nature biotechnology

ARTICLES

https://doi.org/10.1038/s41587-022-01419-7

Check for updates

High-fidelity Cas13 variants for targeted RNA degradation with minimal collateral effects

Huawei Tong[1,8], Jia Huang[2,3,8]✉, Qingquan Xiao[2,3,8], Bingbing He[2,8], Xue Dong[2,8], Yuanhua Liu[2,8], Xiali Yang[4], Dingyi Han[2,3], Zikang Wang[2], Xuchen Wang[2,3], Wenqin Ying[2], Runze Zhang[2,3], Yu Wei[2,3], Chunlong Xu[2,5], Yingsi Zhou[2], Yanfei Li[6], Minqing Cai[2], Qifang Wang[2], Mingxing Xue[2], Guoling Li[2], Kailun Fang[2], Hainan Zhang[2,7]✉ and Hui Yang[1,2,5,7]✉

CRISPR–Cas13 systems have recently been used for targeted RNA degradation in various organisms. However, collateral degradation of bystander RNAs has limited their in vivo applications. Here, we design a dual-fluorescence reporter system for detecting collateral effects and screening Cas13 variants in mammalian cells. Among over 200 engineered variants, several Cas13 variants including Cas13d and Cas13X exhibit efficient on-target activity but markedly reduced collateral activity. Furthermore, transcriptome-wide off-targets and cell growth arrest induced by Cas13 are absent for these variants. High-fidelity Cas13 variants show similar RNA knockdown activity to wild-type Cas13 but no detectable collateral damage in transgenic mice or adeno-associated-virus-mediated somatic cell targeting. Thus, high-fidelity Cas13 variants with minimal collateral effects are now available for targeted degradation of RNAs in basic research and therapeutic applications.

The CRISPR and Cas systems have enabled genome editing in various types of cells and organisms[1,2]. CRISPR–Cas13, the class 2 type VI RNA endonuclease with two HEPN (higher eukaryotes and prokaryotes nucleotide-binding) domains for RNA cleavage, provides *Escherichia coli* with programmable immunity against the lytic, single-stranded RNA MS2 bacteriophage[3,4] and has been used for RNA manipulation in eukaryotic cells[5–8]. Based on their natural RNase activity against target RNAs, together with collateral cleavage of nonspecific single-stranded RNAs in vitro, the Cas13 proteins were recently used for nucleic acid detection[9–11].

The Cas13 protein family has been shown to have high efficiency and specificity in programmable RNA targeting and has been widely used for the cleavage and subsequent degradation of RNAs in yeast, plants, flies, zebrafish and mammals[5,7,12–20]. An ortholog of CRISPR–Cas13d, RfxCas13d (also known as CasRx), can mediate RNA knockdown in vivo and effectively alleviate disease phenotypes in various mouse models[21–24]. RNA-targeting CRISPR–Cas systems thus provide a promising approach for transcriptome engineering in basic research and therapeutic applications. However, in vivo application of CRISPR–Cas13 systems is hindered by the potential existence of collateral effects (Cas13 degrades bystander RNAs once activated by interaction with the target RNA)[3,25,26], one of the fundamental features of CRISPR–Cas immunity[27]. Recently, collateral RNA degradation induced by Cas13 has been found in flies[28], mammalian cells[19,20,29–34] and mammals[34]. Therefore, comprehensive analysis of collateral RNA degradation and its elimination are required for in vivo applications of Cas13 systems.

Previous structural studies[35–39] have identified the mechanisms underlying collateral RNA degradation. Upon binding of Cas13 to the target RNA, the two HEPN domains undergo distinct conformational changes to form a catalytic site on the protein surface that can degrade both target and nontarget RNAs at random, a process termed *cis*/*trans* cleavage (or target/collateral cleavage). However, it is not clear whether there are any distinct binding sites for the target and nontarget RNA substrates near the catalytic site of activated Cas13. If there are such sites, one may be able to selectively remove the nontarget RNA binding through mutagenesis of Cas13, thereby eliminating the collateral effects.

We established a rapid and sensitive dual-fluorescence reporter system for detecting collateral effects and found that Cas13 could induce substantial collateral effects in HEK293T cells when either exogenous or endogenous genes were targeted. Furthermore, we engineered many variants with mutations in HEPN domains and found several variants with markedly reduced collateral activity and comparable on-target cleavage efficiency. The collateral effects revealed by transcriptome-wide RNA sequencing (RNA-seq) and cell proliferation analysis were essentially eliminated when a mutated Cas13 variant Cas13d-N2V8 (967 amino acids) or Cas13X-M17YY (775 amino acids) was used for editing. Thus, by diminishing or eliminating promiscuous RNA binding through mutagenesis, we obtained several high-fidelity Cas13 variants for targeted RNA degradation with minimal collateral effects.

Results

Collateral effects of Cas13 editing in mammalian cells. To evaluate the collateral effects of Cas13 in mammalian cells, we first cotransfected a plasmid coding for EGFP and Cas13a (from *Leptotrichia wadei*, LwaCas13a) or Cas13d (from *Ruminococcus flavefaciens*

[1]HuiGene Therapeutics Co., Ltd., Shanghai, China. [2]Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China. [3]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China. [4]Key Laboratory of Stem Cell Engineering and Regenerative Medicine of Fujian Provincial Colleges and Universities, Department of Human Anatomy, Histology and Embryology, School of Basic Medical Sciences, Fujian Medical University, Fuzhou, China. [5]Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai, China. [6]Zhoupu Hospital Affiliated to Shanghai Health Medical College and Shanghai Key Laboratory of MolecularImaging, Shanghai, China. [7]HuiEdit Therapeutics Co., Ltd., Shanghai, China. [8]These authors contributed equally: Huawei Tong, Jia Huang, Qingquan Xiao, Bingbing He, Xue Dong, Yuanhua Liu. ✉e-mail: jiahuang2022@163.com; zhanghainan@huiedit.com; huiyang@ion.ac.cn

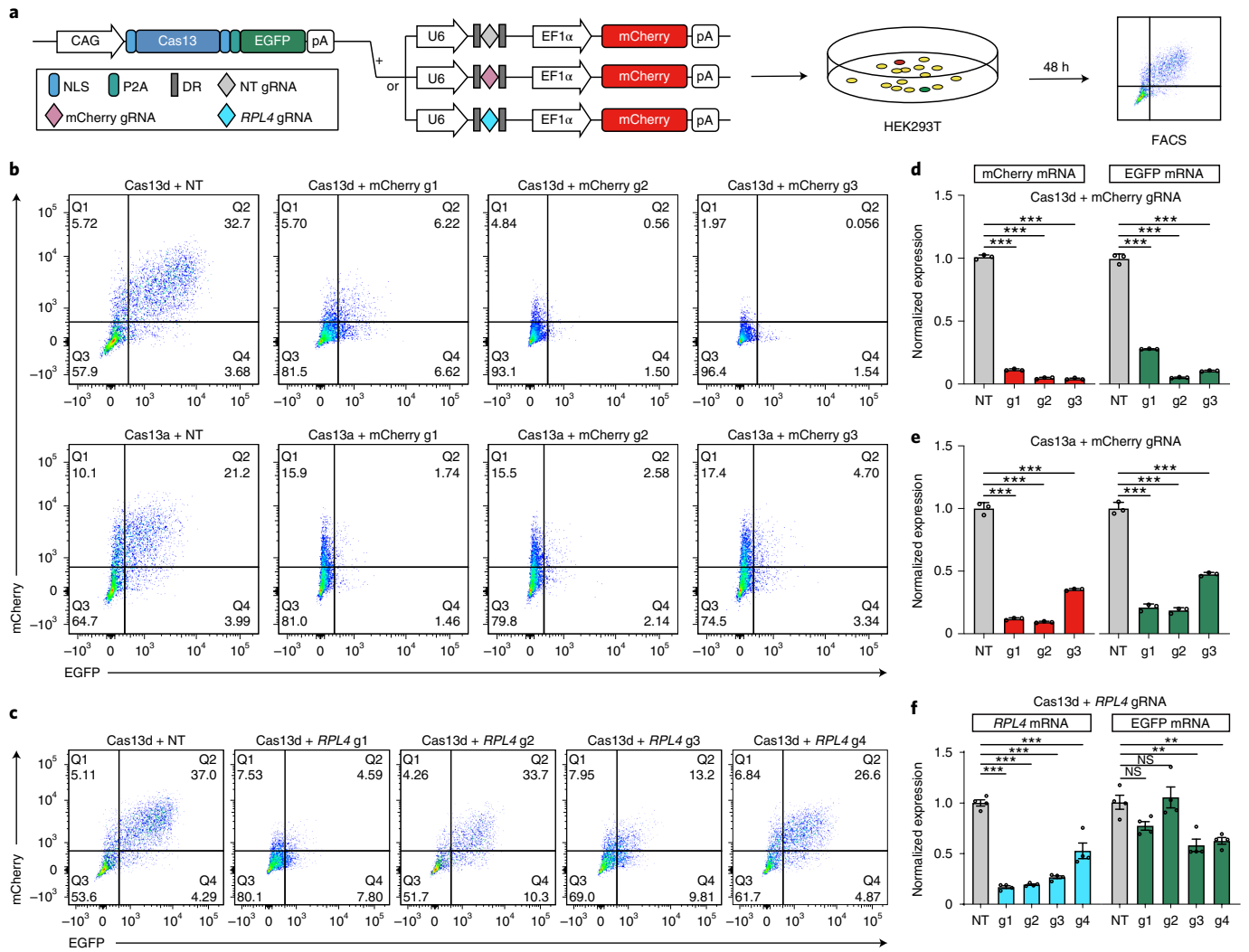


Fig. 1 | Evaluation of collateral effects in transiently transfected HEK293T cells using a dual-fluorescence reporter. a, Schematic diagram of the mammalian dual-fluorescence reporter system used to evaluate collateral effects induced by Cas13 (Cas13d or Cas13a)-mediated RNA knockdown. The dual-fluorescence reporter contained one plasmid encoding Cas13 and EGFP and one plasmid encoding gRNA and mCherry. DR, direct repeat; NLS, nuclear localization signal; NT, nontarget; P2A, 2A peptide from porcine teschovirus-1. **b**, Representative FACS analysis of both mCherry and EGFP fluorescence degradation mediated by Cas13d or Cas13a targeting three different mCherry gRNAs in HEK293T cells, compared with the NT gRNA case. **c**, Representative FACS analysis of mCherry and EGFP fluorescence degradation induced by Cas13d with four different *RPL4* gRNAs in HEK293T cells, compared with the NT gRNA case. **d**,**e**, Relative degradation of mCherry and EGFP transcripts induced by Cas13d (**d**) or Cas13a (**e**) with three different mCherry gRNAs. Degradation relative to NT gRNA was determined by qPCR; $n=3$, $^{***}P<0.0001$ in **d** and **e**. **f**, Relative degradation of *RPL4* and *EGFP* transcripts induced by Cas13d with four different *RPL4* gRNAs; $n=4$, $^{***}P<0.0001$. For EGFP mRNA levels, from left to right, $P=0.1431$, $P=0.9639$, $P=0.0019$ and $P=0.0047$. $^{**}P<0.01$, NS, not significant. All values are presented as mean ± s.e.m. Dunnett's multiple comparisons test after one-way ANOVA was used to compare the NT gRNA control with other gRNA conditions in **d**–**f**.

XPD3002, RfxCas13d) with a plasmid encoding mCherry and a guide RNA (gRNA) targeting mCherry (or a nontargeting gRNA) into HEK293T cells. Expression levels of EGFP and mCherry were measured 48 h after transfection (Fig. 1a). Using fluorescence-activated cell sorting (FACS), we found that cotransfection of mCherry gRNAs (gRNA g1 to g3) with either Cas13a or Cas13d induced significant reductions in not only mCherry fluorescence but also EGFP fluorescence compared with the nontargeting gRNA case (Fig. 1b and Supplementary Fig. 1a–d). Quantitative PCR (qPCR) analysis also showed significant reductions in levels of both EGFP and mCherry transcripts (Fig. 1d,e). These findings demonstrate substantial collateral effects of Cas13-mediated RNA degradation when targeting transiently overexpressed exogenous genes in HEK293T cells.

Next, a dual-fluorescence reporter (EGFP and mCherry) system was used to examine whether Cas13d could induce collateral effects when endogenous genes in HEK293T were targeted. We observed dramatic collateral degradation when targeting the *RPL4* transcript (Fig. 1c and Supplementary Fig. 2a) and slight collateral degradation when targeting the *PKM* and *PFN1* transcripts with Cas13d (Supplementary Fig. 2b–f). Furthermore, Cas13d induced robust knockdown of the *RPL4* transcript when any one of the four gRNAs (gRNA g1 to g4) was used, and collateral effects as indicated by reduction of EGFP transcript were observed for two of the four gRNAs used (Fig. 1f). This observation is consistent with previous reports that the extent of collateral effects differed when different gRNAs were designed to target the same gene[29]. This may be attributed to variation in the stability of the activated Cas13–gRNA

complexes. Thus, Cas13-mediated RNA degradation results in substantial collateral effects in mammalian cells when targeting either exogenous or endogenous genes.

Eliminating collateral effects of Cas13 through mutagenesis. To eliminate the collateral effects of Cas13, we sought to engineer Cas13 via mutagenesis and screen for variants free of collateral effects. First, using the dual-fluorescence approach, we constructed an unbiased screening system with EGFP, mCherry and an EGFP-targeting gRNA together with each Cas13 variant in one plasmid and used FACS to select variants with high-efficiency on-target degradation (based on low EGFP fluorescence) and low collateral degradation activity (based on high mCherry fluorescence) (Fig. 2a). Based on structural analysis and biochemical characterization of Cas13 (refs. [7,26,35–39]), we hypothesized that changing the RNA-binding cleft proximal to the RxxxxH catalytic sites in the HEPN domains would selectively reduce promiscuous RNA binding and collateral degradation while maintaining on-target RNA degradation (Fig. 2b). Thus, we designed and generated a mutagenesis library of more than 100 Cas13d variants, each containing mutations in one of 21 segments (N1–N21, 36 amino acids each). The 21 segments covered the HEPN1-I (N1–N6), helical-1 (N7), HEPN1-II (N8–N10), helical-2 (N10–N14) and HEPN2 (N14–N21) domains (Fig. 2c and Supplementary Fig. 3). The mutation in each segment consisted of four or five random amino acid substitutions (replacing all nonalanine with alanine, X>A, and alanine with valine, A>V).

We then transfected these variants individually into HEK293 cells and analyzed the reporter fluorescence by FACS (Supplementary Fig. 4). The inactive dead Cas13d (dCas13d, carrying R239A, H244A, R858A and H863A mutations in the HEPN domains) was used as a no-degradation control (100%). The reduction in the percentage of fluorescence cells, relative to that for dCas13d, indicated the degradation efficiency of the variants. We found that variants with mutation sites in the N1, N2, N3 or N15 segments (particularly N1V7, N2V7, N2V8, N3V7 and N15V4) exhibited relatively low percentages of $EGFP^+$ cells but high percentages of $mCherry^+$ cells, indicating high on-target activity but low collateral activity (Fig. 2d,f and Supplementary Table 1). Based on the predicted structures of Cas13d variants (predicted by I-TASSER[40]), we found that the mutation sites of various effective variants were mainly located in the α-helix proximal to the catalytic sites of the two HEPN domains (RxxxxH-1 and RxxxxH-2) (Supplementary Fig. 5). We also examined the effects of reducing the number of mutations in the N2 region from four to three, two or one; the variant with four mutations, Cas13d-N2V8 (carrying A134V, A140V, A141V and A143V), had the highest specificity in targeting RNA for degradation (Fig. 2e). Thus, Cas13d-N2V8, hereafter termed high-fidelity Cas13d (hfCas13d), was used in subsequent experiments.

We next targeted the *EGFP* transcript with three additional gRNAs (gRNAs 2 to 4) and found essentially no collateral effects of hfCas13d with gRNA2 or gRNA4 (Supplementary Fig. 6a,b). Moreover, we found that the inclusion or exclusion of the nuclear localization signal (NLS) did not influence the collateral activity of Cas13d, and there was a notable reduction in collateral effects when hfCas13d without NLS was used with various EGFP gRNAs (Supplementary Fig. 6c,d). We also performed a standard 1:2 serial dilution to dilute Cas13d or hfCas13d plasmids targeting the *EGFP* transcript with gRNA2 in HEK293 cells. Cas13d and hfCas13d showed a similar tendency of on-target degradation activity, whereas hfCas13d showed greatly reduced collateral effects compared with Cas13d (Supplementary Fig. 7). However, hfCas13d with gRNA3 still induced collateral effects to some extent, although less than wild-type Cas13d, indicating that gRNA selection is necessary to achieve absence of collateral effects with hfCas13d (Supplementary Fig. 6).

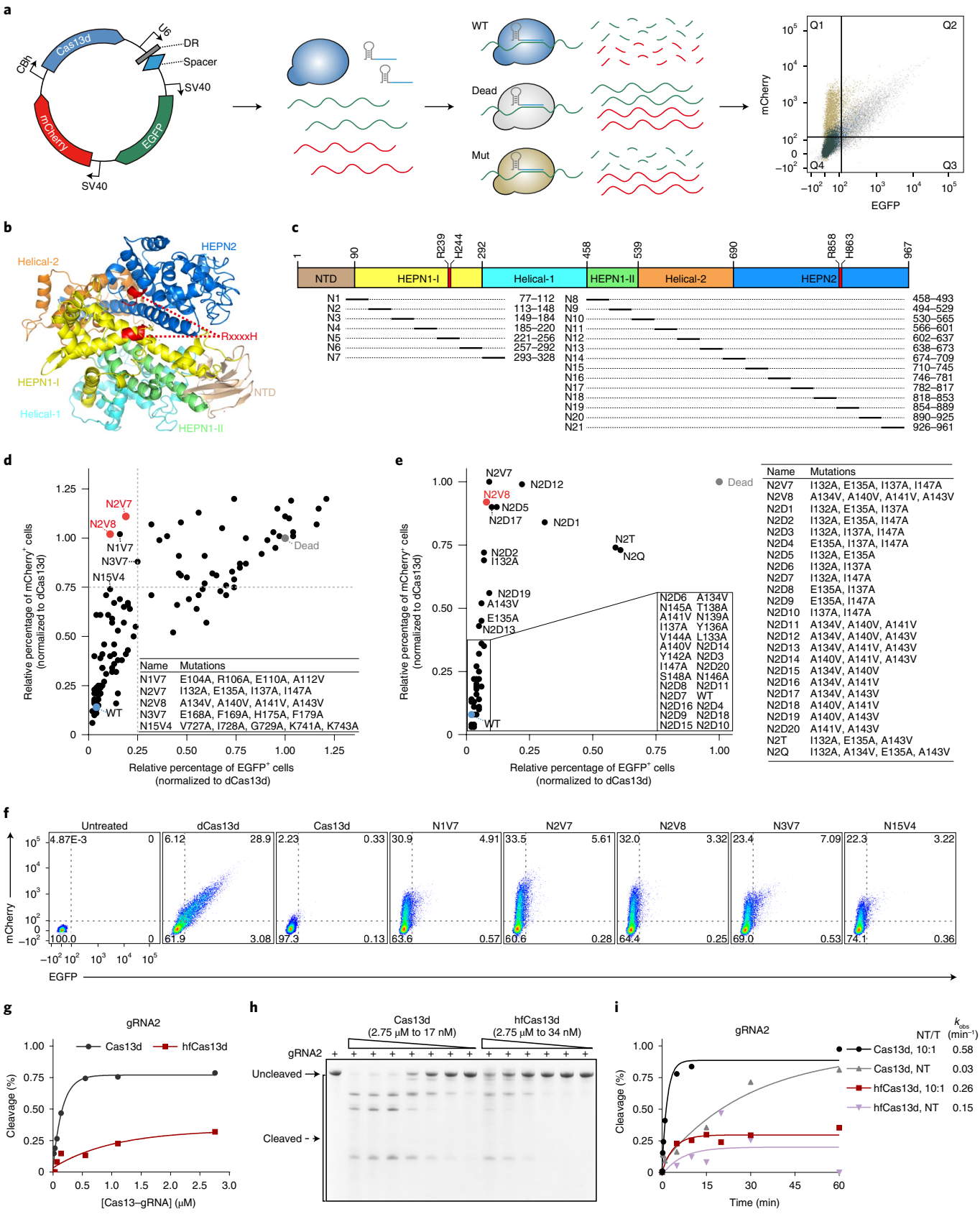
To further confirm that hfCas13d had minimal collateral activity, we performed in vitro collateral cleavage assays with purified wild-type Cas13d or hfCas13d protein (Supplementary Fig. 8). With different gRNAs (gRNAs 1 to 5), we found obviously reduced collateral cleavage activity for hfCas13d at different concentrations compared with Cas13d under all conditions (Fig. 2g,h and Supplementary Fig. 9a–e). Among the five gRNAs tested, we found very low collateral cleavage with gRNA2, gRNA3 and gRNA4, even at an extremely high concentration (2.75 μM, compared with the typical 100 nM) of hfCas13d. We also incubated hfCas13d with gRNA2 at a high concentration (550 nM) for a longer time and still found quite low collateral cleavage (Fig. 2i).

Taken together, compared with wild-type Cas13d, hfCas13d showed remarkably reduced collateral cleavage activity while retaining on-target activity.

Efficacy and specificity of hfCas13d in mammalian cells. Given that the expression levels of *EGFP* transcripts after transient transfection were much higher than those of endogenous transcripts, we next targeted endogenous genes to examine the collateral effects of Cas13d and hfCas13d in HEK293 cells. To evaluate whether the extent of collateral effects induced by Cas13d was correlated with the expression levels of endogenous genes, we selected a panel of 23 endogenous genes with diverse roles and differential expression levels in mammalian cells and designed 1–7 gRNAs for each gene (Fig. 3a). We transfected HEK293 cells with a construct encoding Cas13d, EGFP, mCherry and an on-target gRNA for each endogenous gene or a nontarget gRNA, together with another construct coding for blue fluorescent protein, which was used to confirm the uniformity of transfection efficiency (Fig. 3b,c). We examined the EGFP and mCherry fluorescence to determine the collateral effects 48 h after transfection. Overall, we found that higher expression levels of endogenous target genes or larger amounts of Cas13d induced more collateral degradation by Cas13d (Fig. 3b,c and Supplementary Figs. 10 and 11a,b). By contrast, we found no detectable collateral degradation induced by hfCas13d with any of the tested on-target gRNAs (Fig. 3d,e). To compare the RNA degradation activity of Cas13d and hfCas13d, we transfected Cas13d, hfCas13d or dCas13d, together with gRNAs targeting each transcript, into HEK293 cells and performed a qPCR assay for each

Fig. 2 | Rational mutagenesis of Cas13d to eliminate collateral activity. a, Schematic diagram of the dual-fluorescence reporter system containing Cas13d, EGFP, mCherry and EGFP-targeting gRNA in one plasmid. FACS analysis was performed to select mutant variants with low percentages of $EGFP^+$ cells and high percentages of $mCherry^+$ cells. Mut, mutant; WT, wild type. **b**, View of predicted RfxCas13d structure in ribbon representation. RxxxxH motifs define the catalytic site, shown in red. NTD, N-terminal domain. **c**, The HEPN1, HEPN2, helical-2 and partial helical-1 domains were selected and divided into 21 segments, with each spanning 36 amino acids. **d**,**e**, Quantification of relative percentage of EGFP and/or mCherry positive cells among 118 screened Cas13d mutants (**d**) and mutants with different combinations of mutation sites (**e**). WT and dCas13d were used as controls; relative percentages of fluorescence-positive cells were all normalized to dCas13d. Each dot represents the mean of three biological replicates of every mutant variant. **f**, Representative FACS analysis of Cas13d variants with EGFP-targeting gRNA1. **g**, Quantification of collateral cleavage activity of Cas13d or hfCas13d with gRNA2 in the presence of varying concentrations of Cas13–gRNA complex. Exponential fits are shown as solid lines. **h**, Representative denaturing gel depicts cleavage reactions incubated at 37 °C for 15 min by Cas13d or hfCas13d with gRNA2. **i**, Quantified time-course data of collateral cleavage by Cas13d or hfCas13d with gRNA2. Exponential fits are shown as solid lines, and the calculated pseudo-first-order rate constants (k_{obs}, mean) of different nontarget (NT) to target (T) RNA molar ratios are shown on the right.

transcript 48 h after transfection. We found that both Cas13d and hfCas13d exhibited robust degradation of targeted transcripts, as further confirmed by qPCR analysis (Fig. 3f–h), although the efficiency of targeted RNA degradation was slightly decreased for hfCas13d compared with Cas13d (77 ± 3% and 80 ± 4%, respectively), indicating that hfCas13d retained a high efficiency of RNA degradation for many endogenous genes (Fig. 3f–i). In addition, to achieve RNA degradation of targeted transcripts more efficiently by



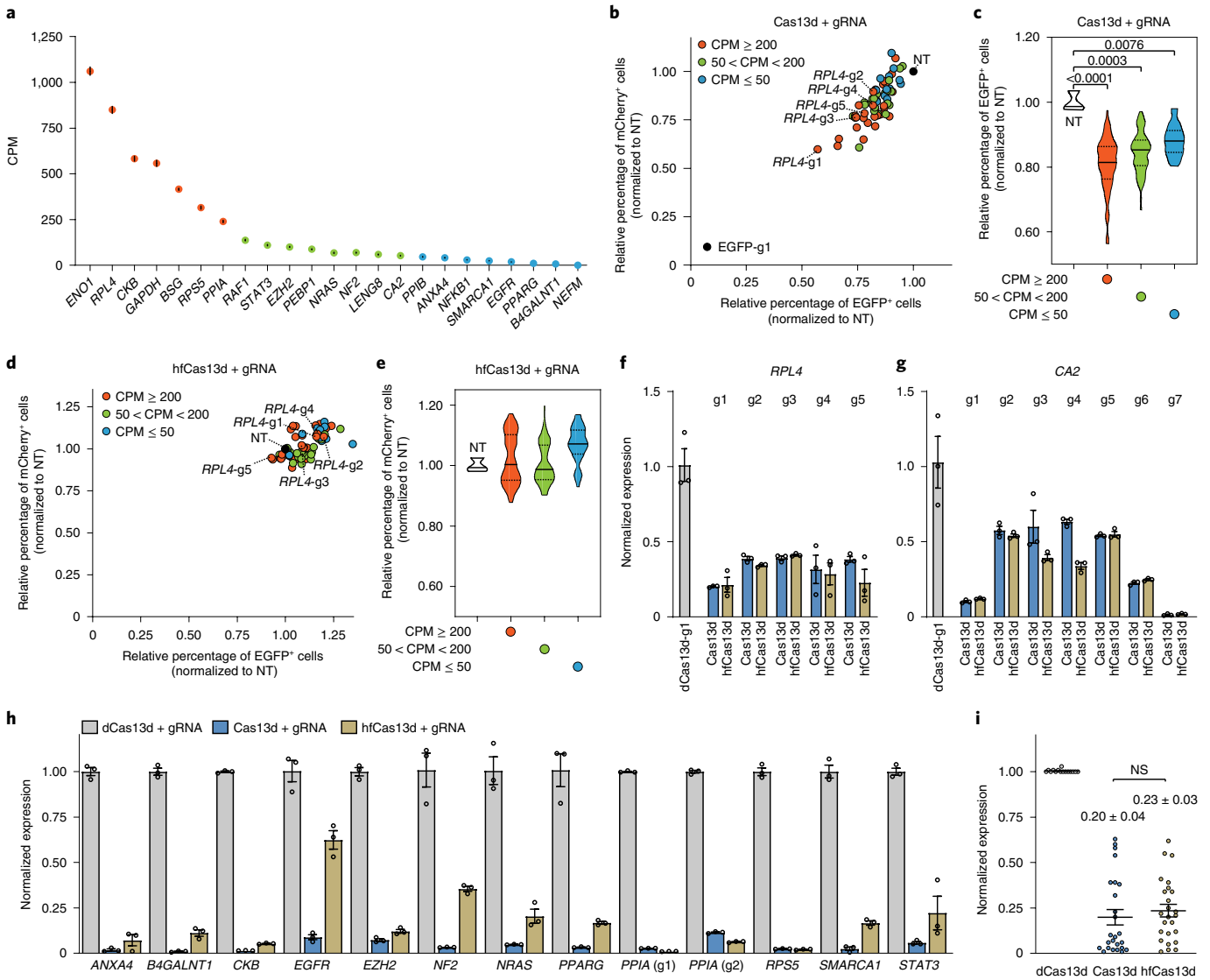
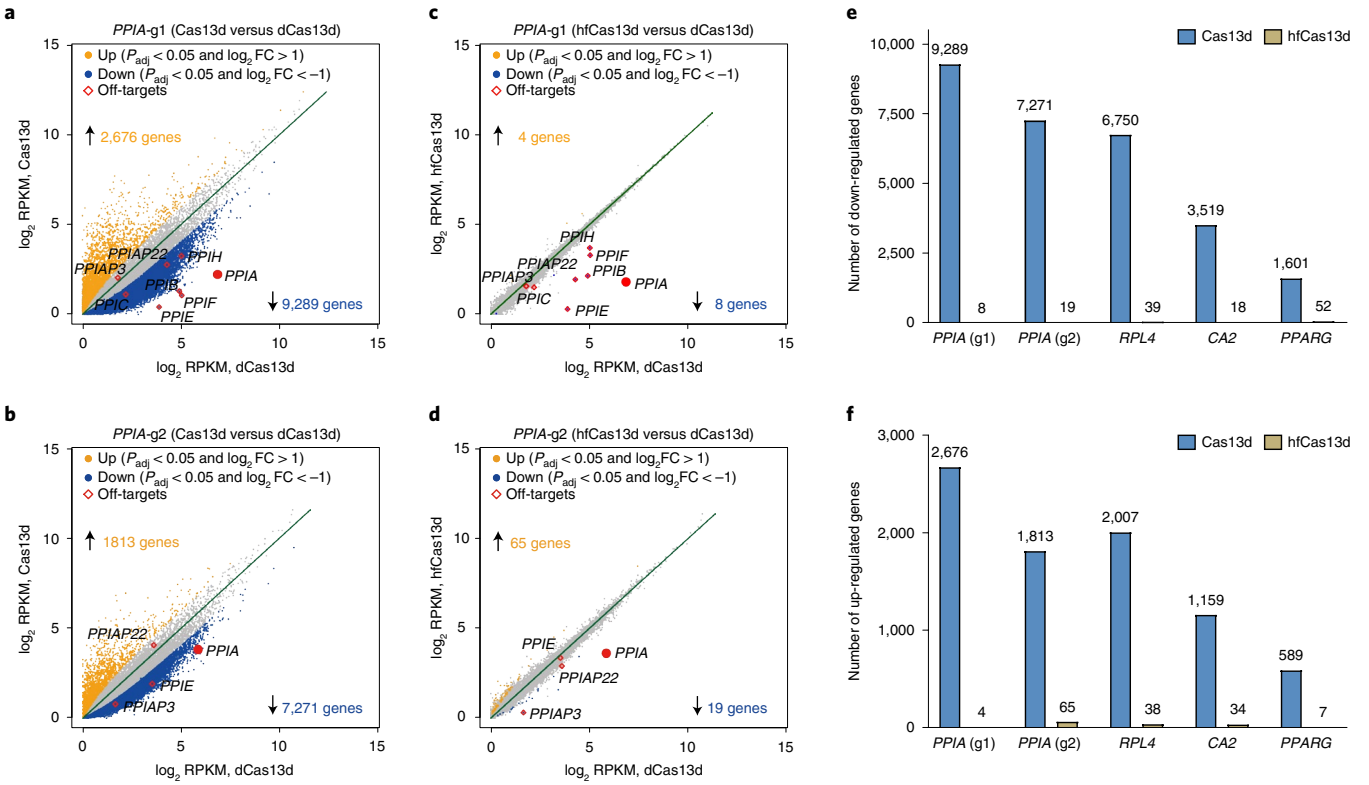
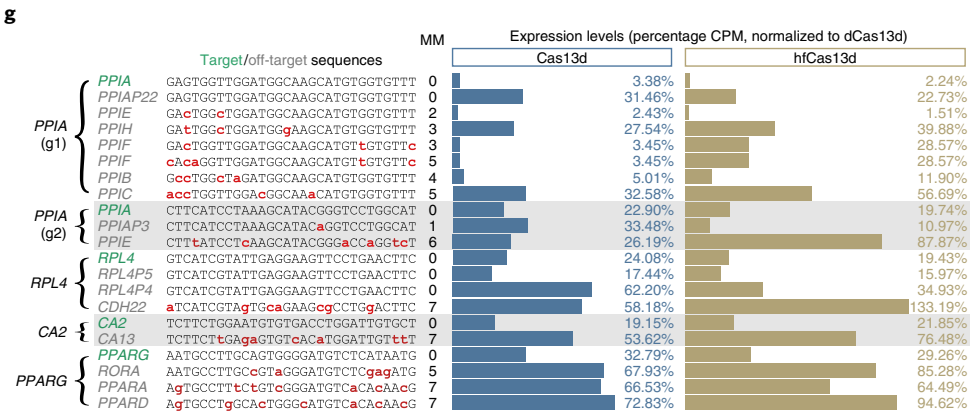


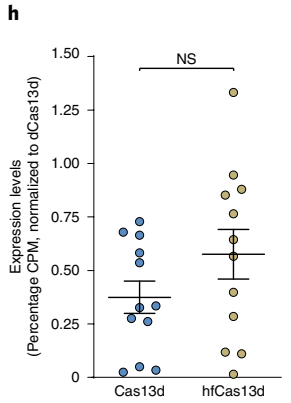
Fig. 3 | Efficient and specific interference activity of hfCas13d in HEK293 cells. a, Expression levels of 23 endogenous genes in HEK293 cells from RNA-seq of dCas13d groups. Each dot represents the mean ± s.e.m. of CPM (counts per million) for each endogenous gene in the dCas13d groups (*PPIA*-g2, *RPL4*-g3, *CA2*-g1 and *PPARG*-g1) from Fig. 4. **b**, Differential reduction in relative percentages of EGFP$^+$ and/or mCherry$^+$ cells was induced by Cas13d targeting 22 endogenous genes, with 1–7 gRNAs for each transcript, compared with nontargeting gRNA (NT). **c**, Statistical quantification of results from **b**. **d**, Differential reduction in relative percentage of EGFP$^+$ and/or mCherry$^+$ cells was induced by hfCas13d. **e**, Statistical quantification of results from **d**. Twenty-nine (six genes, CPM ≥ 200), 24 (eight genes, 50 < CPM < 200) and 13 (eight genes, CPM ≤ 50) gRNAs were used, respectively, in **b**–**e**. **f**,**g**, Relative target RNA knockdown by individual gRNAs targeting *RPL4* (**f**) or *CA2* (**g**) induced by Cas13d or hfCas13d. Relative expression levels were normalized to *RPL4*-g1 or *CA2*-g1 induced by dCas13d, respectively. $n = 3$. **h**, Cas13d and hfCas13d targeted 13 endogenous transcripts. Transcript levels are relative to dCas13d as a control, $n = 3$. **i**, Statistical analysis of data from **f**–**h**. Each dot represents the mean of three biological replicates for each gRNA in **b**,**d** and **i**. All values are presented as mean ± s.e.m. In **c** and **e**, violin plots are centered around the median (black line) with quartiles (lower and upper dashed lines). Minima and maxima are shown as the bottom and top of the violin plots, respectively. Dunnett's multiple comparisons test was used after one-way ANOVA.

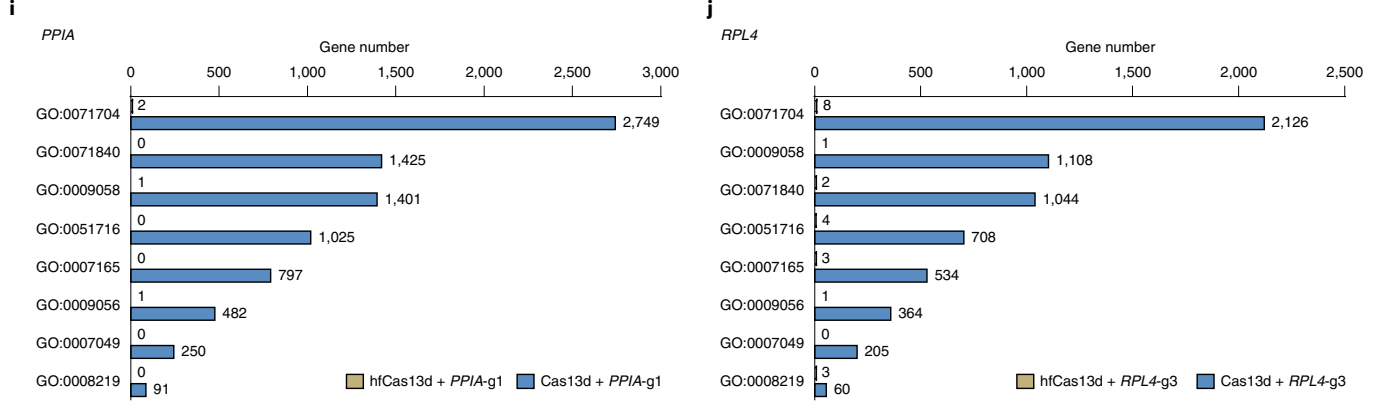
hfCas13d, we screened several gRNAs targeting the same transcript. We found that hfCas13d with a specific gRNA (for example, gRNA *RPL4*-g1 or *CA2*-g7) exhibited the most efficient degradation of target transcripts and induced no detectable collateral degradation (Fig. 3b,d,f,g). By contrast, Cas13d with different gRNAs exhibited different extents of collateral degradation (Fig. 3b). Further analysis showed a positive correlation between on-target activity and collateral activity when targeting transcripts with high expression levels (*RPL4*, $R^2 = 0.81$, Supplementary Fig. 11c), whereas there was no significant correlation when targeting transcripts with low expression levels (*CA2*, $R^2 = 0.08$, Supplementary Fig. 11d). These results indicate that the collateral effects induced by Cas13-mediated RNA degradation are correlated with the expression level of endogenous transcripts and are markedly reduced for mutated variant hfCas13d.

Absence of transcriptome-wide collateral effect for hfCas13d. To comprehensively detect the collateral effects accompanying targeted RNA degradation by Cas13d and hfCas13d, we performed total RNA integrity analysis and transcriptome-wide RNA-seq of HEK293 cells 48 h after transfection with Cas13d, hfCas13d or dCas13d. A significant reduction in RNA integrity was observed in the Cas13d panels, consistent with a previous report[29], but was absent in the hfCas13d panels (Supplementary Fig. 12). We observed widespread off-target transcript degradation in cells expressing Cas13d with *PPIA* gRNA1.









GO:0008219	Cell death	GO:0009056	Catabolic process	GO:0009058	Biosynthetic process	GO:0071704	Organic substance metabolic process
GO:0007049	Cell cycle	GO:0007165	Signal transduction	GO:0051716	Cellular response to stimulus	GO:0071840	Cellular component organization or biogenesis

Fig. 4 | Transcriptome-wide collateral effect analysis for Cas13d and hfCas13d. a,**b**, Scatter plots of differential expression levels between Cas13d-mediated and dCas13d-mediated *PPIA* degradation using gRNA1 (**a**) or gRNA2 (**b**). **c**,**d**, Scatter plot of differential expression levels between hfCas13d-mediated and dCas13d-mediated *PPIA* degradation using gRNA1 (**c**; $n=2$ for hfCas13d, $n=3$ for dCas13d) or gRNA2 (**d**). **e**,**f**, Summary of significantly downregulated (**e**) and upregulated (**f**) genes for Cas13d-mediated and/or hfCas13d-mediated *PPIA* (g1 or g2), *RPL4*-g3, *CA2*-g1 or *PPARG*-g1 degradation. **g**, Sites and relative expression levels of gRNA-dependent off-target transcripts from gRNAs targeting *PPIA* (g1), *PPIA* (g2), *RPL4*-g3, *CA2*-g1 and *PPARG*-g1 were identified in the Cas13d and hfCas13d groups. MM, mismatch number of off-target sites. Bar value, mean of the biological replicates. **h**, Statistical analysis of data from **g**; off-target sites with one or more mismatches were analyzed, and each point represents the mean of biological replicates for the corresponding gene. **i**,**j**, Biological processes of genes significantly downregulated by Cas13d-mediated or hfCas13d-mediated *PPIA* (**i**) or *RPL4* (**j**) degradation. All values are presented as mean ± s.e.m., $n=3$, unless otherwise noted. Two-tailed unpaired two-sample *t* test. P_{adj}, adjusted *P* value.

Along with high efficiency of *PPIA* on-target degradation, 9289 and 2676 genes were significantly downregulated and upregulated compared with the dCas13d control, respectively (Fig. 4a,e,f). In addition, we predicted 96 *PPIA* gRNA-dependent off-target genes using a sequence-based method (Methods) and found six downregulated off-target genes, most of which were paralogs of *PPIA* (Fig. 4a,g and Supplementary Table 2). Furthermore, downregulation of most of these gRNA-dependent off-target genes was eliminated when targeting *PPIA* with a different gRNA (Fig. 4b).

Compared with the dCas13d control, we also found numerous off-target changes induced by Cas13d when targeting *RPL4*, *CA2* or *PPARG* transcripts (Fig. 4e–g and Supplementary Fig. 13a–c). In addition, targeting transcripts with relatively high expression levels (*RPL4* and *PPIA*) in HEK293 cells induced more changes in nontargeted genes than those expressed at low levels (*CA2* and *PPARG*) among the significantly downregulated genes (Fig. 4e,f and Supplementary Fig. 13b,c), in agreement with the results described above (Figs. 2g,h and 3b,c). Further analysis showed that the genes downregulated by Cas13d targeting with *PPIA* and *RPL4* gRNAs were highly distributed in metabolic and biosynthetic processes (Fig. 4i,j and Supplementary Fig. 13g,h). In addition, most of the upregulated genes induced by Cas13d targeting *RPL4* and *PPIA* were enriched in nucleosome assembly and gene expression pathways, which are related to cellular stress regulation after cleavage events, but were not enriched in apoptosis pathways (Supplementary Fig. 14). These results are consistent with previous reports that massive host transcript degradation induced by Cas13 results in retarded growth and dormancy of cells[4,28,29,41].

Compared with Cas13d, use of hfCas13d resulted in a marked reduction in the number of downregulated off-target genes when targeting *PPIA* (9289 versus eight for g1; 7271 versus 19 for g2), *RPL4* (6750 versus 39), *CA2* (3519 versus 18) and *PPARG* (1601 versus 52) (Fig. 4a–f and Supplementary Fig. 13a–f). In addition, hfCas13d degraded some of the predicted gRNA-dependent off-target genes, although the knockdown efficiency was slightly lower than that of Cas13d (Fig. 4a–d,g,h), indicating that mutations in hfCas13d mainly reduced collateral off-target degradation rather than gRNA-dependent off-target degradation.

Taken together, these RNA-seq results fully confirm that hfCas13d exhibits high specificity of on-target RNA degradation with no collateral effects.

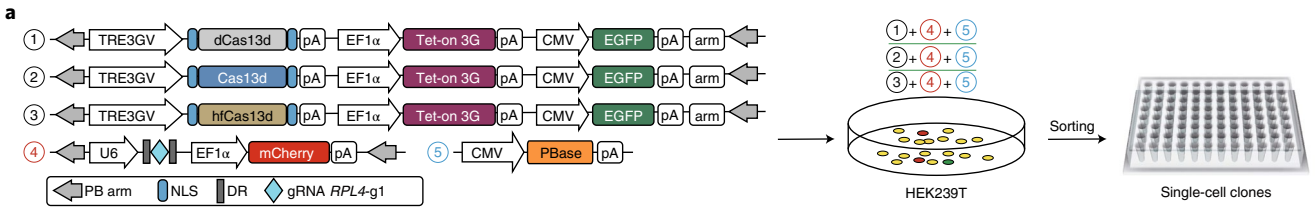
No collateral effects of hfCas13d in cells or in vivo. To further examine the impact on cellular functions due to collateral effects of Cas13d-mediated RNA degradation in vivo, we established doxycycline (Dox)-inducible stable cell lines for targeting *RPL4* using Cas13d, hfCas13d or dCas13d (Fig. 5a). Following Dox treatment, we found that the cell clone carrying Cas13d exhibited significant growth retardation and a notable reduction in *RPL4* transcripts. By contrast, the cell clone carrying hfCas13d exhibited no change in cell growth and showed a similar reduction in *RPL4* transcripts (Fig. 5b–f). These findings show that collateral effects induced by Cas13d-mediated RNA degradation in HEK293T cells lead to severe growth retardation and that hfCas13d can target specific RNAs without affecting cell growth.

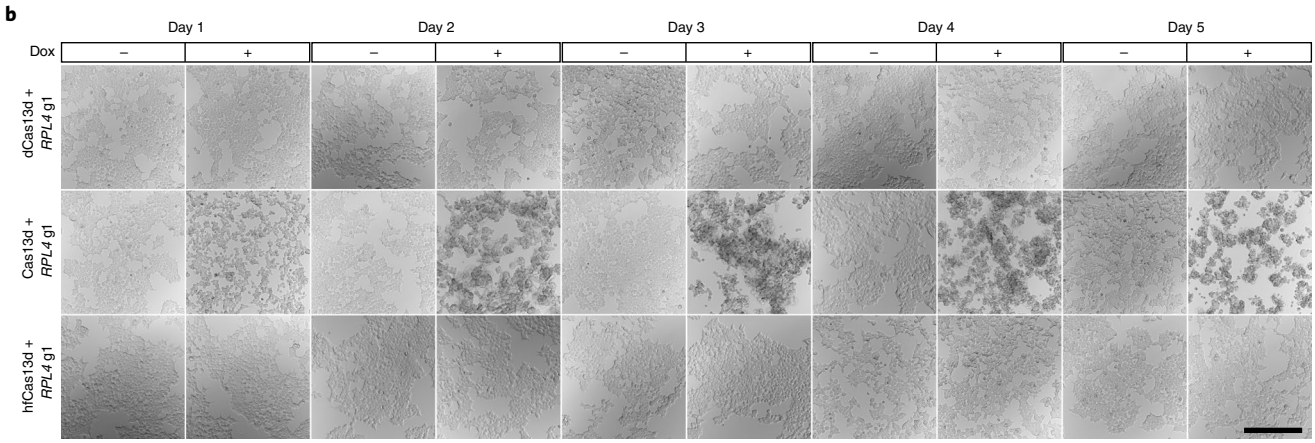
Moreover, we investigated the collateral effects of Cas13d in vivo by generating transgenic mice and examined the RNA-targeting capabilities of Cas13d for endogenous genes, designing gRNAs targeting *Tyr* (for pigmentation). After coinjection of *PBase* messenger RNA (mRNA) and a *piggyBac* vector containing the gRNAs, Cas13d or hfCas13d and the EGFP reporter into zygotes, we obtained F0 mice carrying Cas13d or hfCas13d and crossed these mice with wild-type C57 mice (Fig. 5g). Two weeks after birth, a white hair phenotype of F1 mice with *Tyr* gRNAs was observed in bright-field images, in contrast to the wild-type F1 mice, which had black hair (Fig. 5h). These findings demonstrated the RNA-targeting activity of Cas13d and hfCas13d in vivo. However, compared with the wild type ($n=5$), neither mice with only Cas13d ($n=8$) nor mice with Cas13d and *Tyr* gRNAs ($n=19$) survived for more than 8 weeks. Cas13d with nontargeting gRNA showed collateral activity in the in vitro cleavage assay (Fig. 2i); thus, the lethal phenotype of Cas13d mice may be caused by the collateral activity of Cas13d. Moreover, Cas13d with *Tyr* gRNA targeting led to more lethality compared with the case without gRNA (Fig. 5i). By contrast, all the mice with hfCas13d and *Tyr* gRNAs ($n=5$) survived to adulthood and showed no abnormalities (Fig. 5h,i), and the albino phenotypes were confirmed by the white coat color and the knockdown of *Tyr* expression (Fig. 5h,j). The lethality conferred by Cas13d may not be due to a high copy number of transgenes, as there was no significant difference in copy number between Cas13d and hfCas13d transgenic mice (Fig. 5k). To avoid unexpected effects in F0 mice, we constructed conditional expression mice, in which Cas13d expression could be released by Cre recombinase. We crossed mice containing CAG-LoxP-Stop-LoxP-Cas13d with Cre mice containing *Tyr*

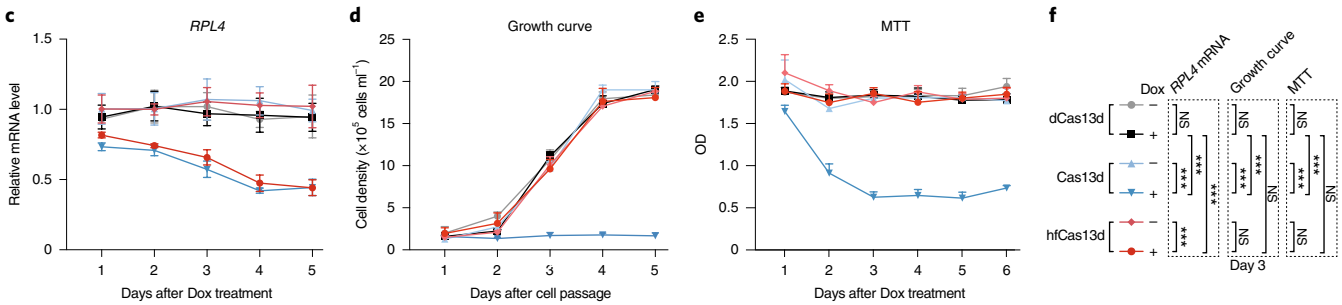
Fig. 5 | Cellular and in vivo consequences of collateral cleavage and their elimination. a, Schematic diagram of the Dox-inducible dCas13d, Cas13d or hfCas13d and *RPL4*-targeting system in HEK293T cells. **b**, Representative bright-field images of different cell clones during the 5 days after Dox treatment. $n=4$ independent biological replicates per group. Scale bar, 300 μm. **c**–**e**, Relative *RPL4* mRNA expression (**c**; $n=3$), growth curve (**d**; $n=3$) and MTT assay (**e**; $n=5$) for different cell clones with and without Dox treatment during 5 or 6 days. OD, optical density. **f**, Statistical analysis of data from **c**–**e**. For **c**, *RPL4* mRNA expression, $^{***}P<0.0001$; for **d**, growth curve, left to right, $P=0.0003$ and $^{***}P<0.0001$; for **e**, MTT assay, $^{***}P<0.0001$. Tukey's multiple comparisons test was used after one-way ANOVA. **g**, Schematic diagram of Cas13d-mediated or hfCas13d-mediated *Tyr* transcript degradation in mice using the *piggyBac* system. **h**, Representative F1-generation albino mice resulting from Cas13d-mediated or hfCas13d-mediated *Tyr* knockdown compared with WT (wild-type) mice at 2 weeks. **i**, Kaplan–Meier curve of F1 generation of WT mice ($n=5$) and mice expressing only Cas13d ($n=8$), Cas13d with *Tyr* gRNAs ($n=19$) or hfCas13d with *Tyr* gRNAs ($n=5$). Statistical differences in survival were evaluated by two-sided log-rank test. **j**, Levels of *Tyr* mRNA in 3-week-old to 8-week-old mice, $n=4$. **k**, Copy numbers of Cas13d and hfCas13d transgenic mice, $n=5$. All values are presented as mean ± s.e.m. Dunnett's multiple comparisons test was used after one-way ANOVA.

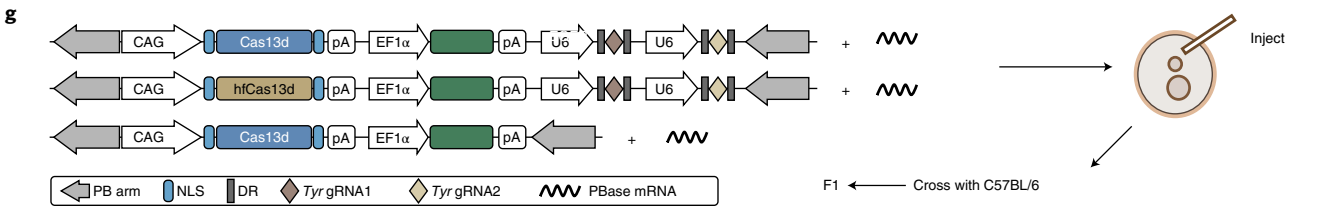
gRNAs and the mCherry reporter (Supplementary Fig. 15a); the results were similar to those obtained for Cas13d transgenic mice when targeting *Tyr* transcripts (Supplementary Fig. 15b,c). Overall, these findings suggest that Cas13d not only induces off-target lethality by itself but causes even worse lethal effect in the presence of targeting gRNAs and that these adverse effects are eliminated when using high-fidelity hfCas13d.

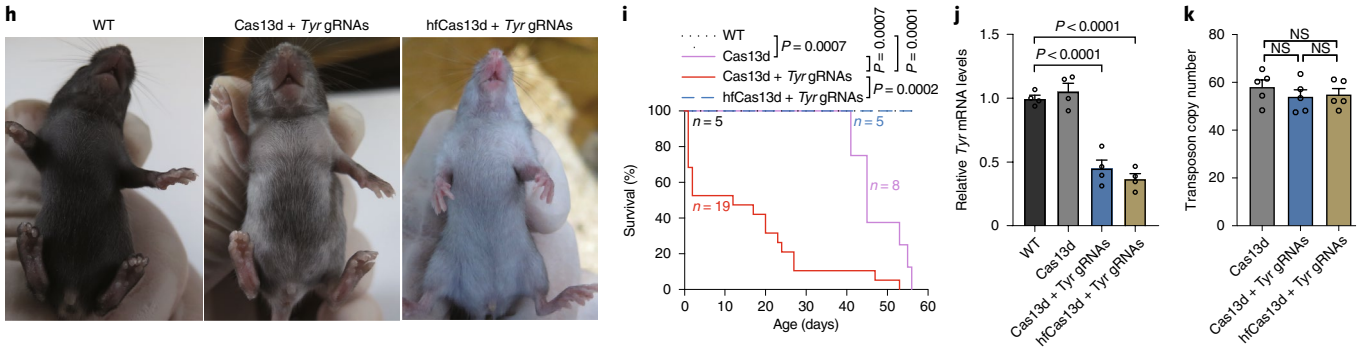
We also examined the collateral effects of Cas13d in somatic cells by intravenous injection of AAV (adeno-associated virus)-packaged Cas13d. PCSK9 is secreted by hepatocytes and has shown great promise as a candidate drug target among all regulators of serum cholesterol[42–44]. AAV8, an efficient liver-targeted gene delivery system[45], was used for in vivo *Pcsk9* knockdown. To test the collateral effects of Cas13d targeting, Cas13d or hfCas13d along with











Pcsk9 gRNA1 and gRNA2 was delivered into mouse livers through AAV8. Levels of *Pcsk9* mRNA in hepatocytes and PCSK9 protein in serum were significantly reduced in both Cas13d-treated and hfCas13d-treated mice compared with AAV8-EGFP-injected mice 4 weeks after AAV infection (Supplementary Fig. 15d,e). Liver injuries, as indicated by elevated levels of serum aspartate aminotransferase and alanine aminotransferase, were detected in mice with AAV-Cas13d targeting *Pcsk9* (Supplementary Fig. 15f,g). By contrast, we did not observe obvious liver injuries in mice with AAV-EGFP or AAV-hfCas13d targeting *Pcsk9* (Supplementary Fig. 15f,g). Thus, hfCas13d-mediated RNA knockdown in vivo carries significant potential for use in disease modeling and therapies.

Generation of hfCas13X by mutagenesis. To generalize the strategy that we had successfully used to markedly diminish the collateral effects of Cas13d, we chose to engineer Cas13X, a newly identified miniature Cas13 protein (775 amino acids) that is more applicable for in vivo applications[20]. Based on the construction of Cas13d mutant variants and the predicted structure of Cas13X (Methods), we also developed a mutagenesis library for Cas13X. Mutations were mainly located in the HEPN1 and HEPN2 domains (Fig. 6a,b and Supplementary Table 3), similar to the case of Cas13d. After screening these mutant variants, we found that Cas13X-M17YY (carrying Y672A and Y676A) exhibited low and high percentages of EGFP$^+$ and mCherry$^+$ cells, respectively, indicating high on-target activity but low collateral activity (Fig. 6c,d). Further studies using different EGFP gRNAs showed that Cas13X-M17YY, abbreviated hereafter as hfCas13X, exhibited efficient on-target degradation activity with no detectable collateral effects (Supplementary Fig. 16). In addition, the in vitro cleavage assay revealed that Casl3X exhibited lower collateral cleavage activity than Cas13d, and hfCas13X showed essentially no collateral cleavage activity (Fig. 6e–g and Supplementary Fig. 17). Moreover, when targeting different endogenous transcripts, hfCas13X exhibited robust degradation of targeted transcripts (Fig. 6h) but with minimal collateral effects (Supplementary Fig. 18).

Similarly, hfCas13X could efficiently degrade endogenous transcripts without affecting cell growth and eliminated the adverse collateral effects induced by Cas13X (Fig. 6i–l). Cas13X overexpression in transgenic mice did not lead to lethality, but these mice had lower body weight compared with wild-type mice; this side effect was eliminated in hfCas13X transgenic mice (Supplementary Fig. 19). Notably, collateral damage in cell lines and transgenic mice with Cas13d overexpression was more severe than that in those with Cas13X overexpression, probably owing to the high collateral activity of Cas13d.

Together, these results demonstrate that hfCas13X could be generated by mutagenesis of the HEPN1 or HEPN2 domain of Cas13X, similar to the generation of hfCas13d, and that hfCas13X has minimal collateral effects.

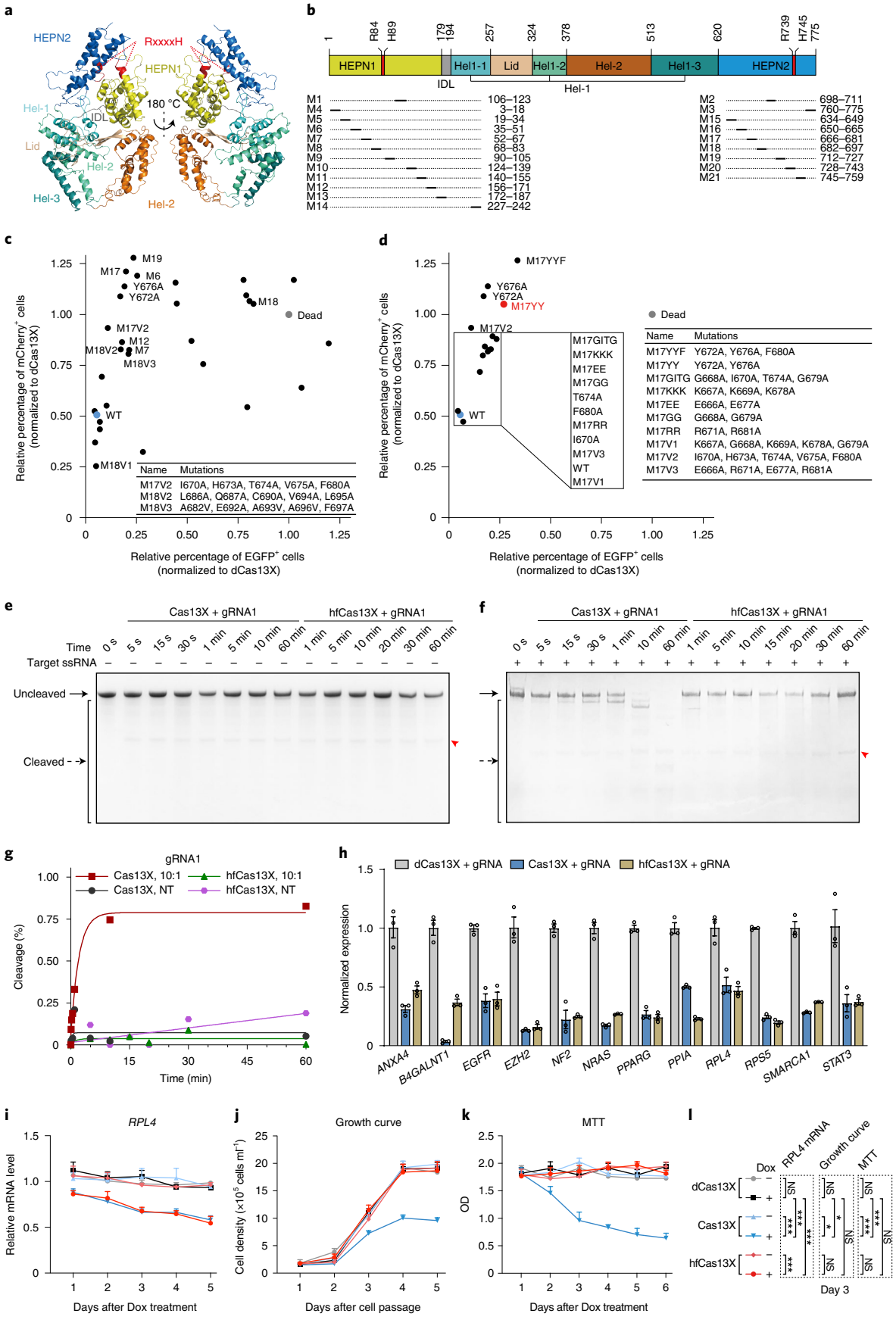
Comparison of different Cas13 proteins. To further support the use of hfCas13d and hfCas13X, we systematically compared the high-fidelity versions hfCas13d and hfCas13X with the wild-type versions Cas13d, Cas13X, PspCas13b and RanCas13b (Supplementary Fig. 20a) with respect to their on-target and collateral activity. First, using the dual-fluorescence approach as described above, we transfected these variants individually into HEK293T cells and analyzed the reporter fluorescence by FACS (Supplementary Fig. 20b,c). We found that hfCas13d and hfCas13X showed higher on-target activity on EGFP transcripts than PspCas13b or RanCas13b. Moreover, hfCas13X induced lower collateral activity on mCherry transcripts than Cas13d, Cas13X, PspCas13b or RanCas13b, and hfCas13d induced lower collateral activity than Cas13d, Cas13X or PspCas13b. We also targeted different endogenous transcripts using these Cas13 proteins (Supplementary Fig. 20d,e). In general, the Cas13 proteins showed comparable on-target activity across endogenous transcripts. However, hfCas13X and hfCas13d showed lower collateral activity than Cas13d, Cas13X, PspCas13b or RanCas13b. In addition, hfCas13X induced even lower collateral activity than hfCas13d in some loci. Overall, these results suggest that hfCas13d and hfCas13X are superior to Cas13d, Cas13X, PspCas13b and RanCas13b.

Discussion

Based on the dual-fluorescence reporter system, transcriptome-wide RNA-seq and cell growth assays, our study demonstrates severe Cas13-induced collateral RNA cleavage in mammalian cells, leading to significant retardation of cell growth. Moreover, we found that Cas13d alone was toxic in mice, possibly owing to the toxicity of the protein or to its collateral activity; Cas13d with a nontarget gRNA showed collateral activity in the in vitro cleavage assay, and some small RNAs in the cellular context may function as nontarget gRNAs. Therefore, current versions of the CRISPR–Cas13 system have a critical deficiency for in vivo applications. After a comprehensive mutagenesis screening of Cas13 variants, we obtained several high-fidelity Cas13 variants that retained high on-target RNA cleavage activity but showed minimal collateral cleavage activity. We found that many variants exhibited either low dual cleavage activity (Fig. 2d, upper right) or high on-target cleavage activity but low collateral cleavage activity (Fig. 2d, upper left). However, we found no variant exhibiting low on-target cleavage but high collateral cleavage activity (Fig. 2d, bottom right). Previous structural studies have elucidated the mechanism of target-activated RNA degradation[35–39], but the precise mechanism of dual cleavage remains unclear.

Our findings suggest a distinct binding mechanism for dual cleavage of Cas13. As depicted in the model for targeted and collateral cleavage (Supplementary Fig. 21), we propose that Cas13 contains two types of separated binding domain proximal to HEPN domains, one specifically for on-target cleavage, both of which are required for collateral cleavage. In support of this model, mutations in N1V7, N2V7, N2V8 and N15V4 variants that surround the cleavage site cause steric hindrance or changes in electrostatic interaction (Supplementary Fig. 5), weakening binding between activated Cas13 and promiscuous RNAs without affecting binding

Fig. 6 | Generation of hfCas13X by rational mutagenesis and its efficacy in vivo. a, Predicted Cas13X structure in ribbon representation. RxxxxH motifs define the catalytic site, shown in red. **b**, The 21 regions selected for subsequent mutagenesis. Hel, helical. **c**,**d**, Quantification of relative percentage of EGFP$^+$ and/or mCherry$^+$ cells among initially screened Cas13X mutants (**c**) and mutants with different combinations of mutation sites (**d**). Relative percentages of positive cells were normalized to dCas13X (dead Cas13X). WT, wild-type Cas13X. Each dot represents the mean of three biological replicates. **e**,**f**, Representative denaturing gel depicting Cas13X or hfCas13X cleavage activity without target RNA (**e**) and with target RNA (**f**). ssRNA, single-stranded RNA. Red arrows show the gRNA band. **g**, Quantified time-course data of collateral cleavage by Cas13X or hfCas13X. Exponential fits are shown as solid lines with 10:1 molar ratio of nontarget (NT) to target (T) RNA, or NT RNA only. **h**, Relative expression of 12 endogenous transcripts for Cas13X and hfCas13X. dCas13X was used as a control, $n=3$. **i**–**k**, Relative *RPL4* mRNA expression (**i**), growth curve (**j**) and MTT assay (**k**) for dCas13X, Cas13X and hfCas13X cell clones with and without Dox treatment during 5 or 6 days, $n=3$. **l**, Statistical analysis of data from **i**–**k**. For **i**, *RPL4* mRNA expression, $P=0.0003$ for hfCas13X (Dox−) versus hfCas13X (Dox+), $P<0.0001$ for dCas13X (Dox+) versus hfCas13X (Dox+), otherwise $^{***}P=0.0001$; for **j**, growth curve, left to right, $P=0.0490$ and $^{*}P=0.0301$; for **e**, MTT assay, $^{***}P<0.0001$. Tukey's multiple comparisons test was used after one-way ANOVA. All values are presented as mean ± s.e.m.


a
HEPN2
RxxxxH
HEPN1
Hel-1
180 °C
IDL
Lid
Hel-2
Hel-3
Hel-2
b
1
R84
H89
179
194
257
324
378
513
620
R739
H745
775
HEPN1
Hel1-1
Lid
Hel1-2
Hel-2
Hel1-3
HEPN2
IDL
Hel-1
M1 106–123
M4 3–18
M5 19–34
M6 35–51
M7 52–67
M8 68–83
M9 90–105
M10 124–139
M11 140–155
M12 156–171
M13 172–187
M14 227–242
M2 698–711
M3 760–775
M15 634–649
M16 650–665
M17 666–681
M18 682–697
M19 712–727
M20 728–743
M21 745–759
c
Relative percentage of mCherry+ cells (normalized to dCas13X)
Relative percentage of EGFP+ cells (normalized to dCas13X)
M19
M17
M6
Y676A
Y672A
M18
Dead
M17V2
M12
M18V2
M7
M18V3
WT
M18V1
Name Mutations
M17V2 I670A, H673A, T674A, V675A, F680A
M18V2 L686A, Q687A, C690A, V694A, L695A
M18V3 A682V, E692A, A693V, A696V, F697A
d
M17YYF
Y676A
Y672A
M17YY
M17V2
WT
M17GITG
M17KKK
M17EE
M17GG
T674A
F680A
M17RR
I670A
M17V3
WT
M17V1
Dead
Name Mutations
M17YYF Y672A, Y676A, F680A
M17YY Y672A, Y676A
M17GITG G668A, I670A, T674A, G679A
M17KKK K667A, K669A, K678A
M17EE E666A, E677A
M17GG G668A, G679A
M17RR R671A, R681A
M17V1 K667A, G668A, K669A, K678A, G679A
M17V2 I670A, H673A, T674A, V675A, F680A
M17V3 E666A, R671A, E677A, R681A
e
Cas13X + gRNA1
hfCas13X + gRNA1
Time 0 s 5 s 15 s 30 s 1 min 5 min 10 min 60 min 1 min 5 min 10 min 20 min 30 min 60 min
Target ssRNA
Uncleaved
Cleaved
f
Cas13X + gRNA1
hfCas13X + gRNA1
0 s 5 s 15 s 30 s 1 min 10 min 60 min 1 min 5 min 10 min 15 min 20 min 30 min 60 min
g
gRNA1
Cas13X, 10:1
hfCas13X, 10:1
Cas13X, NT
hfCas13X, NT
Cleavage (%)
Time (min)
h
dCas13X + gRNA
Cas13X + gRNA
hfCas13X + gRNA
Normalized expression
ANXA4
B4GALNT1
EGFR
EZH2
NF2
NRAS
PPARG
PPIA
RPL4
RPS5
SMARCA1
STAT3
i
RPL4
Relative mRNA level
Days after Dox treatment
j
Growth curve
Cell density (×10^5 cells ml^-1)
Days after cell passage
k
MTT
OD
Days after Dox treatment
l
Dox
RPL4 mRNA
Growth curve
MTT
dCas13X
Cas13X
hfCas13X
NS
*
***
Day 3

between activated Cas13 and target RNAs. Consistent with this model, some variants exhibited increased dual cleavage activity (bottom left of Fig. 2d and Fig. 6c). These variants with high collateral cleavage activity could facilitate nucleic acid detection[9–11]. Although our screening strategy was effective, it was far from saturating the potential mutant repertoire. It would be worth determining whether there might exist mutations that reduce collateral activity in other positions of Cas13 (that is, helical-1 or helical-2) that could be screened by optimizing our mutagenesis scheme or using other high-throughput approaches. Furthermore, a crystal structural model would help to further elucidate the mechanism of high-fidelity Cas13 variants.

Collateral effects have been reported in vitro and in bacterial cells, although multiple studies have claimed there is no observable collateral effect in mammalian cells for Cas13a, Cas13b or Cas13d[5–7]. In addition, Powell et al. recently showed that RfxCas13d could knock down SOD1 expression and alleviate amyotrophic lateral sclerosis without apparent toxicity in mice carrying ~25 copies of human mutant $SOD1^{G93A}$ transgenes[24]. Transient administration of Cas13d ribonucleoprotein and mRNA was also shown to be effective without obvious toxicity in zebrafish embryos[14]. By contrast, several published studies and recent preprints reported collateral activity elicited by Cas13 in flies[28], mammalian cells[19,20,29–34] and mammals[34], supporting our findings with different Cas13 orthologs in cell and mice. This controversial phenomenon may be explained by the differences in target genes with different expression levels, gRNA sequences, Cas13 proteins (Cas13 orthologs with different collateral activity, expression levels and activation duration), and cell-type-specific susceptibility to collateral activity. Although collateral activity can happen, the molecular mechanisms that modulate this response are still unclear. In this study, we markedly reduced the collateral activity of Cas13d and Cas13X by mutagenesis and demonstrated the feasibility of hfCas13d and hfCas13X for efficient on-target RNA degradation with almost no collateral damage in cell lines, transgenic animals and somatic cells. Although hfCas13d showed mild collateral cleavage activity under conditions where hfCas13d or targeted transcripts were expressed at extremely high concentrations (Fig. 2g,h and Supplementary Figs. 6 and 9), we could bypass this side effect in most applications.

In short, the hfCas13 variants, Cas13d-N2V8 and Cas13X-M17YY, with minimal collateral effects developed in this study are expected to be more competitive for in vivo RNA editing and future therapeutic 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/s41587-022-01419-7.

Received: 17 December 2021; Accepted: 1 July 2022;
Published online: 11 August 2022

References

1. Knott, G. J. & Doudna, J. A. CRISPR-Cas guides the future of genetic engineering. *Science* **361**, 866–869 (2018).
2. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507 (2019).
3. Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
4. Meeske, A. J., Nakandakari-Higa, S. & Marraffini, L. A. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* **570**, 241–245 (2019).
5. Abudayyeh, O. O. et al. RNA targeting with CRISPR-Cas13. *Nature* **550**, 280–284 (2017).
6. Cox, D. B. T. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
7. Konermann, S. et al. Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell* **173**, 665–676.e614 (2018).
8. Yang, L. Z. et al. Dynamic imaging of RNA in living cells by CRISPR-Cas13 systems. *Mol. Cell* **76**, 981–997.e987 (2019).
9. Gootenberg, J. S. et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* **356**, 438–442 (2017).
10. Gootenberg, J. S. et al. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **360**, 439–444 (2018).
11. Ackerman, C. M. et al. Massively multiplexed nucleic acid detection with Cas13. *Nature* **582**, 277–282 (2020).
12. Aman, R. et al. RNA virus interference via CRISPR/Cas13a system in plants. *Genome Biol.* **19**, 1 (2018).
13. Jing, X. et al. Implementation of the CRISPR-Cas13a system in fission yeast and its repurposing for precise RNA editing. *Nucleic Acids Res.* **46**, e90 (2018).
14. Kushawah, G. et al. CRISPR-Cas13d induces efficient mRNA knockdown in animal embryos. *Dev. Cell* **54**, 805–817.e807 (2020).
15. Huynh, N., Depner, N., Larson, R. & King-Jones, K. A versatile toolkit for CRISPR-Cas13-based RNA manipulation in Drosophila. *Genome Biol.* **21**, 279 (2020).
16. Abbott, T. R. et al. Development of CRISPR as an antiviral strategy to combat SARS-CoV-2 and influenza. *Cell* **181**, 865–876.e812 (2020).
17. Li, S. et al. Screening for functional circular RNAs using the CRISPR-Cas13 system. *Nat. Methods* **18**, 51–59 (2021).
18. Blanchard, E. L. et al. Treatment of influenza and SARS-CoV-2 infections via mRNA-encoded Cas13a in rodents. *Nat. Biotechnol.* **39**, 717–726 (2021).
19. Ozcan, A. et al. Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature* **597**, 720–725 (2021).
20. Xu, C. et al. Programmable RNA editing with compact CRISPR-Cas13 systems from uncultivated microbes. *Nat. Methods* **18**, 499–506 (2021).
21. He, B. et al. Modulation of metabolic functions through Cas13d-mediated gene knockdown in liver. *Protein Cell* **11**, 518–524 (2020).
22. Zhou, C. et al. CasRx-mediated RNA targeting prevents choroidal neovascularization in a mouse model of age-related macular degeneration. *Natl Sci. Rev.* **7**, 835–837 (2020).
23. Zhou, H. et al. Glia-to-neuron conversion by CRISPR-CasRx alleviates symptoms of neurological disease in mice. *Cell* **181**, 590–603.e516 (2020).
24. Powell, J. E. et al. Targeted gene silencing in the nervous system with CRISPR-Cas13. *Sci. Adv.* **8**, eabk2485 (2022).
25. Porto, E. M., Komor, A. C., Slaymaker, I. M. & Yeo, G. W. Base editing: advances and therapeutic opportunities. *Nat. Rev. Drug Discov.* **19**, 839–859 (2020).
26. East-Seletsky, A. et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270–273 (2016).
27. Varble, A. & Marraffini, L. A. Three new Cs for CRISPR: collateral, communicate, cooperate. *Trends Genet.* **35**, 446–456 (2019).
28. Buchman, A. B. et al. Programmable RNA targeting using CasRx in flies. *CRISPR J.* **3**, 164–176 (2020).
29. Wang, Q. et al. The CRISPR-Cas13a gene-editing system induces collateral cleavage of RNA in glioma cells. *Adv. Sci.* **6**, 1901299 (2019).
30. Wang, L., Zhou, J., Wang, Q., Wang, Y. & Kang, C. Rapid design and development of CRISPR-Cas13a targeting SARS-CoV-2 spike protein. *Theranostics* **11**, 649–664 (2021).
31. Ai, Y., Liang, D. & Wilusz, J. E. CRISPR/Cas13 effectors have differing extents of off-target effects that limit their utility in eukaryotic cells. *Nucleic Acids Res.* **50**, e65 (2022).
32. Kelley, C.P., Haerle, M.C. & Wang, E.T. Negative autoregulation mitigates collateral RNase activity of repeat-targeting CRISPR-Cas13d in mammalian cells. Preprint at *bioRxiv* https://doi.org/10.1101/2021.12.20.473384 (2021).
33. Shi, P. et al. RNA-guided cell targeting with CRISPR/RfxCas13d collateral activity in human cells. Preprint at *bioRxiv* https://doi.org/10.1101/2021.11.30.470032 (2021).
34. Li, Y. et al. Collateral cleavage of 28s rRNA by RfxCas13d causes death of mice. Preprint at *bioRxiv* https://doi.org/10.1101/2022.01.17.476700 (2022).
35. Zhang, C. et al. Structural basis for the RNA-guided ribonuclease activity of CRISPR-Cas13d. *Cell* **175**, 212–223.e217 (2018).
36. Liu, L. et al. Two distant catalytic sites are responsible for C2c2 RNase activities. *Cell* **168**, 121–134.e112 (2017).
37. Zhang, B. et al. Two HEPN domains dictate CRISPR RNA maturation and target cleavage in Cas13d. *Nat. Commun.* **10**, 2544 (2019).
38. Slaymaker, I. M. et al. High-resolution structure of Cas13b and biochemical characterization of RNA targeting and cleavage. *Cell Rep.* **26**, 3741–3751. e3745 (2019).
39. Liu, L. et al. The molecular architecture for RNA-guided RNA cleavage by Cas13a. *Cell* **170**, 714–726.e710 (2017).
40. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).

41. Kiga, K. et al. Development of CRISPR-Cas13a-based antimicrobials capable of sequence-specific killing of target bacteria. *Nat. Commun.* **11**, 2934 (2020).
42. Steinberg, D. & Witztum, J. L. Inhibition of PCSK9: a powerful weapon for achieving ideal LDL cholesterol levels. *Proc. Natl Acad. Sci. USA* **106**, 9546–9547 (2009).
43. Rossidis, A. C. et al. In utero CRISPR-mediated therapeutic editing of metabolic genes. *Nat. Med.* **24**, 1513–1518 (2018).
44. Chan, J. C. et al. A proprotein convertase subtilisin/kexin type 9 neutralizing antibody reduces serum cholesterol in mice and nonhuman primates. *Proc. Natl Acad. Sci. USA* **106**, 9820–9825 (2009).
45. Shen, X., Storm, T. & Kay, M. A. Characterization of the relationship of AAV capsid domain swapping to liver transduction efficiency. *Mol. Ther.* **15**, 1955–1962 (2007).

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

Methods

Construction of plasmids. The Cas13d (*RfxCas13d*) and Cas13a (*LwaCas13a*) genes and gRNA backbone sequences were synthesized by the HuaGene Company. Then, vectors CAG-Cas13d/Cas13a-p2A-EGFP and U6-DR-BpiI-BpiI-DR-EF1α-mCherry were generated to knock down target genes by transient transfection. The gRNA oligos were annealed and ligated into BpiI sites.

Unbiased all-in-one vectors Cbh promoter-NLS-Cas13 (PspCas13b/dPsp Cas13b/RanCas13b/dRanCas13b/Cas13d/dCas13d/Cas13X/dCas13X)-NLS-SV40 polyA-U6-gRNA-SV40 promoter-EGFP-SV40 polyA-SV40 promoter-mCherry-SV40 polyA were generated to evaluate the on-target activity and collateral activity of different Cas13 proteins and used throughout the study, except in the experiments generating the results shown in Fig. 6c,d,h and Supplementary Fig. 6c,d. In the experiments referred to in Fig. 6c,d,h, the plasmids used were constructed as follows: CMV promoter-NLS-Cas13X-NLS-bGH polyA-SV40 promoter-EGFP-SV40 polyA-SV40 promoter-mCherry-SV40 polyA-U6-DR-gRNA-DR. In the experiments referred to in Supplementary Fig. 6c,d, Cbh promoter-Cas13d/dCas13d/hfCas13d-SV40 polyA-U6-gRNA-SV40 promoter-EGFP-SV40 polyA-SV40 promoter-mCherry-SV40 polyA were used to determine the influence of NLS exclusion on collateral activity. The spacer sequences of gRNAs are listed in Supplementary Table 4.

Design and construct of Cas13 mutants. Unbiased all-in-one vectors encoding Cas13 (Cas13d or Cas13X), EGFP, mCherry and an EGFP-targeting gRNA were generated first. I-TASSER[40] was used for protein structure prediction. To design and generate a Cas13d mutagenesis library for screening, we initially divided Cas13d into 21 segments of 36 amino acids in length. Twenty-one BpiI-harboring Cas13d mutants were introduced via site-directed mutagenesis by PCR and the Gibson assembly method using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). The 21 segments covered the HEPN1-I (N1–N6), helical-1 (N7), HEPN1-II (N8–N10), helical-2 (N10–N14) and HEPN2 (N14–N21) domains. We first singly mutated all Y to A. Other amino acids (KSGTDNVQLCHMIFEPWR) were replaced with A. To cover all the residues in the segments mentioned here, we also mutated A to V. Then, we randomly chose four or five mutation sites distributed near-uniformly in distance for each variant. A schematic of the mutagenesis and screening approach is provided in Supplementary Fig. 3. For Cas13d, more than 150 mutants with four or five random amino acid substitutions (replacing all nonalanine with alanine, X>A, and alanine with valine, A>V) were designed and generated by ligating two pairs of phosphorylated oligos (one pair of wild-type oligos and another pair of mutant oligos) treated with PNK (New England Biolabs) into corresponding BpiI-digested backbones. To identify roles of the 17 amino acid residues at and near the mutant sites in both N2V8 and N2V7 variants, one BpiI-harboring mutant construct named N2C was generated; then, single, double, triple or quadruple mutations were introduced by ligating annealed mutant oligos with the corresponding BpiI-digested backbones. For Cas13X, rationally designed mutants with four or five random amino acid substitutions in two regions (M17 and M18) were generated by ligating annealed mutant oligos into the corresponding BpiI-digested backbones. The mutant sequences and corresponding nucleotide sequences used are listed in Supplementary Table 1 (for Cas13d) and Supplementary Table 3 (for Cas13X).

Cell culture, transfection and flow cytometry analysis. HEK293T and HEK293 cell lines were purchased from the Stem Cell Bank, Chinese Academy of Sciences, and cultured with DMEM (Gibco) supplemented with 10% fetal bovine serum (Gibco), 1% penicillin–streptomycin (Thermo Fisher Scientific) and 0.1 mM nonessential amino acids (Gibco) in an incubator at 37 °C with 5% CO_2.

Cas13 mutant screening was conducted in 48-well plates, and consolidation was performed in 24-well plates. The day before transfection, 3×10^4 HEK293 cells per well were plated in 0.25 ml of complete growth medium. After 12 h, 0.5 μg plasmids were transfected into cells with 1.25 μg Polyethylenimine (PEI) (DNA/PEI ratio of 1:2.5). In the 24-well plates, 2×10^5 cells were plated per well in 0.5 ml of complete growth medium, and 0.8 μg plasmids were transfected into HEK293 cells with 2.5 μg PEI. Forty-eight hours after transfection, EGFP expression and mCherry fluorescence were analyzed by BD FACS Aria III, BD LSRFortessa X-20 or Beckman CytoFLEX S. For the experiment to examine the collateral effects of Cas13d when targeting endogenous genes in HEK293 cells (Fig. 3b,c), we introduced another construct coding for blue fluorescent protein (BFP), driven by the CMV promoter. The BFP vector was cotransfected (in a 1:1 molar ratio) with a dual-fluorescence reporter (EGFP and mCherry) vector carrying Cas13 to confirm uniform transfection efficiency among the different gRNA groups. Flow cytometry results were analyzed with FlowJo v.10.5.3.

Harvest of total RNA and quantitative PCR. Forty-eight hours after transfection, 50,000 of both $EGFP^+$ and $mCherry^+$ cells were sorted by BD FACS Aria III for RNA extraction. For the mCherry knockdown groups, total cells from the 12-well plate were collected for RNA extraction. For the in vivo *Tyr* knockdown groups, RNA was extracted from mouse skin tissue. Total RNA was extracted by adding 500 μl TRIzol (Invitrogen) and 200 μl chloroform to the cells. After centrifuging at 12,000 r.p.m. for 15 min at 4 °C, the supernatant was transferred to a 1.5 ml RNase-free tube. Then, 100% isopropanol and 75% alcohol were added to precipitate and purify the RNA. Complementary DNA was prepared using HiScript Q RT SuperMix for qPCR (Vazyme, Biotech) according to the manufacturer's instructions. qPCR reactions were performed with AceQ qPCR SYBR Green Master Mix (Vazyme, Biotech). All the reagents were precooled. qPCR results were analyzed by Roche LC 480 II with the $-\Delta\Delta$Ct method. All the primers used for qPCR reactions are listed in Supplementary Table 5.

Purification of Cas13 protein. Cas13 protein purification was performed according to a protocol previously described[46]. Humanized codon-optimized genes for Cas13d, hfCas13d, Cas13X and hfCas13X were synthesized (Huagene) and cloned into bacterial expression vectors (pC013-Twinstrep-SUMO-huLwCas 13a from the laboratory of F. Zhang; plasmid no. 90097) after plasmid digestion by BamHI and NotI with an NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs). The expression constructs were transformed into BL21(DE3) cells (TIANGEN). One liter of LB growth medium (Tryptone 10.0 g, yeast extract 5.0 g, NaCl 10.0 g; Sangon Biotech) was inoculated with 10 ml for 12 h culture. Cells were then grown to a cell density OD600 of 0.6 at 37 °C, and then SUMO-Cas13 protein expression was induced by supplementation with 0.5 mM IPTG. The induced cells were grown at 16 °C for 16–18 h before being harvested by centrifugation (4,000 r.p.m., 20 min). The collected cells were resuspended in Buffer W (Strep-Tactin Purification Buffer Set, IBA) and lysed using an ultrasonic homogenizer (Scientz). Cell debris was removed by centrifugation, and the clear lysate was loaded onto a Strep-Tactin Sepharose High Performance Column (StrepTrap HP, GE Healthcare). Nonspecific binding proteins and contaminants were flowed through. The target proteins were eluted with elution buffer (Strep-Tactin Purification Buffer Set, IBA). The amino-terminal 6x His/Twinstrep-SUMO tag was removed by SUMO protease (4 °C, >20 h). Then, Cas13 proteins were subjected to a final polishing step by gel filtration (Sephacryl 200, GE Healthcare). The purity of >95% was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis. Proteins were collected and concentrated to about 5 mg ml^{-1} and stored at −80 °C until use.

In vitro transcription and purification of RNA. The gRNA, target RNA and nontarget RNA were transcribed in vitro. For gRNA transcription, equimolar concentrations of complementary oligos were mixed in RNase-free water and heated to 95 °C for 5 min and then slowly cooled to 4 °C by a touch-down procedure using a Bio-Rad PCR instrument. PCR products were used as templates for in vitro transcription using a MEGA shortscript T7 kit (Life Technologies) and then purified using a MEGA clear kit (Life Technologies). In vitro transcription and purification of target and nontarget RNAs followed the same protocols as those used for gRNA, except that the template for transcription was produced by adding series of primers by a PCR program. All the target RNAs, nontarget RNAs and gRNAs were eluted in RNase-free water. Sequences of the target RNAs, nontarget RNAs, gRNAs and primers used for RNA transcription are listed in Supplementary Table 6.

In vitro RNA cleavage assay. Cleavage assays were conducted in cleavage buffer (20 mM HEPES pH 7.0, 50 mM KCl, 2 mM MgCl2, 5 mM DTT, 5% glycerol) at 37 °C. Cas13 proteins and gRNAs, at a molar ratio of 1:1, were incubated at 37 °C for 20 min in cleavage buffer before reactions. In all cleavage assays, a reaction mixture without Cas13–gRNA complex was used as a noncleavage blank control. For the concentration course of cleavage assays, 1 μg nontarget RNA was added to the target RNA–Cas13–gRNA mixture (1:1:1) with final concentrations of Cas13 protein of 17 nM, 34 nM, 68 nM, 136 nM, 550 nM, 1.10 μM and 2.75 μM. Then, the cleavage reaction was performed at 37 °C for 15 min. Reactions were terminated by adding 2× RNA loading dye and quenched at 95 °C for 10 min. For the time course of cleavage assays with different gRNAs, final concentrations of Cas13 protein and gRNAs were 550 nM. Nontargeted RNA (1 μg) was added to the Cas13–gRNA complex, together with target RNAs at different molar ratios of nontarget to target. Then, the mixture was incubated for different time periods at 37 °C. Samples were analyzed using 15% denaturing TBE-urea gels. For data analysis, products were quantified with ImageJ (National Institutes of Health). The background for each measured substrate was first normalized using ImageJ. The cleavage percentage was calculated as the ratio of residual band intensity to the uncleaved band intensity of the blank control. The minimum calculated negative values were replaced by zero as no observed cleavage event occurred. Kinetics data were fitted with a one-phase exponential association curve using Prism (GraphPad).

RNA-seq and off-target analysis. For transcriptome sequencing, 35 μg all-in-one plasmids containing Cas13 (dCas13d, Cas13d or hfCas13d), EGFP, mCherry and a targeting gRNA for each endogenous gene were transfected into HEK293 cells cultured in 10-cm dishes. Forty-eight hours after transfection, 600,000 dual-positive $EGFP^+$ $mCherry^+$ cells (those in the top 15% for fluorescence intensity) were sorted to form a pool for sequencing by BD FACS Aria III, MoFlo Astrios EQ or Moflo XDP. Total RNA was extracted with a TRIzol-based method, fragmented and reverse transcribed to cDNAs with HiScript Q RT SuperMix for qPCR (Vazyme, Biotech) according to the manufacturer's instructions. Total RNA integrity was quantified using an Agilent 2100 Bioanalyzer. The RNA-seq library was qualified using the Illumina NovaSeq 6000 platform (Novogene Co. Ltd.) or GENEWIZ. Illumina adapter sequences and low-quality sequences (phred

score <20) at the 3′ end were trimmed, and paired reads were removed if either of the two reads did not meet the minimum length (30 bp). Cleaned reads after trimming were aligned to a human genome (GRCH38) using HISAT2 (v.2.2.1). For each sample, we count the reads hit on each gene using htseq-count. Differential analysis among cell groups (*PPIA* gRNA1, *PPIA* gRNA2, *RPL4* gRNA3, *CA2* gRNA1 and *PPARG* gRNA1) was performed using limma, a count-based method implemented in R with voom for normalization[47,48]. Benjamini–Hochberg adjusted $P < 0.05$ and at least twofold change in expression were the criteria used to screen the significantly expressed genes. Functional analysis was performed using DAVID to identify the enriched biological terms associated with the significantly changed genes[49,50]. Biological processes from the Gene Ontology database with false discovery rate values less than 0.05 were selected as significantly enriched biological terms. Functional classification was performed through the Term Enrichment Service offered by amiGO (based on the Panther database v.16.0). Scatter plots were produced by basic function plot in R. Bubble plots of the enriched terms were produced using ggplot2 in R.

Off-targets of the gRNAs were predicted using a sequence-based approach. First, gRNAs with length of 30 base pairs (bp) were aligned to the human transcriptome (GRCH38 precursor mRNA and cDNA sequences of the genes from Ensembl release 100) using blastn (megablast). Flexible options were set to capture potential off-targets, tolerating up to seven mismatches in this step, that is, max_target_seqs = 10000, evalue = 10000, word_size = 5, perc_identity = 0.6. Second, the candidates were further filtered with the requirement that a perfect match of at least 10 bp should exist in each alignment. HOMER (v.4.11) was used to annotate the off-targets. Code for predicting off-targets is provided as 'offinder_cas13.sh' in Supplementary Note 1.

Growth curve. Single-cell clones with dCas13d, Cas13d or hfCas13d and *RPL4* gRNA were plated on a 24-well plate at 2×10^5 cells ml^{-1} with or without Dox treatment (1 μg ml^{-1}). Cells were collected at 24, 48, 72, 96 and 120 h. Cell numbers were determined using an automated cell counter (C10311, Invitrogen). Experiments were performed for three replicates.

Determination of cell proliferation. Cell proliferation was assessed by colorimetric thiazolyl blue (MTT) assay. Briefly, single-cell clones with dCas13d, Cas13d or hfCas13d and *RPL4* gRNA were treated with or without Dox (1 μg ml^{-1}) for 0, 24, 48, 72, 96 or 120 h. Then, each group of cells was collected and further plated on a 24-well plate at 2×10^5 cells ml^{-1} with or without Dox treatment (1 μg ml^{-1}). After an incubation period of 24 h at 37 °C, the tetrazolium salt MTT (Sigma-Chemie) was added to a final concentration of 2 μg ml^{-1}, and incubation was continued for 4 h. Cells were washed three times and finally lysed with dimethyl sulfoxide. Metabolization of MTT directly correlates with cell number and was quantitated by measuring the absorbance at 550 nm (reference wavelength, 690 nm) using a microplate reader (type 7500, Cambridge Technology). Experiments were performed for five replicates.

Generation of Cas13 transgenic mice. The zygotes injected to generate transgenic mice were obtained from 8-week-old female B6D2F1 (C57BL/6 X DBA2J) mice. Female ICR mice (8 weeks of age) were used as recipients. Donor plasmids (100 ng μl^{-1}) PB arm-U6 promotor-DR- *Tyr* gRNA 1-DR- *Tyr* gRNA2-DR- CAG promotor -Cas13 (Cas13d, hfCas13d, Cas13X or hfCas13X)-p2A-GFP-PB arm and PBase mRNA (80 ng μl^{-1}) were injected into mouse zygotes. All mice used in this study were housed in a room with a 12 h light/dark cycle at a temperature of 20–24 °C and relative humidity of 45–65%. All animal experiments were approved by the Animal Care and Use Committee of the Institute of Neuroscience, Chinese Academy of Sciences, Shanghai, China.

Relative quantification of transposon copy numbers of transgenic mice. qPCR on DNA from mouse tails was used to determine the transposon copy numbers of individual mouse lines. En2SA DNA amounts were quantified and normalized to β-actin DNA levels. Transposon copy numbers were determined by normalizing values obtained from Cas13d, hfCas13d, Cas13X or hfCas13X mice to those from wild-type mice that only possessed endogenous En2SA (two copies). Primer sequences are shown in Supplementary Table 5.

Hydrodynamic tail vein injection, serum biochemistry and hepatocyte isolation. Male C57BL/6 (SLAC laboratory) mice aged 8 weeks were subjected to hydrodynamic tail vein injection. Mice were infected with 1.0×10^{11} transducing units of AAV in 150 μl phosphate-buffered saline by intravenous injection. Mice were euthanized 4 weeks after injection of AAV. Before whole blood collection, mice were fasted for 4 h. Whole blood was collected at the time of euthanasia. Whole blood was allowed to stand at room temperature for 1 h and centrifuged at 2,000*g* for 20 min. The serum was transferred into new tubes for further analysis. Serum PCSK9 protein was measured with a Mouse Proprotein Convertase9/PCSK9 Quantikine ELISA Kit (R&D Systems) according to the manufacturer's instructions. Serum parameters of liver function, including alanine aminotransferase and aspartate transaminase, were measured using an automatic biochemical analyzer at Adicon Clinical Laboratories Inc. (Shanghai, China). Mouse primary hepatocytes were isolated by standard two-step collagenase perfusion and purified using 40% Percoll (Sigma) with low-speed centrifugation (1,000 r.p.m., 10 min). Hepatocytes were resuspended in DMEM plus 10% fetal bovine serum for FACS, and specific cell populations were used for RNA extraction.

Statistical analysis. Statistical tests were performed using GraphPad Prism 8 and included two-tailed paired two-sample *t*-tests, two-tailed unpaired two-sample *t*-test, Dunnett's multiple comparisons test after one-way analysis of variance (ANOVA), Tukey's multiple comparisons test after one-way ANOVA and log-rank Mantel–Cox test. The statistical test used for the data shown in each figure is noted in the corresponding figure legend, and significant statistical differences are noted as $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. All values are reported as mean ± s.e.m.

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

Data availability

Publicly available datasets used in this study were as follows: GRCh38.p5. RNA-seq data are available under GEO accession number GSE168246. Source data are provided with this paper. Any other data can be obtained from the corresponding author upon reasonable request.

Code availability

Code that supports the findings of this study is available in the Supplementary Information.

References

46. Smargon, A. A. et al. Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* **65**, 618–630.e617 (2017).
47. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
48. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
49. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
50. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

Acknowledgements

We thank M.-m. Poo for helpful discussions and insightful comments on this manuscript; H. Yang and B. Wang from Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, for helpful discussions and technical assistance; L.-L. Chen from CAS Center for Excellence in Molecular Cell Science for the gift of the RanCas13b and dRanCas13b plasmids; and Y. Wang, Y. Zhang and Q. Hu from the Optical Imaging facility and S. Qian, H. Wu and L. Quan from the FACS facility of the Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences for technical support. We thank N. Zhong and L. Xie for technical assistance. This work was supported by HUIGENE Therapeutics Co., Ltd. (H.T.), Lingang Laboratory (LG202106-01-02) (H.Y.), Chinese National Science and Technology major project R&D Program of China (2018YFC2000101) (H.Y.), Strategic Priority Research Program of Chinese Academy of Science (XDB32060000) (H.Y.), National Natural Science Foundation of China (31871502, 31901047, 31925016, 91957122 and 82021001) (H.Y.), Basic Frontier Scientific Research Program of Chinese Academy of Sciences From 0 to 1 original innovation project (ZDBS-LY-SM001) (H.Y.), Shanghai Municipal Science and Technology Major Project (2018SHZDZX05) (H.Y.), Shanghai City Committee of Science and Technology Project (18411953700, 18JC1410100, 19XD1424400 and 19YF1455100) (H.Y.) and the International Partnership Program of Chinese Academy of Sciences (153D31KYSB20170059) (H.Y.).

Author contributions

H.T., J.H. and H.Y. jointly conceived the project. H.T., J.H., Q.X., B.H. and X.D. designed and conducted experiments. Y.L. performed bulk RNA-seq analysis. J.H. and X.Y. performed microinjection and counted the mice every day. Q.X., X.W., R.Z. and Y.W. performed qPCR assays and participated in FACS. X.D. and D.H. participated in protein purification and in vitro cleavage assays. W.Y performed mouse embryo transfer. Y.L., M.C., Q.W. and M.X. assisted with plasmid construction. Z.W., C.X., Y.Z., G.L. and K.F. assisted with cell experiments. H.Y. supervised the whole project. H.T., H.Z. and H.Y. wrote the manuscript.

Competing interests

H.T. discloses a patent application (PCT/CN2021/121926) related to the Cas proteins described in this manuscript. H.T. is an employee of HuiGene Therapeutics Co., Ltd. H.Z. is now an employee of HuiEdit Therapeutics Co., Ltd. H.Y. is a founder of HuiGene

Therapeutics Co., Ltd. and HuiEdit Therapeutics Co., Ltd. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-022-01419-7.

Correspondence and requests for materials should be addressed to Jia Huang, Hainan Zhang or Hui Yang.

Peer review information *Nature Biotechnology* thanks Jonathan Gootenberg 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.

nature research

Corresponding author(s): Hui Yang

Last updated by author(s): Jun 30, 2022

Reporting Summary

Nature Research 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 Research 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	RNA-seq data are available with the GEO accession number: GSE168246. Publicly available datasets used in this study are as follows: GRCh38.p5.
Data analysis	FlowJo V10.5.3, Graphpad Prism(v8.3.0), Image J(v1.52a), R(v4.0), I-TASSER(v5.1), Cutadapt(v0.6.6), hisat2(v2.2.1), htseq-count(HTSeq, v0.12.4), HOMER(v4.11), DAVID(v6.8), blast(v2.5.0), amiGO(based on the Panther database v16.0); Codes that support the findings of this study are available in the Supplementary Information.

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

RNA-seq data are available with the GEO accession number: GSE168246. Source data are provided with this paper. Any other data can be obtained from the corresponding author upon reasonable request. Publicly available datasets used in this study are as follows: GRCh38, release 32, https://www.gencodegenes.org/human/release_32.html

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	No sample-size calculation was performed. Experiments were performed in triplicates(n=3), unless otherwise moted, with the sample size numbers were listed in the corresponding figure legends. Sample sizes for these experiments were chosen based upon field standards and prior knowledge of experimental variation. (PMID: 29551272, 28976959)
Data exclusions	No data were excluded.
Replication	Data were obtained in the fashion of biological replications. We tested experimental conditions using different gRNAs to ensure robustness. All the in vitro and in vivo experimental results could be successfully reproduced. At least three independent experiments were performed.
Randomization	For transfection, wells were randomly assigned. For AAV injection, mice were randomly assigned into different groups.
Blinding	Blinding was not relevant to our study because it is not a subjective trial and the results presented here were purely based on objective description of our novel experimental technology.

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

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T, HEK293 cell lines were purchased from Stem Cell Bank, Chinese Academy of Sciences
Authentication	Cell lines were authenticated via STR analysis by the supplier.
Mycoplasma contamination	Cell lines have been tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell line listed in the database of ICLAC was used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Male DBA (SLAC Laboratory) and female C57BL/6 (SLAC Laboratory) mice at the age of 8 weeks were mated to obtain B6D2F1 zygotes. ICR female mice (SLAC Laboratory) at the age of 8 weeks were used for recipients.
Wild animals	The study did not involve wild animals.

Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All animal experiments were performed and approved by the Animal Care and Use Committee of the Institute of Neuroscience, Chinese Academy of Sciences, Shanghai, China.

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	To isolate cells, the transfected or non-transfected were dissociated enzymatically in an incubation solution of 100 µL Trypsin-EDTA (0.05%) at 37°C for 5 min. The digestion was stopped by adding 1 ml of DMEM medium with 10% Fetal Bovine Serum (FBS). The cell suspension was centrifuged for 3 min (1000 rpm), and the pellet was resuspended in DMEM medium with 10% FBS. Finally, the cell suspension was filtered through a 40-µm cell strainer, and mCherry+/EGFP+ cells were isolated by FACS.
Instrument	BD FACS Aria III, BD LSRFortessa X-20, Beckman CytoFLEX S, MoFlo Astrios EQ, or Moflo XDP.
Software	FlowJo V10.5.3
Cell population abundance	Samples were found to be >95% pure when assessed with a second round of flow cytometry analysis.
Gating strategy	Gating strategy: 1) in FSC-A/SSC-A gate for living cells, 2) using the non-transfected cells to define the gate for EGFP+ and/or mCherry+ cells, 3) apply this gate to all samples.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.