E. Davis**[2], **Pingting Liu**[2], **Hannah R. Kempton**[1], **Xiaowei Zhang**[1], **Augustine Chemparathy**[1], **Baokun Gu**[1], **Xueqiu Lin**[1], **Draven A. Rane**[1], **Ryan M. Jamiolkowski**[1], **Yang Hu**[2], **Sui Wang**[2,†], **Lei S. Qi**[1,3,4,†]

[1]Department of Bioengineering, Stanford University, Stanford, CA 94305, United States

[2]Department of Ophthalmology, Stanford University, Stanford, CA 94305, United States

[3]Department of Chemical and Systems Biology, Stanford University, Stanford, CA 94305, United States

[4]Stanford ChEM-H Institute, Stanford University, Stanford, CA 94305, United States

Abstract

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 crRNAs from a single transcript, however, their low efficiency has hindered applications *in vivo*. Through structure-guided protein engineering, we develop a hyper-efficient LbCas12a variant, termed hyperCas12a, with its catalytically dead version hyperdCas12a showing significantly enhanced efficacy for gene activation, particularly at low crRNA conditions. We demonstrate that hyperdCas12a has minimal off-target effects compared to the wildtype system and exhibits enhanced activity for gene editing and repression. Delivery of the hyperdCas12a-activator and a single crRNA array simultaneously activating endogenous *Oct4, Sox2,* and *Klf4* genes in the retina of postnatal 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.

[†]Correspondence to: L.Y.G. (lucieguo@stanford.edu), S.W. (suiwang@stanford.edu), L.S.Q. (stanley.qi@stanford.edu).
[*]Co-first authors

Author Contributions

L.Y.G. and L.S.Q. conceived of the idea. 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.R., and R.M.J. performed *ex vivo* experiments. L.Y.G., J.B., A.E.D., and P.L. performed *in vivo* experiments with guidance from S.W. and Y.H. L.Y.G., J.B., S.W., and L.S.Q. analyzed the experimental data. A.C. and X.L. performed computational analysis of sequencing data, and X.Z. and R.M.J. analyzed imaging data. L.Y.G. and L.S.Q. wrote the manuscript with input from all authors.

Reporting summary.

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

Code availability

Gene-level fragments per kilobase of transcript per million mapped reads (FPKM) were calculated using a custom Python script available at https://github.com/QilabGitHub/FPKMcalculation. Semi-Fluorescence intensities of individual cells for microscopy were quantitated with a semi-automated image analysis pipeline based on MATLAB (version R2019a) available at https://github.com/QilabGitHub/dCas12a-microscopy.

Competing Interests Statement

The authors have filed a provisional patent via Stanford University related to the work (US patent# 63/148,652). L.S.Q. is a founder and scientific advisory board member of Epicrispr Biotechnologies.

Keywords

CRISPR; Cas12a; hyperCas12a; CRISPRa; CRISPRi; gene editing; off-target; in vivo; multiplexed; gene regulation; retina differentiation

INTRODUCTION

CRISPR-Cas nucleases and their nuclease-deactivated dCas variants have revolutionized the field of genome editing and regulation[1,2]. 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, including *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 thus can process multiple functional 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 screening[8].

Despite these advantages, Cas12a has not been adopted *in vivo* as readily as its Cas9 counterpart, in part due to less ideal editing or regulation efficiency and more variable target-dependent indel efficiencies compared to Cas9[9-11]. As the wildtype Cas12a exhibits a more restricted protospacer adjacent motif (PAM) requirements (TTTV) compared to Cas9 orthologs[5] (e.g., *Streptococcus pyogenes* Cas9, NGG)[12], there have been efforts to engineer Cas12a variants[6,8,10,13-19] with expanded applications in combinatorial screening[8] and *ex vivo* cell therapy[15].

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 performance of Cas systems *in vivo* is often linked to the low copy numbers of Cas and crRNA molecules[20,21]. We speculate that utilizing Cas12a *in vivo* requires optimizing its performance under restricting molecular concentrations. To address this key challenge, here we apply synthetic biology to engineer Cas12a variants with higher performance for diverse genome engineering applications and demonstrate its utility for *in vivo* multiplexed gene modulation in mammalian retina.

RESULTS

Development of hyperdCas12a for enhanced CRISPR activation

We focused on LbdCas12a, since previous studies showed that the wildtype LbdCas12a-fused transcription activator (LbdCas12a-VPR) achieved ~5-fold higher performance than the wildtype AsdCas12a-VPR for single gene activation[5,13]. To enhance the activity of the native LbdCas12a, we adopted structure-guided protein engineering. We focused on negatively charged (aspartate or glutamate) residues of LbdCas12a that reside within 10Å of the target DNA (PDB 5XUS)[22] 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[5,23] to increase targeting range, we tested residues both proximal and distant to the PAM (Fig. 1a). To systematically characterize the library of variants for their performance, we engineered a HEK293T cell line with genomically integrated TRE3G-GFP reporter[24] (Fig. 1b). Co-transfection of plasmids encoding the mutated dCas12a variants fused to a miniaturized VPR (miniVPR)[25] and the TRE3G promoter-targeting crRNA into reporter cells allowed quantitative comparison of the performance of variants through flow cytometry (Extended Data Fig. 1a-d).

While many single mutations worsened dCas12a-mediated gene activation, a few mutants (D122R, E125R, D156R, E159R, D235R, E257R, E292R, D350R, E894R, D952R, and E981R) exhibited enhanced activation compared to the wildtype (WT) dCas12a (Fig. 1c and Extended Data Fig. 1e-f). While WT dCas12a produced ~383x-fold change over non-targeting control, the single D156R mutation enabled >700-fold activation, and several other single mutations enabled ~400 to 600-fold activation. We next examined the performance of these variants after gating for a low BFP cell population, which serves as a proxy for cells with lower crRNA concentration that is particularly relevant for *in vivo* applications where there would be less of the crRNA-Cas12a complex compared *in vitro* (Fig. 1d). Among the low BFP cell population, WT dCas12a exhibited a significant 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, while several others enabled 90-to 200-fold activation (Fig. 1d).

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

We note that D156 of LbCas12a is homologous to E174 on AsCas12a (Extended Data Fig. 2h-i) which is one of the mutated residues in the previously published enhanced AsCas12a[5], while the other 3 mutations have not been reported to date. We also tested combinations of homologous mutations based on enAsdCas12a (E174R/S542R/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 to 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. Here, GFP activation

using WT dCas12a was reduced using a CAG promoter-driven crRNA compared to a U6 promoter-driven crRNA (compare GFP fluorescence of WT in Fig. 1c vs. Fig. 1f), but the single and combinatorial mutants significantly enhanced the level of activation. Notably, the quadruple mutant (D156R/D235R/E292R/D350R) achieved the highest level of activation, ~60-fold above the level achieved by the WT protein (Fig. 1f, **left**). We then tested the mutants in a condition with limiting crRNA quantity (crRNA:dCas12a ratio of 0.2:1). The quadruple mutant outperformed all other mutants, at >300-fold above the level achieved by the WT protein (Fig. 1f, **right**). Additionally, it also outperformed enAsdCas12a at low-crRNA concentration, for CAG-driven crRNA and for a dual-crRNA array (Extended Data Fig. 2k-m). We heretofore refer to this quadruple mutant as “hyper-efficient dCas12a” (hyperdCas12a) for further characterization and *in vivo* gene targeting.

Even though our mutagenesis focused on increasing efficiency (instead of broadening targeting range as in previous studies[5,23]), 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 showed that hyperdCas12a outperformed WT dCas12a for all 3 canonical PAMs (TTTA, TTTC, TTTG) as well as several of the non-canonical PAMs (TTTT, CTTA, TTCA, TTCC) (Fig. 1g). Since out of the 4 mutated residues of hyperdCas12a, only the D156R mutation is proximal to the PAM (Fig. 1a), it is logical that several of these PAMs are also accessible by the homologous E174R mutant of AsdCas12a[5], and that the PAM range of hyperdCas12a may be stricter than that of enAsdCas12a[5].

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 SV40-GFP cassette[1], fusing hyperdCas12a to a KRAB repressor domain with crRNA targeting the SV40 promoter enabled improved repression over non-targeting control, compared to its wildtype equivalent (Fig. 2a-b). This suggests that hyperdCas12a can be modularly coupled to effectors and exhibit enhanced effects for transcriptional and epigenetic modulation.

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 indel formation frequency at each nucleotide position showed that editing by hyperdCas12a peaked around positions 18-23 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 study[26] (265 bp), a truncated WPRE (245 bp)[27], and a small synthetic poly-A tail (49 bp) (Fig. 2e). In transgenic mice expressing Thy1-YFP[28], we co-delivered AAV-hyperCas12a by intravitreal injection along with AAV-crRNA (YFP) in one eye, and its wildtype counterpart in the contralateral eye as a side-by-side control (Fig. 2f). For all mice tested, hyperCas12a showed improved YFP knockout compared to WT Cas12a (Fig 2g-i). Despite using minimal versions of all regulatory elements, the AAV containing hyperdCas12a (4743 bp) teetered on the AAV packaging limit (~4.7bp); by being 234bp

larger, enAsdCas12a exceeded this limit (Fig. 2e). This highlights the utility of hyperCas12a for enhanced AAV-based *in vivo* gene-editing.

CRISPR activation by hyperdCas12a is specific

Cas12a is known to be highly precise in human cells[29], with likely higher specificity compared to Cas9[9,13]. To evaluate the specificity of CRISPR activation by hyperdCas12a compared to WT Cas12a on a genome-wide scale, we carried out whole-transcriptome 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 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 to the WT dCas12a in Fig. 1c (Fig. 2j). Comparing the targeting vs. non-targeting crRNAs, both WT dCas12a and hyperdCas12a showed similar specificity (Fig. 2j, Supplementary Table 4). Two biological replicates were analyzed 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 (~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 hr (Fig. 3a), which resulted in ~90% cells that contain 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 endogenous transcription factor genes *Oct4 (also known as Pou5f1), Sox2*, and *Klf4*, given their known synergistic regenerative role in multiple contexts[30,31]. 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 cells[32]. We performed immunostaining to visualize target protein expression 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 by using paired crRNAs (even though target sequences for *Klf4* crRNAs were >500 bp apart) and further synergy in *Sox2* activation by using a “triplet” of three separate *Sox2* crRNAs (Extended Data Figs. 6-7). In contrast, we did not achieve synergy with paired crRNA for *Oct4*, possibly due to higher basal level 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 vs. hyperdCas12a. All crRNAs tested, including paired and triplet crRNAs, exhibited enhanced activation using hyperdCas12a compared to WT dCas12a (Fig. 3b-d).

HyperdCas12a drives enhanced multiplex activation

Since *Oct4, Sox2*, and *Klf4* are known to work synergistically, there is strong rationale for their multiplexed activation[33]. To simultaneously activate these three targets, we generated a single crRNA array driven by the U6 promoter encoding 6 crRNAs (Fig. 3e), based on our crRNA screening in P19 cells. The crRNA combination that achieved highest activation in 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 Fig. 6-8). With dCas12a (D156R) and a double mutant (D156R + E292R), we achieved significantly enhanced activation than WT dCas12a, and further enhancement with hyperdCas12a which achieved ~5-fold activation of *Oct4*, ~8-fold activation of *Sox2*, and ~70-fold activation of *Klf4* (Fig. 3f-h) and which also outperformed enAsdCas12a (Extended Data Fig. 9). Interestingly, with hyperdCas12a, we observed compelling Oct4 activation in P19 cells despite its location as the 6th crRNA, despite prior studies with WT dCas12a showing decreased expression of crRNAs at and beyond the 4th position in the crRNA array[4,34,35]. We note decreased activation of each target gene compared to the level achieved by single crRNAs (compare Fig. 3f to Fig. 3b-d), consistent with previous observations[35,36], possibly due to decreased copies of the longer pre-crRNA array expressed by the U6 promoter compared to shorter individual crRNAs. Nevertheless, hyperdCas12a performed robustly in using a single CRISPR array to activate multiple endogenous targets. Additionally, 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[37-39], 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 vivo*[37]. We constructed single plasmids consisting of hyperdCas12a with an optimized nuclear-targeting sequence (NLS) (Extended Data Fig. 2a-d) and single crRNAs to either *Sox2, Klf4*, or *Oct4* and validated its ability to drive *in vivo* CRISPR activation of single gene targets at 21 days after in vivo electroporation at postnatal day 0 (P0) (Fig. 4a-d, Extended Data Fig. 10a-d), whereas non-targeting crRNA did not show significant activation (Extended Data Fig. 10e). The CAG-GFP plasmid was co-electroporated to serve as the electroporation efficiency control. Within the electroporated GFP+ patches in the retina, we observed numerous HA+ cells, indicating successful delivery and expression of hyperdCas12a (Fig. 4).

Overexpression of *Sox2, Oct4* and *Klf4* individually have been shown to direct the differentiation of retinal progenitor cells[40-42] (RPCs) towards specific fates, and the synergistic co-activation of three ectopic transcription factors can induce formation of iPSCs *in vitro*[43] and rejuvenate mature retinal ganglion cells *in vivo*[31]. We tested whether our hyperdCas12a system can synergistically activate endogenous *Sox2, Klf4* and *Oct4* in postnatal RPCs in vivo, and whether this manipulation affects the 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* (Fig. 4e, 5a) and *Klf4* (Fig. 4e, 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), enAsdCas12a (Fig. 5g-i, j-l), 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 have received the hyperdCas12a and poly-crRNA array plasmid. The in vivo electroporation technique delivers DNA mainly to mitotic cells, and at postnatal day 0, mitotic RPCs give rise to rod photoreceptors, Müller glia, and bipolar and amacrine neurons[44], which migrate to and reside in the ONL (outer nuclear layer) or INL (inner nuclear layer), 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 GCL and inner plexiform layer (IPL). This effect was less pronounced with enAsdCas12a and was almost absent with WT dCas12a (Fig. 6a-b). Furthermore, the enrichment of HA+ cells in GCL and IPL was absent when hyperdCas12a was used to activate individual *Sox2, Klf4*, or *Oct4*, suggesting the importance of simultaneous multiplexed gene activation for phenotypic functions (Fig. 6a-b, Extended Data Fig. 10a-d).

In most of the HA+ cells that migrated into GCL, we observed expression of Pax6 (marker for displaced amacrine and ganglion cells in 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-like and ganglion-like cells and drive their migration into the GCL and IPL, and support the conclusion that hyperdCas12a can activate multiple endogenous genes to alter phenotypes *in vivo*.

DISCUSSION

In this work, we developed an optimized *Lachnospiraceae bacterium* Cas12a variant termed “hyper-efficient Cas12a” (hyperdCas12a) for *in vivo* CRISPR multiplexed genome modulation. HyperdCas12a enables simultaneous activation of endogenous targets in postnatal retina and altered the differentiation of retinal progenitor cells. HyperdCas12a outperforms 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 improves CRISPR-mediated gene repression, and its nuclease-active version (hyperCas12a) achieves more effective gene editing *in vivo* in retinal ganglion cells. Additionally, hyperdCas12a-mediated endogenous gene activation outperformed WT dCas12a or enAsdCas12a *in vivo*.

The enhanced dCas12a system is broadly useful both for *in vitro* and *in vivo* applications that require multiple genetic manipulations, which is currently limited using existing tools. Our system enables simultaneous modulation at multiple genomic loci, thus paving the way

for CRISPR-based modulation of multiple pathways or synergistic targets, such as in the case of non-monogenic diseases which consist of 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 desired phenotypic changes would depend on context. For example, CRISPRa has been useful as a tool to rescue haploinsufficiency by rescuing the expression of the endogenous functional allele, in which only a 2-fold increase is sufficient to reverse the disease phenotype[45]. In other instances, especially for directing cell differentiation, higher levels of activation may be needed to induce 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 pluripotency[32]. 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 multi-gene modulation efficiency (Fig. 6g). While the Cas12a system has other intrinsic limitations (e.g., uneven array processing[6,46]) 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[6,10,19,47] to achieve even greater potential of the Cas12a system.

CRISPR-mediated transcriptional activation of endogenous genes at their native chromatin loci would minimize 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. Additionally, targeted endogenous gene regulation may also remodel chromatin status and epigenetic modifications at endogenous loci and lead to persistent gene regulation effects compared to overexpression of corresponding cDNA[48-50]. Also, 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 activate potentially all isoforms. CRISPR activators can also be useful in cases of large cDNAs[51,52] (or in the future, for large intergenic noncoding RNAs) that may not fit in 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.

METHODS

Cell culture

HEK293T cells (ATCC) and the reporter cell line[24] with stable expression of TRE3G-GFP were cultured in DMEM + GlutaMAX (Thermo Fisher) supplemented with 10% FBS (Alstem) and 100U/mL of penicillin and streptomycin (Life Technologies). P19 cells (ATCC) were cultured in alpha-MEM with nucleosides (Invitrogen) with same FBS and pen/strep as above. Cells were maintained at 37°C and 5% CO2 and passaged using standard cell culture techniques. For transient transfection of HEK293T cells, cells were seeded the day before transfection at $1x10^5$ cells/mL. Transient transfections were performed using 3 mL of

TransIT-LT1 transfection reagent (Mirus) per mg of plasmid. HEK293T cells were analyzed 2-5 days post transfection, as specified by each experiment. For transient transfection of P19 cells, cells were seeded the day before transfection at density of $2x10^5$ cells/mL. Transient transfections were performed using 3 μl of Mirus X2 transfection reagent (Mirus) per μg of plasmid. For double-selection, cells were treated with 500 ug/ml of hygromycin and 2 μg/ml of puromycin. P19 cells were analyzed 3 days post transfection, as indicated.

Plasmid cloning

Standard molecular cloning techniques were used to assemble constructs in this paper. Nuclease-dead dCas12a from *Lachnospiraceae* bacterium and its crRNA backbone were modified from previous publication[24]. The step-by-step protocol for generating crRNA arrays is provided at Nature Protocol Exchange[53].

Flow cytometry

Cells were dissociated using 0.05% Trypsin-EDTA (Life Technologies), resuspended in PBS+10% FBS, and analyzed for fluorescence using a CytoFLEX S flow cytometer (Beckman Coulter). 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 and analyzed using FlowJo (v10).

Targeted deep sequencing and data analysis

Targeted deep sequencing primers were designed to amplify a total 223-bp amplicon encompassing the crYFP target with generic adaptors. YFP-R, GACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGGTGGTGCAGATGAACTTCAGG; YFP-F, ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTCATGGTGAGCAAGGGCG. DNA was extracted after PCR and gel purification with Qiagen Gel Extraction kit. Samples were pooled in equal amounts and a mixed barcoded library sequenced by GENEWIZ Amplicon-EZ sequencing service (GENEWIZ). More than 350,000 reads were generated with each sample using Illumina platform. Data analysis was performed with CRISPResso2[53].

qRT-PCR

RNA was isolated from transfected cells using Qiagen RNeasy plus kit (Qiagen) followed by reverse transcription of 100 ng RNA into cDNA using iScripst kit (Biorad). A Quantitative real-time PCR (qRT-PCR) reaction was performed using SYBR master mix (Biorad) according to the manufacturer's protocol (primers used are indicated in the supplementary information (qPCR primer list)). Quantification of RNA expression was normalized based on expression of glyceraldehyde 3-phosphate dehydrogenase and calculated using ΔΔCt. Please refer to Supplementary Table 3 for qPCR primers.

Immunostaining and quantitation

P19 cells were seeded onto black flat-bottom 96-well plates at 48hr after transfection (continuing in dual selection media), fixed with 1xDPBS/4% formaldehyde 24hr after

seeding. Each well was permeabilized with 1x DPBS/0.25% Triton X-100 and blocked with 1x DPBS/5% donkey serum, then incubated at 4C overnight with primary antibodies diluted in 1x DPBS/5% donkey serum: mouse anti-Oct4 (1:200, BD bioscience, 611203), rabbit anti-Sox2 (1:200, Cell signaling, 14962), and goat anti-Klf4 (1:200, R&D system, AF3158). Each well was washed 3x with 1xDPBS then incubated for 1hr with Alexa Fluor-conjugated 488 or 647 donkey secondary antibodies (Life Tech) at 1:500 diluted in same buffer as primary antibodies. Each well was then washed 3x with 1xPBS, and each well is immersed in 1xPBS in each well. No nuclear dye was used. Imaging was done with Leica DMi8 inverted microscope with 20x objective and a Leica DFC9000 CT camera. Cell fluorescence intensities were quantitated with 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, mean fluorescent intensities of all corresponding channels for every cell were collected for further statistical analysis. For display, one representative image was used per condition, with roughly equal number of cells between conditions to facilitate qualitative comparison.

RNA-seq

HEK reporter cell line stably expressing TRE3G-GFP were seeded in a 6 well plate at density of $2*10^5$/ml and were co-transfected next day with Tet crRNA or LacZ non-target crRNA with dCas12aWT or hyperdCas12a, in duplicates. One day after transfection, transfected cells were placed in antibiotic selection (hygromycin 500μg/ml and puromycin 2μg/ml) for 2 days before harvest. Total RNA was isolated by using RNeasy Plus Mini Kit (QIAGEN). Library preparation and next-generation sequencing were performed by Novogene (Chula Vista, CA) as described previously[30]. Spliced Transcripts Alignment to a Reference (STAR) software[55] was used to index hg19 genome and GFP sequence, and then to map paired end reads to the genome. HTSeq-Count[54] 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 available at https://github.com/QilabGitHub/FPKMcalculation. Pearson correlation coefficients were obtained using QR decomposition and regression.

Animals

For in vivo electroporation experiments, wild-type neonatal mice were obtained from timed pregnant CD1 mice (Charles River Laboratories). For AAV experiments, Thy1-YFP-17 transgenic mice were originally generated by Drs. Guoping Feng and Josh Sanes[28] and were acquired from Dr. Zhigang He; male mice age 6-8 weeks 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

AAV2s were produced by AAVnerGene (North Bethesda, MD) using previously described approaches[26]. AAV titers were determined by real-time PCR. AAV-Cas12a and AAV-crYFP were mixed at a ratio of 2:1. AAV-Cas12a was diluted to 4.5×10^{12} vector genome

(vg)/ml and AAV-crYFP was diluted to 2.25 x 10^{12}. For intravitreal injection, mice were anesthetized by xylazine and ketamine based on their body weight (0.01 mg xylazine/g + 0.08 mg ketamine/g). 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 x 10^9 vg/retina of Cas12a and 4.5 x 10^9 vg/retina of crYFP. Mice were sacrificed 10 weeks after AAV injection. Transcardiac perfusion was performed as described[26]. For retina wholemount, 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 publications[56] and used at 1:4000, and rat HA (clone 3F10, 1:200, Roche) was diluted in the same staining buffer. Floating retinas were incubated with primary antibodies overnight at 4°C and washed three times for 30 min each with PBS. Secondary antibodies (Cy2, Cy3, or Cy5 conjugated) were then applied (1:200; Jackson ImmunoResearch) and incubated for 1 h at room temperature. Retinas were again washed three times for 30 min each with PBS before a cover slip was attached with Fluoromount-G (SouthernBiotech). Quantitation of 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 analysis on mouse retina wet mount, threshold-based segmentation was performed based on the fluorescent channel representing crRNA, which had highest signal-to-noise ratio and distributes evenly throughout the cytoplasm. Morphological operations were then applied to remove noise and thus yields masks for single cells. Based on the masks, mean fluorescent intensities of all corresponding channels for every cell were collected for further statistical analysis.

In vivo plasmid electroporation

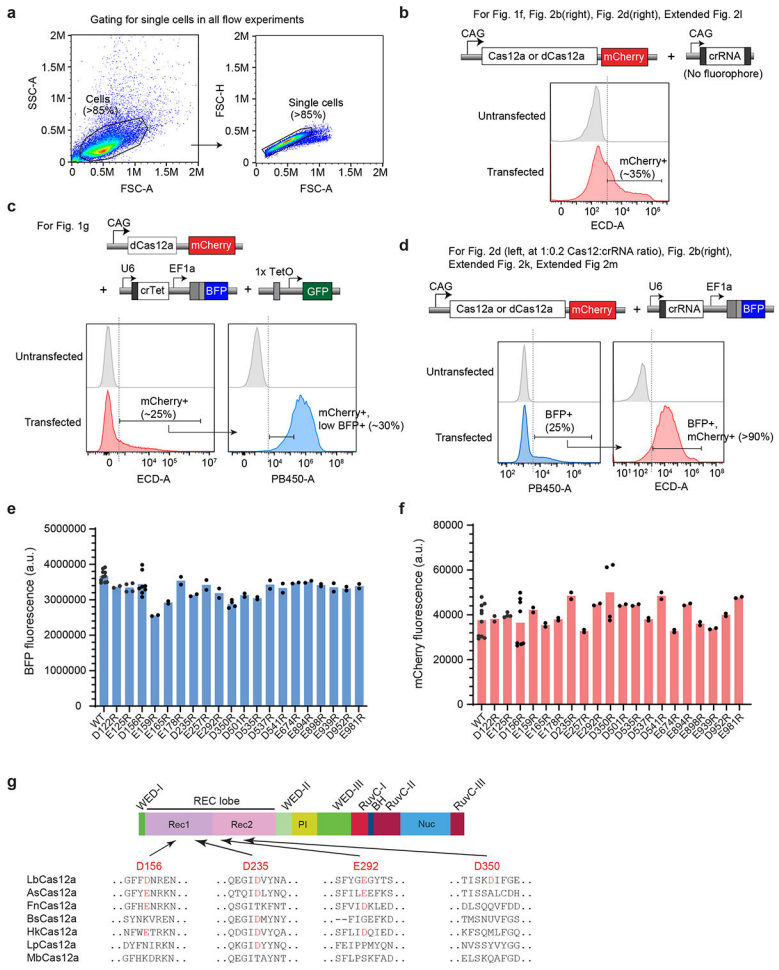
Plasmid DNA was injected into the subretinal space of neonatal mice, and electrical pulses were applied with tweezer-style electrodes as described[37,38]. More details are provided at Nature Protocol Exchange[53]. Plasmid with wildtype dCas12a was mixed with CAG-GFP construct in ~5:1 ratio and electroporated at a concentration of up to 2μg/μl 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 described[39]. Eyeballs were fixed in 4% paraformaldehyde (PFA) in 1×PBS (pH 7.4) for 2hr at room temperature. 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 hr; 1:1 mixed solution of OCT and 30% sucrose in PBS, 4°C, overnight), frozen and stored at −80°C. A Leica CM3050S cryostat (Leica Microsystems) was used to prepare 20 μm cryosections. Retinal cryosections were washed in 1× PBS briefly, incubated in 0.2% Triton, 1× PBS for 20 min, and blocked for 30 min in blocking solution of 0.1% Triton, 1% bovine serum albumin and 10% donkey serum (Jackson ImmunoResearch Laboratories) in 1× PBS. Slides were incubated with primary antibodies diluted in blocking solution in a humidified chamber at room temperature at 4°C overnight. After washing in 0.1% Triton 1× PBS three times, slides were incubated with secondary antibodies and DAPI (Sigma-Aldrich; D9542) for 1-2 hr, washed three times with 0.1% Triton, 1× PBS and mounted in Fluoromount-G (Southern Biotechnology Associates). Primary antibodies for Oct4, Sox2 and Klf4 are as described in above “immunostaining” section. Additional primary antibodies used were rat anti-HA

(Roche; 3F10), guinea pig anti-RBPMS (PhosphoSolutions; 1832), and rabbit anti-Pax6 (Thermo; 42-6600). Retinal slices were imaged with the LSM 710 Confocal inverted laser scanning microscope, with Plan Apochromat objective 40x.1.4 Oil (FWD=0.13mm) with 405, 488, 561 and 633 lasers. Quantitation was performed as described[38] using Fiji software.

Statistics and Reproducibility

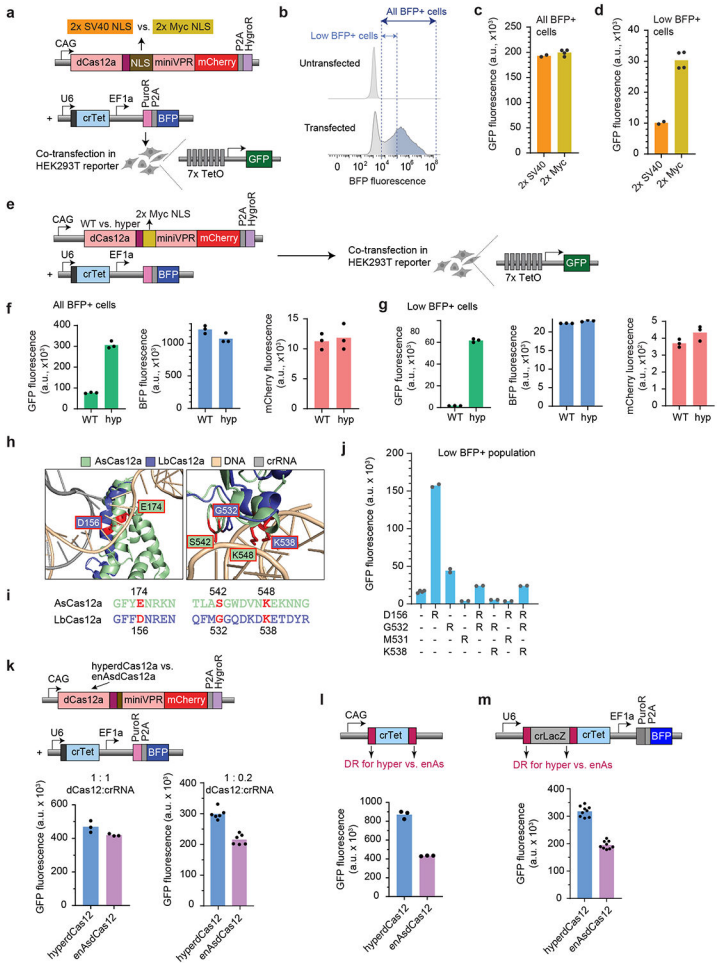
Most in vitro data are represented as bar graphs with individual data points and the bar at mean, without error bars. For *in vivo* experiments, data are presented as bar graphs of mean ± s.d. or box-and-whisker plots. The statistical analysis was based on sample size (n), indicating the number of biologically independent experiments or animals, as described in the respective figure legends. No statistical method was used to predetermine 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 Fig. 6-8), and the whole-transcriptome RNAseq (two independent biological replicates, in 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 analysis. Data collection and analysis were not blinded. Data were compared between groups of animals using two-tailed Student t-tests. Statistical tests were performed using Prism 9.1.0 (GraphPad Software). For RNA-seq analysis, Pearson correlation coefficients were obtained using QR decomposition and regression.

Extended Data



Extended Data Fig. 1. Gating strategies for flow cytometry

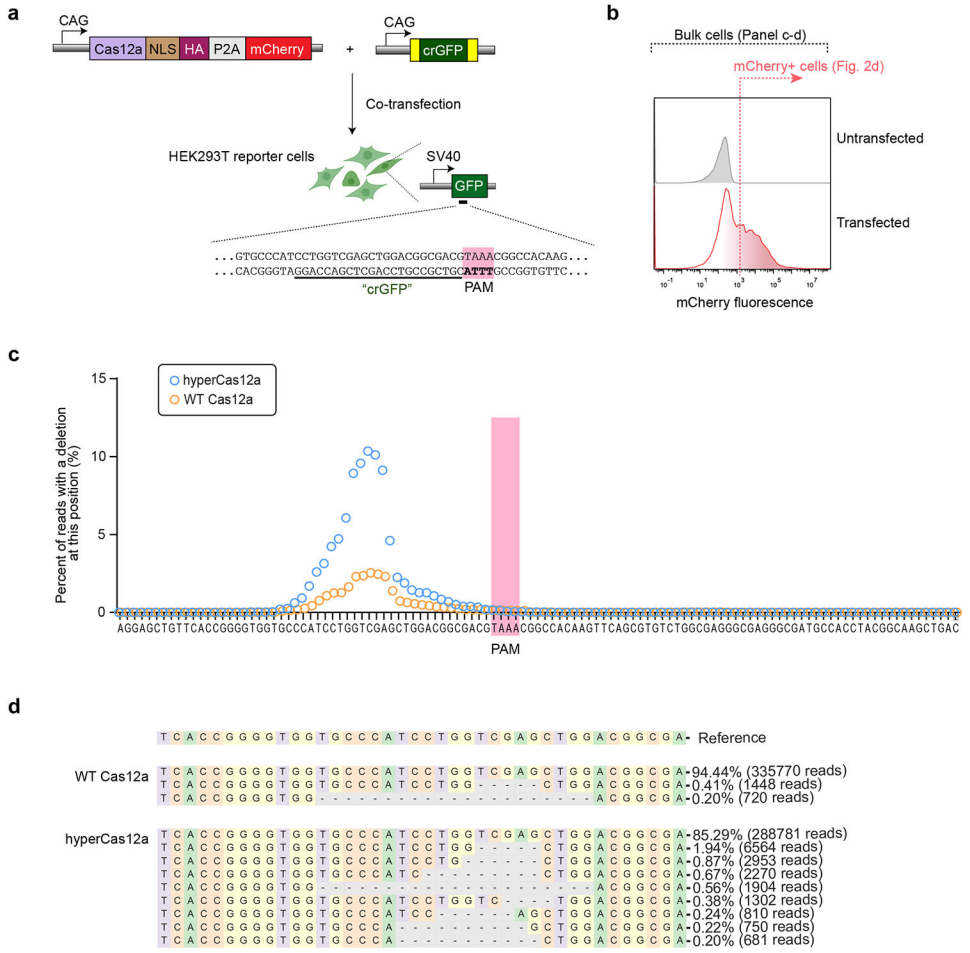
a, Standard strategy for gating cells based on forward scatter (FSC) and side scatter (SSA) (left), then further gating for singlets based on FSC-height (FSC-H) and FSC-area (FSC-A) (right). To analyze transfected cells, further gating is applied to the singlet population based on fluorescence intensity. Please note that for Fig. 1c-e, and Extended 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. Optimizing the nuclear localizing signal and comparing to enAsdCas12a

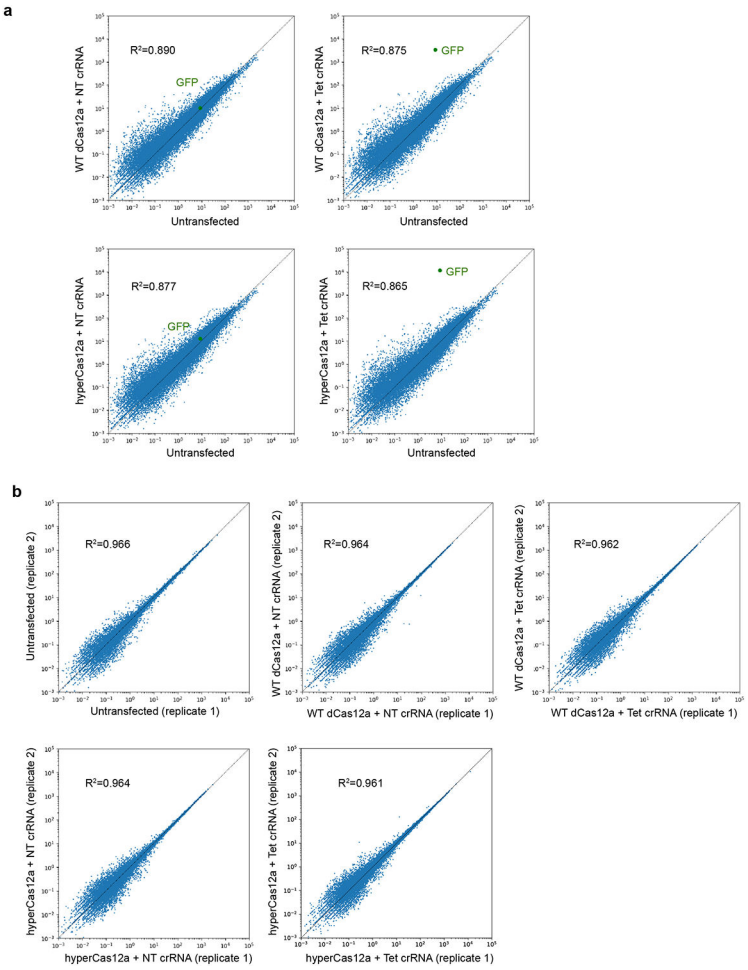
a. Schematic to test two different nuclear localization signals. Constructs containing either 2×SV40 or 2×Myc NLS fused with WT dCas12a are co-transfected with Tet crRNA in TRE3-GFP HEK293T reporter cells. **b**, Representative flow cytometry histogram of BFP intensity, showing threshold for BFP+ cells, 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 vs. hyperCas12a (hyp) with 2×Myc NLS, as well as BFP and mCherry average fluorescence in each gated BFP group. **h-i,** Alignment of the structure of LbCas12a vs. AsCas12a proteins (h) and alignment of peptide sequences (i) encompassing mutations harbored by enAsCas12a, a reported enhanced variant of Cas12a from Acidaminococcus with the E174R/S542R/K548R mutations[5] 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 AsCas12a[5]. **k,** Comparison of hyperdCas12a vs. enAsdCas12a with a single crRNA driven by U6 promoter in 1:1 vs. 1:0.2 ratio of dCas12:crRNA, in TRE3G-GFP HEK293T cells. **l,** Comparison of hyperdCas12a vs. enAsdCas12a with single crRNA driven by CAG promoter

flanked by direct repeats (DR) specific to LbCas12a vs. AsCas12a. **m,** Comparison of hyperdCas12a vs. 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≥3 independent experiments; bar graph in j shows the mean of n≥2 independent experiments; each data point represents value of an independent experiment.



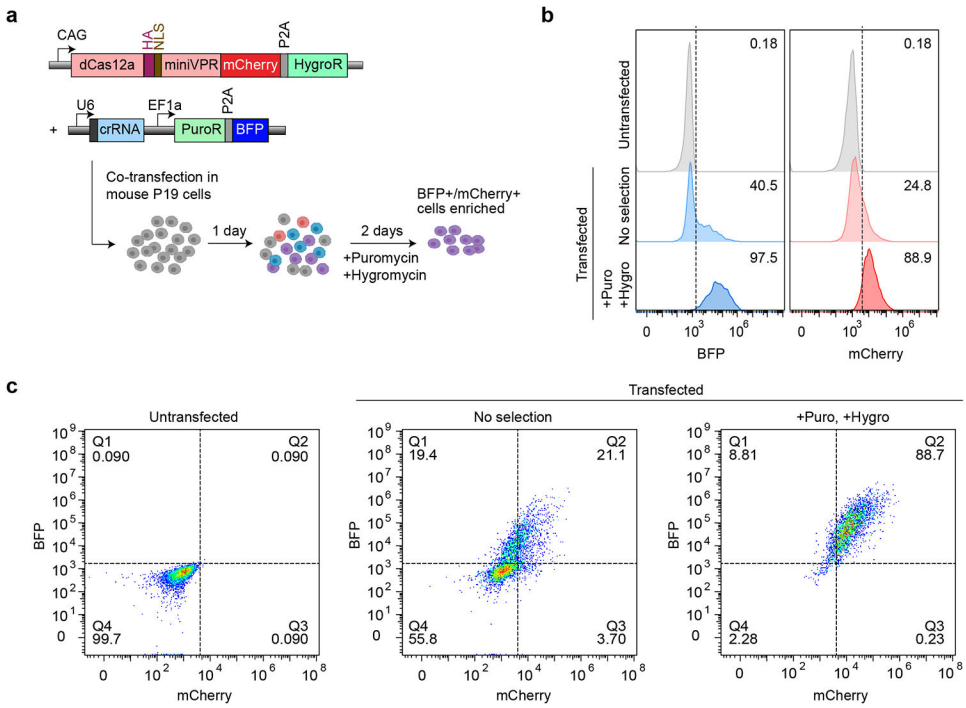
Extended Data Fig. 3. Improved gene editing by hyperCas12a

a. Nuclease-active WT Cas12a vs. 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 analyzed by CRISPResso2[52].



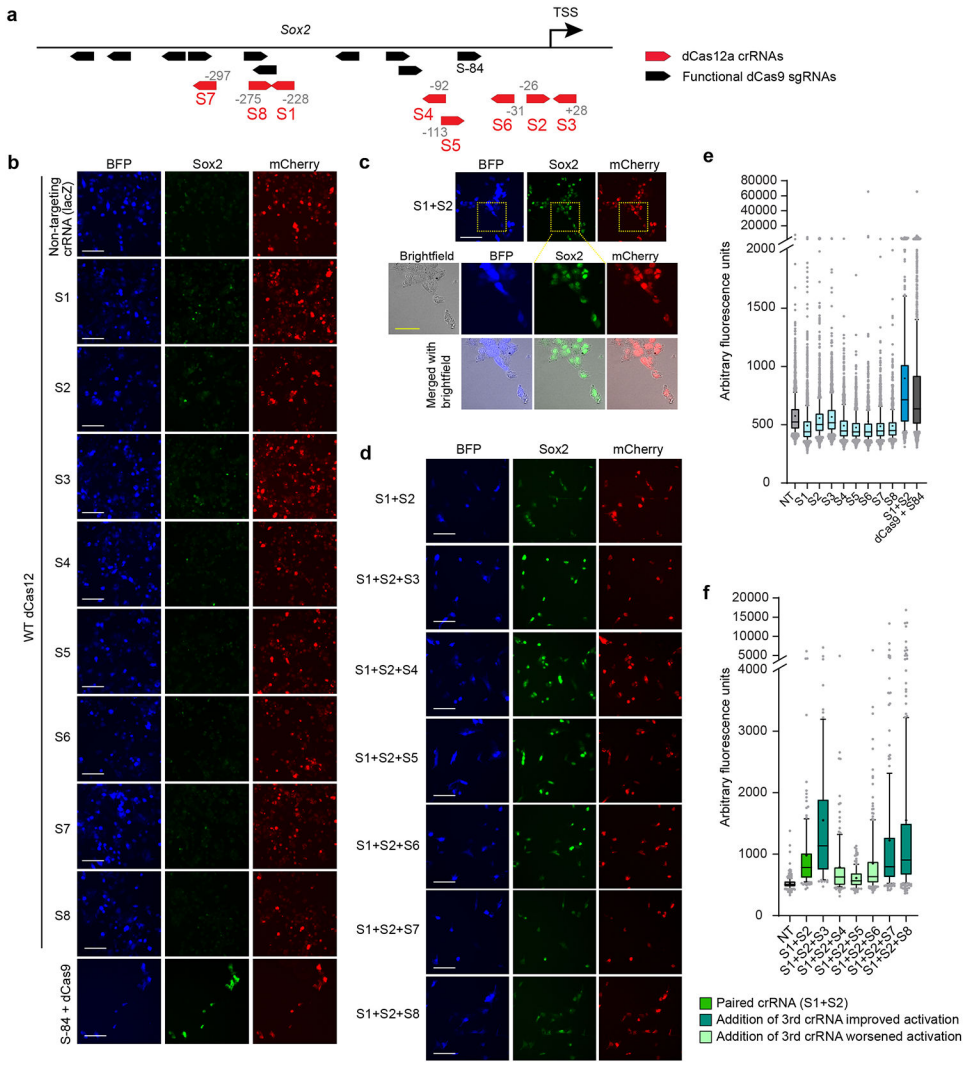
Extended Data Fig. 4. Characterization of off-target effects of hyperdCas12a

a. RNA-seq FPKMs (fragments per kilobase million fragments mapped) are plotted as transfected vs. 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-transfection in mouse P19 cells

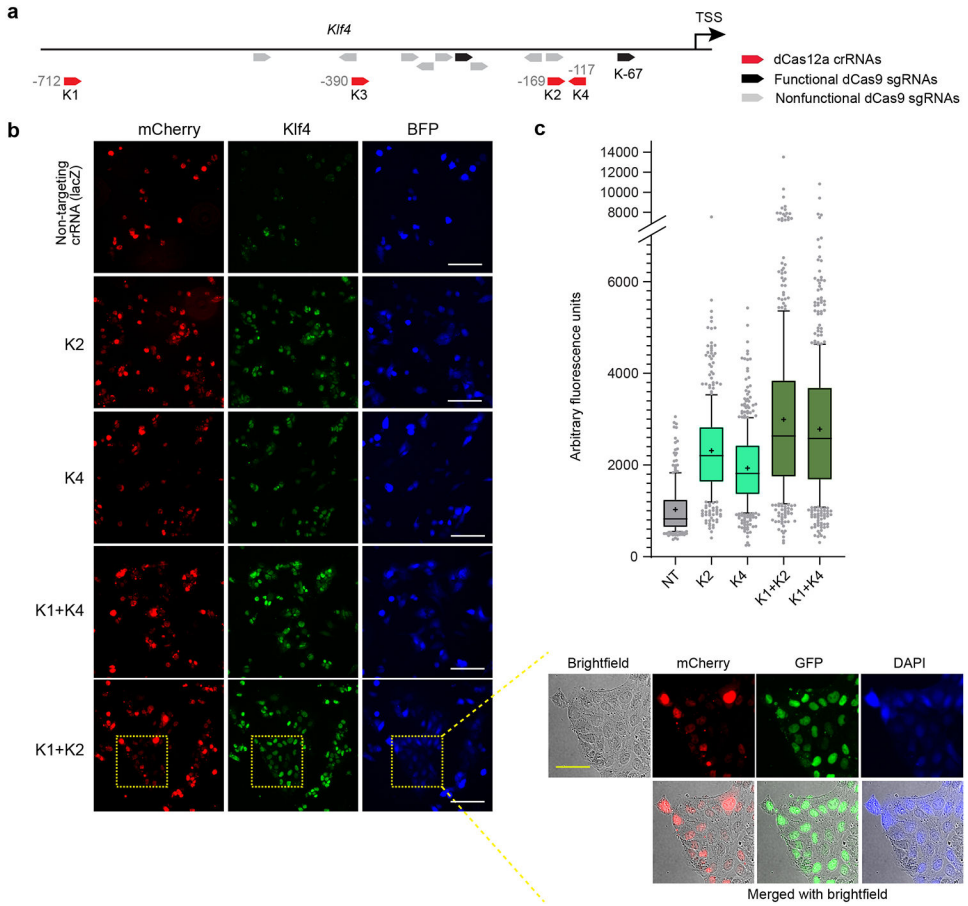
a, Mouse P19 cells were co-transfected with constructs expressing puromycin resistance (PuroR) and hygromycin resistance (HygroR), then selected with puromycin and hygromycin at 24 hr after transfection. Cells were collected for analysis 72 hr after transfection. **b,** Histograms showing percentage 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. Screening dCas12a crRNAs for activating endogenous *Sox2*

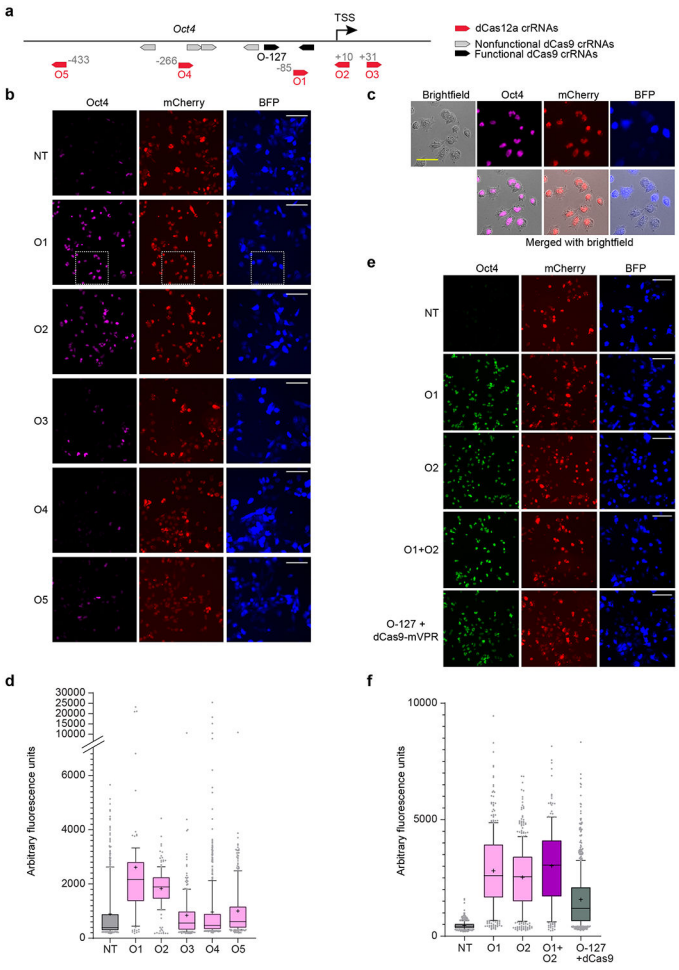
a, Schematic of dCas12a crRNAs (red) targeting promoter of *Sox2*, and their relative positions to validated dCas9 sgRNAs[32] that are functional (black) or non-functional (grey) for activating *Sox2*. Arrows indicate sense or antisense binding of crRNAs/sgRNAs to target DNA. The genomic position of 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)[32]. Scale bar, 100 μm. **c-d,** Immunostaining of Sox2 expression and colocalization 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 μm; yellow scale bar (within inset), 50 μm **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*

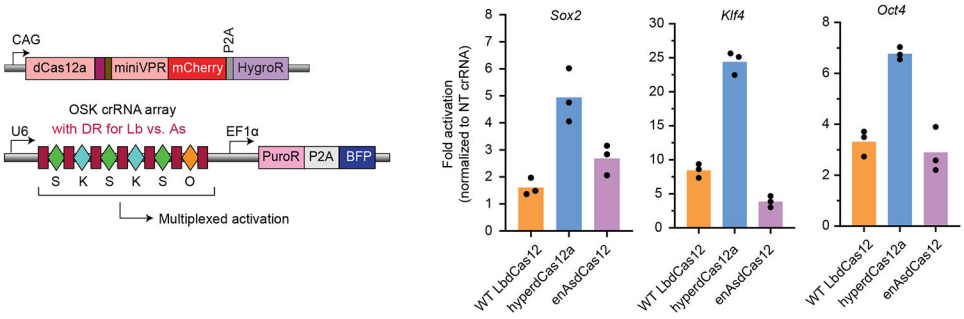
a, Schematic of dCas12a crRNAs (red) targeting promoter of *Klf4* and their relative positions to known dCas9 sgRNAs[32] that are functional (black) or non-functional (grey) for activating *Klf4*. Arrows indicate sense or antisense binding of crRNAs/sgRNAs to the target DNA. The genomic position of the first "T" in PAM (relative to TSS, which is "0") 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 μm; yellow scale bar (within inset), 50 μm **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. Screening dCas12a crRNAs for activating endogenous *Oct4*

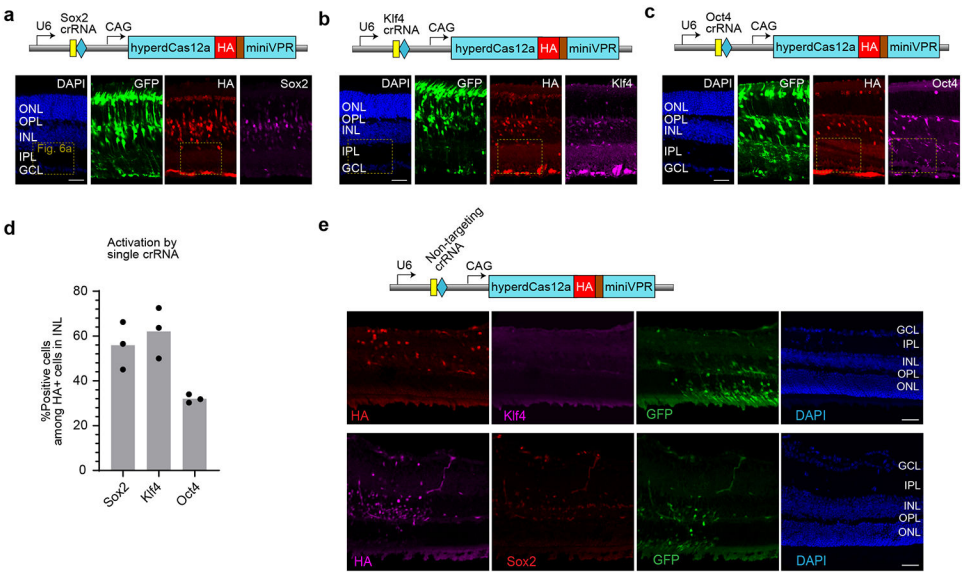
a, Schematic of dCas12a crRNAs (red) targeting promoters of *Oct4* and their relative positions to known dCas9 sgRNAs[32] that are functional (black) or non-functional (grey) for activating *Oct4*. Arrows indicate sense or antisense binding of crRNAs/sgRNAs to the target DNA. The genomic position of 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 μm. **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 μm. **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 μm. **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 vs. enAsdCas12a and 6-crRNA array 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.



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 postnatal mouse retina and representative immunofluorescence images. CAG-GFP is used to mark the electroporated patch. Scale bar, 50 μm. **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 μm.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

L.Y.G. acknowledges support from Stanford's National Eye Institute (NEI) T32 Vision Training Grant (5T32EY020485), the Knights Templar Eye Foundation (KTEF), and the VitreoRetinal Surgery Foundation (VRSF). L.S.Q. acknowledges support from Li Ka Shing Foundation, National Science Foundation CAREER award (Award #2046650), National Institutes of Health (Grant # 1U01DK127405), Stanford Maternal & Child Health Research Institute (MCHRI), Stanford-Coulter Translational Research Grant, and California Institute for Regenerative Medicine (CIRM, DISC2-12669). S.W. is supported by funding from American Diabetes Association (1-16-INI-16), NIH 1R01EY03258501, NIH R01NS109990, NIH-NEI P30-EY026877 and Research to Prevent Blindness, Inc. The work is supported by the Li Ka Shing Foundation, the Knight Templar Eye Foundation, Stanford-Coulter Translational Research Grants Program, National Science Foundation CAREER award (Award #2046650), California Institute for Regenerative Medicine (CIRM, DISC2-12669), NIH 1U01DK127405, and the Stanford Maternal and Child Health Research Institute through the Uytengsu-Hamilton 22q11 Neuropsychiatry Research Award Program. BioRender.com was used to generate illustrations.

Data availability

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/Stanley_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.

REFERENCES

1. Qi LS et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152, 1173–1183 (2013). [PubMed: 23452860]
2. Doudna JA & Charpentier E The new frontier of genome engineering with CRISPR-Cas9. Science (80-.). 346, (2014).
3. Zetsche B et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell (2015) doi:10.1016/j.cell.2015.09.038.
4. Zetsche B et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol 35, 31–4 (2017). [PubMed: 27918548]
5. Kleinstiver BP 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). [PubMed: 30742127]
6. Campa CC, Weisbach NR, Santinha AJ, Incarnato D & Platt RJ Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. Nat. Methods 16, 887–893 (2019). [PubMed: 31406383]
7. Fonfara I, Richter H, BratoviÄ M, Le Rhun A & Charpentier E The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature (2016) doi:10.1038/nature17945.
8. Gier RA et al. High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. Nat. Commun 11, 1–9 (2020). [PubMed: 31911652]
9. Kim D et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. Nat. Biotechnol 34, 863–868 (2016). [PubMed: 27272384]
10. Bin Moon S et al. Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3′-overhang. Nat. Commun 9, (2018).
11. Zetsche B et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol 35, 31–34 (2017). [PubMed: 27918548]
12. Li F et al. Comparison of CRISPR/Cas Endonucleases for in vivo Retinal Gene Editing. Front. Cell. Neurosci 14, 1–9 (2020). [PubMed: 32038177]
13. Tak YE et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. Nat. Methods (2017) doi:10.1038/nmeth.4483.

14. Ling X et al. Technology Improving the efficiency of CRISPR-Cas12a-based genome editing with site-specific covalent Cas12a- crRNA conjugates Technology Improving the efficiency of CRISPR-Cas12a-based genome editing with site-specific covalent Cas12a-crRNA conjugates. Mol. Cell 1–10 (2021) doi:10.1016/j.molcel.2021.09.021. [PubMed: 33417852]
15. Zhang L et al. AsCas12a ultra nuclease facilitates the rapid generation of therapeutic cell medicines. Nat. Commun 12, (2021).
16. Jones SK et al. Massively parallel kinetic profiling of natural and engineered CRISPR nucleases. Nat. Biotechnol 39, 84–93 (2021). [PubMed: 32895548]
17. Liu P et al. Enhanced Cas12a editing in mammalian cells and zebrafish. Nucleic Acids Res. 47, 4169–4180 (2019). [PubMed: 30892626]
18. Kocak DD et al. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. Nat. Biotechnol 37, 657–666 (2019). [PubMed: 30988504]
19. Nguyen LT, Smith BM & Jain PK Enhancement of trans-cleavage activity of Cas12a with engineered crRNA enables amplified nucleic acid detection. Nat. Commun 11, 1–13 (2020). [PubMed: 31911652]
20. Wang D, Zhang F & Gao G CRISPR-Based Therapeutic Genome Editing: Strategies and In Vivo Delivery by AAV Vectors. Cell 181, 136–150 (2020). [PubMed: 32243786]
21. Zhang Y et al. Enhanced CRISPR-Cas9 correction of Duchenne muscular dystrophy in mice by a self-complementary AAV delivery system. Sci. Adv 6, 1–12 (2020).
22. Yamano T et al. Structural Basis for the Canonical and Non-canonical PAM Recognition by CRISPR-Cpf1. Mol. Cell 67, 633–645.e3 (2017). [PubMed: 28781234]
23. Gao L et al. Engineered Cpf1 variants with altered PAM specificities. Nat. Biotechnol 35, 789–792 (2017). [PubMed: 28581492]
24. Kempton HR et al. Multiple Input Sensing and Signal Integration Using a Split Cas12a System. Mol. Cell 1–8 (2020) doi:10.1016/j.molcel.2020.01.016.
25. Vora S et al. Rational design of a compact CRISPR-Cas9 activator for AAV-mediated delivery. bioRxiv 9, (2018).
26. Wang Q et al. Mouse gamma-Synuclein Promoter-Mediated Gene Expression and Editing in Mammalian Retinal Ganglion Cells. J. Neurosci 40, JN-RM-0102-20 (2020).
27. Levy JM 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). [PubMed: 31937940]
28. Feng G et al. Imaging Neuronal Subsets in Transgenic Mice Expressing Multiple Spectral Variants of GFP. Neuron 28, 41–51 (2000). [PubMed: 11086982]
29. Kleinstiver BP et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. Nat. Biotechnol 34, 869–874 (2016). [PubMed: 27347757]
30. Liu Y et al. CRISPR Activation Screens Systematically Identify Factors that Drive Neuronal Fate and Reprogramming. Cell Stem Cell 23, 758–771.e8 (2018). [PubMed: 30318302]
31. Lu Y et al. Reprogramming to recover youthful epigenetic information and restore vision. Nature 588, (2020).
32. Liu P, Chen M, Liu Y, Qi LS & Ding S CRISPR-Based Chromatin Remodeling of the Endogenous Oct4 or Sox2 Locus Enables Reprogramming to Pluripotency. Cell Stem Cell 22, 252–261.e4 (2018). [PubMed: 29358044]
33. Lu Y et al. Reversal of ageing- and injury-induced vision loss by Tet-dependent epigenetic reprogramming. bioRxiv 710210 (2019) doi:10.1101/710210.
34. Gonatopoulos-Pournatzis T et al. Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9-Cas12a platform. Nat. Biotechnol 1–11 (2020) doi:10.1038/s41587-020-0437-z. [PubMed: 31919444]
35. Tak YE et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. Nat. Methods 14, 1163–1166 (2017). [PubMed: 29083402]
36. Breinig M et al. Multiplexed orthogonal genome editing and transcriptional activation by Cas12a. Nat. Methods 16, 51–54 (2019). [PubMed: 30559432]

37. Matsuda T & Cepko CL Controlled expression of transgenes introduced by in vivo electroporation. Proc. Natl. Acad. Sci. U. S. A 104, 1027–1032 (2007). [PubMed: 17209010]
38. Wang S, Sengel C, Emerson MM & Cepko CL A gene regulatory network controls the binary fate decision of rod and bipolar cells in the vertebrate retina. Dev. Cell 30, 513–527 (2014). [PubMed: 25155555]
39. Chan CSY et al. Cell type- And stage-specific expression of Otx2 is regulated by multiple transcription factors and cis-regulatory modules in the retina. Dev. 147, 1–13 (2020).
40. Rocha-Martins M et al. De novo genesis of retinal ganglion cells by targeted expression of Klf4 in vivo. Dev. 146, (2019).
41. Sharma P et al. Oct4 mediates Müller glia reprogramming and cell cycle exit during retina regeneration in zebrafish. Life Sci. Alliance 2, 1–21 (2019).
42. Lin YP, 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). [PubMed: 18719084]
43. Takahashi K & Yamanaka S Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell 126, 663–676 (2006). [PubMed: 16904174]
44. Venkatesh A, Ma S, Langellotto F, Gao G, P. C. Retinal gene delivery by rAAV and DNA electroporation. Curr Protoc Microbiol 0, 1–7 (2013).
45. Matharu N et al. CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. Science (80-.). 363, (2019).
46. Liao C et al. Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. Nat. Commun 10, (2019).
47. Magnusson JP, Rios AR, Wu L & Qi LS Enhanced Cas12a multi-gene regulation using a CRISPR array separator. bioRxiv 2021.01.27.428408 (2021) doi:10.1101/2021.01.27.428408.
48. Black JB 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). [PubMed: 27524438]
49. Nuñez JK et al. Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. Cell 184, 2503–2519.e17 (2021). [PubMed: 33838111]
50. Nakamura M, Ivec AE, Gao Y & Qi LS Durable CRISPR-Based Epigenetic Silencing. BioDesign Res. 2021, 1–8 (2021).
51. Kemaladewi DU et al. A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. Nature 572, 125–130 (2019). [PubMed: 31341277]
52. Xu X, et al. Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. Mol Cell. 81(20):4333–4345.e4. (2021). [PubMed: 34480847]

References for Methods

53. Guo LY, et al. "Multiplex CRISPR genome regulation in mouse retina with hyper-efficient Cas12a" Protocol Exchange (2022) DOI: 10.21203/rs.3.pex-1811/v1
54. Clement K et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol 37, 224–226 (2019). [PubMed: 30809026]
55. Dobin A et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
56. Anders S, Pyl PT & Huber W HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015). [PubMed: 25260700]
57. Kwong JMK, 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). [PubMed: 19737887]

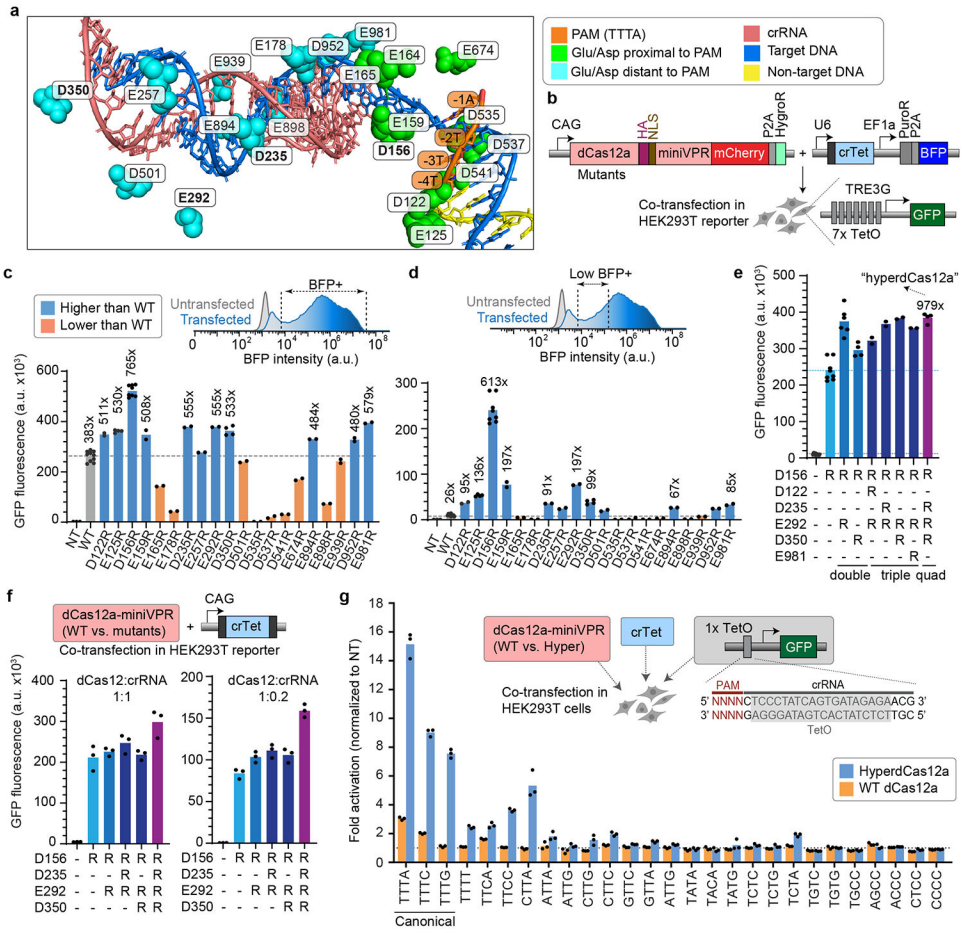


Figure 1 |. Development of combinatorial dCas12a mutants with superior activity at low crRNA conditions

a, Structure of LbCas12a (PDB 5XUS) highlighting Glu and Asp residues within 10Å of the target DNA. Green, residues proximal to PAM. Blue, residues distant to PAM. **b,** Constructs used for co-transfection to test CRISPR activation using a Tet crRNA (crTet) driven by U6 promoter, with dCas12a mutants in HEK293T cells stably expressing GFP driven by the inducible TRE3G promoter, and collected for flow cytometry 2 days after transfection. **c**, GFP fluorescence in reporter cells for WT dCas12a vs. mutants. Fold changes were calculated relative to non-targeting crLacZ. For ease of visualization, dotted line in is drawn at the level of WT. **d**, Representative flow cytometry histogram of BFP intensity, comparing untransfected vs. transfected cells, showing subset of “low BFP” cells. Fold changes of GFP fluorescence in this “low BFP” population were calculated relative to non-targeting crLacZ. Dotted line is drawn at the level of WT. **e**, GFP fluorescence in the “low BFP” cells, comparing WT dCas12, single mutants, and combinatorial variants with several most potent single mutations from **c**. The quadruple mutant (D156R + D235R + E292R + E350R) is heretofore referred to as “hyper-efficient dCas12” (hyperdCas12a). Fold changes were calculated relative to non-targeting crLacZ. Dotted line is drawn at the level of the single D156R mutant. In c-e, each data point represents the mean GFP intensity of an independent experiment, with each bar representing the average of 2 or more independent experiments. **f,** GFP fluorescence for WT dCas12a vs. dCas12 mutants, both at crRNA:dCas12a ratio =

1:1 (left panel), or 0.2:1 (right panel). **g.** In parental HEK293T cells, hyperdCas12a vs. WT dCas12a and crTet were co-transfected with a third plasmid containing a truncated TRE3G promoter that contains a single TetO element preceded by 27 various PAMs. Cells were gated for mCherry+ and low BFP+. Fold changes were calculated relative to non-targeting crLacZ. For ease of visualization, dotted line is drawn at the level of the non-targeting crRNA. In f-g, each data point represents the mean GFP intensity of an independent experiment, with each bar showing the average of 3 independent experiments.

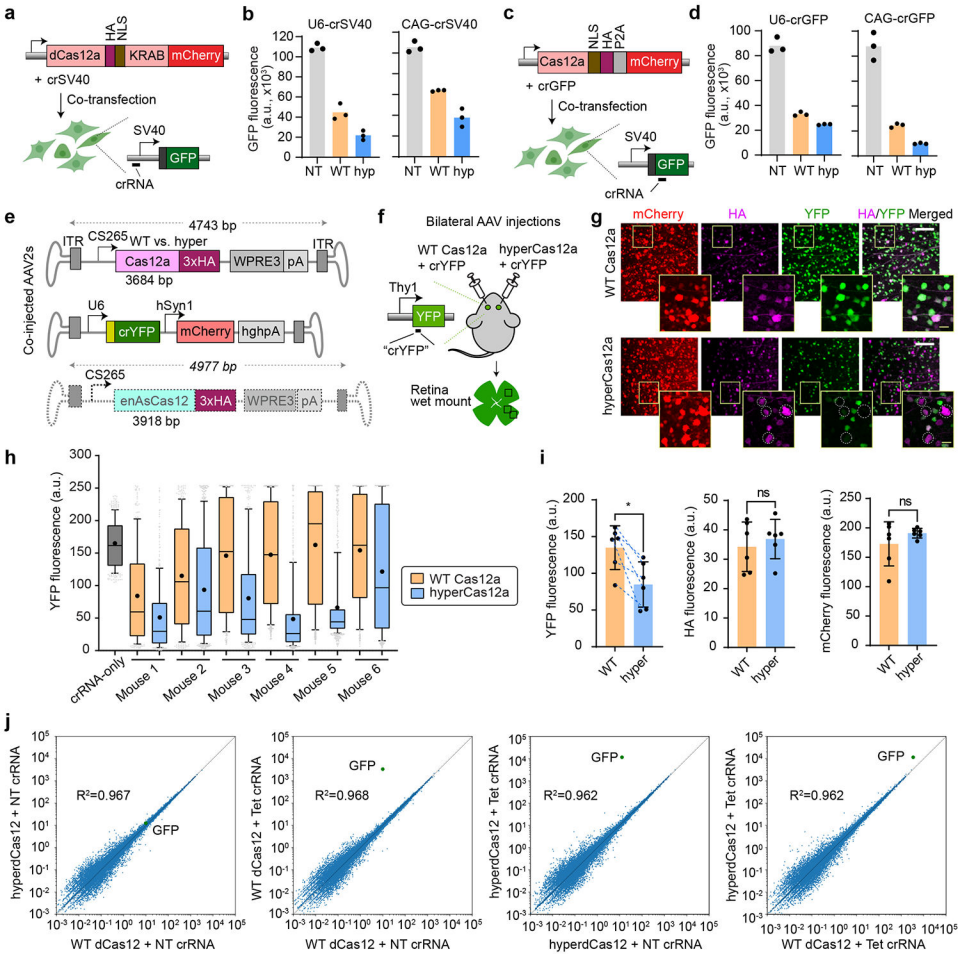


Figure 2 |. HyperCas12a outperforms wildtype dCas12 in CRISPR repression and *in vivo* gene editing

a, Constructs to test CRISPR repression. WT dCas12a or hyperdCas12a fused to the transcriptional repressor KRAB is co-transfected with crRNA (driven by U6 promoter or CAG promoter) targeting SV40 promoter in HEK293T cells stably expressing SV40-GFP, and analyzed at 5 days after transfection. **b**, GFP fluorescence in assay described in panel a. **c**, Constructs to test gene editing. Nuclease-active WT Cas12a or hyperCas12a is co-transfected with crGFP (targeting a coding region of GFP) in HEK293T cells stably expressing SV40-GFP, and analyzed at 5 days after transfection. **d**, GFP fluorescence in the gene editing assay described in panel c. In b and d, each data point represents the mean GFP intensity of an experiment, with each bar showing the average of 3 independent experiments. NT, non-targeting crRNA; WT, wildtype dCas12a, hyp, hyperdCas12a. **e.** AAV constructs for *in vivo* gene editing. **f**. Schematic of intravitreal injections, where AAV-hyperCas12a + AAV-crYFP is delivered into one eye while AAV-WT Cas12a + AAV-crYFP is delivered to the fellow eye. **g**. Immunohistochemistry of retinal wet mounts. Dotted circles highlight mCherry+/HA+ cells missing YFP expression. White scale bars, 100 μm. Yellow scale bars (within insets), 20 μm. **h.** YFP fluorescence in mCherry+ cells quantitated in each mouse by automated segmentation analyses. Data for all 6 mice are displayed, which are 6 independent biological replicates. For each mouse, n = 250-800 cells were analyzed

with the exact n number shown in Source data for Fig. 2. 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. **i.** The mean YFP fluorescence (left), HA signal (middle) and mCherry fluorescence (right) for each mouse as measured by automated segmentation analysis. Mean± s.d. and individual data points shown for n=6 independent animals. P-values were calculated using a paired two-tailed Student's t-test; **p=0.0078; ns, non-significant. For YFP graph, blue dotted lines are drawn to connect values for each mouse to facilitate comparison of this paired dataset. **j,** Plasmids with dCas12a-miniVPR (WT or hyper) and crRNA were co-transfected into HEK293T cells stably expressing TRE3G-GFP (per Fig. 1b), and collected for genome-scale RNA-sequencing (RNA-seq) 2 days after transfection. *GFP* gene is labeled in green. The plots represent representative results from two independent RNA-seq experiments (for reproducibility plots, refer to Extended Data Fig. 4b).

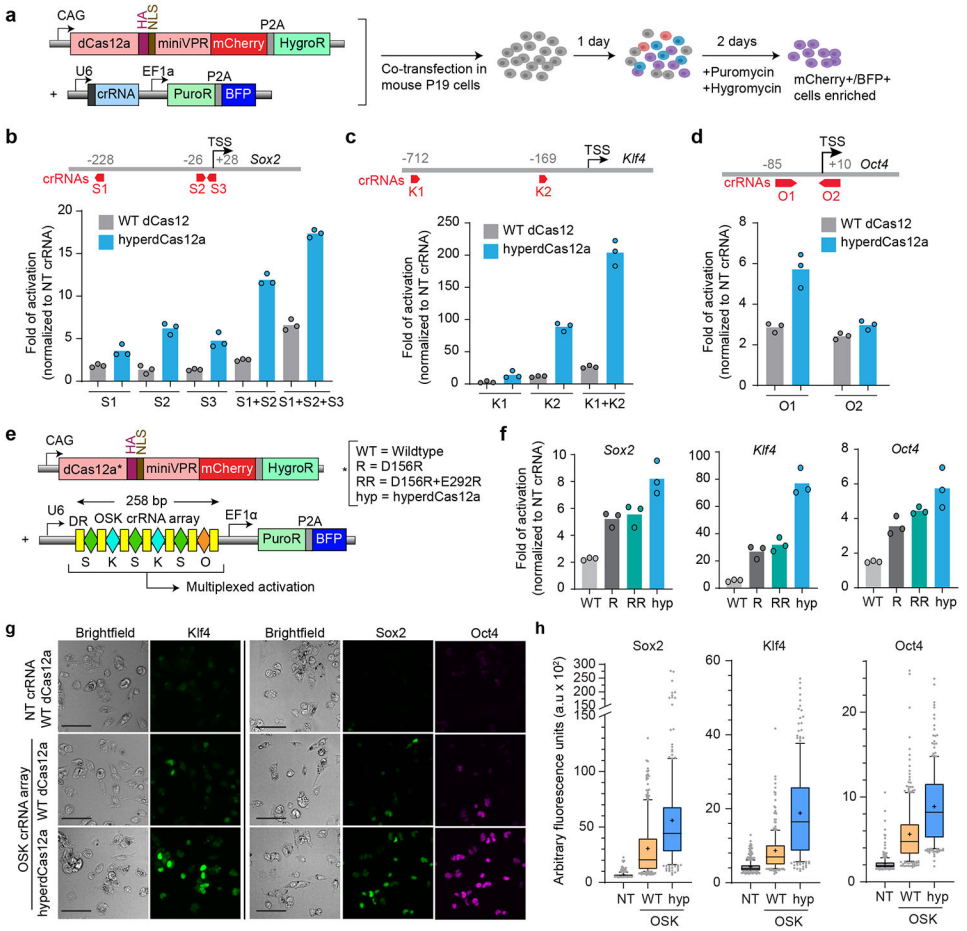


Figure 3 | HyperdCas12a enables multiplex activation of endogenous genes

a, Mouse P19 cells were co-transfected (with plasmids shown in left panel), then selected with puromycin and hygromycin 24hr after transfection to enrich for mCherry+/BFP+ cells. Cells were collected for analysis 72 hours after transfection. **b-d,** Schematics of crRNAs targeting promoters of *Sox2* (**b**), *Klf4* (**c**), and *Oct4* (**d**), as well as transcriptional activation of each target gene by qPCR by WT dCas12 vs. hyperdCas12a, relative to non-targeting crRNA. TSS= transcriptional start site. The genomic position of the first "T" in PAM (relative to TSS, which is "0") are shown for each crRNA targeting to the corresponding promoter. **e**, Constructs used for multiplex activation. **f**, Comparison of dCas12a variants in multiplex transcriptional activation of each target gene by qPCR, relative to non-targeting crRNA. **g**, Immunostaining of cells from experiment in panel e, with antibodies targeting endogenous Sox2, Klf4, or Oct4. Scale bar, 100 μm. **h.** Quantification of Sox2, Klf4 and Oct4 expression in mCherry+ cells from immunofluorescence images in panel g. For each condition, 200-400 individual cells were quantitated. 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. For **b-c, f**, bars represent mean values of three independent biological replicates; **g** shows representative images of one experiment from panel **e-f** with multiple fields of view.

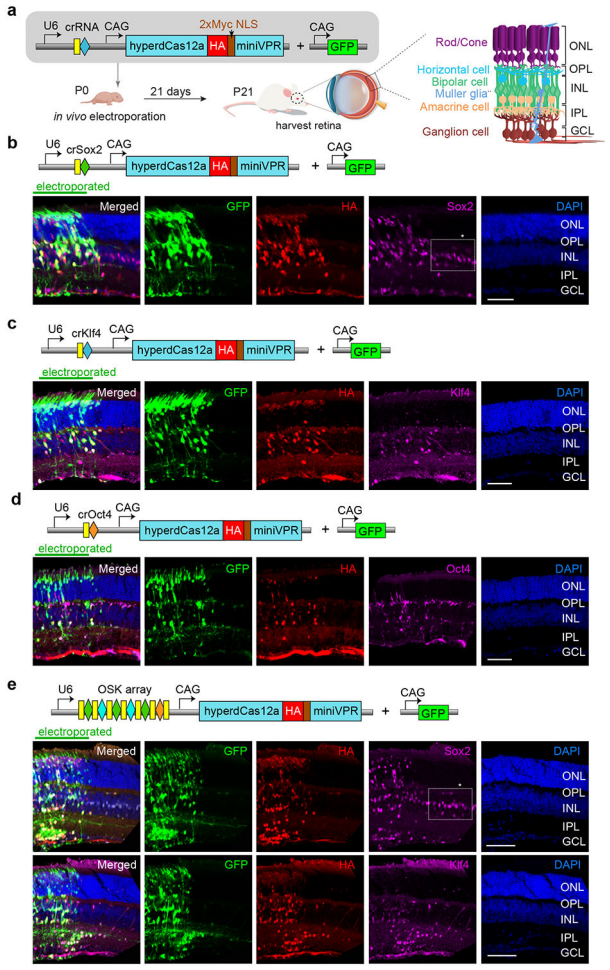


Figure 4 | In vivo gene activation by hyperdCas12a with single crRNA and poly-crRNA array

a, Constructs and experimental schematic of *in vivo* plasmid electroporation in postnatal mouse retina. CAG-GFP is used to mark the electroporated patch. Wildtype CD-1 pups are electroporated on day of birth (P0), and sacrificed at day 21 of life (P21) to access retinal histology. **b-d**, Representative retinal slices from mouse retina electroporated with a plasmid containing a single crRNA and hyperdCas12a to activate endogenous gene *Sox2* (b), *Klf4* (c) or *Oct4* (d) expression. Note that GFP signal marks the boundary of the electroporated patch, thus the area that did not receive electroporated plasmids serves as an internal control that aids in interpreting the specificity of immunostaining. HA marks the cells that received the plasmid with hyperdCas12a and crRNA. Immunostaining was performed with antibody to Sox2 (**panel b**), Klf4 (**panel c**) or Oct4 (**panel d**) indicating cells that achieved CRISPR endogenous gene activation by single crRNA . **e**, Representative retina slices of electroporation of a single plasmid containing a poly-crRNA array and hyperdCas12a driving activation of *Sox2* (top panel) and *Klf4* expression (bottom panel). Box with asterisk (*) marks basal expression of Sox2 in the inner nuclear layer (INL) outside the electroporation boundary. ONL: outer nuclear layer, OPL: outer plexiform layer, IPL: inner plexiform layer, GCL: ganglion cell layer. The images in **b-d** are representative slices from n=3 independent biological replicates (and are quantitated in Extended Data Fig.

10d). Images in **e** are representative slides of n=4-5 independent biological replicates (and are quantitated in Fig. 5j-k). Scale bar, 50 μm.

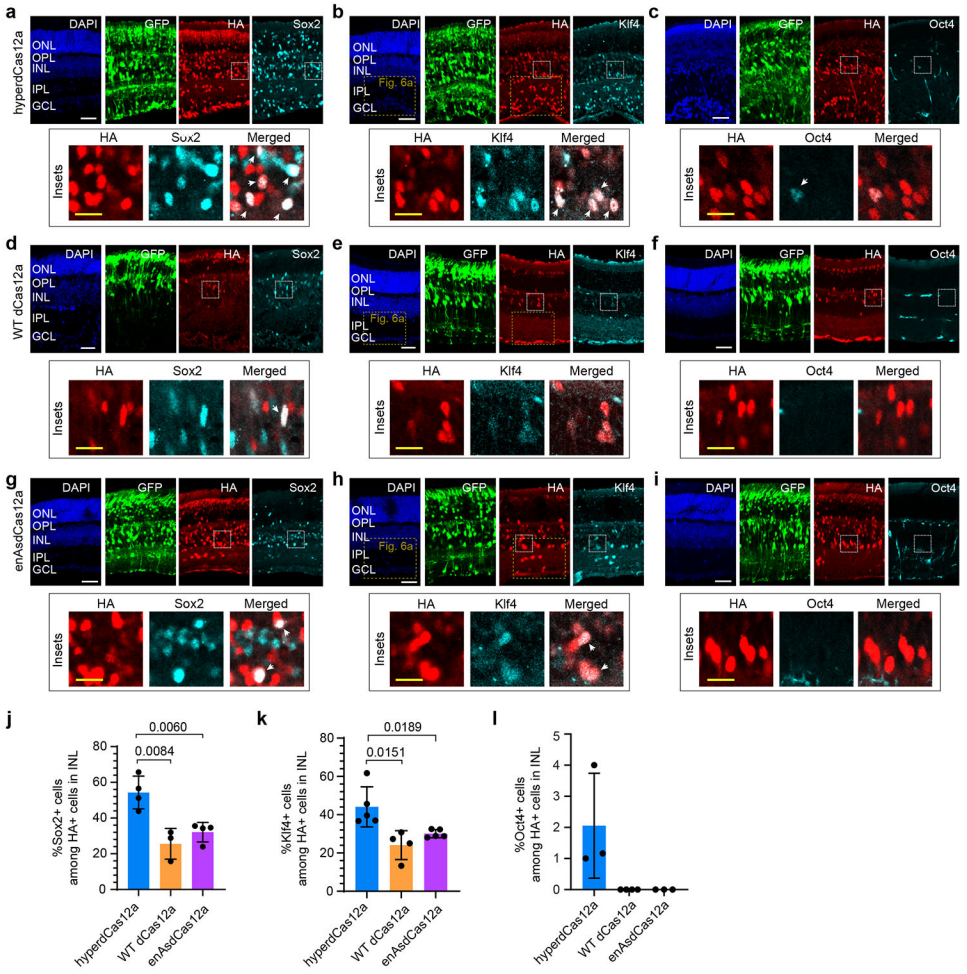


Figure 5 |. ***In vivo*** **multiplex gene activation by hyperdCas12a compared to dCas12a alternatives**

a-i Representative retinal slices after *in vivo* electroporation with crRNA array and hyperdCas12a (**panel a-c**), WT LbdCas12a (**panel d-e**), or enAsdCas12a (**panel g-i**) to activate endogenous *Sox2, Klf4* and *Oct4* expression. Insets highlight HA+ cells in the inner nuclear layer (INL). ONL, outer nuclear layer. OPL, outer plexiform layer. INL, inner nuclear layer. IPL, inner plexiform layer. GCL, ganglion cell layer. White scale bar, 50 μm; yellow scale bar (within insets), 20 μm. **j-l,** Quantitative comparison of the percentage of Sox2+ cells (j), Klf4+ cells (k) and Oct4+ cells (i) among HA+ cells in INL layer in mouse retina electroporated with plasmids containing crRNA array and hyperdCas12a, WT dCas12a or enAsdCas12a. Value represent mean± s.d. and individual data points shown for 3-5 independent biological replicates. For **j-k**, p values were calculated using an unpaired two-tailed Student's t-test and are indicated on the graphs.

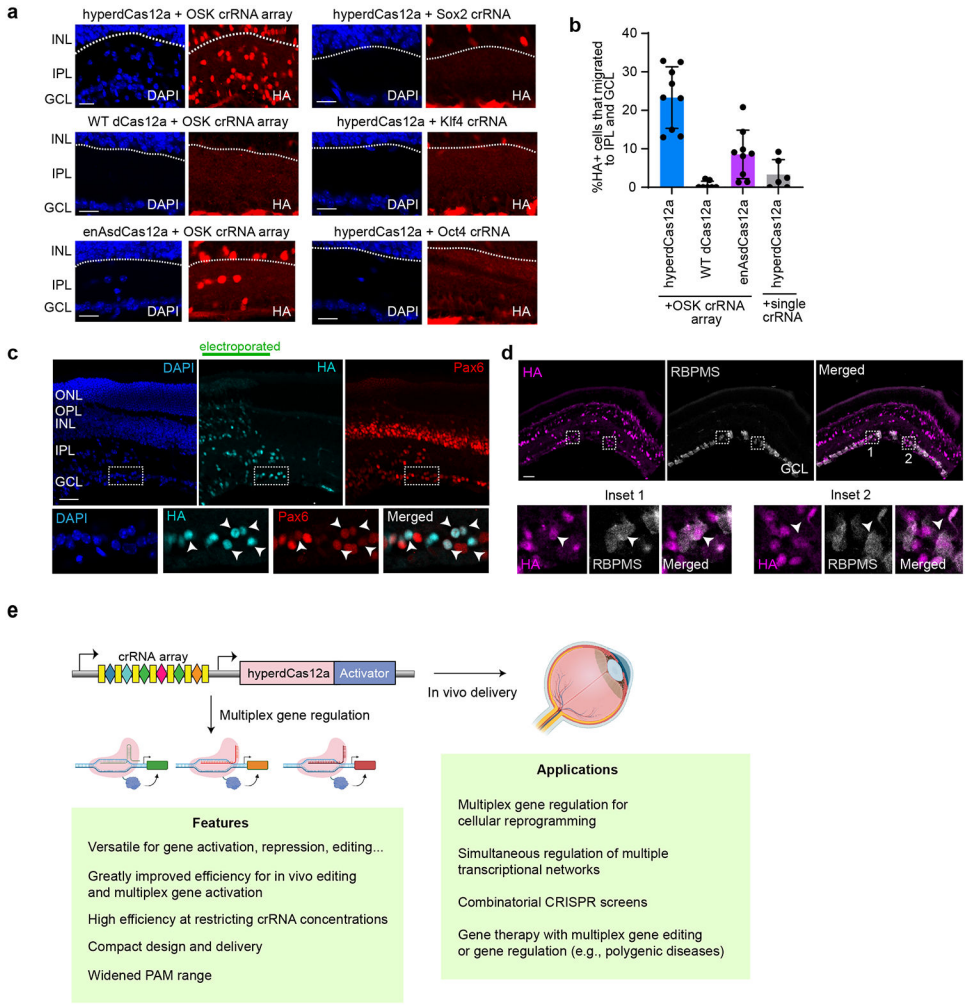


Figure 6 |. Multiplex gene activation by hyperdCas12a induces retina progenitor cells migration and altered differentiation *in vivo*

a, Representative images showing presence of HA+ cells in inner plexiform layer (IPL) and ganglion cell layer (GCL) as induced by hyperdCas12a vs. dCas12 alternatives, for single crRNA and with poly-crRNA array. Scale bars, 20 μm. **b,** Quantitation of experiment in panel a for 6-9 independent biological replicates, shown as individual data points with mean± s.d. Scale bars, 50 μm. **c,** HyperdCas12a-mediated activation of endogenous *Oct4, Sox2, Klf4* in P0 retinal progenitor cells induces formation of Pax6+ cells. Insets show co-localized Pax6, HA and DAPI staining. **d,** HyperdCas12a activation of endogenous *Oct4, Sox2, Klf4* induces formation of RBPMS+ cells. Two insets from the slice are shown on the right. Scale bars, 50 μm. **e.** Illustration of features and potential applications of hyperdCas12a. Illustration created with Biorender.com.