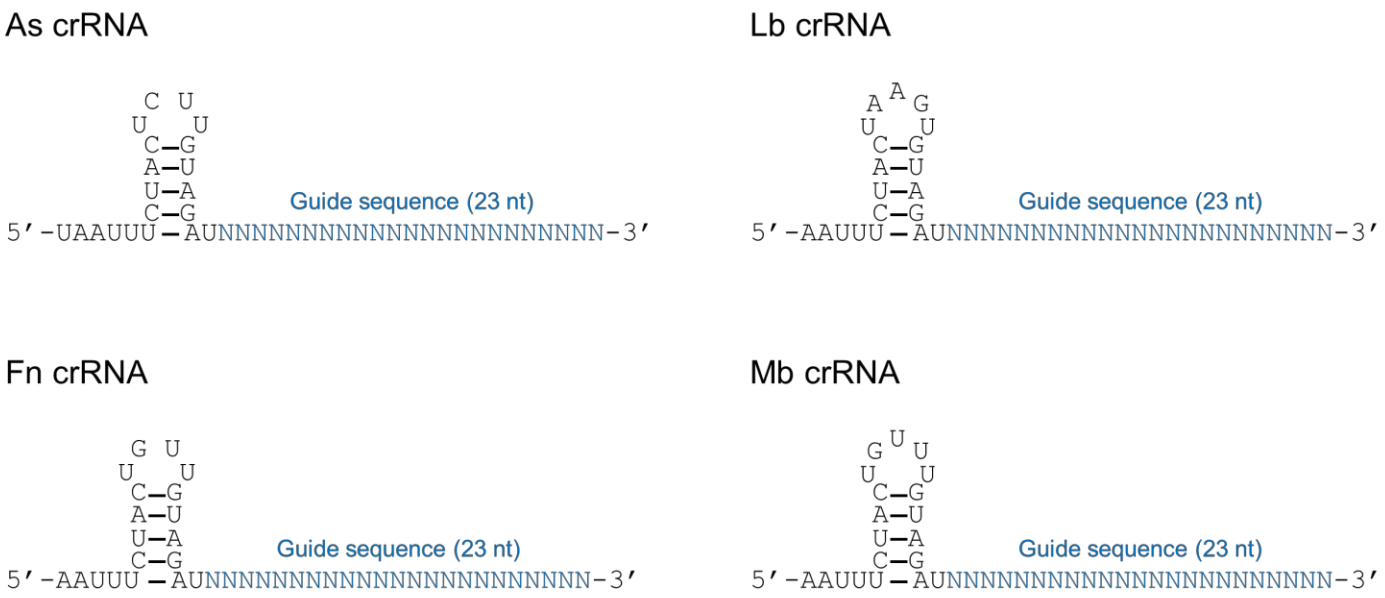
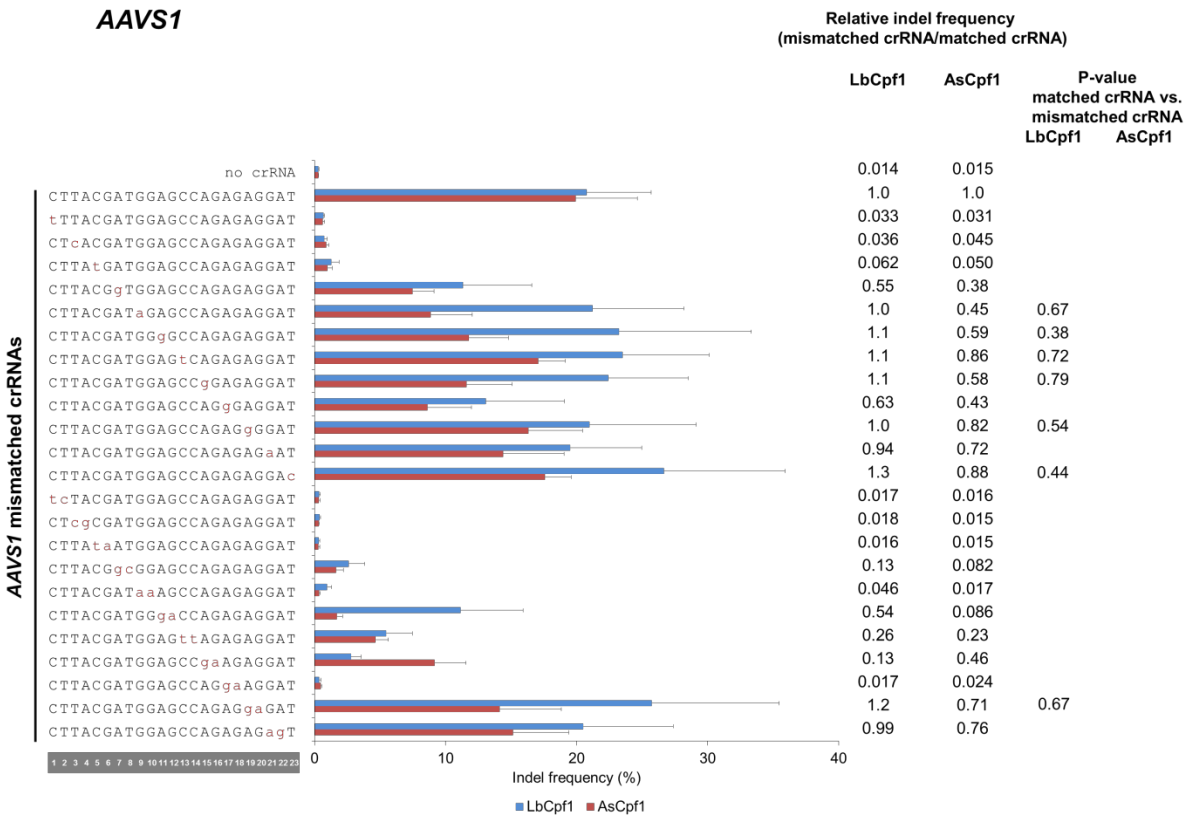


Supplementary Figure 1

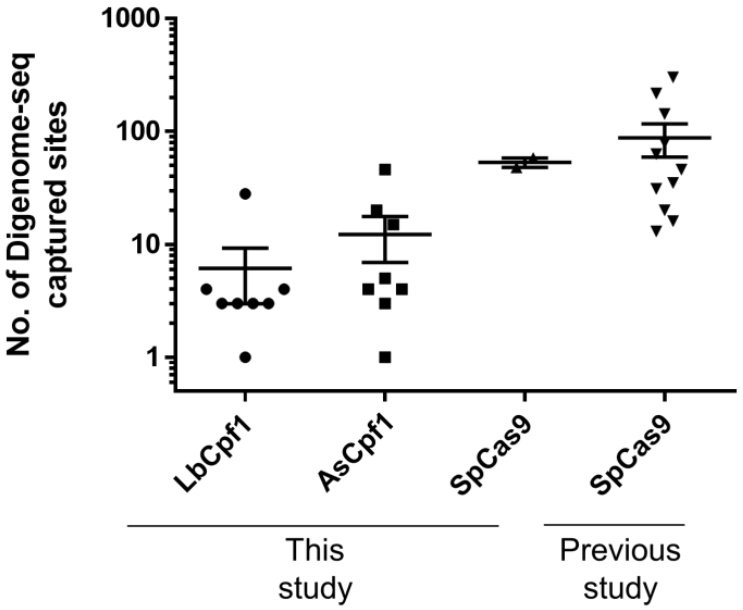
crRNA sequences.



Supplementary Figure 2

Specificity of Cpf1 examined using mismatched crRNAs at the AAVS1 sites.

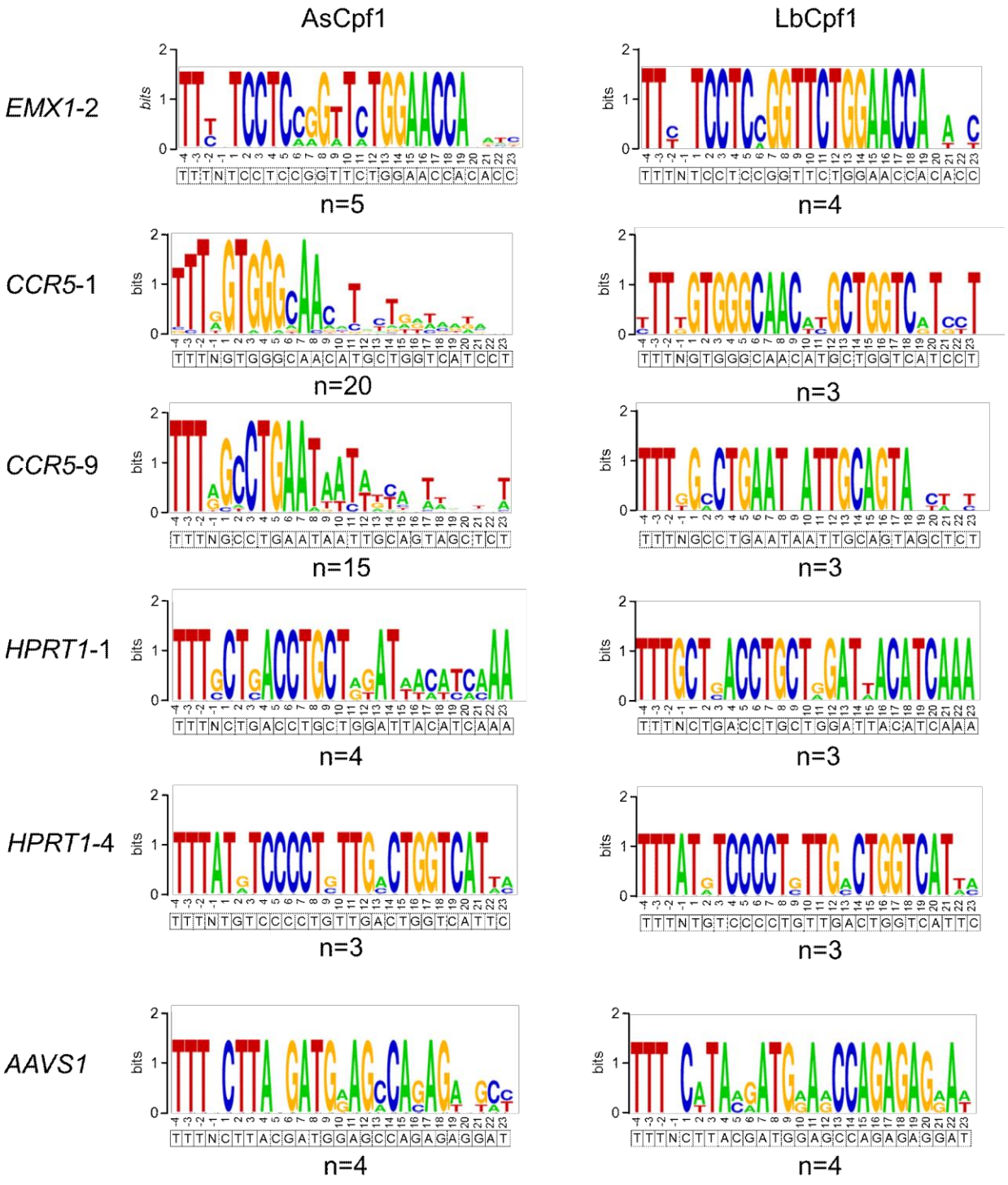
Indel frequencies obtained with mismatched crRNAs were measured by targeted deep sequencing. Error bars represent s.e.m. (n = 3). Some mismatched crRNAs appeared more active than the perfectly-matched crRNA but the differences were not statistically significant (Student's t-test, P > 0.4).



Supplementary Figure 3

The number of Digenome-seq captured sites with LbCpf1, AsCpf1, and SpCas9.

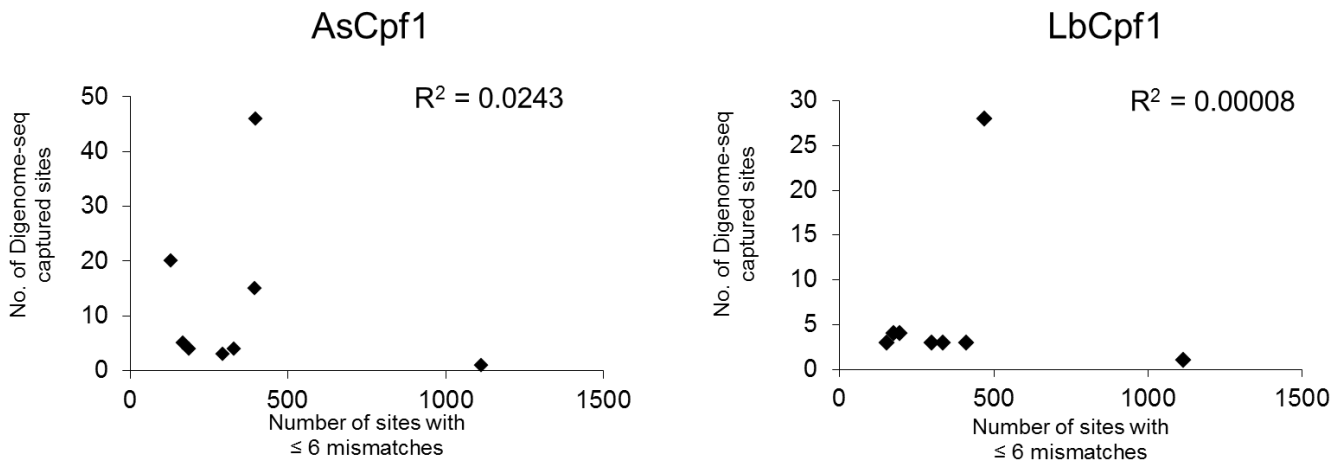
In vitro cleavage sites were determined via mononoplex Digenome-seq with LbCpf1 (n = 8), AsCpf1 (n = 8), and SpCas9 (n = 2) in this study and by multiplex Digenome-seq with SpCas9 (n = 11) in our previous study[11].



Supplementary Figure 4

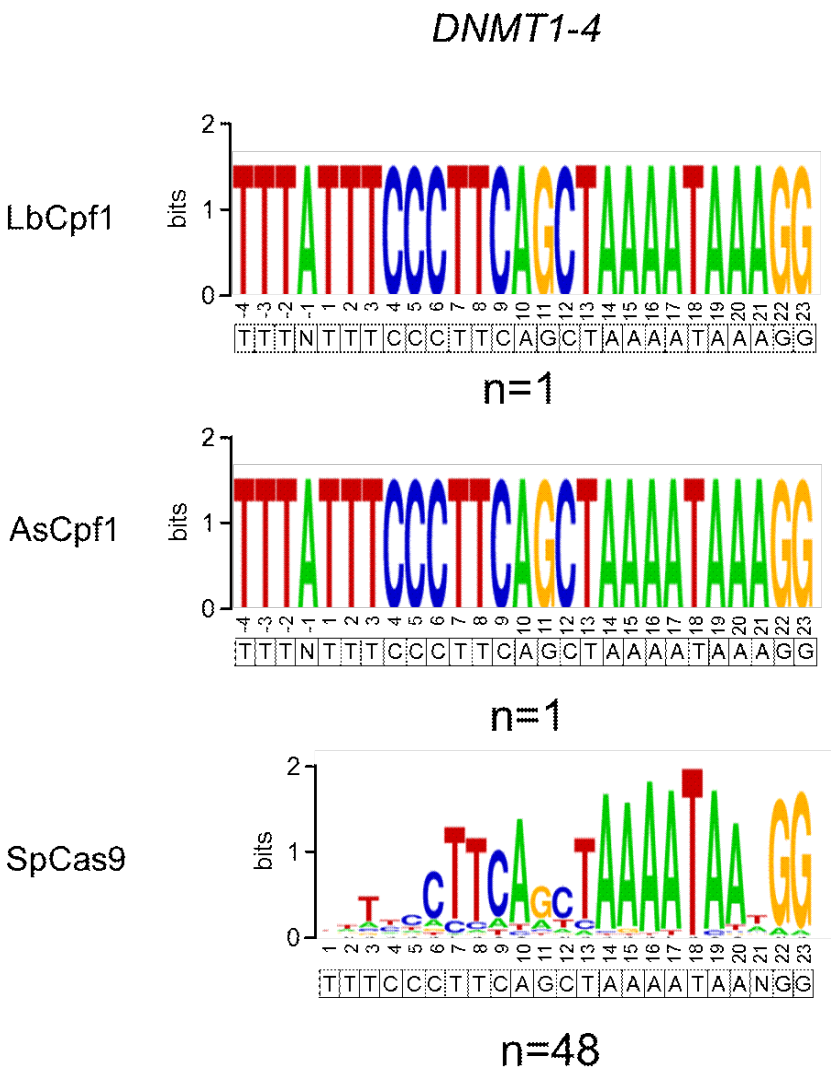
Sequence logos of Cpf1-mediated Digenome-captured sites.

Sequence logos of Digenome-captured sites obtained using AsCpf1 (Left) and LbCpf1 (Right). The number below sequence logos, n, indicates the number of Digenome-captured sites.



Supplementary Figure 5

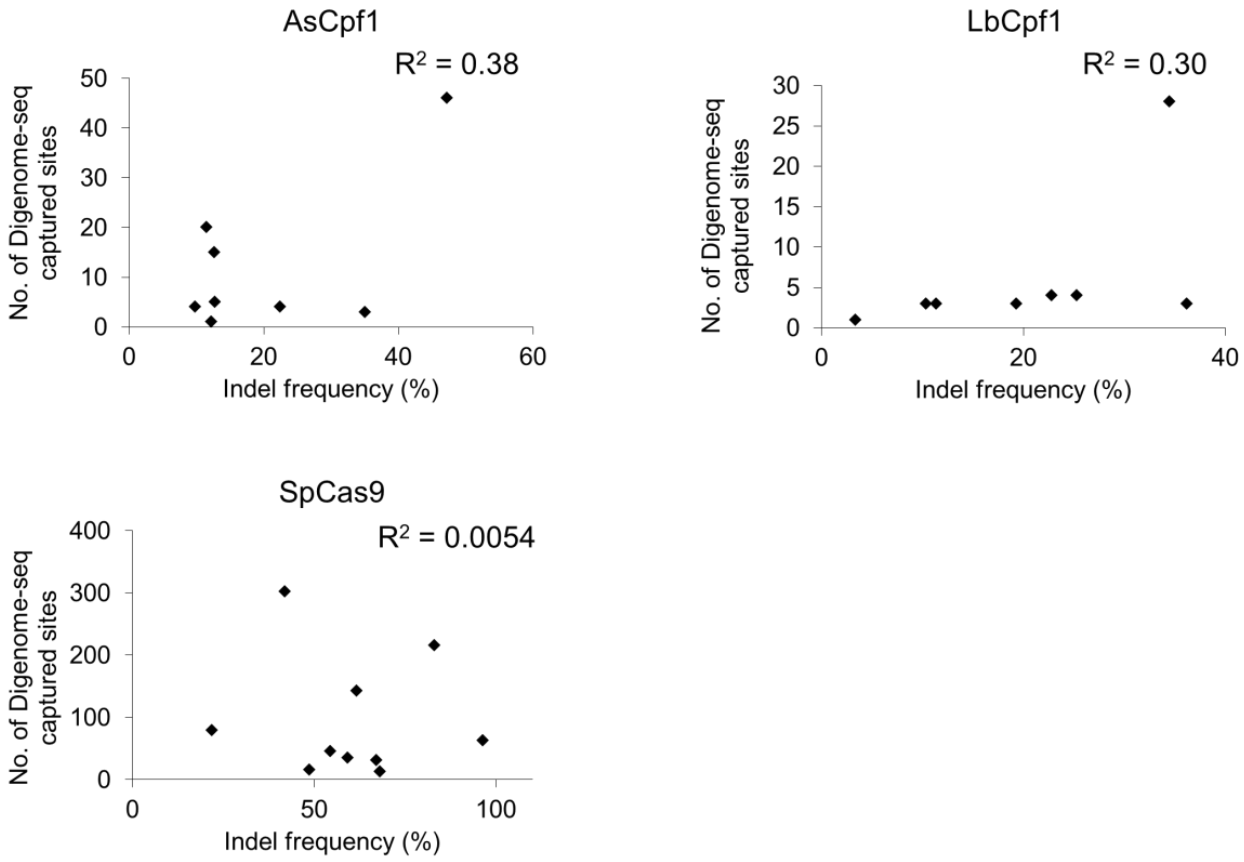
Poor correlation between the number of Digenome-captured sites and the number of homologous sites with 6 or fewer mismatches in the human genome.



Supplementary Figure 6

Sequence logos of Digenome-captured sites.

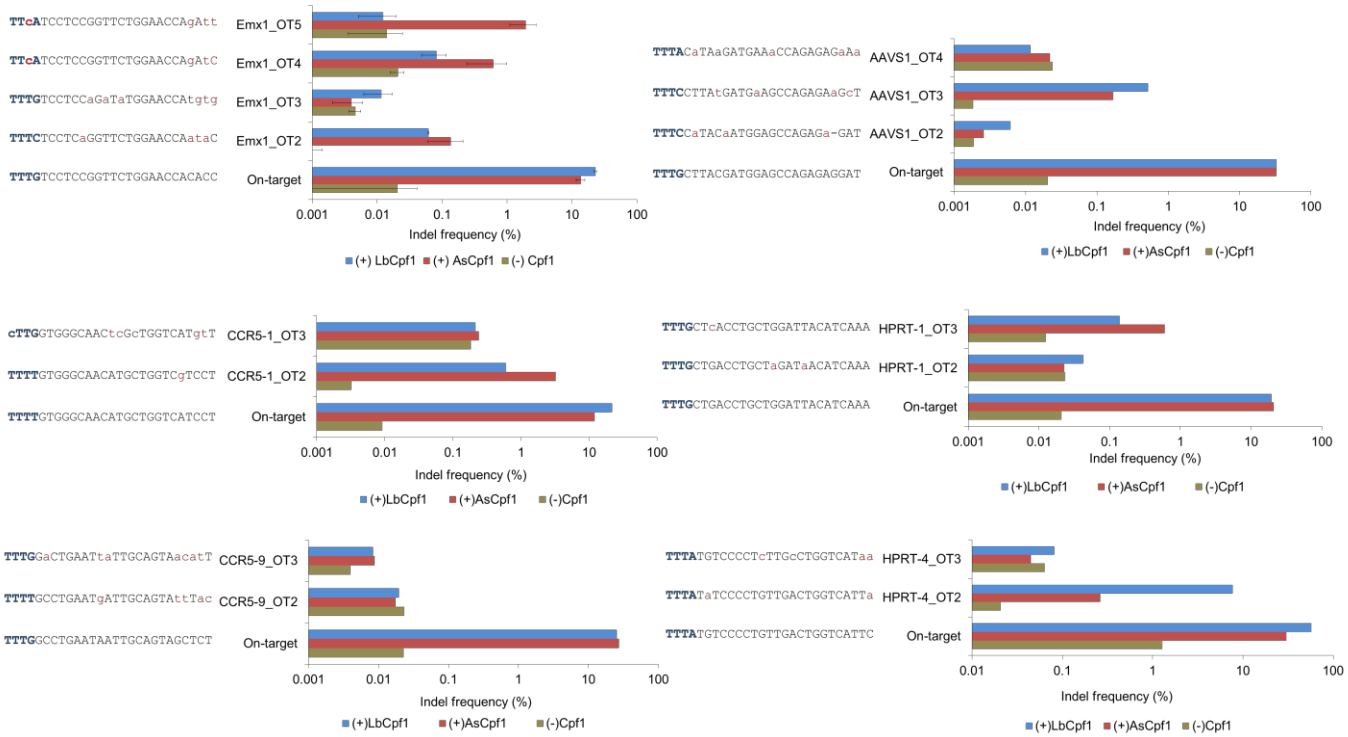
Sequence logos were obtained via WebLogo using Digenome-captured sites. Note that just the single *DNMT1*-4 on-target site was captured with LbCpf1 and AsCpf1.



Supplementary Figure 7

Poor correlations between indel frequencies at on-target sites and genome-wide target specificities of Cpf1 and Cas9.

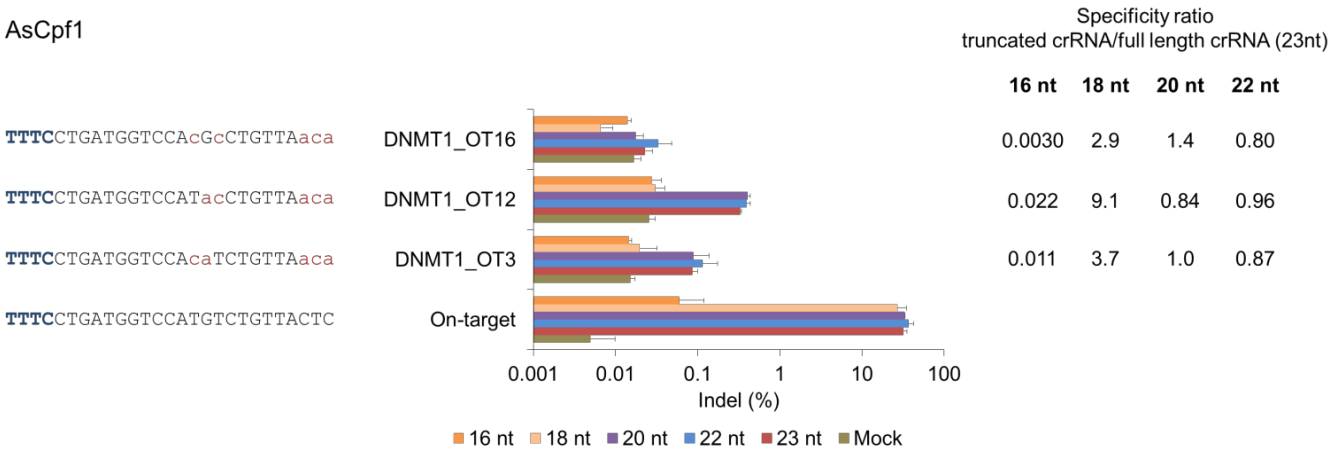
SpCas9 indel frequencies at on-target sites and numbers of Digenome-seq positive sites were from our previous study[11].



Supplementary Figure 8

Indel frequencies at Digenome-captured sites.

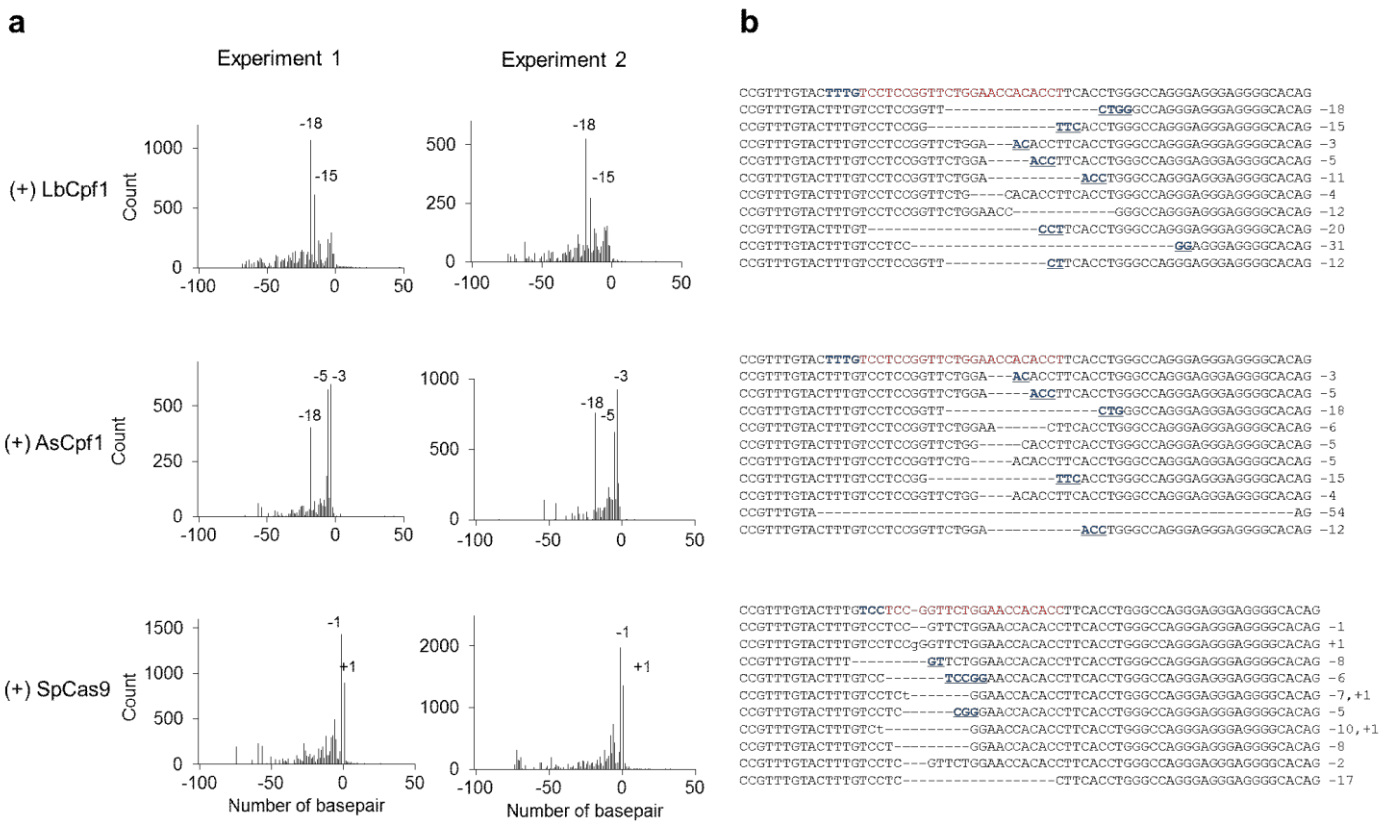
Indel frequencies at Digenome-captured sites were determined via targeted deep sequencing in HEK293T17 cells transfected with AsCpf1 (red) or LbCpf1 (blue) plasmids.



Supplementary Figure 9

Indel frequencies at on-target and off-target sites obtained with truncated and full-length crRNAs.

crRNAs truncated at the 3' end (tru-crRNAs) were designed to match the *DNMT1*-3 target site. Each tru-crRNA was transfected with the AsCpf1 expression plasmid into HEK293T cells using lipofectamine 2000. After 72 hr, genomic DNA was isolated and indel frequencies at the on-target and off-target sites were measured by targeted deep sequencing. Error bars represent mean ± s.e.m.

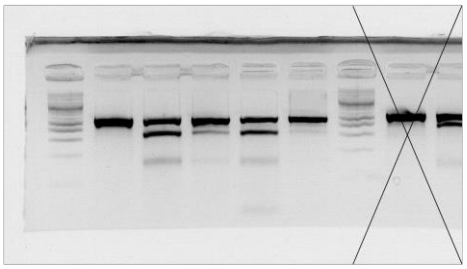


Supplementary Figure 10

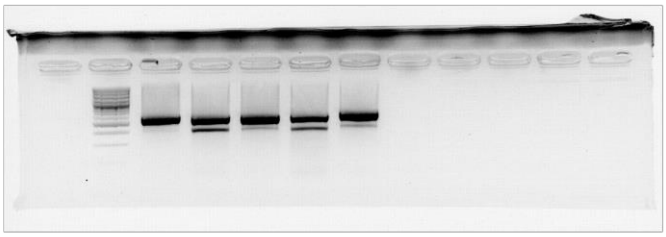
Mutation signatures of LbCpf1, AsCpf1, and SpCas9.

(a) The number of mutant sequence reads binned by the deletion/insertion size in base pairs. Mutation signatures at the *EMX1*-2 site were obtained in HEK293T cells transfected with LbCpf1, AsCpf1, or SpCas9 plasmids using targeted deep sequencing. (b) Mutant sequences induced at the *EMX1*-2 target site. The most frequently-identified sequences are shown. PAM sequences are shown in blue. crRNA/sgRNA target sites are shown in red. Microhomology sequences are underlined and in blue. The number of deleted or inserted bases are shown on the right.

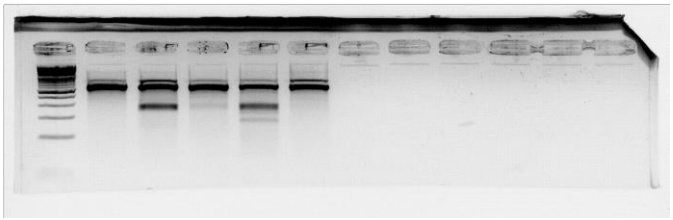
***DNMT1*-3**



***DNMT1*-4**



AAVS1



Supplementary Figure 11

Full-length gel images

Full-length gel images of Fig. 1a. Unrelated lanes are indicated by cross.

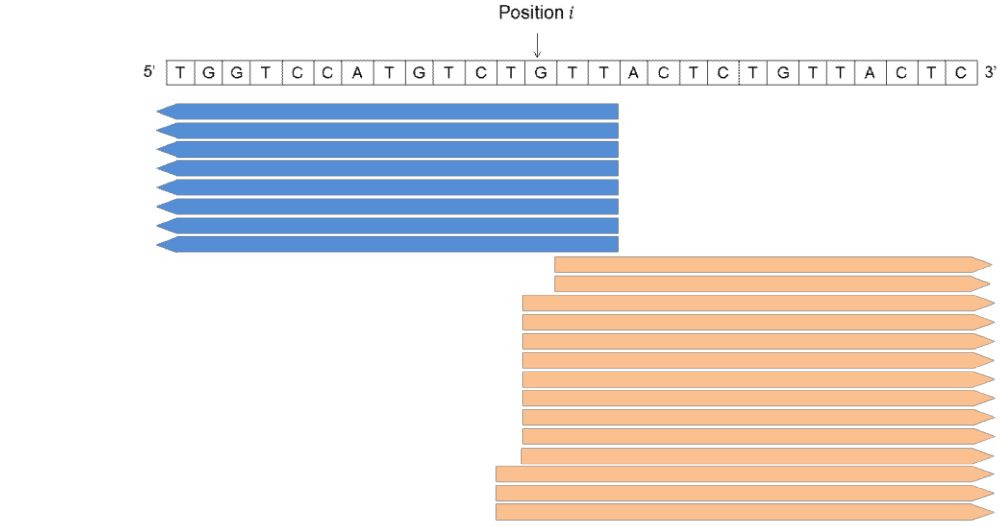
Score at position i =

$$\sum_{a=1}^{5} \frac{(F_i - 1)}{D_i} X \frac{(R_{i-1+a} - 1)}{D_{i-1+a}} \times (F_i + R_{i-1+a} - 2) + \sum_{a=1}^{5} \frac{(R_{i+2} - 1)}{D_{i+2}} X \frac{(F_{i-3+a} - 1)}{D_{i-3+a}} \times (R_{i+2} + F_{i-3+a} - 2)$$

F_i : Number of forward sequence reads starting at position i

R_i : Number of reverse sequence reads starting at position i

D_i : Sequencing depth at position i



Forward strand count	0	0	0	0	0	0	0	0	0	0	0	3	9	2	0	0	0	0	0	0	0	0	0	0	0	0
Reverse strand count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0
Depth	8	8	8	8	8	8	8	8	8	8	8	11	20	22	22	14	14	14	14	14	14	14	14	14	14	14

Score at position i

= (9-1)/20 * [(0-1)/20 * (9+0-2) + (0-1)/22 * (9+0-2) + (8-1)/22 * (9+8-2) + (0-1)/14 * (9+0-2) + (0-1)/14 * (9+0-2)]

+ (8-1)/22 * [(0-1)/8 * (8+0-2) + (3-1)/11 * (8+3-2) + (9-1)/20 * (8+9-2) + (2-1)/22 * (8+2-2) + (0-1)/22 * (8+0-2)]

= 3.46

Supplementary Figure 12

In vitro DNA cleavage scoring system for Cpf1 Digenome-seq analysis.

Nature Biotechnology: doi:10.1038/nbt.3609

Supplementary Information

Genome-wide target specificities of Cpf1 nucleases in human cells

Supplementary Table 1. Cpf1 and Cas9 target sites.

DNA sequences, positions, and indel frequencies of Cpf1 and Cas9 target sites are shown. Cpf1 PAMs and Cas9 PAMs are shown in blue and red, respectively.

Target ID	Gene	Chromosome	Position	Target site	Indel frequency (%)		
					LbCpf1	AsCpf1	SpCas9
DNMT1-3	*DNMT1* (intron)	Chr 19	10133739	taatg**TTTCCT**GATGGTCCATGTCTGTTACTCgcctgtcaagtggcgtg	31 ± 3.5	42 ± 8.1	51 ± 7.2
DNMT1-4	*DNMT1*	Chr 19	10133686	ttcct**TTTA**TTTCCCTTCAGCTAAAATAA**AGG**aggaggaagctgctaag	6.4 ± 2.6	9.7 ± 1.5	12 ± 1.5
EMX1-2	*EMX1*	Chr 2	72933764	tgtac**TTTG**T**CCT**CCGGTTCTGGAACCACACCttcacctgggccaggga	18 ± 5.8	13 ± 3.0	38 ± 7.8
AAVS1	Intergenic	Chr 19	55115548	taagg**TTTG**CTTACGATGGAGCCAGAG**AGG**ATcctgggagggagagctt	19 ± 3.5	23 ± 4.2	36 ± 6.9
CCR5-1	*CCR5*	Chr 3	46373006	tttgg**TTTT**GTGGGCAACATGC**TGG**TCATCCTcatcctgataaactgca	13 ± 1.5	10 ± 1.3	23 ± 3.5
CCR5-9	*CCR5*	Chr 3	46373664	aattc**TTTG**G**CCT**GAATAATTGCAGTAGCTCTaacaggttggaccaagc	18 ± 3.6	21 ± 5.4	20 ± 4.9
HPRT1-1	*HPRT1*	Chr X	134475241	aattc**TTTG**CTGA**CCT**GCTGGATTACATCAAAgcactgaatagaaatag	20 ± 1.4	20 ± 2.3	36 ± 2.4
HPRT1-4	*HPRT1*	Chr X	134486436	ttact**TTTA**TGT**CCC**CTGTTGACTGGTCATTCtagttaaaaaaaaaaaa	65 ± 1.7	48 ± 4.5	43 ± 2.3
VEGFA-2	*VEGFA*	Chr 6	43770839	gaaac**TTTT**CGT**CCA**ACTTCTGGGCTGTTCTCgcttcggaggagccgtg	1.3 ± 0.20	5.4 ± 0.81	43 ± 6.0
HBB	*HBB*	Chr 11	5226711	ggttc**TTTG**AGTCCTTTGGGGATCTGTCCACTcctgatgctgttat**ggg**	2.3 ± 0.24	3.3 ± 0.53	19 ± 3.8

Nature Biotechnology: doi:10.1038/nbt.3609