

Supplementary Figure 1 Related to **Figure 2a**. Evaluation of (a) single amino acid mutations and (b) combination mutants to construct the AsCpf1 RVR variant, which is active at target sites with TATV PAMs. Dots show mean \pm s.e.m. (n = 2).



Supplementary Figure 2 Related to **Figure 2b-d**. Histograms of abundances of 4⁸ PAMs (NNNNNNN) at each *in vitro* cleavage time point for (a) WT AsCpf1, (b) S542R/K607R, and (c) S542R/K548V/N552R. The color of each histogram represents elapsed time. NNNNVRRT sequences, which were used to center the histograms, are shown in black.



Supplementary Figure 3 Data processing pipeline for the *in vitro* cleavage assay used for Figure 2d.



Supplementary Figure 4 (a) Comparison of the activity of WT AsCpf1 to the RR variant at target sites with cytosine-containing PAMs. (b) Activity of the RR variant at TYCV and VYCV sites (V = A, C, or G), demonstrating that the presence of a 5' T in the PAM sequence can be optional in many cases (*i.e.*, NYCV PAMs can be recognized). The data for TYCV sites is the same as that shown in (a). All indel percentages were measured in HEK293T cells. Dots show mean \pm s.e.m. (n = 2-3).



Supplementary Figure 5 Related to **Figure 2e**. Activity of (a) WT AsCpf1, (b) the RR variant, and (c) the RVR variant at target sites with highly active PAMs in HEK293T cells. Dots show mean \pm s.e.m. (n = 3). **Figure 2e** shows these data in aggregate. For the AsCpf1 RR variant, the three CCCC sites are not included in **Figure 2e**.



Supplementary Figure 6 Editing efficiency of the AsCpf1 RR variant at TYCV sites in mouse Neuro2a cells. (a) Diagram of the mouse *PCSK9* locus. Gray boxes represent coding sequences. (b) Indel percentages produced by the RR variant at *PCSK9* target sites with TYCV PAMs. Bars show mean \pm s.e.m. (n = 3). (c) Representative indels at the target site (#2) with the highest editing efficiency. The red triangle represents the putative cleavage site on the top strand.



if either PAM A or PAM B can be recognized by nuclease, count as cut site



Supplementary Figure 7 Related to **Figure 2f**. (a) Definition of targeting range for Cpf1 and Cas9. Comparison of the targeting range of Cpf1 (+RR and RVR variants) to Cas9 (+VQR and VRER variants) in (b) the human genome and (c) coding sequences. Plots show the probability mass function of the distance (in base pairs) to the nearest cleavage site. The boxplots indicate median and interquartile range. Genomic regions that contain Ns or masked repeats were ignored in this analysis.

а





а	RECI	N	.D. ^{III}	<i>q</i> `	
	Maria Maria	All demonstration and the		ية مرا	1. And Martin and Martin and Martin and Andrews and Martin
	— III				BpCpf1 Bu
			11		LD3CpT1 La
			11		PeCnf1 Ca
			-11		
I −−−					
II					
					PbCpf1 Pa
<u> </u>					
			10.	0 01	EeCof1 Ei
				11 II I	CMtCpf1 Ca
<u> </u>					AsCpf1 Ac
					St.
					PmCpf1 Pc
					Pr
			10		LbCof1 La
Н.——				0.01	
IЧ—					
			III I	1 11	
l Lr	⋈ ⋶╡║╌ ⋶╌⋶ ┨╌╞╌┼╌╌╟║┠				
					Lb2Cpf1 La
					Licpti Le
			18 1		
			10	0 01	Balling and Balling an
<u> </u>					PcCpf1 Pc
<u>ا</u>					
				 	
1 <u> </u>					
			and the second s		

andidatus Gristian (goode 4) andidatus Roizmanbacteria bacterium GW2011_GWA2_37_7 arcubacteria group bacterium GW2011_GWC2_44_17 elcococcus kunzii roteocatella sphenisci andidatus Methanomethylophilus alvus achnospira pectinoschiza ubacterium rectale ubacterium eligens andidatus Methanoplasma termitum vnergistes jonesii cidaminococcus sp. BV3L6 uccinivibrio dextrinosolvens orphyromonas macacae revotella disiens naerovibrio sp. RM50 . loraxella bovoculi 237 rancisella tularensis subsp. novicida U112 hiomicrospira sp. XS5 achnospiraceae bacterium ND2006 achnospiraceae bacterium NC2008 utyrivibrio fibrisolvens seudobutyrivibrio ruminis bribacterium sp. NK2B42 utyrivibrio sp. NC3005 achnospiraceae bacterium MA2020 achnospiraceae bacterium MC2017 eptospira inadai chnospiraceae bacterium COE1 mithella sp. SC_K08D17 lavobacterium branchiophilum avobacterium sp. 316 acteroidales bacterium KA00251 orphyromonas crevioricanis revotella brevis acteroidetes oral taxon 274 revotella bryantii revotella albensis

b

а

	/								
_	1								
h		1							
D		1							
Buturivibrio proteoclasticus	BnCnf1	AP					Suppl		
Lashpaspirasaan bastarium MC2017	L b2Coff						Suppl		
Candidatus Fallusubastaria bastarium OM0011, OM00, 41, 14	LUSOPII			PGEK			consei		
Candidatus Paraminihantaria bartarium QW2011_GWA2_41_14	D-O-H						Comu		
Candidatus Peregrinibacteria bacterium GW2011_GWA2_33_10	Peopri	SLA	GGWD	VINKE		K W E D L L L	Seque		
candidate division ws6 bacterium Gw2011_GwA2_37_6		NLL	GGWS	DGQE	KNKA	KWEKLIL	putati		
uncultured bacterium (gcode 4)		Q F L		KDKE	NERL				
Candidatus Hoizmanbacteria bacterium GW2011_GWA2_37_7			NGWD	KNKE	SDNS	MFPKVCF	RECI,		
Parcubacteria group bacterium GW2011_GWC2_44_17	PbCpf1	ALL	KGWD	ENKE	YDFM	DVPRLLL	contai		
Helcococcus kunzii		I L A	DGWS	E <u>SK</u> I	SDNR	SIPNYPF	contai		
Proteocatella sphenisci		TLA	KGWD	ΑΝΚΕ	IDNS	<u>MLPKQF</u> L	mutag		
Candidatus Methanomethylophilus alvus		TLA	DGWD	LNKE	RDNK	MLPKIFV	abbrox		
Lachnospira pectinoschiza		TLA	SGWS	RSKE	FDNN	MLPKVFI	appres		
Eubacterium rectale		TLA	DGWS	KSKE	y s <mark>n</mark> n	MIPKVFL	previo		
Eubacterium eligens	EeCpf1	TLA	NGWS	QSKE	FDNN	MLPKVFL	2015		
Candidatus Methanoplasma termitum	CMtCpf1	TLA	NGWD	QNKV	YDYA	NLPRVFL	2015)		
Synergistes jonesii		ALA	KGWD	ΚΝΚΕ	QNNT	ELPKVFL	(green		
Acidaminococcus sp. BV3L6	AsCpf1	TLA	SGWD	VNKE	KNNG	MIPKCST	(B. 001)		
Succinivibrio dextrinosolvens		QLG	EGFS	KSKE	NDCL	FIPKCSI	mutat		
Porphyromonas macacae	PmCpf1	SLL	SGWD	K N K F		MIPKVFF	altered		
Prevotella disiens		K M	TGWK	SDNG	ΤΟΥĠ	RLNKVII			
Anaerovibrio sp. RM50		TMI	AGWD	YNNF	TÀNG	MIPKVVF	indica		
Moraxella bovoculi 237	MbCpf1	τ Π Γ	NGWD	INKF	KDNF	OFPKVFF	hasos		
Francisella tularensis subsp. novicida U112	FnCpf1	T A	NGWD	KNKF	PDNT	MIPKVFF	00303		
Thiomicrospira sp. XS5		Ť Í	SGWD	ANKF	TANA	MIPKVFF	for cla		
Lachnospiraceae bacterium ND2006	LbCof1	OEM	GGWD	KDKF	TDYR	MIPKVFF	Table		
Lachnospiraceae bacterium NC2008		õ 🖬	GGWD	VNKF	RDCY	MEPKVEE	Table		
Butvrivibrio fibrisolvens		T F I	D G W D	Ý G N F	FAVI	SIONIMV			
Pseudobutvrivibrio ruminis		TIT	NGWD	KNKF		MIPKVFF			
Oribacterium sn NK2B42		TEI	B GWD		FANI	MIPKVFF			
Butvrivibrio sp. NC3005		T III				MIPKVFF			
Lachnospiraceae bacterium MA2020	Lb2Cof1	+	NGWD			MIPKVEE			
Lachnospiraceae bacterium MC2017	Lozopii	수 비 는			KEVM				
Lentospirace Bacentain inegeri	LiCnf1	+		ENRE					
Lachnospiraceae bacterium COE1	Liopii	+							
Smithalla sp. SC. K08D17	ScCof1	누 듣 는							
Elevobactarium branchionbilum	Saobii								
Elavobactorium sp. 216		X H H	NEWE						
Bactoroidalos bactorium KA00251									
Bacteroidales bacteriain (Ado25)	DoCof1	3							
Porpriyronionas crevionicanis Provotella bravis	РССрп								
Postoroidateo eral taxon 274		+ $ -$							
Daciel oldeles of al taxoff 274		부 는 발							
Prevolella Diyantil			GGWD						
Prevolella albensis		Y L L	GGWD	VINKE					
			A	A	A	A			
			Т	Т	Т	Т			
			1						
Mutation in A	Coff.	5	100	E 4 0\/	FEOD	607P			
Wutation in P	sopri:	5	42Π	040V	332n	007 m			

lementary Figure 9 Sequence rvation of Cpf1 orthologs. (a) ence alignment of 43 Cpf1 or ive Cpf1 orthologs, highlighting the WED-II, and PI domains, which in the residues selected for genesis screening. Cpf1 name viations follow conventions we ously reported (Zetsche et al. Cell . (b) Zoom-in of the positions boxes) corresponding to the ed residues in AsCpf1 conferring d PAM specificity. A red line ates an insertion of one or more in the alignment that are omitted arity. See also Supplementary 3.



Supplementary Figure 10 Engineering the PAM recognition of LbCpf1. (a) Crystal structures of AsCpf1 (PBD ID: 5B43) and LbCpf1 (PDB ID: 5ID6), highlighting the corresponding residues mutated to alter PAM specificity. The PAM duplex shown for LbCpf1 is a model. (b) Activity of LbCpf1 G532R/K595R and G532R/K538V/Y542R at TYCV and TATV sites, respectively, in HEK293T cells. Each point represents the mean of three replicates, and the red lines indicate the overall means within each group. The data for AsCpf1 also appears in **Figure 2e**. n.s. *p* > 0.05 (Mann-Whitney); *****p* < 0.0001 (Wilcoxon signed-rank).



Supplementary Figure 11 Related to **Supplementary Figure 10b**. Activity of the (a) LbCpf1 RR variant and (b) LbCpf1 RVR variant at target sites with preferred PAMs in HEK293T cells. Dots show mean \pm s.e.m. (n = 3). **Supplementary Figure 10b** shows these data in aggregate. The target sites are the same as those shown in **Supplementary Figure 5b-c**. For the RR variant, the three CCCC sites are not included in **Supplementary Figure 10b**.