OMTN, Volume *26*

Supplemental information

The comparison of ZFNs, TALENs, and SpCas9 by GUIDE-seq in HPV-targeted gene therapy

Zifeng Cui, Hui Liu, Hongfeng Zhang, Zhaoyue Huang, Rui Tian, Lifang Li, Weiwen Fan, Yili Chen, Lijie Chen, Sen Zhang, Bhudev C. Das, Konstantin Severinov, Inga Isabel Hitzeroth, Priya Ranjan Debata, Zhuang Jin, Jiashuo Liu, Zheying Huang, Weiling Xie, Hongxian Xie, Bin Lang, Ji Ma, Haiyan Weng, Xun Tian, and Zheng Hu

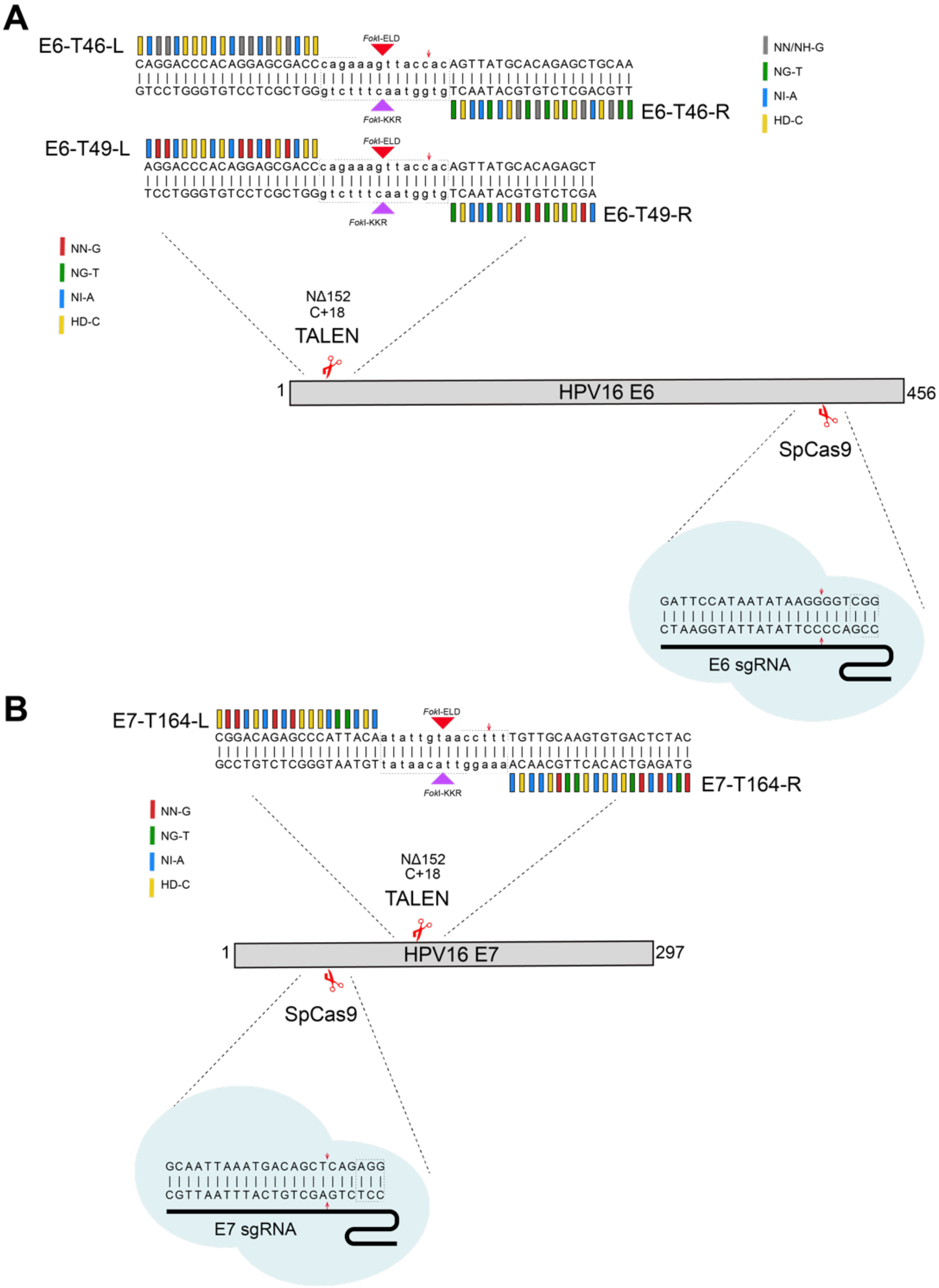


Fig. S1. The cleavage targets of TALENs and SpCas9 in HPV16 E6 **(A)** and E7 **(B)** genes.

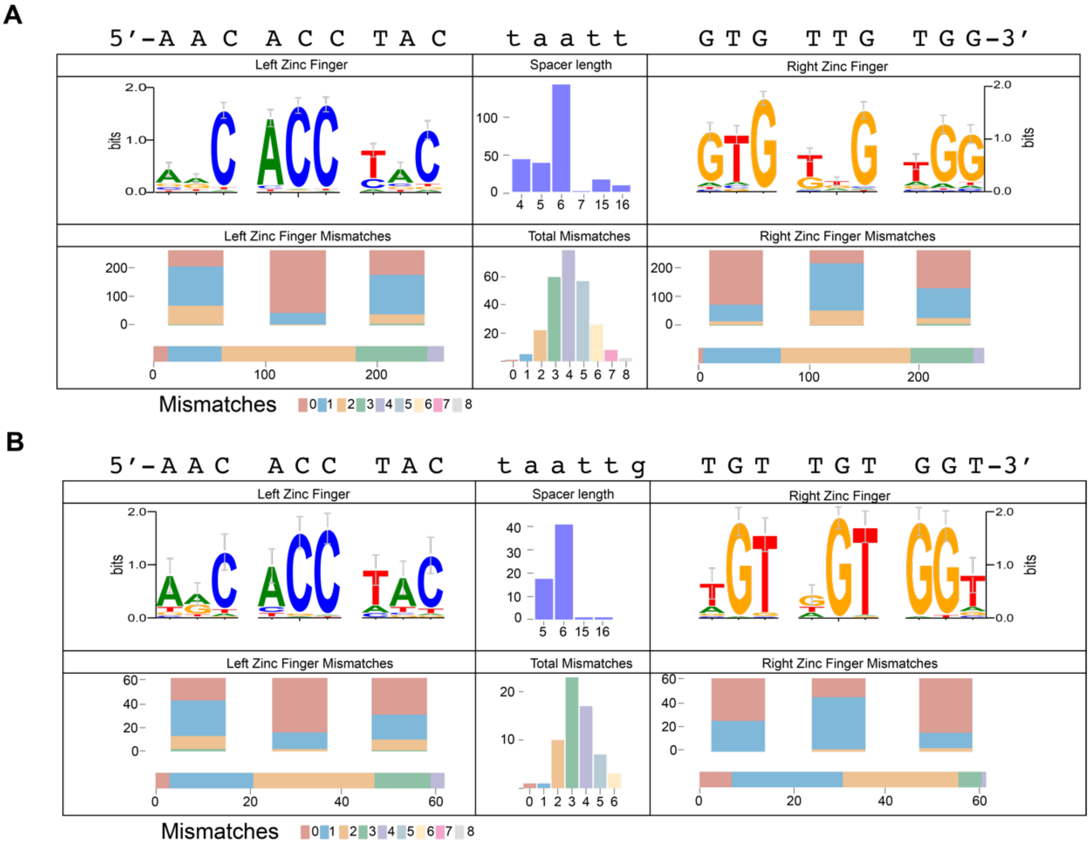


Fig. S2. The web-logos of ZFN Wolfe-162 **(A)** and Wolfe-165 **(B)**.

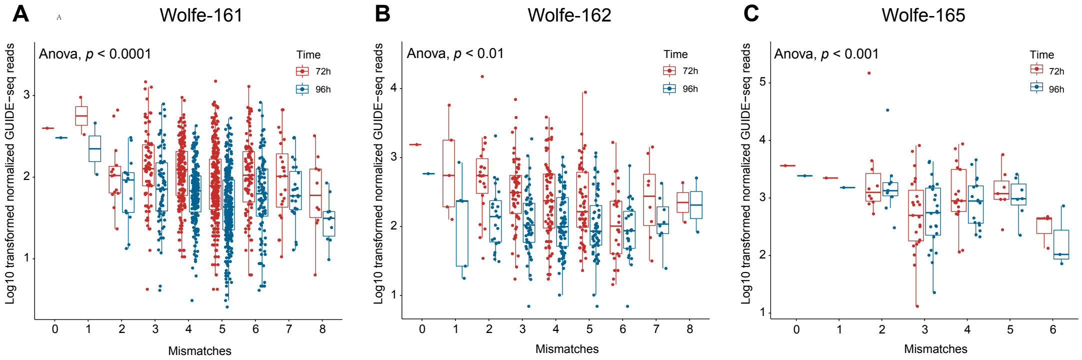
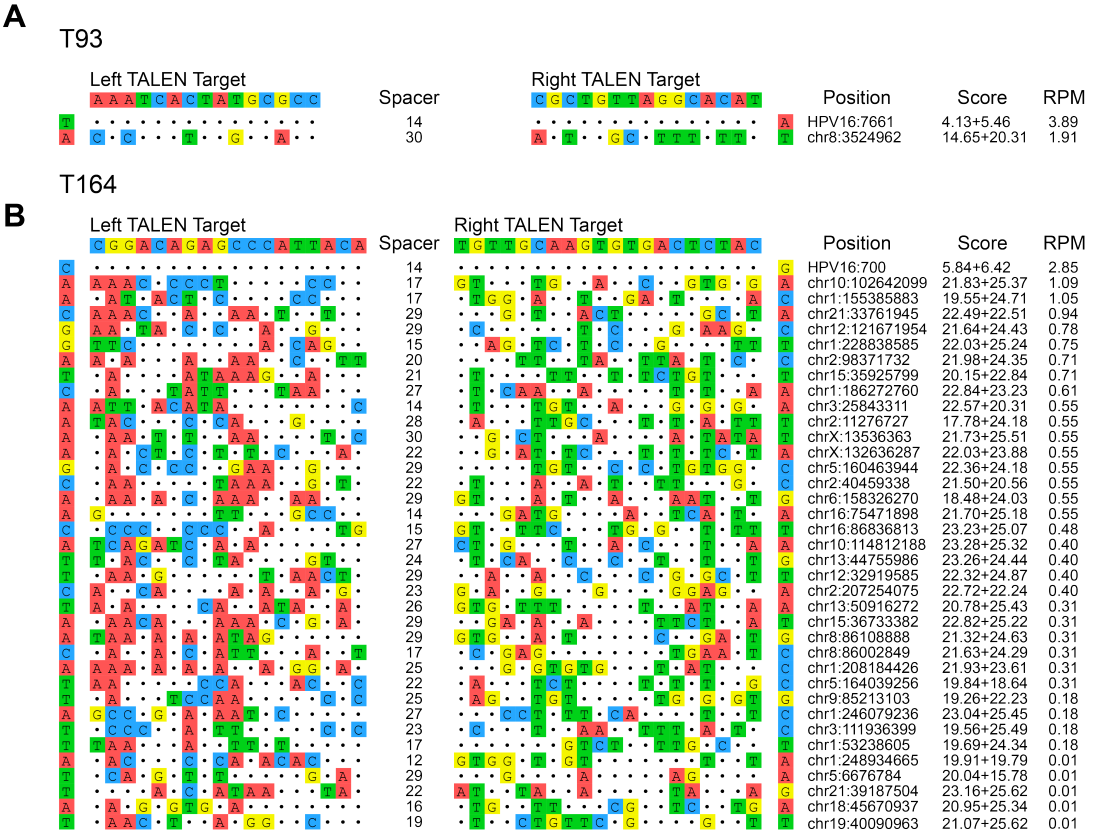


Fig. S3. The relationships between log10 transformed normalized GUIDE-seq reads and mismatches of off-targets for ZFN Wolfe-161 **(A)**, Wolfe-162 **(B)** and Wolfe-165 **(C)**, *P* values from ANOVA analysis are provided.



A

T93

	Left TALEN Target	Spacer	Right TALEN Target		Position	Score	RPM
	AAATCACTATGCGCC		CGCTGTTAGGCACAT				
T	•••••••••••••••	14	•••••••••••••••	A	HPV16:7661	4.13+5.46	3.89
A	C•C•••T••G••A••	30	A•T••GC•TTT•TT•	T	chr8:3524962	14.65+20.31	1.91

B

T164

	Left TALEN Target	Spacer	Right TALEN Target		Position	Score	RPM
	CGGACAGAGCCCATTACA		TGTTGCAAGTGTGACTCTAC				
C	••••••••••••••••••	14	••••••••••••••••••••	G	HPV16:700	5.84+6.42	2.85
A	AAAC•CCCT•••••CC••	17	GT•••TG••A••C••GTG•G	A	chr10:102642099	21.83+25.37	1.09
A	•AT•ACT•C••••CC•••	17	•TGG•A••T••GA•T••••A	C	chr1:155385883	19.55+24.71	1.05
C	AAAC••A••AA•T••T••	29	•••G•T••ACT•••••GC•T	A	chr21:33761945	22.49+22.51	0.94
G	AA•TA•C•C••A••G•••	29	•C••••••T•C•••G•AAG•	C	chr12:121671954	21.64+24.43	0.78
G	TTC••••••••A•CAG••	15	••AG•TC•T•C••G••••TT	T	chr1:228838585	22.03+25.24	0.75
A	A•A•••A••AA••C••TT	20	••••TT••TA••TTA•T•C•	C	chr2:98371732	21.98+24.35	0.71
T	•A••••ATAAAG••A•••	21	•T••••TT••T•TCTGT•••	T	chr15:35925799	20.15+22.84	0.71
C	•A•••TATT•••TAA•••	27	•T•CAA••A•••••T•T••A	A	chr1:186272760	22.84+23.23	0.61
A	ATT•ACATA••••••••C	14	•T•••TGT••A•••G•G•G•	A	chr3:25843311	22.57+20.31	0.55
A	TAC•••C•CA•••G••••	28	•A•••TTGC•••T•T•A•TT	T	chr2:11276727	17.78+24.18	0.55
A	•AA•T•T••AA••••T•C	30	••G•CT•••A••••A•TATA	T	chrX:13536363	21.73+25.51	0.55
A	•A•CT•C•T•T•C•••A•	22	••G•AT•TC•••T•T•TC•T	A	chrX:132636287	22.03+23.88	0.55
G	•A•C•CC••GAA••G•••	29	•••••TGT••C•C•TGTGG•	C	chr5:160463944	22.36+24.18	0.55
C	•AA•••••TAAA••G•T•	22	•T•••A••A•T••TT•••G•	C	chr2:40459338	21.50+20.56	0.55
A	•AA•A•C•AAA••AA•••	29	GT•••A•T••A•••AAT••T	G	chr6:158326270	18.48+24.03	0.55
A	G•••••••TT•••GCC••	14	•••GATG••••A••TCA•T•	A	chr16:75471898	21.70+25.18	0.55
C	•CCC••CCC••A••••TG	15	GT••TTC•••TG•G••T•TT	T	chr16:86836813	23.23+25.07	0.48
A	TCAGATC•A•A•••••••	27	CT•G•••T••A•C•••T••A	A	chr10:114812188	23.28+25.32	0.40
T	T•AC••C•TA••••GT••	24	•T•CA••C••C••T••T••T	G	chr13:44755986	23.26+24.44	0.40
T	•AA•G••••••T•AACT•	29	••A••A••C•••C•G•GC•T	T	chr12:32919585	22.32+24.87	0.40
C	A••CA••••A•A••A•G•	23	G•A••G•••G••••GGAG••	A	chr2:207254075	22.72+22.24	0.40
T	A•A••••CA••ATA••A•	26	GTG•TTT•••••T••AT••A	A	chr13:50916272	20.78+25.43	0.31
A	•AACA•••AAA•C•G•A•	29	••GA•A••A••••TTCT••A	T	chr15:36733382	22.82+25.22	0.31
A	TAA•A•A•ATAG••••••	29	GTG••A•T•••••C••GA•T	G	chr8:86108888	21.32+24.63	0.31
C	•A••A•C•ATT•••A••T	17	•C•GAG••••••••TGAA•T	C	chr8:86002849	21.63+24.29	0.31
A	AAA•A•A•A••A•GG•A•	25	•••G•GTGTG•••T•AT•••	C	chr1:208184426	21.93+23.61	0.31
T	AA•••••CCA•••AC••C	22	•A•••TCT••••T•T•T••G	C	chr5:164039256	19.84+18.64	0.31
T	•A•••TCCAA•••••C•C	25	•AG••TGT•••••TG•G•GT	G	chr9:85213103	19.26+22.23	0.18
A	GCC•G•A•AAT•C•••••	27	•••CCT•TT•CA••••T••T	C	chr1:246079236	23.04+25.45	0.18
T	•CCC••A•TT•••••C•C	23	•C•••T••AA••TTT•A•T•	C	chr3:111936399	19.56+25.49	0.18
T	TAA•••A••TT•T•••••	17	•••••••GTCT••TTG•C••	T	chr1:53238605	19.69+24.34	0.18
A	•AC•••C•CA•ACAC•••	12	GTGG•T•GT•••••••T••T	A	chr1:248934665	19.91+19.79	0.01
T	•CA•G•T•T•••••G•A•	29	•••G••••A•••••AG••••	A	chr5:6676784	20.04+15.78	0.01
T	••••A•C•ATAA•••TA•	22	AT••TA••A•••••TA•••A	G	chr21:39187504	23.16+25.62	0.01
A	•A•G•GTG•A••••••••	16	•TG••TT•••CG••TC••TG	A	chr18:45670937	20.95+25.34	0.01
T	•AAC•T••A•GG••C•••	19	•T••CTGTTC•G••••G••T	T	chr19:40090963	21.07+25.62	0.01

Fig. S4. The schematic of on targets and off-targets of TALENs targeting URR **(A)** and E7 **(B)** in 96h experiments.

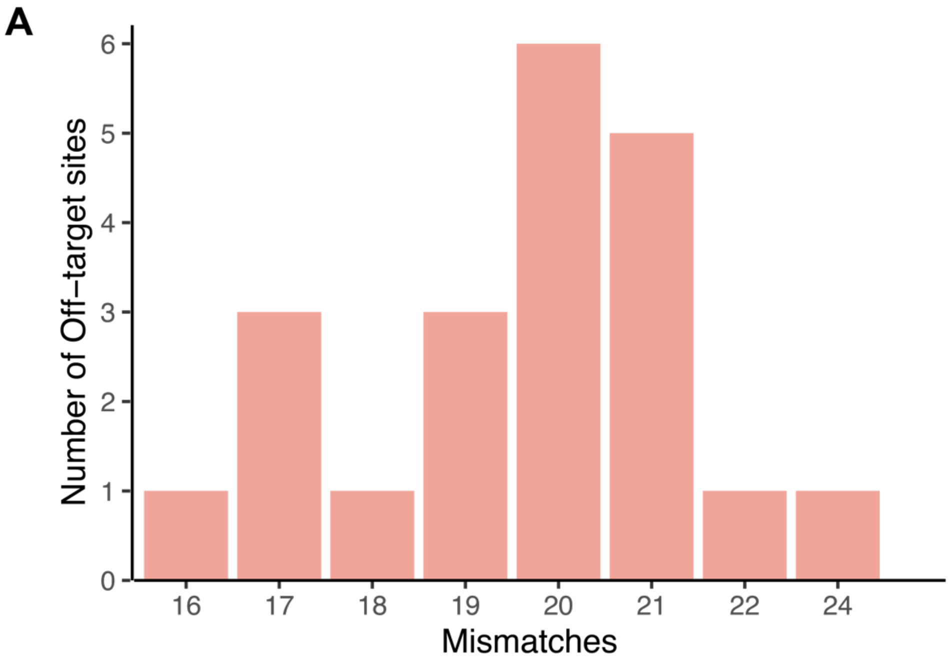


Fig. S5. The histogram exhibited off-target distribution of different mismatches for T49.

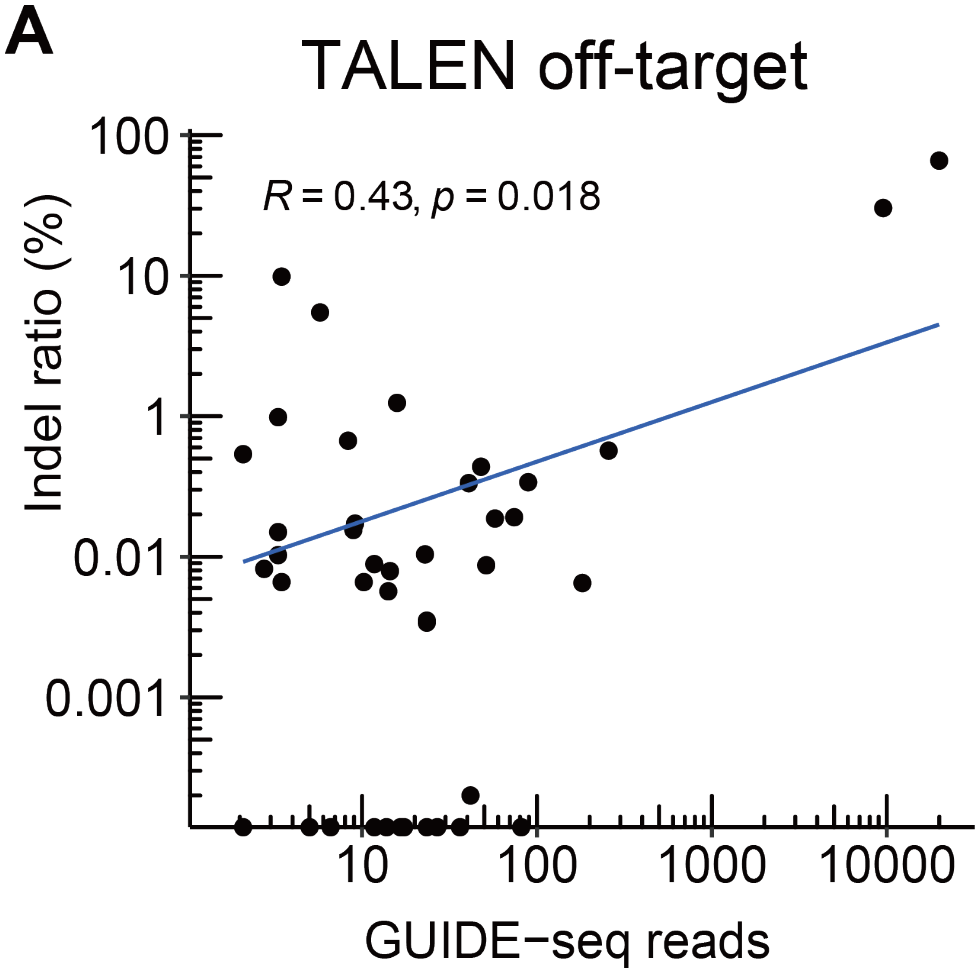


Fig. S6. The correlation between GUIDE-seq reads and validated indel ratios of TALEN off-targets.

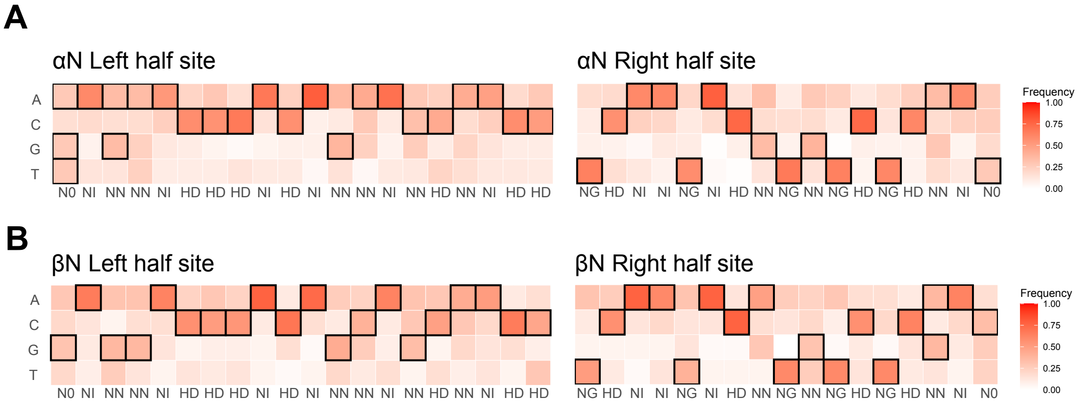


Fig. S7. The nucleotide frequencies of each RVD derived from T49-αN-NN off-targets **(A)** and T49-βN-NN off-targets **(B)**.

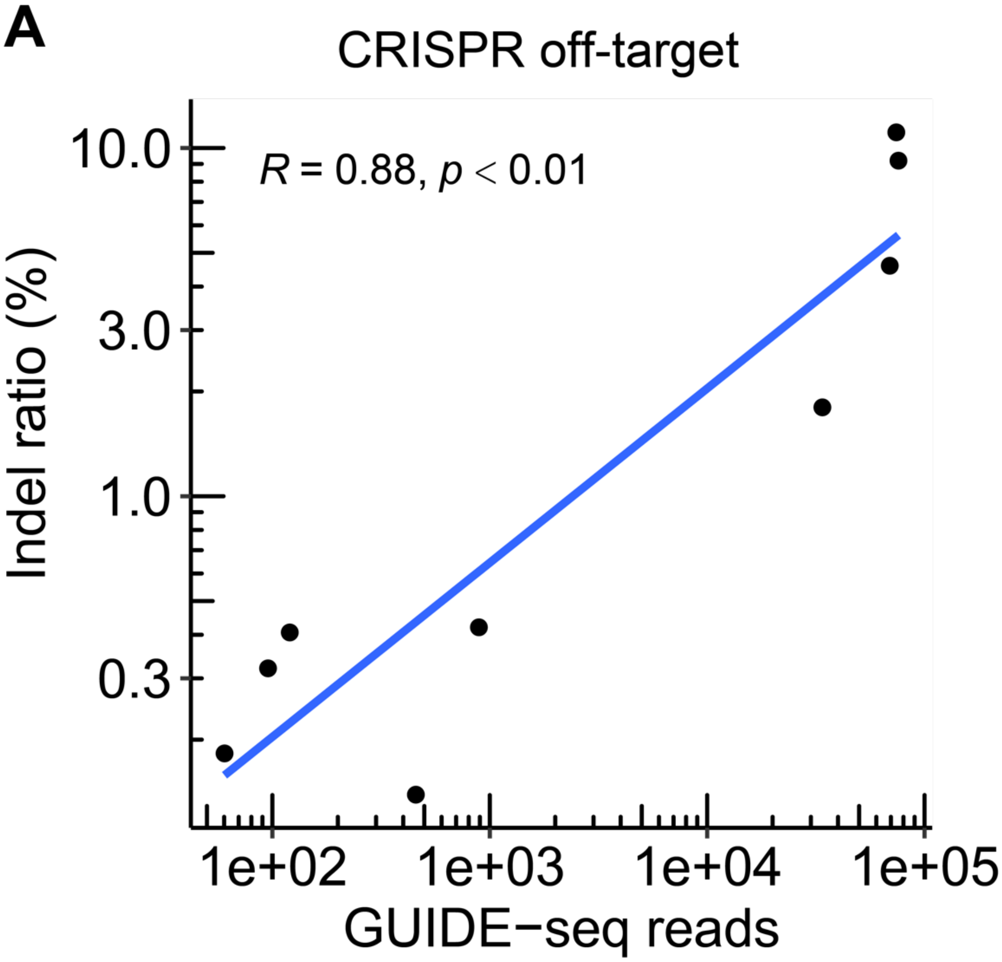
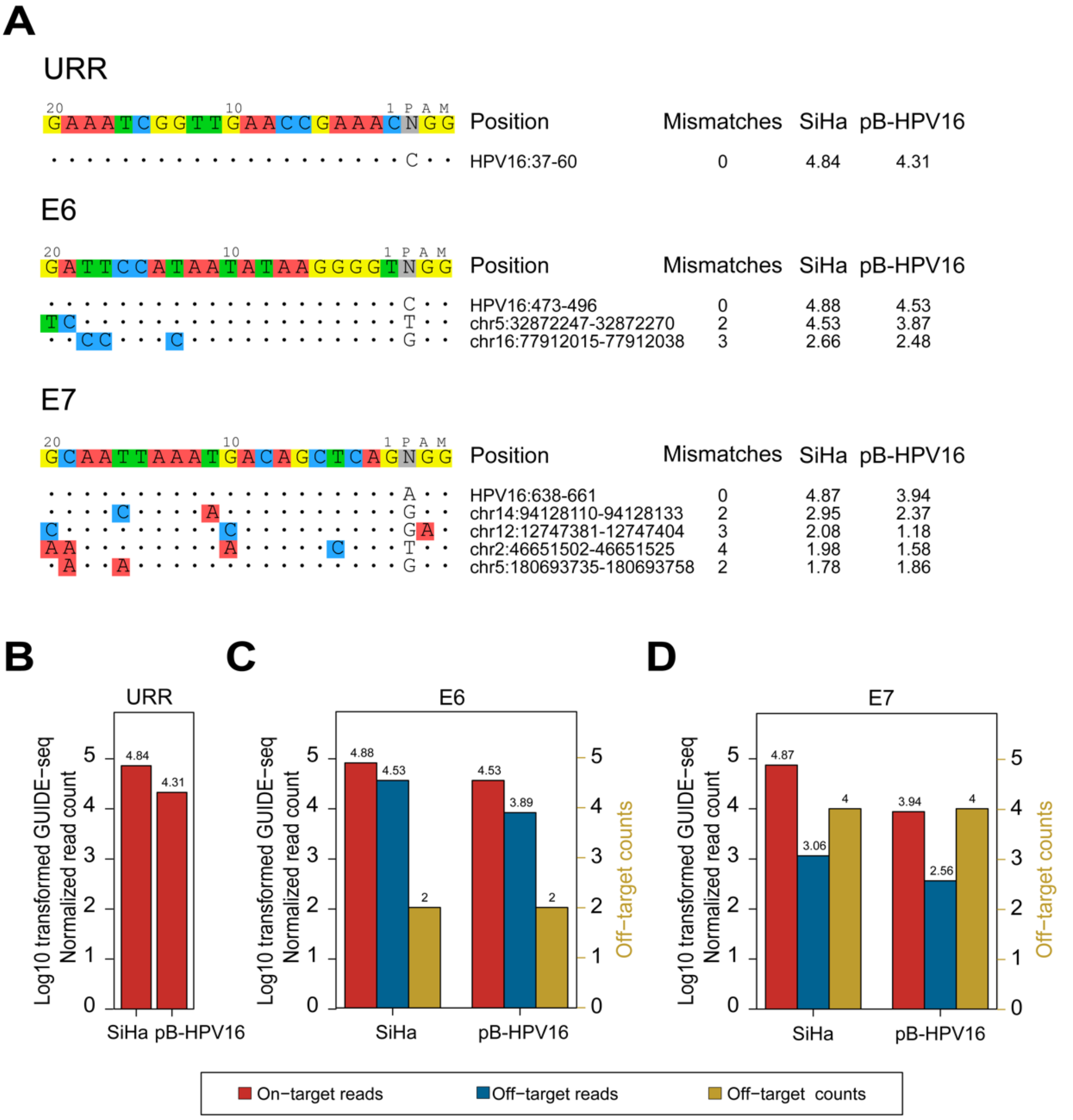


Fig. S8. The correlation between GUIDE-seq reads and validated indel ratios of CRISPR off-targets.



URR

GAAATCGGTTGAACCGAAACNGG	Position	Mismatches	SiHa	pB-HPV16
····················C··	HPV16:37-60	0	4.84	4.31

E6

GATTCCATAATATAAGGGGTNGG	Position	Mismatches	SiHa	pB-HPV16
····················C··	HPV16:473-496	0	4.88	4.53
TC··················T··	chr5:32872247-32872270	2	4.53	3.87
··CC···C············G··	chr16:77912015-77912038	3	2.66	2.48

E7

GCAATTAAATGACAGCTCAGNGG	Position	Mismatches	SiHa	pB-HPV16
····················A··	HPV16:638-661	0	4.87	3.94
····C····A··········G··	chr14:94128110-94128133	2	2.95	2.37
C·········C·········GA·	chr12:12747381-12747404	3	2.08	1.18
AA········A·····C···T··	chr2:46651502-46651525	4	1.98	1.58
·A··A···············G··	chr5:180693735-180693758	2	1.78	1.86

Fig. S9. The comparison of SpCas9 off-targets in SiHa and pB-HPV16 cell lines. **(A)** Schematic of GUIDE-seq on/off-target sites and MIP validated results for 3 sgRNAs in SiHa and pB-HPV16 cell lines. **(B-D)** The comparison of log10 transformed normalized on-target reads, off-target reads and off-target counts for each sgRNA targeting URR **(B)**, E6 **(C)** and E7 **(D)** between SiHa cell lines and pB-HPV16 cell lines in 96h experiments.

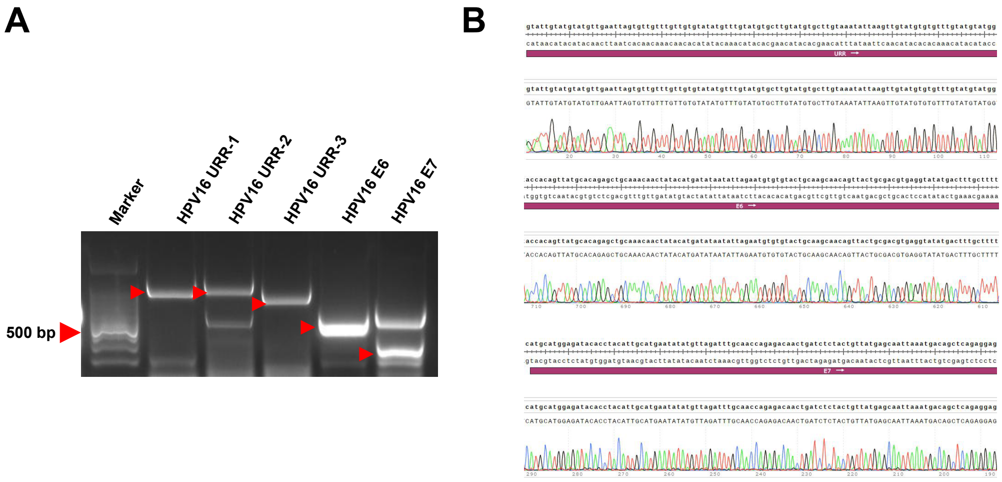


Fig. S10. The validation of HPV16 DNA in pb-HPV16 transfected stable cell lines. **(A)** PCR products of HPV16 DNA, including URR, E6 and E7 genes. **(B)** Sanger sequencing of URR region.

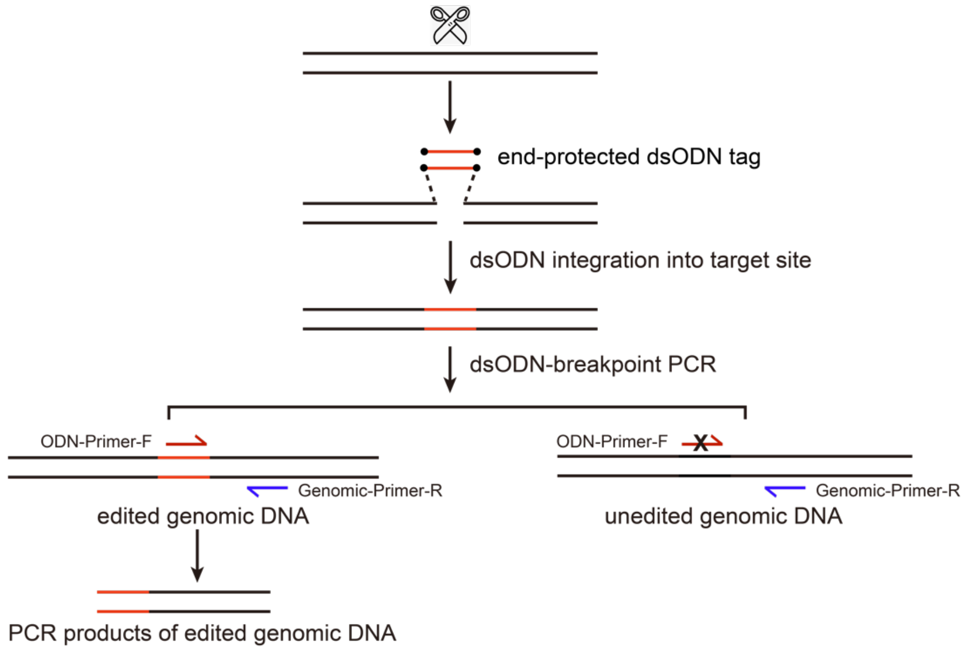


Fig. S11. The schematics illustrating the dsODN-PCR process and the generated products.

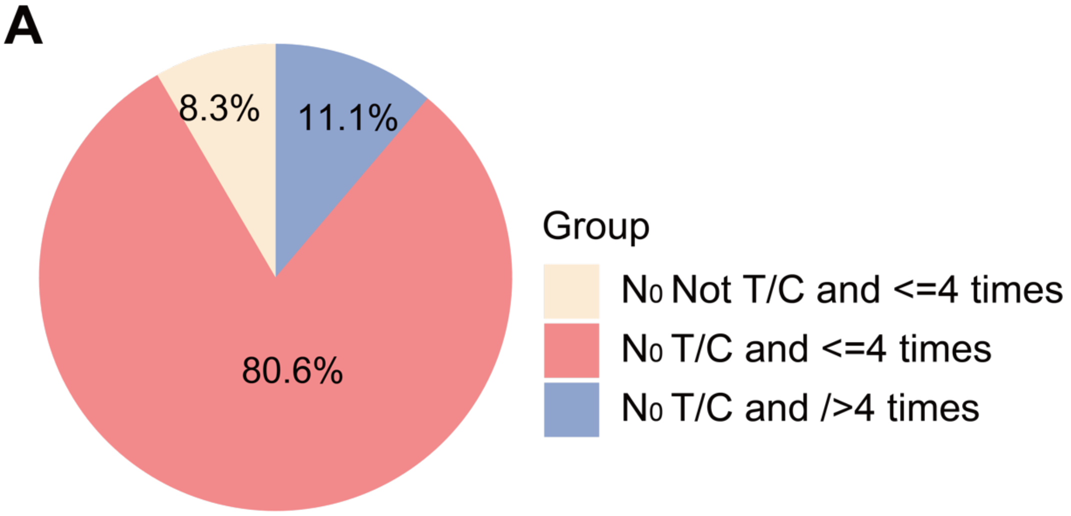


Fig. S12. The verification of modified TALEN-NT 2.0 program in previous reported 36 validated off-targets.

Table S1. The off-targets of three ZFNs in original model.

Table S2. The increased Off-targets of three ZFNs in extended model.

Table S3. The indel rates of Wolfe-165 in 96h by MIP sequencing.

Table S4. The on and off-target summary of 9 TALENs in this study.

Table S5. The indel rates of TALEN T49-WT-NN and T49-βN-NN in 96h by MIP sequencing.

Table S6. The frequency of 4 nucleotides in each target position of three T49 TALENs in 96h.

Table S7. The MIP validated results of off-target sites for 3 SpCas9 sgRNAs.

Table S8. The DSB toxicity evaluation supported by detected dsODN integration sites using GUIDE-seq method.

Table S9. The DSB toxicity sites supported by detected dsODN integration positions using GUIDE-seq method.

Table S10. Three SpCas9 sgRNAs used in this study.

Table S11. Three effective zinc finger sequences targeting HPV URR gene.

Table S12. The RVD and target sequences of each TALEN.

Table S13. The dsODN-tag breakpoint PCR primers used in this study.

Table S14. Previous in vivo validated off-targets of three-finger ZFN targeted to VEGFA.

Table S15. The consistency rate of modified TALEN-NT 2.0 program using validated off-target sites in previous study.