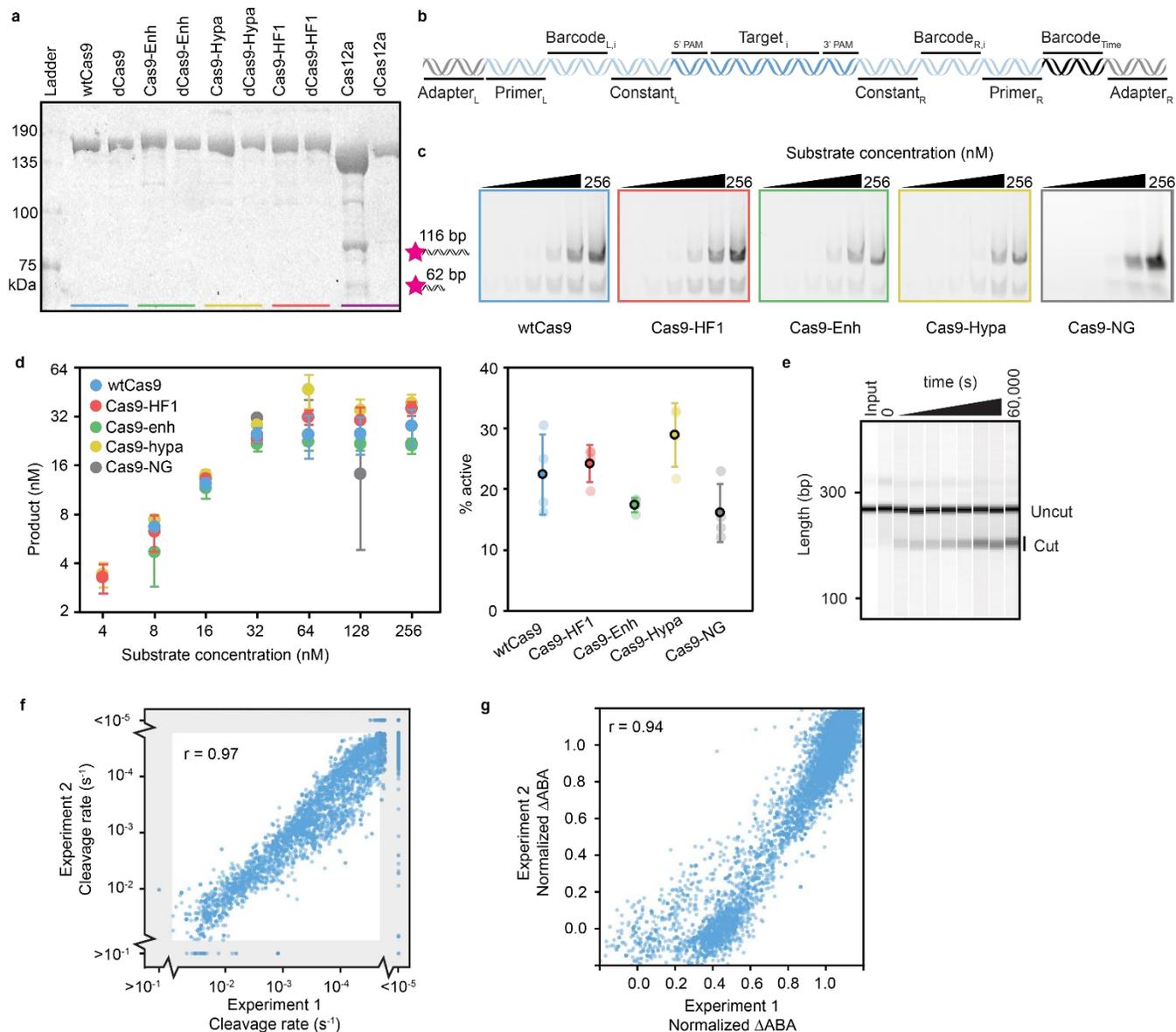

Supplementary information

Massively parallel kinetic profiling of natural and engineered CRISPR nucleases

In the format provided by the authors and unedited

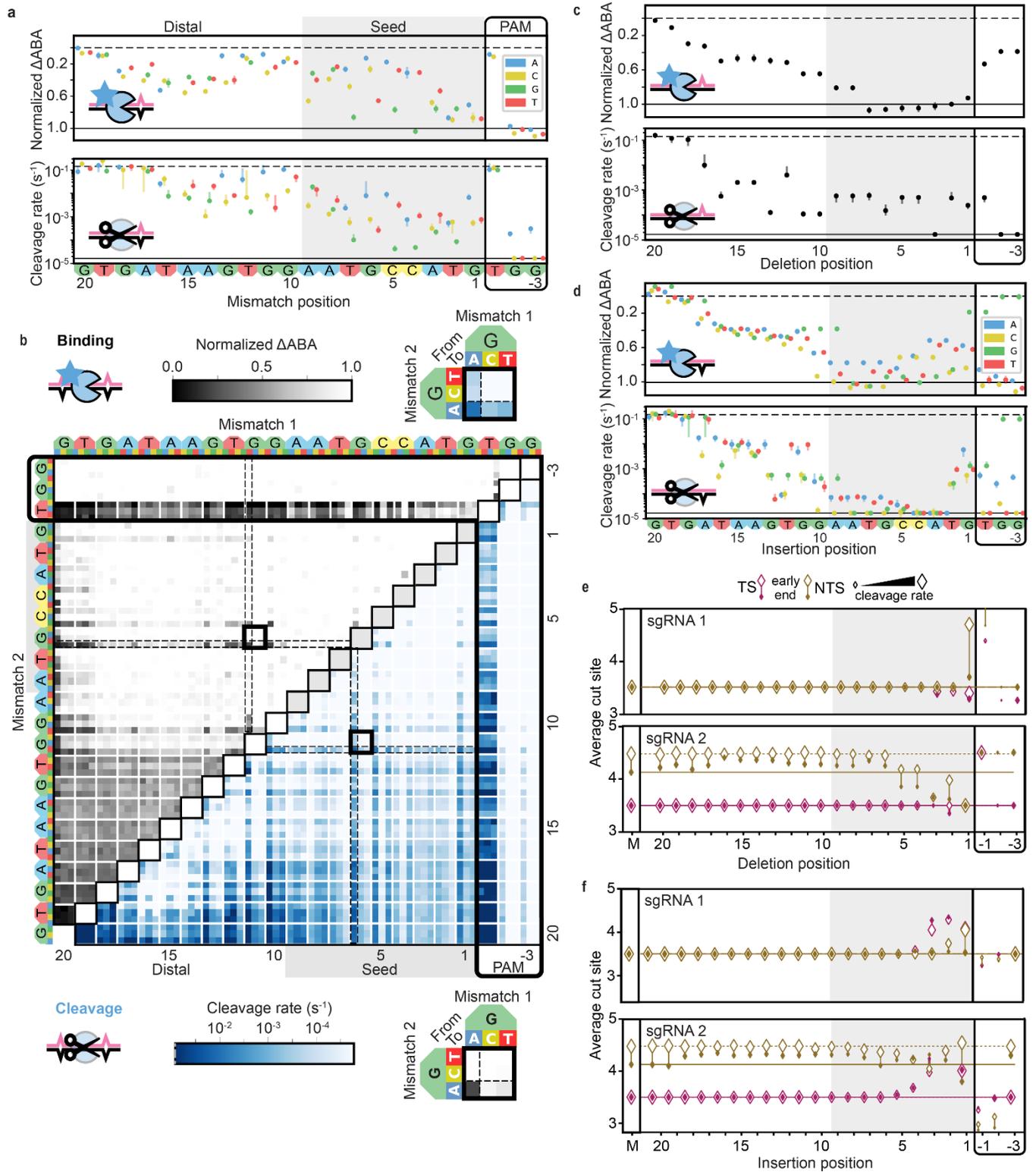
SUPPLEMENTARY INFORMATION

Supplementary Figures



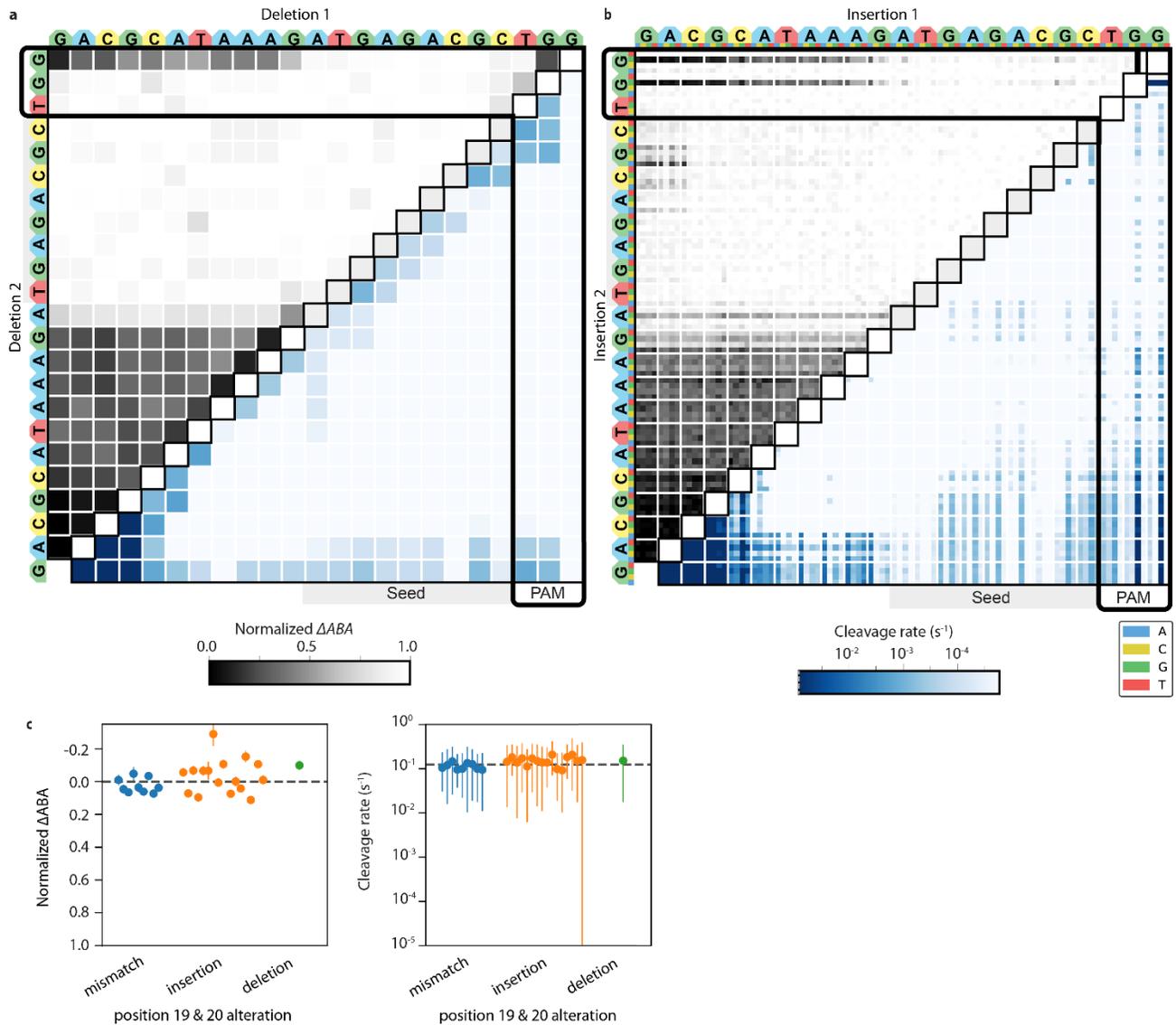
Supplementary Figure 1. Biochemical characterization of CRISPR-Cas nucleases. (a) Coomassie-stained 10% SDS-PAGE gel of purified nucleases and their catalytically-inactive variants. Gel was not repeated. (b) Diagram of a target DNA library member. The regions in blue are purchased as pooled oligonucleotides. Primer regions are for PCR amplification. Barcodes identify digested DNAs. Constant regions assure uniform target sequence context and oligo length. Timestamp barcodes and sequencing adapters are ligated to all DNAs after quenching the cleavage reaction. (c) Representative active site titration cleavage gels (10% native PAGE, TBE) and (d) quantification of three replicates for the indicated

Cas9 variants. 128nM Cas9 RNP was incubated with 2-256nM of ATTO647N-labeled matched DNA (pink star) for 30 minutes. **(d, right)** The active nuclease concentration (mean \pm SD; of at least three replicates) was determined from the concentration of product formed at 64-256 nM input DNA concentrations. **(e)** Representative time course chromatographs from a wtCas9 nuclease reaction (sgRNA2), resolved by capillary electrophoresis. Each sample was run separately; two independent replicates for each. **(f)** Cleavage rate reproducibility for two wtCas12-crRNA3 experiments. The gray area contains targets with rates beyond the experimental dynamic range. r = Pearson correlation coefficient excluding gray area. **(g)** Normalized apparent binding affinity correlation for two dCas9-sgRNA1 experiments. r = Pearson correlation coefficient.

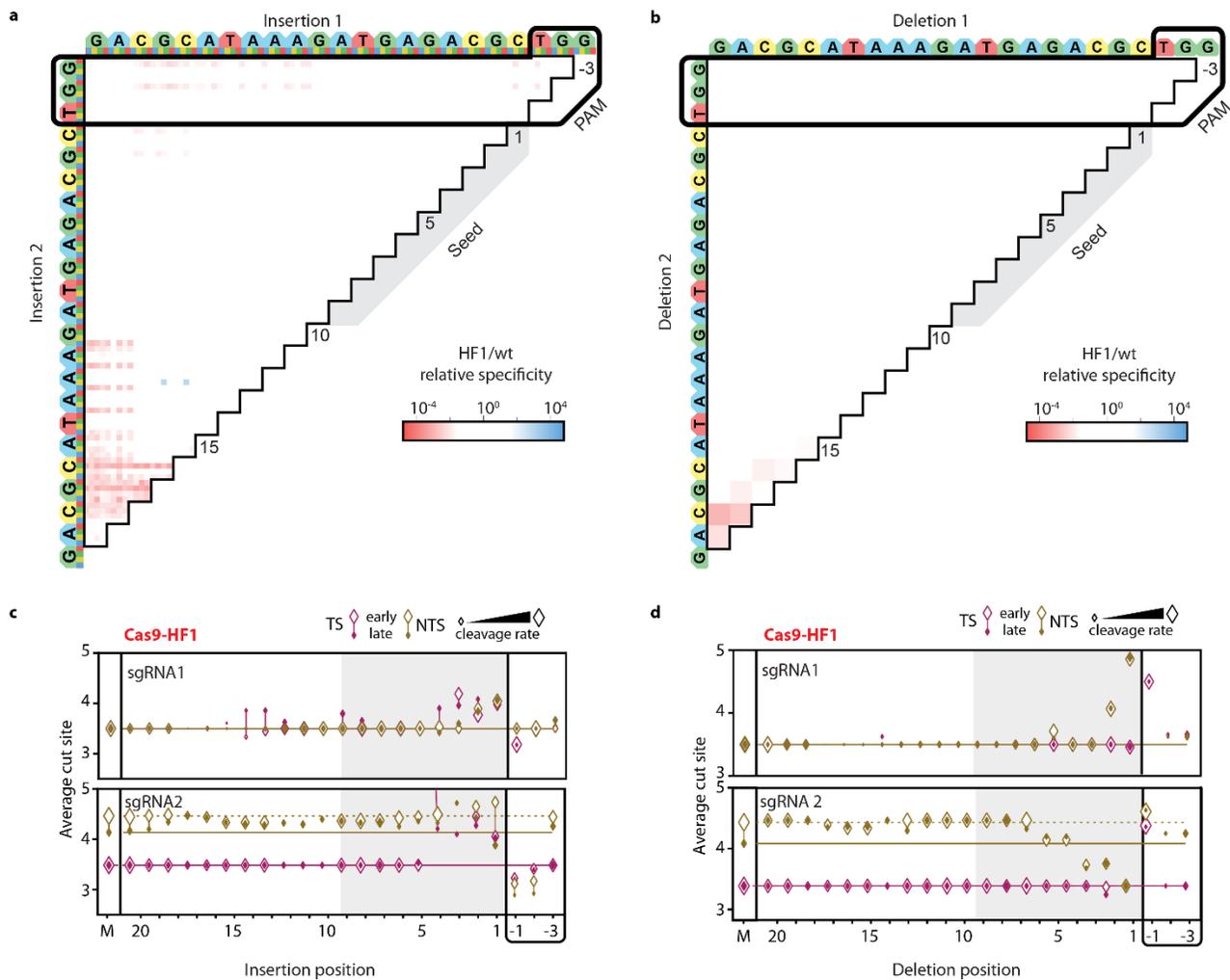


Supplementary Figure 2. Comprehensive analysis of off-target wtCas9 DNA binding and cleavage with sgRNA2. (a) dCas9 Δ ABAs (upper, 0 indicates matched target, median \pm SD from bootstrap analysis

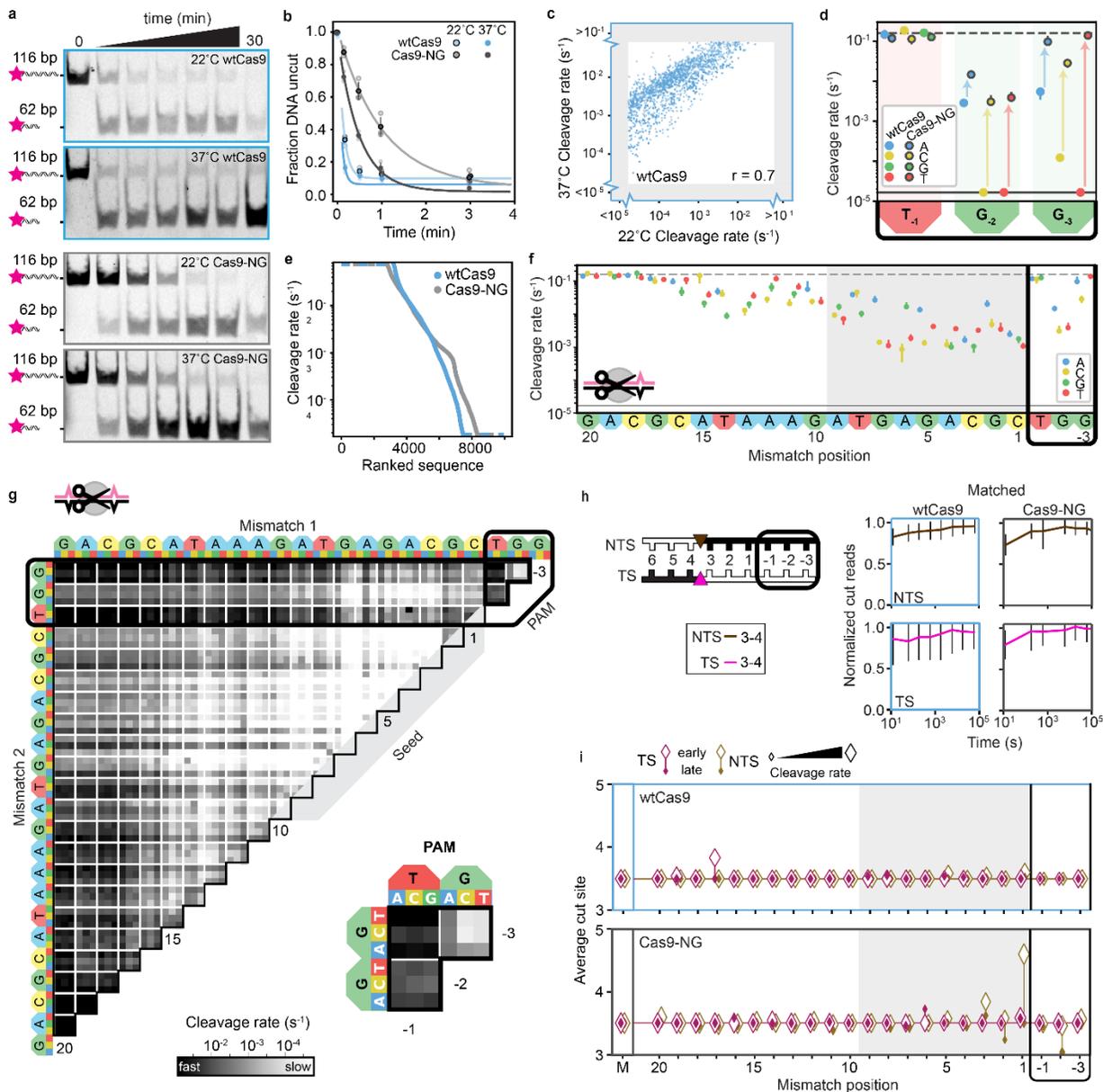
of $n \geq 5$ DNA clusters for each target) and cleavage rates (lower, error bars: SD from 50 bootstrap analysis measurements) for all DNAs with one sgRNA2-relative mismatch. Dashed line: normalized matched target Δ ABA or cleavage rate. Solid line: scrambled DNA Δ ABA (negative control) or limit of detection for the slowest-cleaving targets. **(b)** Δ ABAs (upper, grays) and cleavage rates (lower, blues) for targets containing two sgRNA2-relative mismatches. Black boxes expanded in callouts. **(c)** dCas9 Δ ABAs (upper, median \pm SD from bootstrap analysis of $n \geq 5$ DNA clusters for each target) and Cas9 cleavage rates (lower, error bars: SD from 50 bootstrap analysis measurements) for targets containing one sgRNA2-relative deletion or **(d)** insertion. **(e)** Average cut site positions for each strand (TS, NTS) from targets containing one deletion or **(f)** insertion compared to sgRNA 1 (upper) or 2 (lower). Range: first timepoint with $>33\%$ cut reads (early, open diamonds) to the final time point (end, filled diamonds). Dashed and solid horizontal lines: average cut site positions for matched DNA at early and late time points.



Supplementary Figure 3. Comprehensive analysis of off-target wtCas9 DNA binding and cleavage of targets with insertions or deletions. (a) Normalized ΔABA s (upper) and cleavage rates (lower) for targets containing two sgRNA1-related deletions or (b) insertions. (c) Normalized ΔABA s (left, median \pm SD from bootstrap analysis of $n \geq 5$ DNA clusters for each target) and cleavage rates (right, SD from 50 bootstrap analysis measurements) for DNAs containing sgRNA-related alterations at positions 19 and 20. Dashed line: matched target normalized ΔABA (left) or cleavage rate (right).

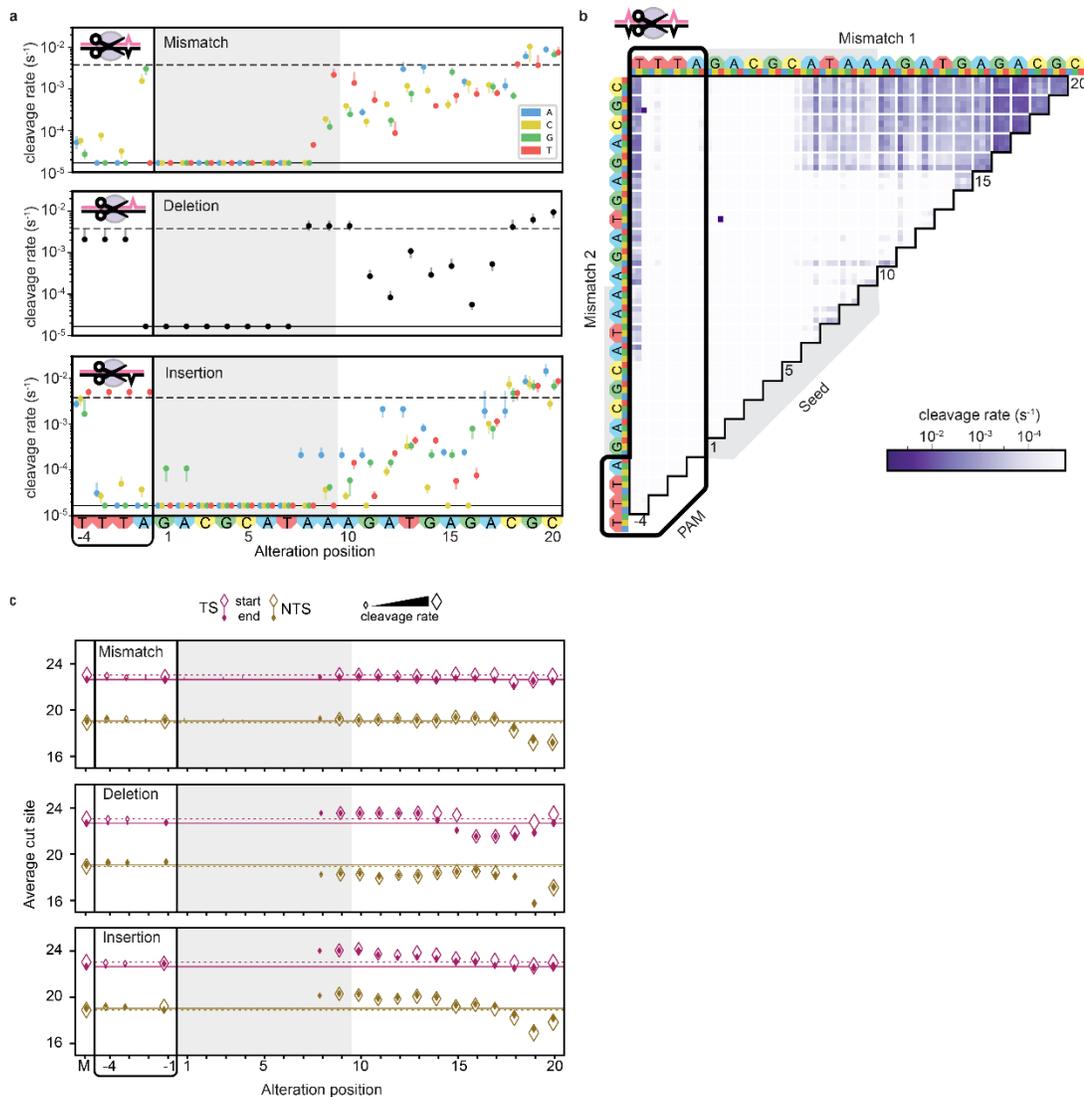


Supplementary Figure 4. Comparison of Cas9-HF1 and wtCas9 nuclease activities. (a) The ratio of Cas9-HF1 to wtCas9 cleavage specificities for all targets containing two sgRNA1-related insertions or (b) deletions. Red: slower cleavage by Cas9-HF1; blue: slower cleavage by wtCas9. (c) Average cut site positions generated by Cas9-HF1 for each strand (TS, NTS) for targets containing one insertion or (d) deletion site compared to sgRNA 1 (upper) or sgRNA 2 (lower). Range: earliest timepoint with >33% cut reads (open diamonds) to final time point (filled diamonds). Dashed and solid horizontal lines: Mean cut site positions for 146 matched DNAs (M) at early and late time points.

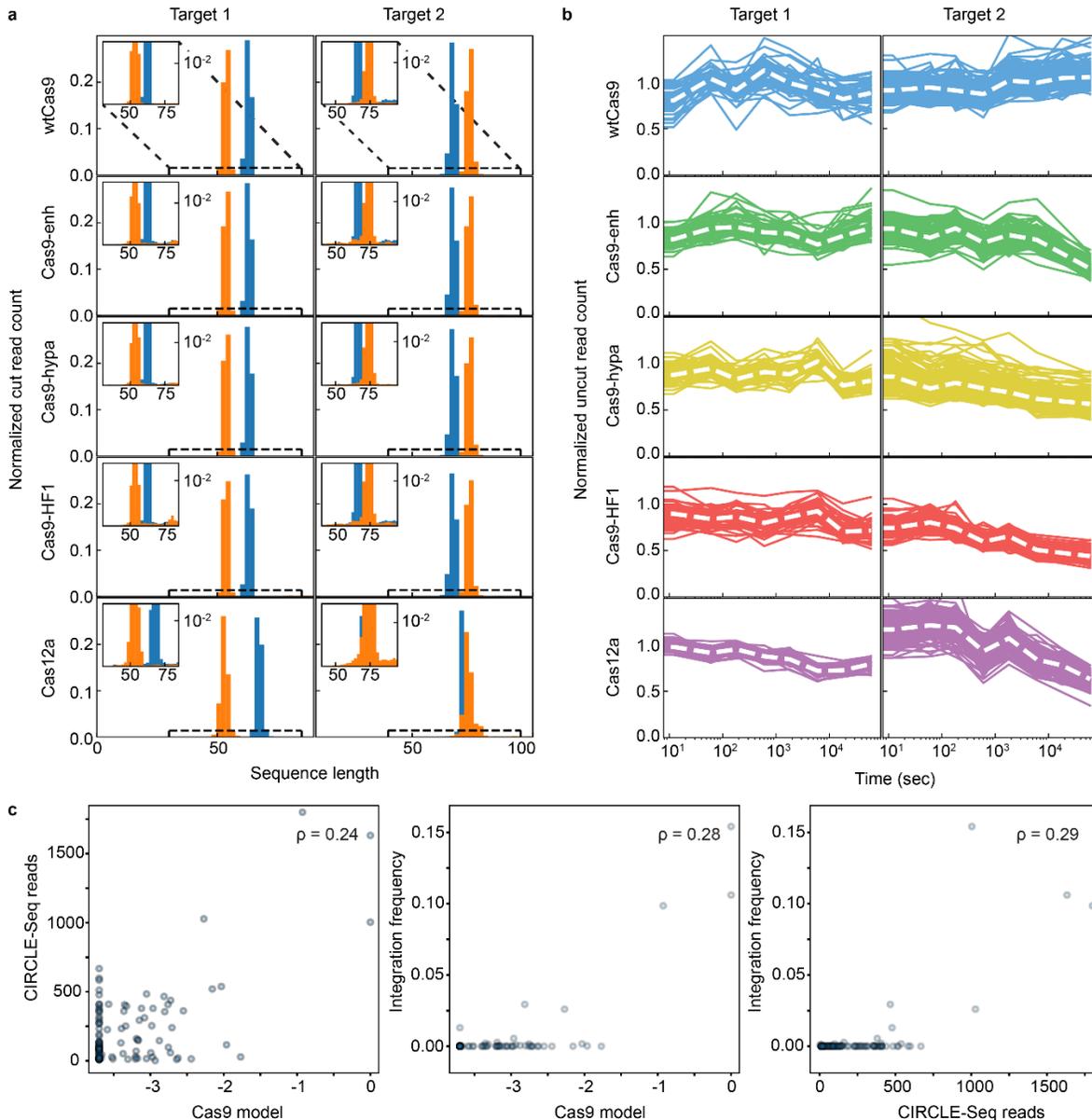


Supplementary Figure 5: Comparison of wtCas9 and Cas9-NG cleavage rates. (a) Representative gels of the cleavage of a matched target (10% native PAGE, TBE). (b) Quantification of three gels for wtCas9 and Cas9-NG cleavage at 22°C and 37°C. 100 nM of the indicated Cas9 RNP was incubated with 10 nM of ATTO647N-labeled matched DNA (pink star) for up to 30 minutes. Cleavage rates were determined by fitting the data to a single exponential: wtCas9 at 22°C: $0.14 \pm 0.001 \text{ s}^{-1}$; wtCas9 at 37°C: $0.22 \pm 0.007 \text{ s}^{-1}$; Cas9-NG at 22°C: $0.018 \pm 0.001 \text{ s}^{-1}$, and Cas9-NG at 37°C: $0.037 \pm 0.001 \text{ s}^{-1}$ (mean \pm SD of three replicates). (c) wtCas9 cleavage rates correlate between 37°C and 22°C experiments. The gray area contains targets with rates beyond the experimental dynamic range. r = Pearson correlation coefficient

excluding gray area. **(d)** Comparison of the PAM-dependent cleavage rates of Cas9-NG and wtCas9 at 37°C. Dashed line: cleavage rate of the matched target. Solid line: limit of detection for the slowest-cleaving targets. Error bars: SD from 50 bootstrap analysis measurements. **(e)** Rank-ordered cleavage rates of all wtCas9 or Cas9-NG targets at 37°C. Cas9-NG is broadly similar to wtCas9, except for non-NGG PAMs. Dashed lines: cleavage rates beyond the limit of detection. **(f)** Cas9-NG cleavage rates for targets with one sgRNA1-relative mismatch. Dashed line: cleavage rate of the matched target; solid line: limit of detection for the slowest-cleaving targets. Error bars: SD from 50 bootstrap analysis measurements. **(g)** Cleavage rates for targets containing two sgRNA1-relative mismatches. Black box expanded in callout. **(h)** Normalized median reads for the target (TS) and nontarget (NTS) strands of matched DNA cleaved by wtCas9 (left) or Cas9-NG (right) at 37°C. Error bars: maximum SD for cut products from cleavage of 146 matched DNA controls. **(i)** Average cut site positions generated by wtCas9 (upper) or Cas9-NG (lower) at 37°C for each strand (TS, NTS) for targets with sgRNA1-relative mismatches. Range: earliest timepoint with >33% cut reads (open diamonds) to final time point (filled diamonds). Dashed and solid horizontal lines: Mean cut site positions for 146 matched DNAs (M) at early and late time points.

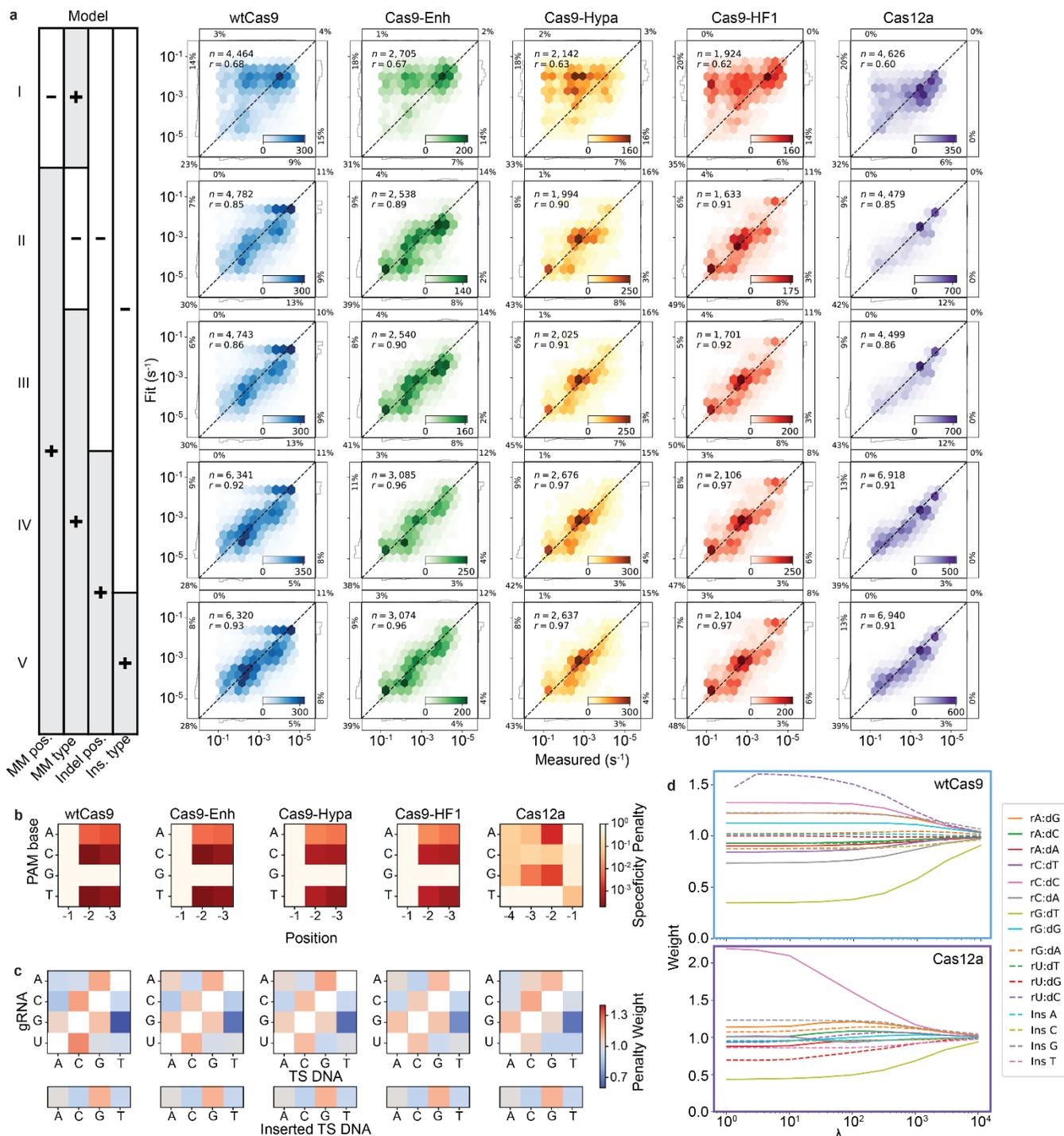


Supplementary Figure 6. Analysis of off-target Cas12a cleavage with crRNA 4. (a) Cleavage rates for all targets with one crRNA4-related mismatch (upper), deletion (middle) or insertion (lower). Dashed line: matched target cleavage rate. Solid line: limit of detection for the slowest-cleaving targets. Error bars: SD from 50 bootstrap analysis measurements. (b) Cleavage rates for targets with two crRNA4-related mismatches. (c) Average cut site positions generated by Cas12a for each strand (TS, NTS) for targets containing crRNA4-related mismatches (upper), deletions (middle) or insertions (lower). Range: earliest timepoint with >33% cut reads (open diamonds) to final time point (filled diamonds). Dashed and solid horizontal lines: Mean cut site positions for 146 matched DNAs (M) at early and late time points.



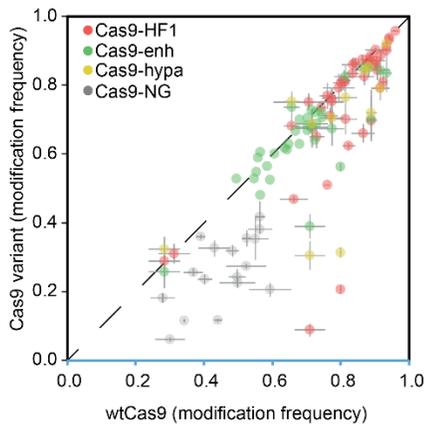
Supplementary Figure 7. Cas12a exhibits limited *trans* cleavage of NucleaSeq libraries. We looked for two signatures of *trans* cleavage activity: **(a)** The distribution of normalized read counts for DNAs with only a left (blue) or right (orange) barcode show that very few DNAs were cut outside of the expected enzyme cleavage sites. Inset: a zoomed-in view shows few indiscriminately cut products outside of the main peaks, which correspond to the canonical cleavage sites. A comparable amount of short DNA for Cas9 and Cas12a suggests that these fragments arise during library preparation and NGS. Cas12a-catalyzed *trans* cleaved DNA is a relatively minor component of our data (within the noise). Total histogram areas are normalized to one. **(b)** The time-dependent read counts of ~150 uncut control

sequences (not complementary to the guide RNAs; overlapping lines) are not depleted more quickly for Cas12a than the other enzymes. These sequences are normalized only for total read counts at each time point, then the zero timepoint value is set to one. These values are proportional to the fraction of the library at each time point, not the absolute read count. Robust *trans* cleavage by Cas12a should deplete these values at a faster rate than other enzymes. Cas12a behaves similarly to the engineered Cas9 variants in all cases. No individual traces are observed to decrease faster than the group, as expected if trans-cleavage showed sequence bias for some control sequences. White dashed line: median.



Supplementary Figure 8. Comparison of biophysical models for Cas nuclease specificity. (a) Correlation between measured and fit cleavage rates for each protein for the indicated simplified model. Histograms: rates beyond the adjacent detection limit, representing the percent of the total fits or measurements as indicated. Color bars indicate sequence count. r : Pearson correlation coefficient. **(b)** PAM position weight matrices computed from the cleavage rates for each protein. **(c)** Penalty weight

values of mismatch (top) and insertion types (bottom) for each protein. **(d)** Transition and insertion weights as a function of the regularization parameter, λ (see Methods).



Supplementary Figure 9. wtCas9 on-target cleavage is more efficient than engineered Cas9 variants in cells. On-target cleavage activity from prior studies comparing wtCas9 with the indicated Cas9 variant (Chen et al., *Nature*, 550(7676):407–410, 2017; Jinek et al., *Science*, 337(6096):816–821, 2012; Kleinstiver et al., *Nature*, 529(7587):490–495, 2016; Nishimasu et al., *Science*, 361(6408):1259–1262, 2018; Slaymaker et al., *Science*, 351(6268):84–88, 2016). Edit efficiency was measured by eGFP disruption or indel formation as scored by deep sequencing of the target site. Error bars as replotted from each publication where available. Data with error bars represent either mean \pm SD (Kleinstiver et al., *Nature*, 529(7587):490–495, 2016; Nishimasu et al., *Science*, 361(6408):1259–1262, 2018) or SEM (Chen et al., *Nature*, 550(7676):407–410, 2017) for $n \geq 3$ replicates.

Supplementary Tables

Supplementary Table 1. Guide RNAs used in this work

Name	RNA Sequence
sgRNA 1 (for all <i>SpCas9</i> variants)	GACGCAUAAAGAUGAGACGCGUUUUAGAGCUAGA AAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCA ACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
sgRNA 2 (for all <i>SpCas9</i> variants)	GUGAUAAGUGGAAUGCCAUGGUUUUAGAGCUAG AAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
crRNA 3 (for <i>AsCas12a</i>)	GUCAAAAGACCUUUUUAAUUUCUACUCUUGUAGA UGUGAUAAGUGGAAUGCCAUGUGGA
crRNA 4 (for <i>AsCas12a</i>)	GUCAAAAGACCUUUUUAAUUUCUACUCUUGUAGA UGACGCAUAAAGAUGAGACGCUGGA

Supplementary Table 2. DNA nucleotides used in this work

Name	Protein	Descriptor	Sequence	Reference
pr238	dCas9	H840A	TCGTTTAAGTGATTATGATGTCGATGC CATTGTTCCACAAAGTTTCCTTAAA	(Jinek et al., Science, 337(6096):816– 821, 2012)
pr239			TTTAAGGAACTTTGTGGAACAATGG CATCGACATCATAATCACTTAAACGA	
pr236		D10A	GAAATACTCAATAGGCTTAGCTATCG GCACAAATAGCGTCG	
pr237			CGACGCTATTTGTGCCGATAGCTAAG CCTATTGAGTATTC	
pr271	Cas9- enh	K848A	CGATCACATTGTTCCACAAAGTTTCCT TGCAGACGATTCAATAGACAATAAG	(Slaymaker et al., Science, 351(6268):84– 88, 2016)
pr272			CTTATTGTCTATTGAATCGTCTGCAAG GAACTTTGTGGAACAATGTGATCG	
pr273		K1003A	CGTTGGAAGCTTTGATTAAGAAAT ATCCAGCACTTGAATCGGAGTTTGT	
pr274			ACAACTCCGATTCAAGTGCTGGATA TTTCTTAATCAAAGCAGTTCCAACG	
pr275		K1060A	ACTTGCAAATGGAGAGATTTCGCAAAG CCCCTCTAATCGAA	

pr276			TTCGATTAGAGGGGCTTTGCGAATCTC TCCATTTGCAAGT	
pr263	Cas9- HF1	N497A	CAGCTCAATCATTATTGAACGCATG ACAGCCTTTGATAAAAATCTTCCAAA TGAAAAAG	(Kleinstiver et al., Nature, 529(7587):490– 495, 2016)
pr264			CTTTTTCATTTGGAAGATTTTTATCAA AGGCTGTCATGCGTTCAATAAATGAT TGAGCTG	
pr269		Q926A	CCAATTGGTTGAAACTCGCGCAATCA CTAAGCATGTGGCA	
pr270			TGCCACATGCTTAGTGATTGCGCGAG TTTCAACCAATTGG	
pr265		R661A	GCCGTTATACTGGTTGGGGAGCTTTGT CTCGAAAATTGATTA	
pr266			TAATCAATTTTCGAGACAAAGCTCCC CAACCAGTATAACGGC	
pr267	Cas9- HF1	Q695A	GTTTTGCCAATCGCAATTTTATGGCGC TGATCCATGATGATAGTTTG	(Kleinstiver et al., Nature, 529(7587):490– 495, 2016)
pr268			CAAACATCATCATGGATCAGCGCCA TAAAATTGCGATTGGCAAAAC	
pr399	Cas9- hypo	H698A, Q695A	GCGCTGATCGCTGATGATAGTTTGAC ATTTAAAGAAGACATTCAAAAAGCAC AA	(Chen et al., Nature, 550(7676):407– 410, 2017)
pr400		M694A, N692A	GCCAAAGCTGCGATTGGCAAAACCAT CTGATTTCAAAAA	
pr698	Cas9- NG	R1335V,	GATGCCACTCTTATCCATCA	(Nishimasu et al., Science, 361(6408):1259– 1262, 2018)
pr699		L1111R,	GCCTGTCTGTACTTCTGTTT	
pr702		D1135V,	TGATGGATAAGAGTGGCATC	
pr703		G1218R,	AAACAGAAGTACAGACAGGC	
pr697		E1219F, A1322R, T1337R	AAACAGAAGTACAGACAGGCGGATTC TCCAAGGAGTCAATTCGCCCAAAAAG AAATTCGGACAAGCTTATTGCTCGTA AAAAAGACTGGGATCCAAAAAATAT	

			GGTGGTTTTGTTAGTCCAACGGTAGCT TATTCAGTCCTAGTGGTTGCTAAGGTG GAAAAAGGGAAATCGAAGAAGTTAA AATCCGTTAAAGAGTTACTAGGGATC ACAATTATGGAAAGAAGTTCCTTTGA AAAAAATCCGATTGACTTTTTTAGAAG CTAAAGGATATAAGGAAGTTAAAAAA GACTTAATCATTAAACTACCTAAATAT AGTCTTTTTGAGTTAGAAAACGGTCGT AAACGGATGCTGGCTAGTGCCCGTTT CTTACAAAAGGAAATGAGCTGGCTC TGCCAAGCAAATATGTGAATTTTTTAT ATTTAGCTAGTCATTATGAAAAGTTG AAGGGTAGTCCAGAAGATAACGAACA AAAACAATTGTTTGTGGAGCAGCATA AGCATTATTTAGATGAGATTATTGAG CAAATCAGTGAATTTTCTAAGCGTGTT ATTTTAGCAGATGCCAATTTAGATAA AGTTCTTAGTGCATATAACAAACATA GAGACAAACCAATACGTGAACAAGCA GAAAATATTATTCATTTATTTACGTTG ACGAATCTTGGAGCTCCCCGTGCTTTT AAATATTTTGATACAACAATTGATCGT AAAGTGTATCGCTCTACAAAAGAAGT TTTAGATGCCACTCTTATCCATCA	
pr309		Matched Target 1	ACGCTCTTCCGATCTTTTAGACGCATA AAGATGAGACGCTGGAGATCGGAAG AGCAC	This work
pr310		GTGCTCTTCCGATCTCCAGCGTTCAT CTTTATGCGTCTAAAAGATCGGAAGA GCGT		
pr477	Library 1	ATAACTAATTGAGCTGAACGCAC - Barcode_{L_i}-CAGAT-Target- TGATCGTCTTAGAGACAAGCACAGTC		

			TCT-Barcode _{R,i} - AGCACAACACTACACGACTATTTCAG Oligonucleotide pool. See Supplementary file 2 and Figure S1B.
pr475	Primers for Library 1		ATAACTAATTGAGCTGAACGCAC
pr476			CTGAATAGTCGTGTAGTTGTGCT
pr307	Matched target 2		ACGCTCTTCCGATCTTTTAGTGATAAG TGAATGCCATGTGGAGATCGGAAGA GCAC
pr308			GTGCTCTTCCGATCTCCACATGGCATT CCACTTATCACTAAAAGATCGGAAGA GCGT
pr371	Library 2		AACCGCCGAATAACAGAGT- Barcode _{L,i} - CAGAT-Target-TGATCAGCATGTC- Barcode _{R,i} -AGTGTGCGAGGCGTTCTT Oligonucleotide pool. See Supplementary file 2 and Figure S1B.
pr364	Primers for library 2		AACCGCCGAATAACAGAGT
pr365			AAGAACGCCTCGCACACT
pr460	ATTO647N Primer for library 2		/5atto647n/AACCGCCGAATAACAGAGT

Supplementary Table 3. Plasmids used in this work.

Name	Protein	Construct	Mutations	References
pIF324	<i>SpCas9</i>	6xHis-MBP-3xFlag- <i>SpCas9</i>		(Jinek et al., Science, 337(6096):816–821, 2012)
pIF335		6xHis-MBP-3xFlag- <i>dSpCas9</i>	D10A, H840A	(Jinek et al., Science, 337(6096):816–821, 2012)
pIF325		6xHis-MBP-3xFlag-enhanced <i>SpCas9</i> -1.1	K848A, K1003A, R1060A	(Slaymaker et al., Science, 351(6268):84–88, 2016)
pIF326		6xHis-MBP-3xFlag-enhanced <i>dSpCas9</i> -1.1	D10A, H840A, K848A, K1003A, R1060A	(Slaymaker et al., Science, 351(6268):84–88, 2016)

pIF329		6xHis-MBP-3xFlag- <i>SpCas9</i> -HF1	R661A, Q695A, Q926A	(Kleinstiver et al., Nature, 529(7587):490–495, 2016)
pIF330		6xHis-MBP-3xFlag- <u>d</u> <i>SpCas9</i> -HF1	D10A, R661A, Q695A, H840A, Q926A	(Kleinstiver et al., Nature, 529(7587):490–495, 2016)
pIF350		6xHis-MBP-3xFlag-hypa <i>SpCas9</i>	N692A, M694A, Q695A, H698A	(Chen et al., Nature, 550(7676):407–410, 2017)
pIF351		6xHis-MBP-3xFlag-hypa <u>d</u> <i>SpCas9</i>	D10A, N692A, M694A, Q695A, H698A, H840A	(Chen et al., Nature, 550(7676):407–410, 2017)
pIF580		6xHis-MBP- <i>SpCas9</i> -NG	R1335V, L1111R, D1135V, G1218R, E1219F, A1322R, T1337R	(Nishimasu et al., Science, 361(6408):1259–1262, 2018)
pIF502	<i>AsCas12a</i>	6xHis-TwinStrep-SUMO- <i>AsCas12a</i> -3xFlag		(Strohkendl, Saifuddin, Rybarski, Finkelstein, & Russell, Mol. Cell, 71(5):816-824.e3, 2018)

Supplementary File Descriptions

Supplementary File 1. Complete DNA sequences as synthesized for DNA libraries used with CHAMP and NucleaSeq methods. From 5' to 3': Primer-binding region for library amplification, first barcode for identifying a cleaved DNA fragment, constant region for uniform target sequence context, 5' Cas12a PAM, target sequence, 3' Cas9 PAM, constant region for uniform target sequence context and length, second barcode for identifying a cleaved DNA fragment, Primer-binding region for library amplification (**Supplementary Fig. 1b and Online Methods**).

Supplementary File 2. Cleavage rates and normalized changes in apparent binding affinity (ΔABA) for all measured DNAs with the indicated RNPs. *Sequence*: the DNA sequence related to the PAM and guide RNA for the tested RNP. *Descriptor*: Relationship between the tested DNA sequence and the intended DNA target matching the guide RNA and PAM. *ndABA*: The change in apparent binding affinity (as measured using CHAMP) for the indicated DNA sequence compared to that of a matched DNA, normalized to the matched DNA (0) and an unmatched negative control DNA (1). *ndABA_unc*: standard deviation of the normalized change in apparent binding affinity as measured by bootstrap analysis. *cleavage_rate_log*: the log-transformed cleavage rate as measured by NucleaSeq for the indicated DNA sequence. *cleavage_rate_log_unc*: standard deviation of the log-transformed cleavage rate as measured by bootstrap analysis.

Supplementary File 3. Cleavage rates and average cleavage sites for all measured DNAs treated with the indicated RNPs. *Sequence*: the DNA sequence related to the PAM and guide RNA for the tested RNP. *Descriptor*: Relationship between the tested DNA sequence and the intended DNA target matching the guide RNA and PAM. *cleavage_rate_log*: the log-transformed cleavage rate as measured by NucleaSeq for the indicated DNA sequence. *cleavage_rate_log_unc*: standard deviation of the log-transformed cleavage rate as measured by bootstrap analysis. *L_##* or *R_##*: Average left and right side (5'→3'NTS) cleavage sites for the indicated DNA sequence with the indicated RNP at the indicated time. *None*: Less than 33% of DNAs with the indicated sequence were cleaved in the overall reaction; no average cleavage site available.