www.sciencemag.org/cgi/content/full/science.aas9129/DC1

Supplementary Materials for

Engineered CRISPR-Cas9 nuclease with expanded targeting space

Hiroshi Nishimasu*, Xi Shi, Soh Ishiguro, Linyi Gao, Seiichi Hirano, Sae Okazaki, Taichi Noda, Omar O. Abudayyeh, Jonathan S. Gootenberg, Hideto Mori, Seiya Oura, Benjamin Holmes, Mamoru Tanaka, Motoaki Seki, Hisato Hirano, Hiroyuki Aburatani, Ryuichiro Ishitani, Masahito Ikawa, Nozomu Yachie, Feng Zhang, Osamu Nureki*

*Corresponding author. Email: nisimasu@bs.s.u-tokyo.ac.jp (H.N.); nureki@bs.s.u-tokyo.ac.jp (O.N.)

Published 30 August 2018 on *Science* First Release
DOI: 10.1126/science.aas9129

This PDF file includes:

Materials and Methods

Sample preparation

The gene encoding SpCas9 (residues 1–1,368) was cloned between the *Nde*I and *Xho*I sites of the modified pET-28a vector (Novagen). The mutations were introduced by a PCR-based method, and the sequences were confirmed by DNA sequencing.

For *in vitro* cleavage experiments, the N-terminally His_6-tagged SpCas9 proteins were expressed in *Escherichia coli* Rosetta2 (DE3) (Novagen). The SpCas9-expressing *E. coli* Rosetta2 (DE3) cells were cultured at 37°C in LB medium (containing 20 mg/l kanamycin) until the OD_{600} reached 0.8, and protein expression was then induced by the addition of 0.1 mM isopropyl β-D-thiogalactopyranoside (Nacalai Tesque). The *E. coli* cells were further cultured at 20°C overnight, and harvested by centrifugation. The *E. coli* cells were resuspended in buffer A (50 mM Tris-HCl, pH 8.0, 20 mM imidazole and 1 M NaCl), lysed by sonication, and then centrifuged. The supernatant was mixed with 1 ml Ni-NTA Superflow (QIAGEN), and the mixture was loaded into a Poly-Prep Column (Bio-Rad). The protein was eluted with buffer B (50 mM Tris-HCl, pH 8.0, 0.3 M imidazole and 0.3 M NaCl). The protein was loaded onto a 1 ml HiTrap SP HP column (GE Healthcare) equilibrated with buffer C (20 mM Tris-HCl, pH 8.0, and 0.3 M NaCl). The protein was eluted with a linear gradient of 0.3–1 M NaCl. The purified proteins were stored at −80°C until use. The 98-nt sgRNA (containing stem loops 1, 2 and 3) was transcribed *in vitro* with T7 RNA polymerase, and was purified with an RNeasy kit (QIAGEN). *Staphylococcus aureus* Cas9 (SaCas9) and its sgRNA were prepared as described previously (*17*). *Acidaminococcus sp. BV3L6* Cas12a (AsCas12a), *Lachnospiraceae bacterium ND2006* Cas12a (LbCas12a), and their crRNAs were prepared as described previously (*18*).

For crystallization, the inactivating mutation (N863A) in the HNH domain was introduced into SpCas9-NG. The SpCas9-NG (N863A) protein was expressed in *E. coli* Rosetta2 (DE3), and was then purified using Ni-NTA and HiTrap SP HP columns. To remove the His_6-tag, the purified protein was mixed with TEV protease, and was dialyzed at 4°C overnight against buffer D (20 mM Tris-HCl, pH 8.0, 40 mM imidazole, 0.3 M NaCl). The protein was passed through the Ni-NTA column equilibrated with buffer D. The protein was further purified by chromatography on a HiLoad Superdex 200 16/60 column (GE Healthcare) equilibrated with buffer E (10 mM Tris-HCl, pH 8.0, 150 mM NaCl and 1 mM DTT). The 81-nt sgRNA (containing stem loops 1 and 2) was transcribed *in vitro* with T7 RNA polymerase, and was purified by 10% denaturing (7 M urea) polyacrylamide gel electrophoresis.

Structure determination

SpCas9-NG was crystallized using a protocol similar to that for the SpCas9 VQR variant (*14*). The SpCas9-NG–sgRNA–DNA complex was reconstituted by mixing the SpCas9-NG protein, the 81-nt sgRNA, the 28-nt target DNA strand (Sigma-Aldrich) and the 8-nt non-target DNA strand (Sigma-Aldrich) (molar ratio, 1:1.5:2.3:2.7). The SpCas9-NG–sgRNA–DNA complex was purified by gel filtration chromatography on a Superdex 200 Increase 10/300 column (GE Healthcare) equilibrated with buffer F (20 mM HEPES-NaOH, pH 7.5, 250 mM KCl, 5 mM $MgCl_2$ and 1 mM DTT). The purified SpCas9-NG–sgRNA–DNA complex was crystallized at 20°C, by the hanging-drop vapor diffusion method. The crystals were obtained by mixing 1 µl of complex solution ($A_{260\ nm}$ = 15) and 1 µl of reservoir solution (0.1 M Tris-acetate, pH 8.0, 0.4 M KSCN and 14–16% PEG 3,350). The crystals were improved by microseeding, using Seed Bead (Hampton Research). The crystals were cryoprotected in buffer (0.1 M Tris-acetate, pH 8.0, 0.4 M KSCN, 30% PEG 3,350 and 10% ethylene glycol). The X-ray diffraction data were collected at 100 K on beamline BL41XU at SPring-8 (Hyogo, Japan), and were processed using XDS (*19*) and AIMLESS (*20*). The structure was determined by molecular replacement with MOLREP (*21*), using the SpCas9 structure (PDB: 4UN3) as the search model. The structure models were built using COOT (*22*) and refined using PHENIX (*23*). Molecular graphic images were prepared using CueMol (http://www.cuemol.org).

***In vitro* cleavage experiment**

The *Bam*HI-linearized pUC119 plasmid, containing the 20-nt target sequence and the PAM, was used as the substrate for *in vitro* cleavage experiments. The linearized plasmid target bearing the TGG PAM (100 ng, 5 nM) was incubated at 37°C for 30 min with the SpCas9–sgRNA complex (200 nM, molar ratio, 1:2), in 10 µl of reaction buffer, containing 20 mM HEPES-NaOH, pH 7.5, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT and 5% glycerol. The reaction was stopped by the addition of quench buffer, containing EDTA (20 mM final concentration) and Proteinase K (10 µ g). The reaction products were analyzed using a MultiNA microchip electrophoresis system (SHIMADZU). To monitor the DNA cleavage time courses, the linearized plasmid target bearing the PAM (300 ng, 5 nM) was incubated at 37°C with the SpCas9–sgRNA complex (50 nM, molar ratio, 1:2), in 30 µl of reaction buffer. Aliquots (5 µl) were taken at 0.5, 1, 2 and 5 min, and mixed with 15 µl of quench buffer. The reaction products were analyzed using a MultiNA microchip electrophoresis system. *In vitro* cleavage experiments were performed at least three times, and representative electropherograms are shown.

The *in vitro* cleavage activity of SaCas9 was examined, using the pUC119 plasmid containing the 21-nt target sequence and the TTGAAT PAM. The *in vitro* cleavage activities of AsCas12a and LbCas12a were examined, using the pUC119 plasmid containing the 24-nt target sequence and the TTTA PAM.

PAM identification assay

The PAM identification assay was performed as described previously (*10*). The PAM library (80 ng) was incubated at 37°C with the purified SpCas9 (WT SpCas9 or SpCas9-NG) (100 nM) and the sgRNA (750 ng), in 10 µl of 1×CutSmart buffer (NEB). The reactions were quenched at the indicated times (<0.1, 1, 5, 15, 30 and 60 min) by the addition of 50 µl Buffer PB (QIAGEN), and were column-purified. The purified DNA was amplified with two rounds of PCR over 24 total cycles, using custom primers containing Illumina adaptors, and sequenced with a 75-cycle NextSeq kit (Illumina).

Indel analysis

The indel analysis was performed as previously described (*14*). Briefly, HEK293FT cells were plated one day prior to transfection in 96-well plates (Corning), at a density of 2.5×10^4 cells per well. Cells were transfected with the Cas9 plasmid (75 ng) and the sgRNA plasmid (25 ng) per well, using Lipofectamine 2000 (Life Technologies) according to the manufacturer's recommended protocol. Cells were harvested approximately 3 days post-transfection. Genomic DNA was extracted using a QuickExtract DNA extraction kit (Epicentre). PCR fragments for targeted deep sequencing were generated in two step PCR reactions, as previously described (*24*). Briefly, genomic regions of interest were amplified using primers with PCR handles for the second round of amplification, and then Illumina P5 adapters as well as unique sample-specific barcodes were attached to the first round PCR products.

GUIDE-seq

The human embryonic kidney (HEK) cell line 293FT (Fisher Scientific) was maintained in Dulbecco's modified Eagle's medium (DMEM, Life Technologies), supplemented with 10% fetal bovine serum (Gibco), at 37°C in a 5% CO_2 atmosphere. GUIDE-seq experiments were performed essentially as previously described (*13*). Briefly, 2×10^5 HEK293FT cells were transfected using a Lonza 4D-Nucleofector X Unit (program CM-137), in 20 µl Solution SE with 0.3 µg Cas9 and 0.2 µg sgRNA plasmids (1:1 molar ratio) targeting the *EMX1* or *VEGFA* locus, along with 10 pmol of a GUIDE-seq end-protected dsODN containing an *Nde*I restriction site. The GUIDE-seq tag integration frequencies at the intended on-target sites were estimated by restriction-fragment length polymorphism (RFLP) assays, using the primers: 5′-CCATCCCCTTCTGTGAATGT-3′ (Fw) and 5′-GGAGATTGGAGACACGGAGA-3′ (Rev) for *EMX1*, and 5′-ACTGACTAACCCCGGAACCA-3′ (Fw) and 5′-AAATTACCCATCCGCCCCC-3′ (Rev) for *VEGFA*. Tag-specific amplification and library preparation were performed before high-throughput sequencing, using an Illumina NextSeq 500/550 High output V2 kit (300 cycles). GUIDE-seq data were analyzed with the GUIDE-seq analysis software (http://www.jounglab.org/guideseq).

Base-editing analysis

HEK293Ta cells were maintained in DMEM (Sigma) supplemented with 10% (v/v) fetal bovine serum (FBS) (Thermo Fisher Scientific) and 1% Penicillin-Streptomycin (Sigma), at 37°C in a 5% CO_2 atmosphere. 5×10^5 HEK293Ta cells/well were transfected with 300 ng of base-editor plasmid and 100 ng of sgRNA plasmid, using 1.2 μl of 1 mg/ml Polyethylenimine (Polyscience). The cells were harvested 3 days after transfection, treated with 200 μl of 50 mM NaOH, incubated at 95°C for 10 min, and then neutralized with 20 μl of 1 M Tris-HCl, pH 8.0. Briefly, two round PCRs were performed for library preparation for high-throughput amplicon sequencing. Genomic regions of interest were first amplified to add the custom adapter sequence. The first PCR products were subjected to the second round PCR to attach Illumina TruSeq adapters and sample specific indices. After quantification using a KAPA Library Quantification Kit Illumina (KAPA Biosystems), the sequencing library was subjected to paired-end sequencing (600 cycles) on MiSeq (Illumina). The sequencing reads were demultiplexed based on sample specific indices and primer sequences and mapped to the target regions, using NCBI BLAST+ (version 2.6.0) with the blastn-short option. The C-to-T editing frequency and mutational spectrum for each target tested were normalized by subtracting the sequencing control samples (null vector transfection experiments), as described previously (*16*).

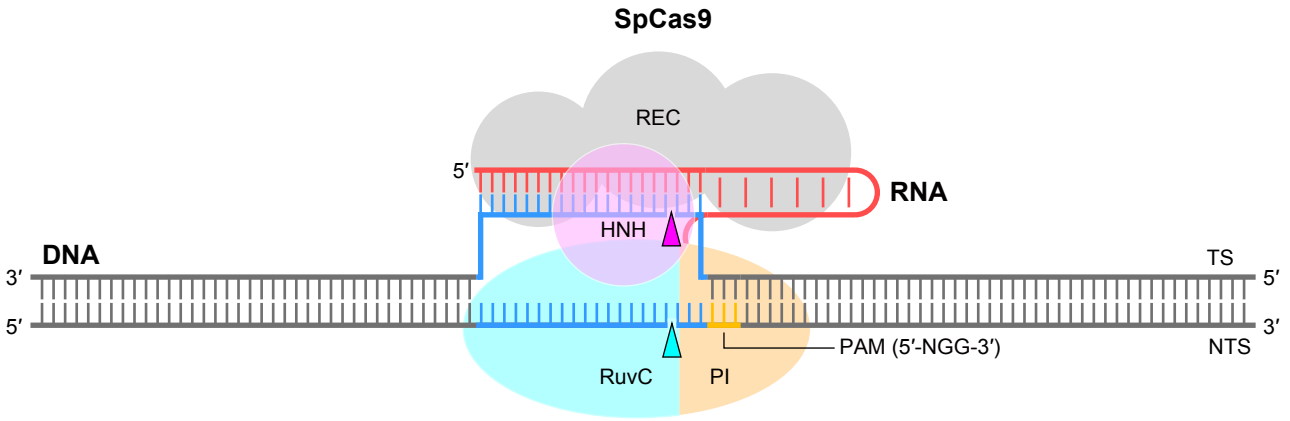


Fig. S1. Schematic of RNA-guided DNA cleavage by SpCas9.
The 20-nt target sequence (protospacer) and the 5′-NGG-3′ PAM sequence (in the non-target strand) are colored blue and yellow, respectively. The sites cleaved by the HNH and RuvC nuclease domains are indicated by magenta and cyan triangles. TS, target strand; NTS, non-target strand.

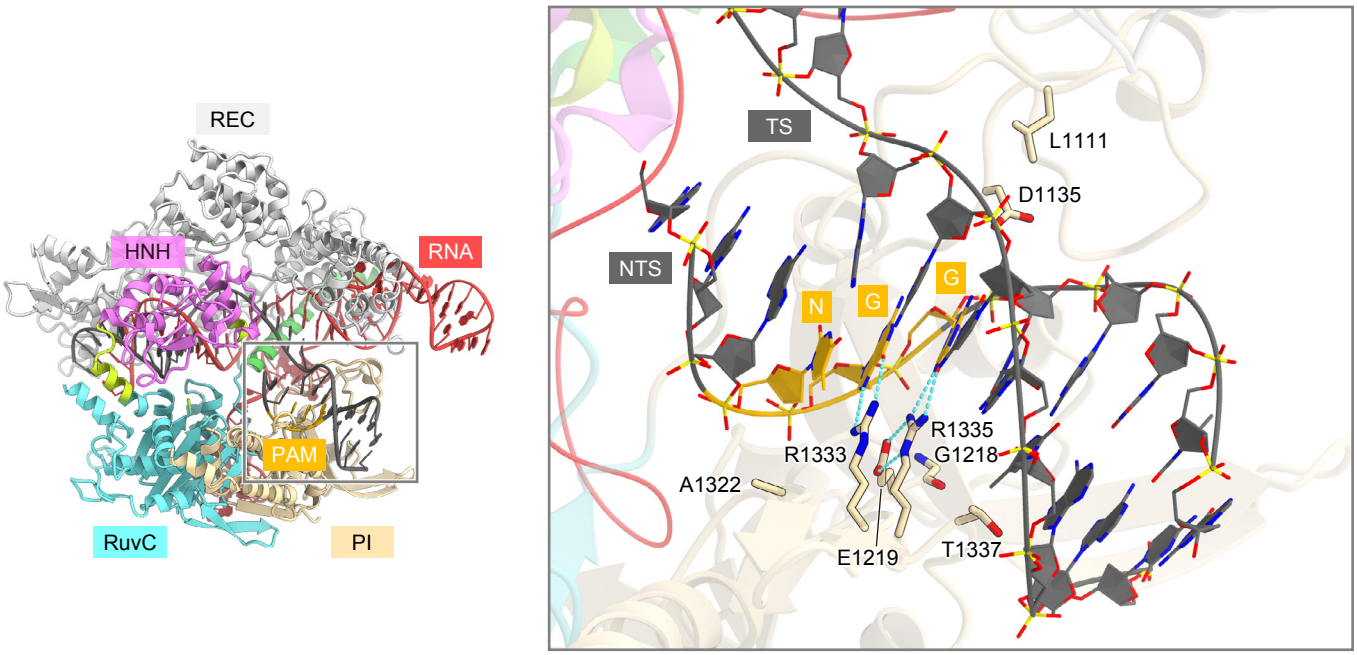


Fig. S2. PAM recognition by SpCas9.
In the crystal structure of SpCas9 in complex with the sgRNA and its target DNA (PDB: 4UN3), the second and third G nucleobases in the PAM are recognized by Arg1333 and Arg1335 in the PAM-interacting (PI) domain, respectively (*7*). TS, target strand; NTS, non-target strand.

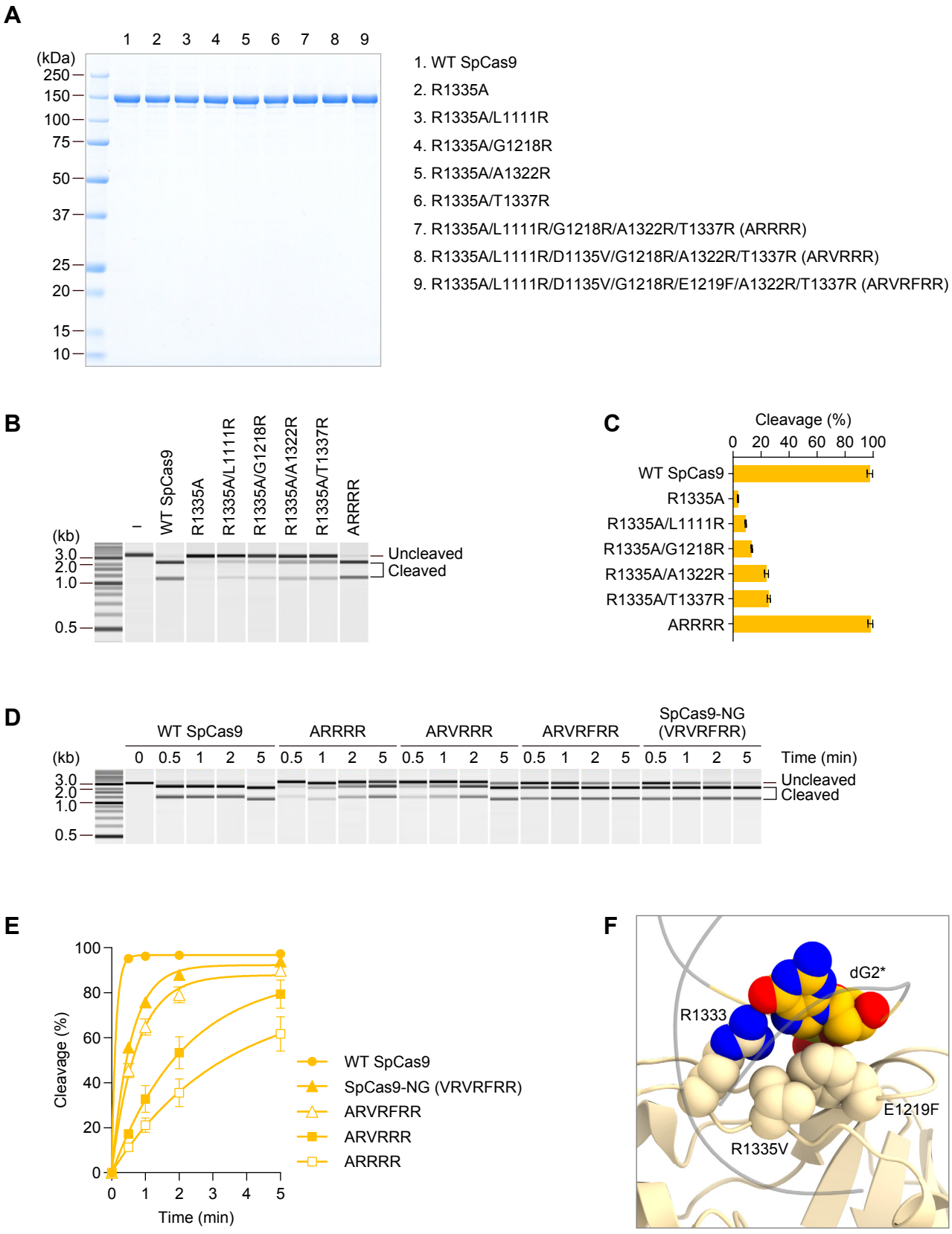


Fig. S3. Engineering of SpCas9-NG.

(**A**) SDS-PAGE analysis of the wild-type (WT) and mutant SpCas9 proteins used for *in vitro* cleavage experiments. The gel was stained with SimplyBlue SafeStain.
(**B**) *In vitro* DNA cleavage activities of WT SpCas9 and the R1335A, R1335A/L1111R, R1335A/G1218R, R1335A/A1322R, R1335A/T1337R and R1335A/L1111R/G1218R/A1322R/T1337R (ARRRR) mutants. A linearized plasmid target bearing the TGG PAM was incubated with the SpCas9–sgRNA complex (200 nM) at 37°C for 30 min, and the reaction products were then analyzed using a MultiNA microchip electrophoresis system.
(**C**) Quantification of the DNA cleavage activities in (**B**). Data are shown as mean ± s.d. (n = 3).
(**D**) *In vitro* DNA cleavage time courses of WT SpCas9 and the R1335A/L1111R/G1218R/A1322R/T1337R (ARRRR), R1335A/L1111R/D1135V/G1218R/A1322R/T1337R (ARVRRR), R1335A/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R (ARVRFRR), and R1335V/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R (VRVRFRR, SpCas9-NG) mutants. A linearized plasmid target bearing the TGG PAM was incubated with the SpCas9–sgRNA complex (50 nM) at 37°C for the indicated time points (0.5, 1, 2 and 5 min), and the reaction products were then analyzed using a MultiNA microchip electrophoresis system.
(**E**) Quantification of the DNA cleavage time courses in (**D**). Data are shown as mean ± s.d. (n = 3).
(**F**) Modeling of Phe1219 (E1219F) and Val1335 (R1335V) into the SpCas9–sgRNA–target DNA complex structure (PDB: 4UN3).

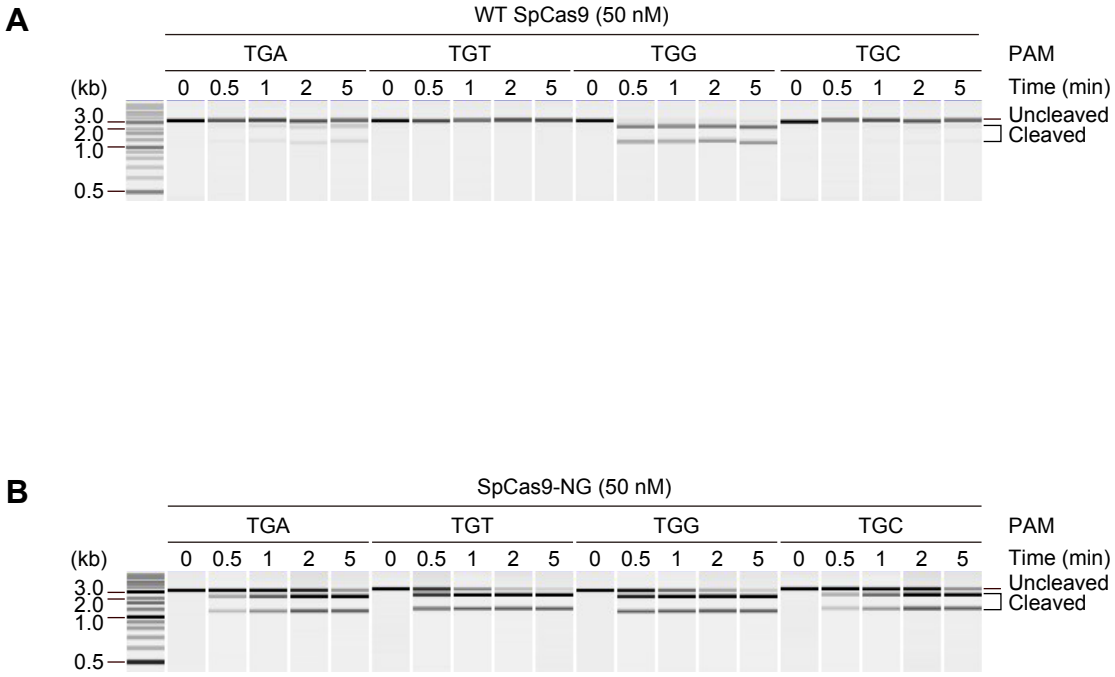


Fig. S4. *In vitro* DNA cleavage activities of WT SpCas9 and SpCas9-NG.
(**A** and **B**) *In vitro* DNA cleavage time courses of WT SpCas9 (**A**) and SpCas9-NG (**B**). A linearized plasmid target bearing the TGN PAM was incubated with the SpCas9–sgRNA complex (50 nM) at 37°C for the indicated time points (0.5, 1, 2 and 5 min), and the reaction products were then analyzed using a MultiNA microchip electrophoresis system.

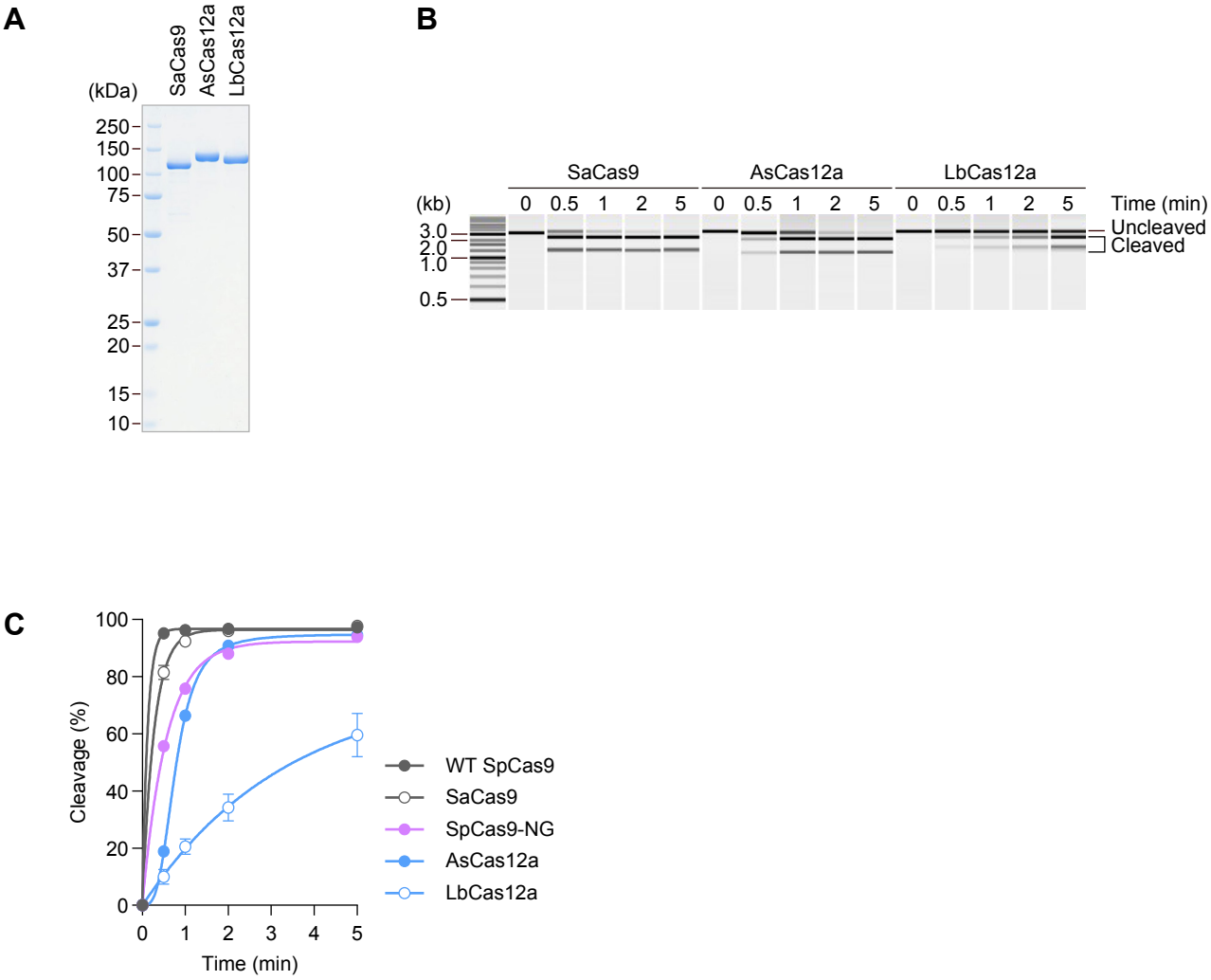


Fig. S5. Comparison of the *in vitro* DNA cleavage activities of WT SpCas9, SpCas9-NG, SaCas9, AsCas12a and LbCas12a.
(**A**) SDS-PAGE analysis of SaCas9, AsCas12a and LbCas12a.
(**B**) *In vitro* DNA cleavage time courses of SaCas9, AsCas12a and LbCas12a. A linearized plasmid target bearing the TTGAAT PAM was incubated with the SaCas9–sgRNA complex (50 nM) at 37°C for the indicated time points (0.5, 1, 2 and 5 min). A linearized plasmid target bearing the TTTA PAM was incubated with the Cas12a–crRNA complex (50 nM) at 37°C for the indicated time points (0.5, 1, 2 and 5 min). The reaction products were then analyzed using a MultiNA microchip electrophoresis system.
(**C**) Quantification of the DNA cleavage activities in (**B**). Data are shown as mean ± s.d. (n = 3). The DNA cleavage time courses of WT SpCas9 and SpCas9-NG for the TGG PAM targets (Fig. 1, B and C) are shown for comparison.

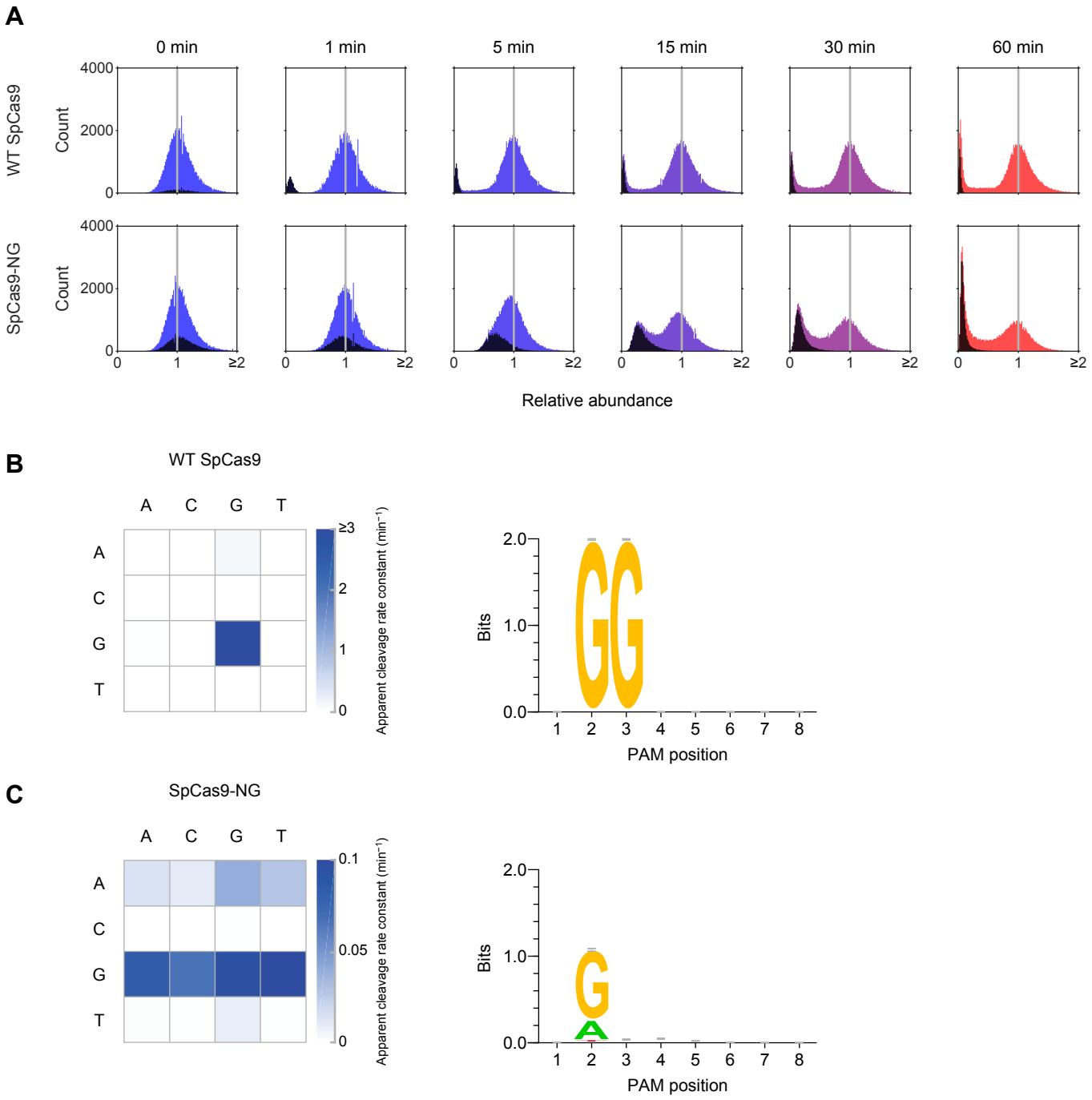


Fig. S6. PAM identification analysis.
(**A**) Histograms of the abundances of 4^8 PAMs (NNNNNNNN) at each *in vitro* cleavage time point for WT SpCas9 and SpCas9-NG. The color of each histogram represents elapsed time. NGG (WT SpCas9) and NGN (SpCas9-NG) sequences are shown in black.
(**B** and **C**) PAM specificities of WT SpCas9 (**B**) and SpCas9-NG (**C**).

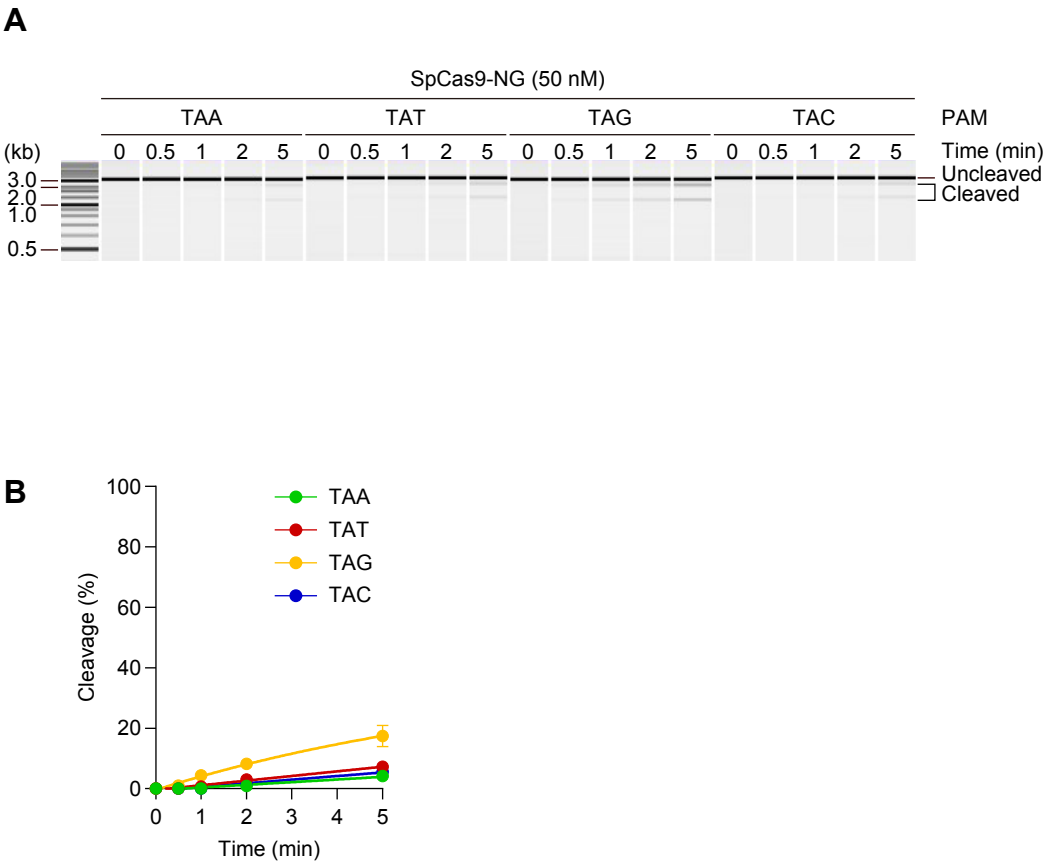


Fig. S7. *In vitro* DNA cleavage activity of SpCas9-NG toward the TAN PAM targets.
(**A**) *In vitro* DNA cleavage time courses of SpCas9-NG toward the TAN PAM targets. A linearized plasmid target bearing the TAN PAM was incubated with the SpCas9–sgRNA complex (50 nM) at 37°C for the indicated time points (0.5, 1, 2 and 5 min), and the reaction products were then analyzed using a MultiNA microchip electrophoresis system. Data are shown as mean ± s.d. (n = 3).
(**B**) Quantification of the DNA cleavage activities in (**A**). Data are shown as mean ± s.d. (n = 3).

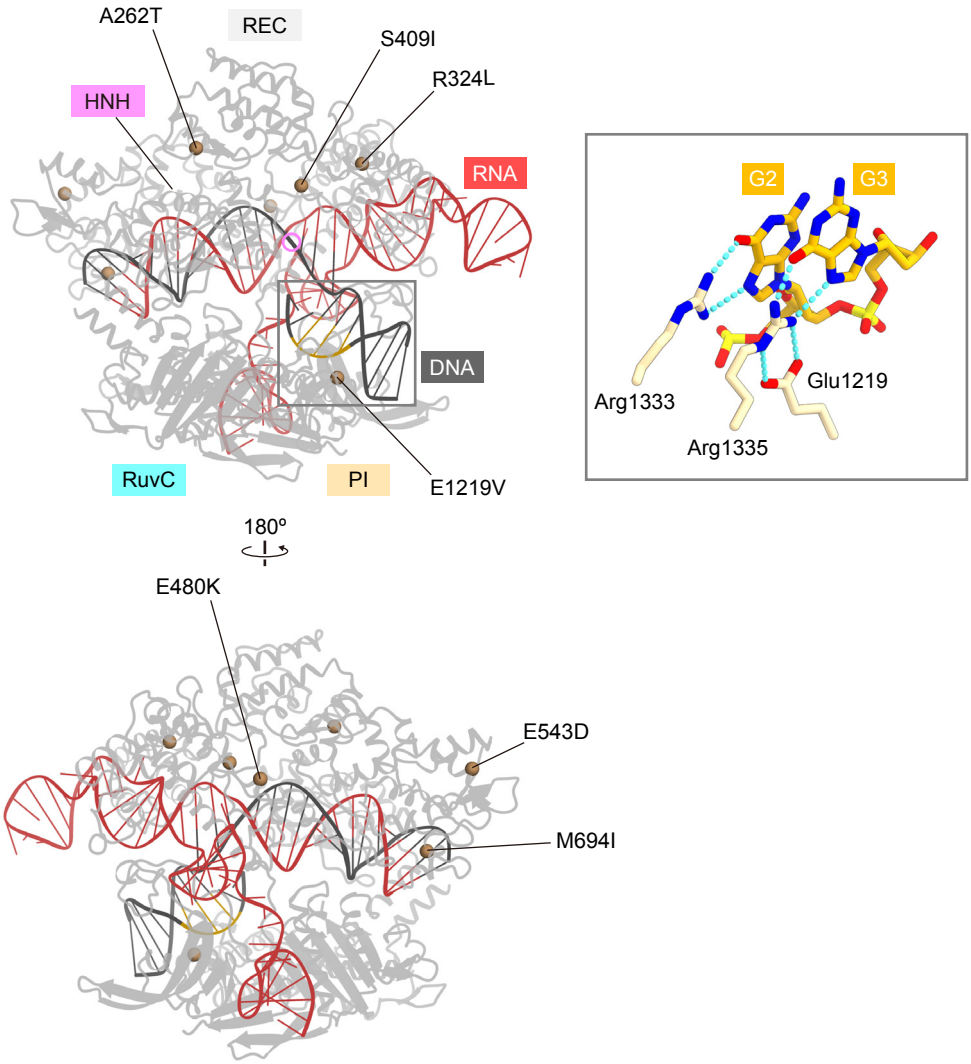


Fig. S8. xCas9 mutations.
The seven mutations in xCas9 (xCas9 3.7) are mapped on the SpCas9–sgRNA–target DNA complex structure (PDB: 4UN3). SpCas9-NG and xCas9 do not share the identical mutation. E1219V and A262T/R324L/S409I/E480K/E543D/M694I are located in the PI and REC domains, respectively. Glu1219 forms salt bridges with Arg1335, suggesting that the E1219V mutation increases the flexibility of Arg1335 to allow the less stringent recognition of the third PAM nucleotide. Ala262, Arg324 and Ser409 are located near the cleavage site in the target DNA strand (magenta circle), suggesting that the A262T/R324L/S409I mutations may facilitate the DNA cleavage by the HNH domain. Met694 forms hydrophobic interactions with the RNA–DNA heteroduplex, suggesting that the M694I mutation could affect the interactions with the RNA–DNA heteroduplex. The roles of the E480K and E543D mutations are unclear, since Glu480 and Glu543 are located farther away from the nucleic acids and exposed to the solvent.

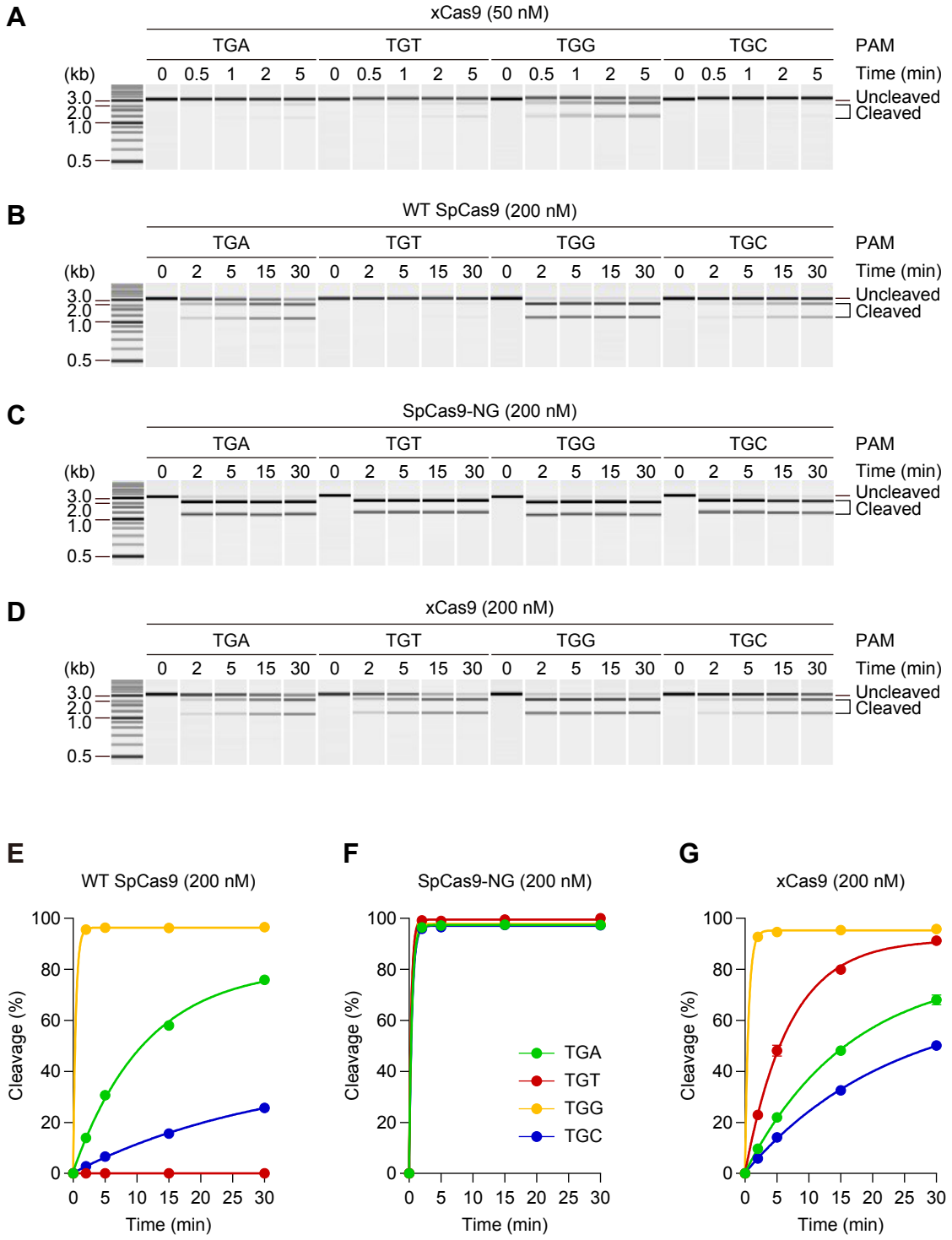


Fig. S9. Comparison of the *in vitro* DNA cleavage activities of WT SpCas9, SpCas9-NG and xCas9. (**A** to **D**) *In vitro* DNA cleavage time courses of xCas9 (50 nM) (**A**), WT SpCas9 (200 nM) (**B**), SpCas9-NG (200 nM) (**C**) and xCas9 (200 nM) (**D**). A linearized plasmid target bearing the TGN PAM was incubated with the SpCas9–sgRNA complex at 37°C for the indicated time points, and the reaction products were then analyzed using a MultiNA microchip electrophoresis system.
(**E** to **G**) Quantification of the DNA cleavage activities in (**B** to **D**). Data are shown as mean ± s.d. (n = 3).

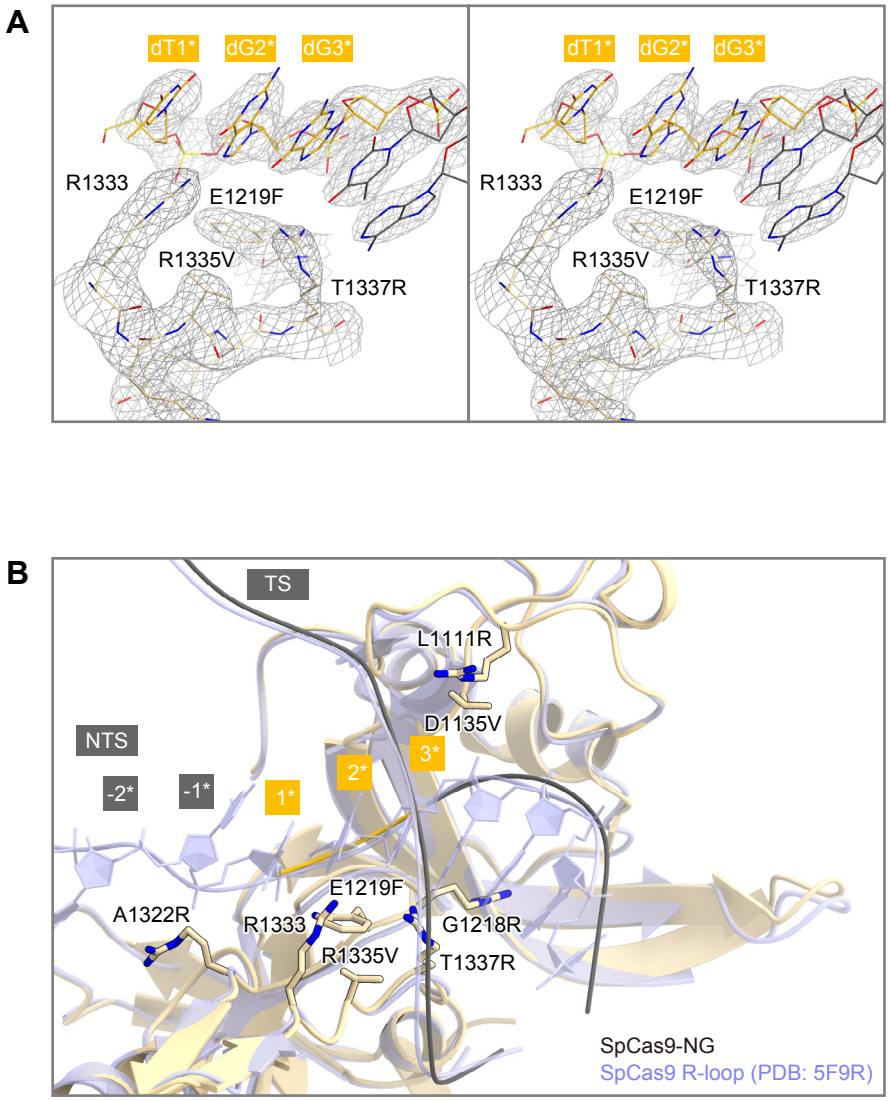


Fig. S10. Crystal structure of SpCas9-NG.
(**A**) $2mF_O - DF_C$ electron density maps for the key residues (contoured at 1.0σ) and the non-target DNA strand (contoured at 2.2σ) (stereo view).
(**B**) Structural comparison of SpCas9-NG with the SpCas9 R-loop complex (PDB: 5F9R) (semi-transparent blue).

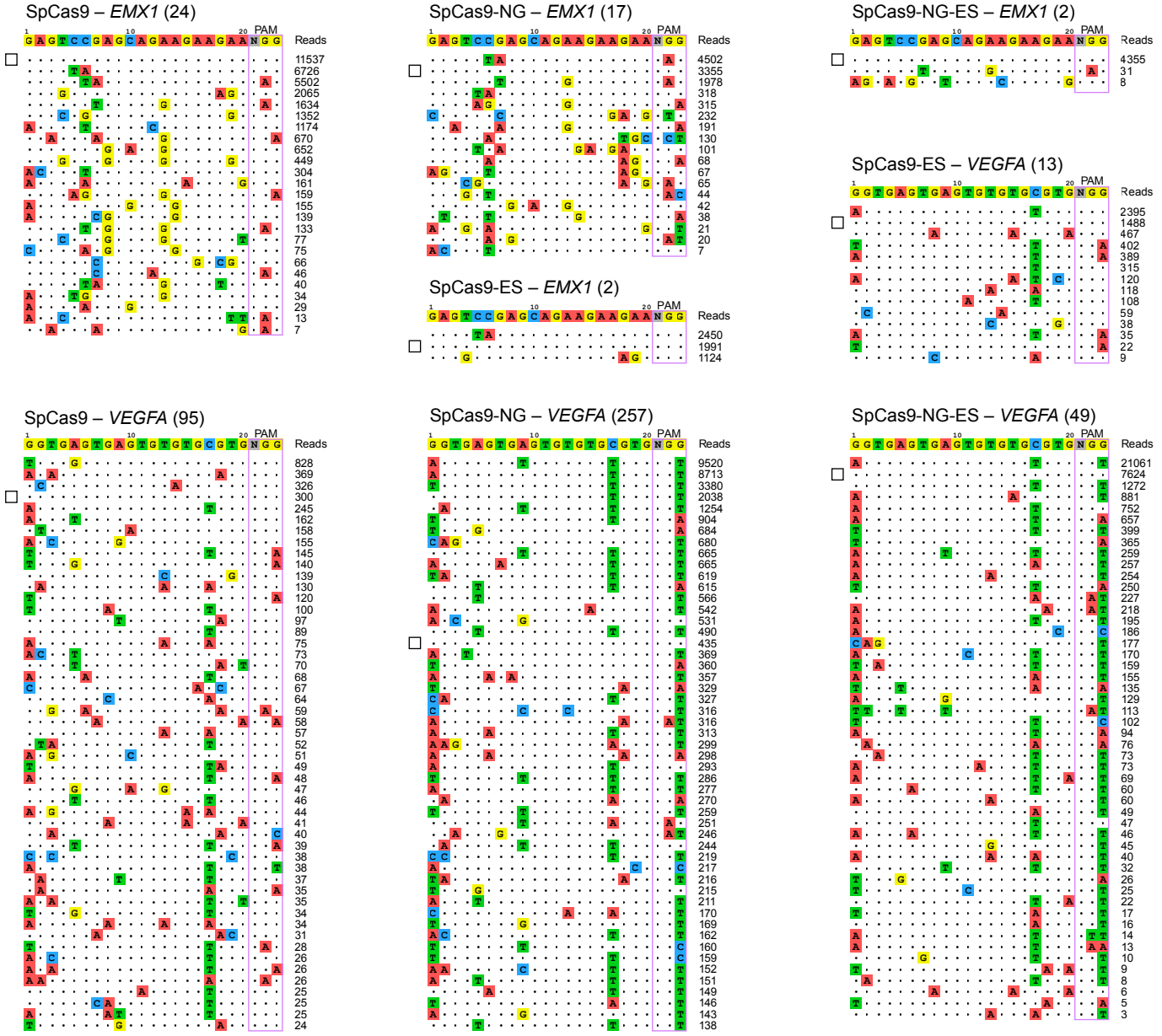


Fig. S11. Off-target sites identified by GUIDE-seq.
GUIDE-seq on-target and off-target reads for the *EMX1* and *VEGFA* sites are shown for SpCas9, SpCas9-ES, SpCas9-NG and SpCas9-NG-ES. The on-target sequences are indicated by open squares. The total numbers of identified off-target sites are shown in parentheses. The top 50 off-target sites are shown for the *VEGFA* target for SpCas9 and SpCas9-NG.

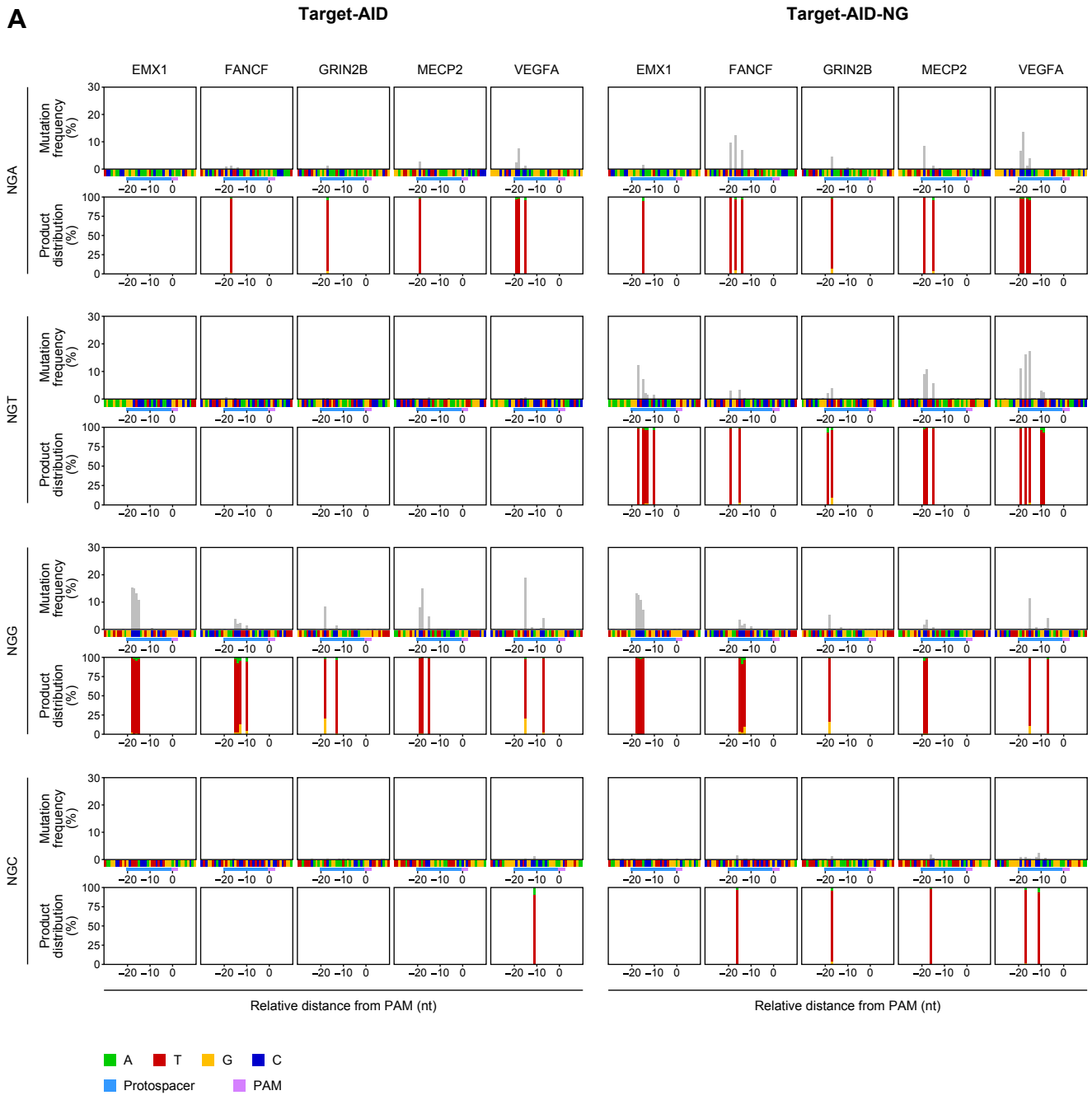
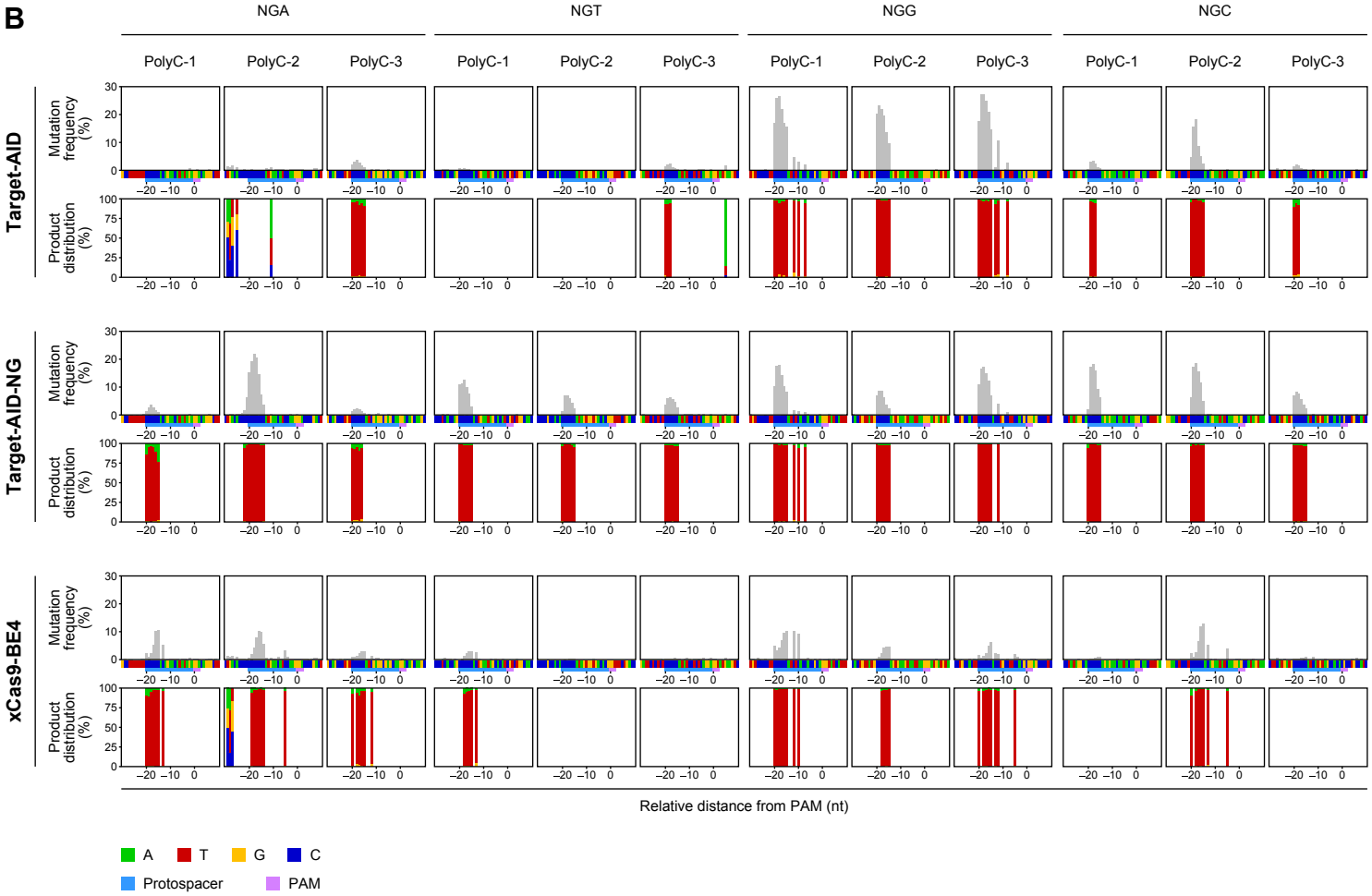


Fig. S12. Base editing in human cells.
(**A**) Base-editing spectra of Target-AID and Target-AID-NG at the 20 endogenous target sites.
(**B**) Base-editing spectra of Target-AID, Target-AID-NG and xCas9-BE4 at the 12 poly-C sites. Relative mutated base occupancies are shown for positions with mutation frequencies of ≥1%.


B
NGA
NGT
NGG
NGC
PolyC-1
PolyC-2
PolyC-3
Target-AID
Target-AID-NG
xCas9-BE4
Mutation frequency (%)
Product distribution (%)
Relative distance from PAM (nt)
A
T
G
C
Protospacer
PAM

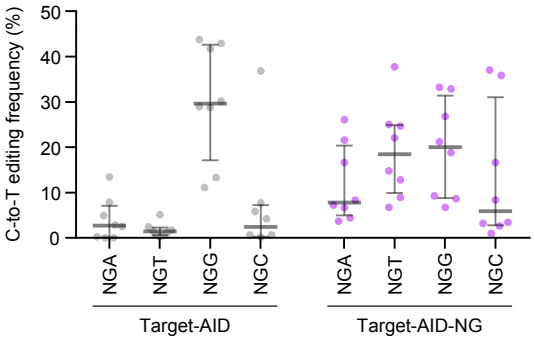


Fig. S13. C-to-T editing efficiencies of Target-AID and Target-AID-NG at the 32 endogenous target sites in HEK293T cells.

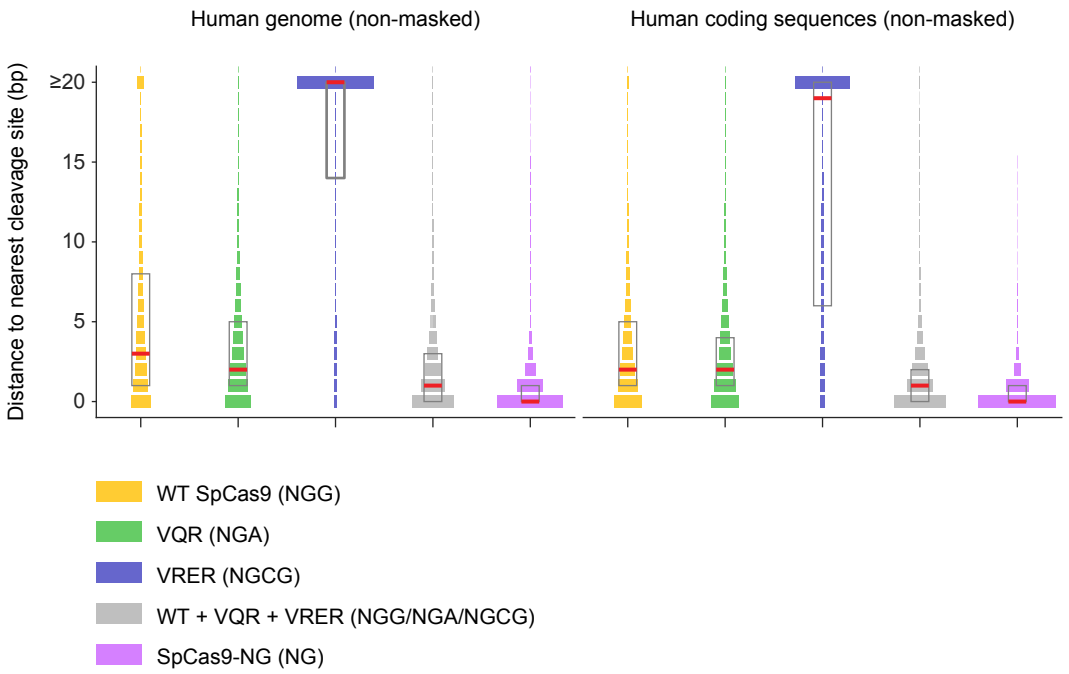


Fig. S14. Targeting range of SpCas9-NG.
Comparison of the targeting ranges of WT SpCas9, SpCas9-NG, and the previously described SpCas9 variants (VQR, VRER, and WT + VQR + VRER) in the human genome and coding sequences. Plots show the probability mass function of the distance (in base pairs) to the nearest cleavage site. The boxplots indicate median and interquartile range. Genomic regions that contain Ns or masked repeats were ignored in this analysis. SpCas9-NG increases the targeting range of SpCas9 to one cleavage site for every ~2.2 bp in the human coding sequence (~3.1-fold increase relative to WT SpCas9 alone). Importantly, SpCas9-NG can efficiently target NGT sites that are inaccessible by WT SpCas9 or the VQR and VRER variants (*4*).

Table S1 Data collection and refinement statistics.

Data collection	
Beamline	SPring-8 BL41XU
Wavelength (Å)	1.000
Space group	*C*2
Cell dimensions	
a, *b*, *c* (Å)	177.4, 68.6, 188.0
α, *β*, *γ* (°)	90, 111.5, 90
Resolution (Å)*	48.17–2.70 (2.77–2.70)
R_{merge}	0.099 (1.202)
R_{pim}	0.063 (0.778)
$I/\sigma I$	9.5 (1.7)
Completeness (%)	100 (100)
Multiplicity	3.5 (3.4)
CC(1/2)	0.992 (0.553)
Refinement	
Resolution (Å)	48.17–2.70
No. reflections	58,087
R_{work}/R_{free}	0.218/0.244
No. atoms	
Protein	10,535
Nucleic acid	2,464
Ion	13
Solvent	48
B-factors (Å^2)	
Protein	89.0
Nucleic acid	85.1
Ion	87.5
Solvent	51.7
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.452
Ramachandran plot (%)	
Favored region	97.18
Allowed region	2.74
Outlier region	0.08

*Values in parentheses are for the highest resolution shell.

Table S2 Target sequences of the indel analysis.

Locus	PAM	Site #	Sequence (5′ to 3′)
EMX1	GGG	1	GGCCCCAGTGGCTGCTCTGG
	TGG	2	GTCACCTCCAATGACTAGGG
	GGT	1	GCAGGGATCCAAGCACACAA
	GGT	2	GGGCTCCCATCACATCAACC
	GGA	1	GCAGCCAGCGACGTGCCCCA
	AGA	2	GAGGACAAAGTACAAACGGC
	GGC	1	GTCTCTCTCTTAATGACACG
	AGC	2	GCTTTACCCAGTTCTCTGGG
VEGFA	GGG	1	GATGTCTGCAGGCCAGATGA
	GGG	2	GCTGAGGTGCAGAATCCAGG
	CGT	1	GCTCTCAGGCCCTGTCCGCA
	GGT	2	GGTTCCGTGCCAGCTCCCAA
	CGA	1	GGGAAGCTGGGTGAATGGAG
	TGA	2	GCCACCACAGGGAAGCTGGG
	AGC	1	GCGCTCGGCCACCACAGGGA
DYRK1A	CGG	1	GTTGCCCTCATTATTCAGCA
	TGT	1	GTACCTATCTGAGCATACCG
	GGT	2	GTCAGTGGAGAGCACCATAA
	CGA	1	GCATCTCTGTGTATACCAAA
	AGA	2	GGACTGTTTGAAAGGTGCAG
	CGC	1	GCTCCAGCCTAATTTCCCAA
GRIN2B	CGG	1	GCAGGAAGAACAGTTCAAGA
	AGT	1	GCAAATACCAGAGATAAGAG
	GGT	2	GCTCTGTCAGAAGATCTCCA
	AGA	1	GCTGTAACAGGAGGGCCAGG
	AGA	2	GAGCAAATACCAGAGATAAG
	GGC	1	GCATTGCTGTCATCCTCGTG
	GGC	2	GATCAGGAAATAGAGCCACA
MECP2	GGG	1	GCAGAGCTAGGGGTTCAGAG
	AGG	2	GCCTGCATATTACAACCAAG
	AGT	1	GCCTTCATAGAATTGAAGAG
	GGT	2	GGAACCTGTGAGCTGCACCA
	CGA	1	GCGTACTTCGAAAAGGTAGG
	GGA	2	GGTCCTGGAGAAAAGTCCTG
	CGC	1	GTATCCGGAGAGAATTTGCA
FANCF	AGG	1	GGACTCTCTGATGAAGACCC
	TGG	2	GGAATCCCTTCTGCAGCACC
	CGT	1	GCGGTCTCAAGCACTACCTA
	AGA	1	GCGCTTCAATGGCTATAGAG
	GGA	2	GAGAACCCAAATCTCCAGGA
	GGC	1	GCGGCTGCACAACCAGTGGA
	CGC	2	GTGACGTCCTGCTCTCTCTG
PTEN	GGG	1	GGAGGCTATCAACAAAGAAT
	AGT	1	GCAGATTTCTTGAGTCCAGG
	GGT	2	GTTTGCAGTTGGCTAAGAGA
	GGA	1	GTACCTAATGGACTTCAGGG
	TGA	2	GTACTCATACCAAAAAAGAG
	GGC	1	GTTCTGTCACCAACTGAAGT
	GGC	2	GCTACGCTAGCAGTGAGCAA
SHANK3	CGG	1	GCCCGCTGCCGTATCCCGAG
	CGG	2	GAGCCCTCCCCGACCCACCG
	CGT	1	GCTCCTTTAACCTTGACGAA
	AGT	2	GTACGGTGTTGAATACCAGT
	CGA	1	GCGCTTCCCGCGGAGCACCT
	GGA	2	GGTTGCCCCCAACAGCCCCA
	TGC	1	GCAGAGGTAGAATCAGCAGG
	CGC	2	GGACCATGAGATAGAAGGCG
CUL3	TGG	1	GTAAACCTGGAATAACACGA
	AGG	2	GCCCAGTTCGAACTTCCTGG
	GGT	1	GCAAGGGCTGAGCAGCAGTA
	GGA	1	GATGGCCCCTAGAAAAGCAA
UBE3A	TGG	1	GTACAGTTAGTACTCAGCAG
	AGG	2	GCAGTTGTAGGGAAATAGGG
	GGT	1	GCTTTGCTACAACCCCAGCA
	GGT	2	GCTTCGAAGTGCTTGAAAAT
	GGA	1	GTAAGCATAGAGGTGCTATG
	AGA	2	GTAGACCAGGAGAATAAAGG
	AGC	1	GCCCATCCCTGAGTCCAGCG
	GGC	2	GCACTTGTCCGGGTAAGTTG

Table S3 Target sequences of the base-editing analysis.

Locus	PAM	Sequence (5′ to 3′)
EMX1	GGA	GAGGACAAAGTACAAACGGC
	GGT	GGGCTCCCATCACATCAACC
	GGG	GGCCCCAGTGGCTGCTCTGG
	AGC	GCTTTACCCAGTTCTCTGGG
VEGFA	TGA	GCCACCACAGGGAAGCTGGG
	CGT	GCTCTCAGGCCCTGTCCGCA
	GGG	GATGTCTGCAGGCCAGATGA
	AGC	GCGCTCGGCCACCACAGGGA
GRIN2B	AGA	GAGCAAATACCAGAGATAAG
	GGT	GCTCTGTCAGAAGATCTCCA
	GGG	GTCTGACCGGAAGATCCAGG
	GGC	GATCAGGAAATAGAGCCACA
FANCF	AGA	GCGCTTCAATGGCTATAGAG
	CGT	GCGGTCTCAAGCACTACCTA
	TGG	GGAATCCCTTCTGCAGCACC
	CGC	GTGACGTCCTGCTCTCTCTG
MECP2	CGA	GCGTACTTCGAAAAGGTAGG
	AGT	GCCTTCATAGAATTGAAGAG
	AGG	GCCTGCATATTACAACCAAG
	CGC	GTATCCGGAGAGAATTTGCA
PolyC-1	GGA	CCCCCCACAGAACATATAGA
	CGT	CCCCCCACTGAATAGCACAT
	GGG	CCCCCCATCTCACCACCGCA
	GGC	CCCCCCAGAGAACAATACAA
PolyC-2	GGA	CCCCCCCATGAACATCCAAG
	GGT	CCCCCCATAGTGGTGACACG
	GGG	CCCCCCAGTGAAGTAATTAA
	AGC	CCCCCCACAGAATATCAAGG
PolyC-3	GGA	CCCCCCATCGAGGCGAGCAA
	GGT	CCCCCCATATCCTCTCAGGA
	CGG	CCCCCCACCTGCCTTCCACG
	CGC	CCCCCCATCATGTAGCCACA

References

1. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). doi:10.1126/science.1225829 Medline

2. L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, F. Zhang, Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013). doi:10.1126/science.1231143 Medline

3. S. H. Sternberg, S. Redding, M. Jinek, E. C. Greene, J. A. Doudna, DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014). doi:10.1038/nature13011 Medline

4. B. P. Kleinstiver, M. S. Prew, S. Q. Tsai, V. V. Topkar, N. T. Nguyen, Z. Zheng, A. P. W. Gonzales, Z. Li, R. T. Peterson, J.-R. J. Yeh, M. J. Aryee, J. K. Joung, Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015). doi:10.1038/nature14592 Medline

5. F. A. Ran, L. Cong, W. X. Yan, D. A. Scott, J. S. Gootenberg, A. J. Kriz, B. Zetsche, O. Shalem, X. Wu, K. S. Makarova, E. V. Koonin, P. A. Sharp, F. Zhang, *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–191 (2015). doi:10.1038/nature14299 Medline

6. B. Zetsche, J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin, F. Zhang, Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015). doi:10.1016/j.cell.2015.09.038 Medline

7. C. Anders, O. Niewoehner, A. Duerst, M. Jinek, Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014). doi:10.1038/nature13579 Medline

8. C. Anders, K. Bargsten, M. Jinek, Structural plasticity of PAM recognition by engineered variants of the RNA-guided endonuclease Cas9. *Mol. Cell* **61**, 895–902 (2016). doi:10.1016/j.molcel.2016.02.020 Medline

9. S. Hirano, H. Nishimasu, R. Ishitani, O. Nureki, Structural basis for the altered PAM specificities of engineered CRISPR-Cas9. *Mol. Cell* **61**, 886–894 (2016). doi:10.1016/j.molcel.2016.02.018 Medline

10. L. Gao, D. B. T. Cox, W. X. Yan, J. C. Manteiga, M. W. Schneider, T. Yamano, H. Nishimasu, O. Nureki, N. Crosetto, F. Zhang, Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35**, 789–792 (2017). doi:10.1038/nbt.3900 Medline

11. J. H. Hu, S. M. Miller, M. H. Geurts, W. Tang, L. Chen, N. Sun, C. M. Zeina, X. Gao, H. A. Rees, Z. Lin, D. R. Liu, Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018). doi:10.1038/nature26155 Medline

12. F. Jiang, D. W. Taylor, J. S. Chen, J. E. Kornfeld, K. Zhou, A. J. Thompson, E. Nogales, J. A. Doudna, Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351**, 867–871 (2016). doi:10.1126/science.aad8282 Medline

13. S. Q. Tsai, Z. Zheng, N. T. Nguyen, M. Liebers, V. V. Topkar, V. Thapar, N. Wyvekens, C. Khayter, A. J. Iafrate, L. P. Le, M. J. Aryee, J. K. Joung, GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015). doi:10.1038/nbt.3117 Medline

14. I. M. Slaymaker, L. Gao, B. Zetsche, D. A. Scott, W. X. Yan, F. Zhang, Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016). doi:10.1126/science.aad5227 Medline

15. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016). doi:10.1038/nature17946 Medline

16. K. Nishida, T. Arazoe, N. Yachie, S. Banno, M. Kakimoto, M. Tabata, M. Mochizuki, A. Miyabe, M. Araki, K. Y. Hara, Z. Shimatani, A. Kondo, Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729 (2016). doi:10.1126/science.aaf8729 Medline

17. H. Nishimasu, L. Cong, W. X. Yan, F. A. Ran, B. Zetsche, Y. Li, A. Kurabayashi, R. Ishitani, F. Zhang, O. Nureki, Crystal Structure of *Staphylococcus aureus* Cas9. *Cell* **162**, 1113–1126 (2015). doi:10.1016/j.cell.2015.08.007 Medline

18. T. Yamano, B. Zetsche, R. Ishitani, F. Zhang, H. Nishimasu, O. Nureki, Structural basis for the canonical and non-canonical PAM recognition by CRISPR-Cpf1. *Mol. Cell* **67**, 633–645.e3 (2017). doi:10.1016/j.molcel.2017.06.035 Medline

19. W. Kabsch, XDS. *Acta Crystallogr. D* **66**, 125–132 (2010). doi:10.1107/S0907444909047337 Medline

20. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013). doi:10.1107/S0907444913000061 Medline

21. A. Vagin, A. Teplyakov, Molecular replacement with MOLREP. *Acta Crystallogr. D* **66**, 22–25 (2010). doi:10.1107/S0907444909042589 Medline

22. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004). doi:10.1107/S0907444904019158 Medline

23. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R.

Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010). doi:10.1107/S0907444909052925 Medline

24. P. D. Hsu, D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Cradick, L. A. Marraffini, G. Bao, F. Zhang, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013). doi:10.1038/nbt.2647 Medline