Supplemental Materials:

Genome editing HPV-E6/E7 by TALENs to treat HPV-related cervical malignancy

Zheng Hu[†1], Wencheng Ding[†1], Da Zhu[1], Lan Yu[1], Xiaohui Jiang[1], Xiaoli Wang[1], Changlin Zhang[1], Liming Wang[1], Teng Ji[1], Dan Liu[1], Dan He[2], Xi Xia[3], Tao Zhu[1], Juncheng Wei[1], Peng Wu[1], Changyu Wang[1], Ling Xi[1], Qinglei Gao[1], Gang Chen[1], Rong Liu[1], Kezhen Li[1], Shuang Li[1], Shixuan Wang[1], Jianfeng Zhou[1], Ding Ma*[1] and Hui Wang*[1]

Supplemental Table 1. Characteristics of TALENs targeting HPV16 and HPV18 E6 and E7 ongogenes.

Gene	ID	target sequence*	spacer lenth	TAL RVDs
HPV16E6	T3	GCACCAAAAGAGAACTgcaatgtttcaggacccacaggagcgacccAGAAAGTTACCACAGTTATGC	30	NN HD NI HD HD NI NI NI NI NN NI NN NI NI HD NG NN HD NI NG NI NI HD NG NN NG NN NN NG NI NI HD NG NG NG HD NG
	T18	GCAATGTTTCAGGACCCAcaggagcgacccagaAAGTTACCACAGTTATGC	15	NN HD NI NI NG NN NG NG NG HD NI NN NN NI HD HD HD NI NN HD NI NG NI NI HD NG NN NG NN NN NG NI NI HD NG NG
	T27	CAGGACCCACAGGAGcgacccagaaagttaCCACAGTTATGCACAG	15	HD NI NN NN NI HD HD HD NI HD NI NN NN NI NN HD NG NN NG NN HD NI NG NI NI HD NG NN NG NN NN
	T272	ATGGAACAACATTAGAACAGCAATacaacaaaccgttgtgtgATTTGTTAATTAGGTGT	18	NI NG NN NN NI NI HD NI NI HD NI NG NG NI NN NI NI HD NI NN HD NI NI NG NI HD NI HD HD NG NI NI NG NG NI NI HD NI NI NI NG
	T502	GCATGAATATATGTTAGatttgcaaccagagaCAACTGATCTCTACTGTT	15	NN HD NI NG NN NI NI NG NI NG NI NG NN NG NG NI NN NI NI HD NI NN NG NI NN NI NN NI NG HD NI NN NG NG NN
HPV16E7	T512	ATGTTAGATTTGCAACCAGAGACAACTgatctctactgttatgagcAATTAAATGACAGCTCAGAGGAGG	19	NI NG NN NG NG NI NN NI NG NG NG NN HD NI NI HD HD NI NN NI NN NI HD NI NI HD NG HD HD NG HD HD NG HD NG NN NI NN HD NG NN NG HD NI NG NG NG NI NI NG NG
	T514	GTTAGATTTGCAACCagagacaactgatctCTACTGTTATGAGCA	15	NN NG NG NI NN NI NG NG NG NN HD NI NI HD HD NG NN HD NG HD NI NG NI NI HD NI NN NG NI NN
	T518	AGATTTGCAACCAGAGACAACTgatctctactgttatgagcAATTAAATGACAGCTCAGAGGAGG	19	NI NN NI NG NG NG NN HD NI NI HD HD NI NN NI NN NI HD NI NI HD NG HD HD NG HD HD NG HD NG NN NI NN HD NG NN NG HD NI NG NG NG NI NI NG NG
HPV18E6	T12	GAGGATCCAACACGGcgaccctacaagctaCCTGATCTGTGCACGG	15	NN NI NN NN NI NG HD HD NI NI HD NI HD NN NN HD HD NN NG NN HD NI HD NI NN NI NG HD NI NN NN
	T34	ACAAGCTACCTGATCtgtgcacggaactgaACACTTCACTGCAAG	15	NI HD NI NI NN HD NG NI HD HD NG NN NI NG HD HD NG NG NN HD NI NN NG NN NI NI NN NG NN NG
	T194	GCCATAAATGTATAGATttttattctagaattagagaattaAGACATTATTCAGACTCTGTGT	24	NN HD HD NI NG NI NI NI NG NN NG NI NG NI NN NI NG NI HD NI HD NI NN NI NN NG HD NG NN NI NI NG NI NI NG NN NG HD NG
	T257	ATGGAGACACATTGGAAAAACTaactaacactgggttatacAATTTATTAATAAGGTGCCTGCGGTGCC	19	NI NG NN NN NI NN NI HD NI HD NI NG NG NN NN NI NI NI NI NI HD NG NN NN HD NI HD HD NN HD NI NN NN HD NI HD HD NG NG NI NG NG NI NI NG NI NI NI NG NG
	T480	TAAGTATGCATGGACctaaggcaacattgcAAGACATTGTATTGC	15	NG NI NI NN NG NI NG NN HD NI NG NN NN NI HD NN HD NI NI NG NI HD NI NI NG NN NG HD NG NG
HPV18E7	T492	GGACCTAAGGCAACATTGCAAGACATtgtattgcatttagagccccAAAATGAAATTCCGGTTGACCTTCT	20	NN NN NI HD HD NG NI NI NN NN HD NI NI HD NI NG NG NN HD NI NI NN NI HD NI NG NI NN NI NI NN NN NG HD NI NI HD HD NN NN NI NI NG NG NG HD NI NG NG NG NG
	T497	AAGGCAACATTGCAAgacattgtattgcatTTAGAGCCCCAAAATG	15	NI NI NN NN HD NI NI HD NI NG NG NN HD NI NI HD NI NG NG NG NG NN NN NN HD NG HD NG NI NI
	T519	GTATTGCATTTAGAGCCCCAAAATgaaattccggttgaccttctatgtcACGAGCAATTAAGCGACTCAG	25	NN NG NI NG NG NN HD NI NG NG NG NI NN NI NN HD HD HD HD NI NI NI NI NG HD NG NN NI NN NG HD NN HD NG NG NI NI NG NG NN HD NG HD NN NG
eGFP	TAL-eGFP	CCGGCGAGGGCGAGGgcgatgccacctacgGCAAGCTGACCCTGA	15	HD HD NN NN HD NN NI NN NN NN HD NN NI NN NN NG HD NI NN NN NN NG HD NI NN HD NG NG NN HD

* Uppercase letters: left and right target sequences of TALENs, lowercase letters: spacer squences.

Supplemental Table 2. Detection of the top ten predicted off target sites in mouse genome

#ID	Genomic	Predicted off target sequences*	Score	Primers#	size (bp)	T512-Treated		Untreated	
						Reads	Mutation	Reads	Mutation
1	chr10	TCCTCCTCTTGGCTGGCATTTTCCAcagatggcaagtagggtAATTT CATGCCTGATTAGAGGAGGA	-1.77	ttccctccttgcttcaacag ggaagggagtcaaagttttcc	226	23	0	19	0
2	chr10	CCATCTTCTGAGATCTTCTTTAATTgagtctgaagaaagaAATTAA AGAAGATCTCAGAAGATAA	-1.83	ggggattgcattgaatctgt actgaagagtcggggttcct	217	14	0	21	0
3	chr8	TCTGCTACATATGCAGCTAGAGATACGAgctctggagtacccATT TATATGACAGCTCAGAGGAAAA	-1.91	cacttcttggccctgacatt ccctgacagtcatttttccaa	203	14	0	17	0
4	chr2	TCTGTTCACTTTACAGCCAGAGCCATCTtctcatcttctcattccatAG ACAAAGTACTTCTCATAGGATGA	-1.94	ctcggagctccctatcactg ttcacgctgtttctttgtgc	222	23	0	15	0
5	chr8	TCATCCTTTAAAGTCTCAATAAATTgtgcagaagatacagagaaAAT AAGATAAGAGATGGGAGGTGGA	-1.96	ccagagaagccctgttttca tgaggctgctgtgagttcat	231	24	0	20	0
6	chr2	TCCTTCTCTGAGAAGGTATTGAAATctctcatagcaattagAGTTTT GTCTATTTTCAAAACTATCATT	-1.97	gacagtctgttccttccctctc tgttactacgttagttaagtgctctgt	180	23	0	19	0
7	chr10	TTTGATACATATGCAGCTAGAGTCAAGTccctgaagaaagaAATT AAAGAAGATCTCAGAAGATGG	-1.98	aatgggcctctctttccagt ggggattgcattgaatctgt	196	15	0	17	0
8	chr9	TCCCCACATCAGCTGCCTTTTAATTaaataaaacctatgcctgtcATTG TTCTATTGCTGCAGAAATACCATA	-1.98	ggtctccaggatcagatttcc tgctgtcatgttcccagcta	240	22	0	19	0
9	chr7	TATAATAAACTCAAAAAAATAGACAACAccaaagcaaataaccca ACTTAAAATAGGGTACAGAGGGGGA	-1.99	ggacaaaaaggcagcctaca cccatttgtcgattctccat	228	19	0	21	0
10	chr10	AACTACTCTGAGATGTTATCTTATGtcagtaagatggccaagataAAA TGAGTGACAGCTCATGGGGGTG	-2	ttcaacatcttgaggcatgg gccacactggcttccatact	186	14	0	17	0
Total						191	0	185	0

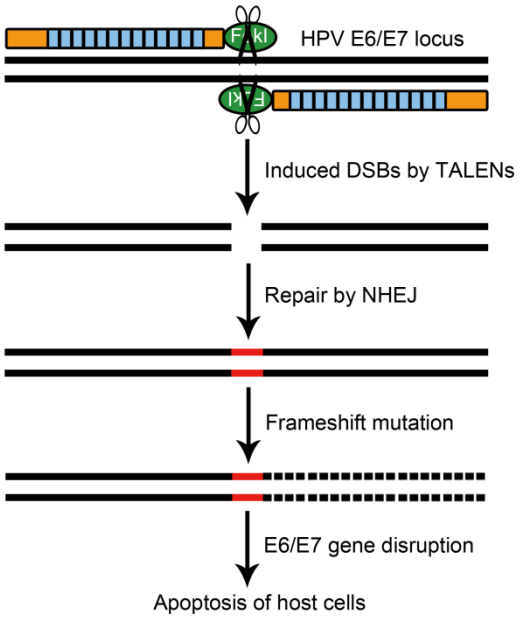
* Uppercase letters: predicted left and right target sequences of TALENs, lowercase letters: spacer sequences.

The primers were designed based on the genomic sequences that contained the predicted off target sites.

Supplemental Table 3. Primer sequences used in T7EI experiments.

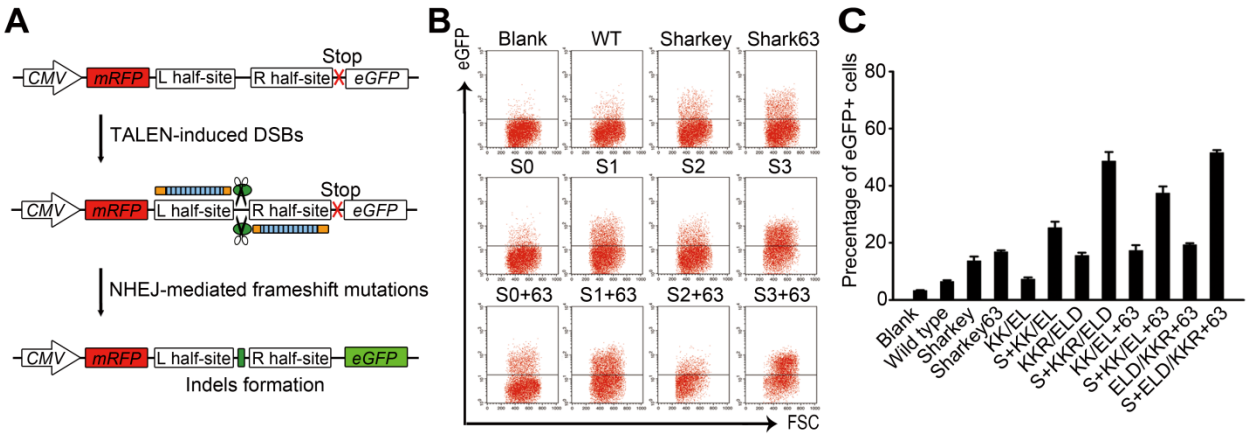
Primers ID	Target of TALENs	Sites of HPV		primer sequences
T7EI-T27	HPV16-E6-T27	HPV16-E6	F	aaactgcacatgggtgtgtg
			R	tttgcttttcttcaggacacag
T7EI-T512	HPV16-E7-T512	HPV16-E7	F	tgtcaaaagccactgtgtcc
			R	taaaatctaccaaatcttcacctgt
T7EI-T34	HPV18-E6-T34	HPV18-E6	F	gctaattgcatacttggcttg
			R	gtgcccagctatgttgtgaa
T7EI-T519	HPV18-E7-T519	HPV18-E7	F	gtgccagaaaccgttgaatc
			R	cctccccgtctgtaccttct

Supplemental Figure 1



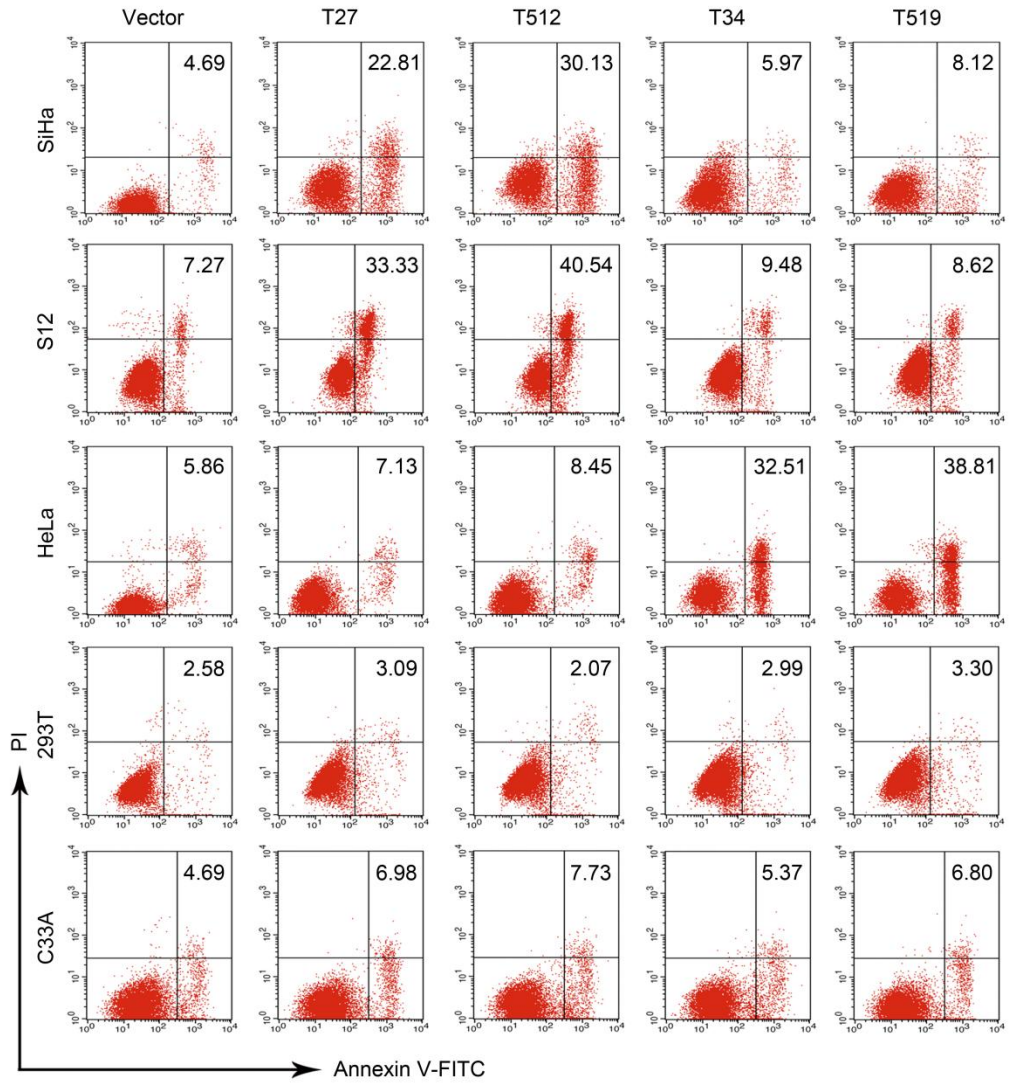
Supplemental Figure 1 A schematic diagram of TALEN-mediated disruptions of HPV oncogenes. TALENs generated DSBs in the coding sequence regions of HPV E6/E7. In response to the DNA damage, the host cells should react and repair the DSBs through the NHEJ pathway, resulting in frameshifts of the viral oncogenes and ablation of their functions.

Supplemental Figure 2



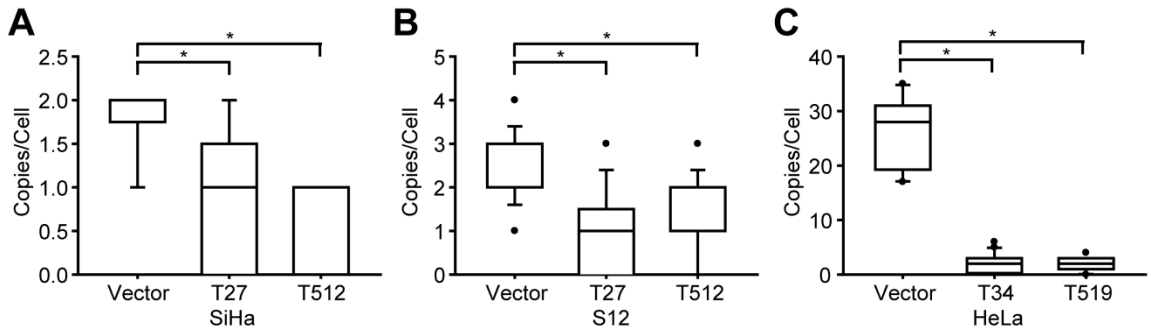
Supplemental Figure 2 Screening of the best TALENs FokIs using the surrogate reporter system. (**A**) The schematic diagram of the surrogate reporter system. The reporter consists of the mRFP gene, the target sequences of T512 (left and right half-sites) and the eGFP gene. mRFP is constitutively expressed from the CMV promoter, while eGFP is not expressed because its sequence is out of frame. When TALENs cleave and induce a DSB into the target sequence, the DSB is repaired by error-prone NHEJ, which often causes frameshift mutations. And such mutations can render eGFP in frame with mRFP, which induces the expression of the mRFP-eGFP fusion protein. (**B**) eGFP fluorescence was detected using FACS after cotransfection of the reporter and T513 with the FokI variants. (**C**) The statistical results of the triple-separated experiments in **B**.

Supplemental Figure 3



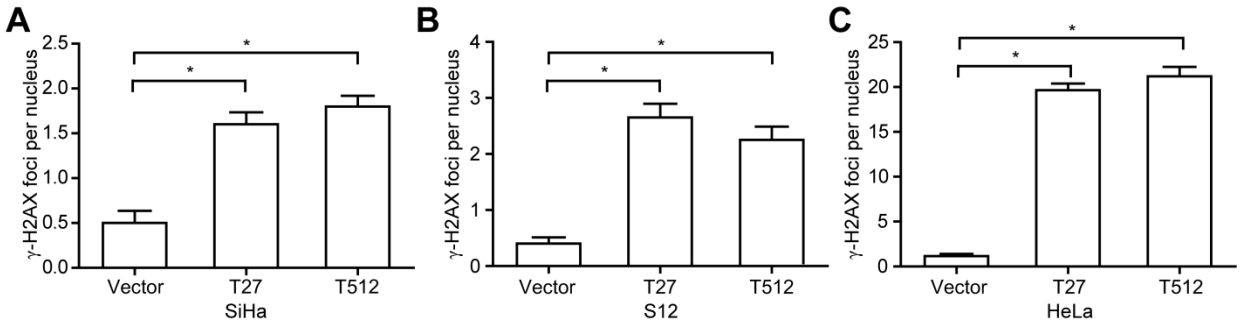
Supplemental Figure 3 The representative dot-plots illustrating the apoptotic cells of SiHa, S12, HeLa, 293T and C33A cell lines after treatment with Vector, T27, T512, T34 and T519 for 48 h detected by flow cytometry after Annexin V-FITC/PI staining. PI, propidium iodide. These experiments were performed triplicate and the average apoptotic rates were performed on the upper right corners.

Supplemental Figure 4



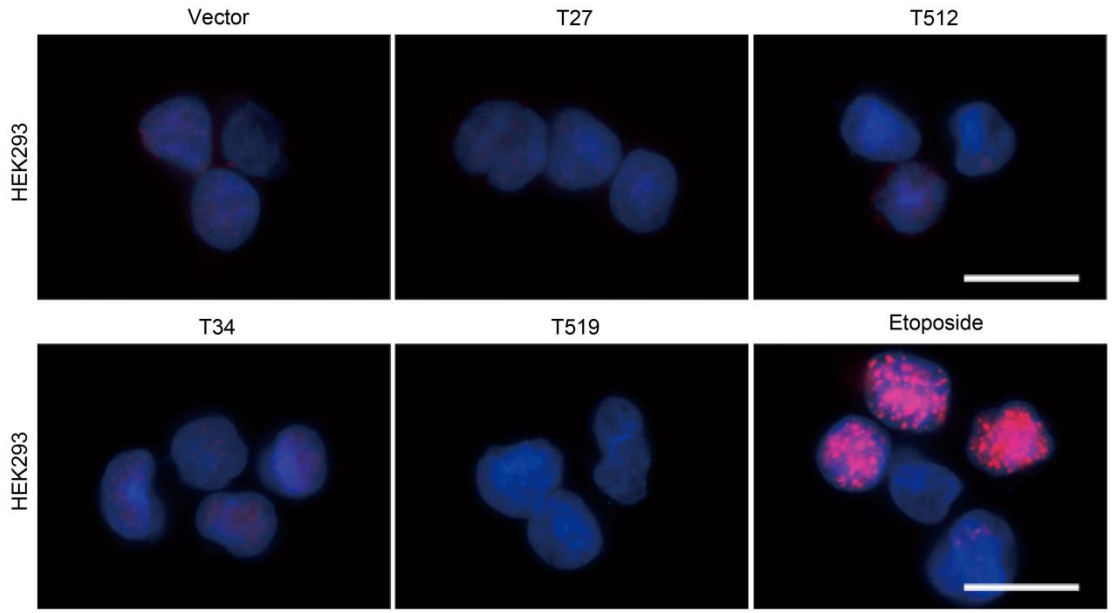
Supplemental Figure 4 The average HPV copies number in cells treated with TALENs detected by FISH. Average HPV16 copies per cell in SiHa (**A**) and S12 (**B**) after treatment with T27 and T512 and in HeLa (**C**) after treated with T34 and T519 were presented as box plot. *, $p < 0.01$.

Supplemental Figure 5



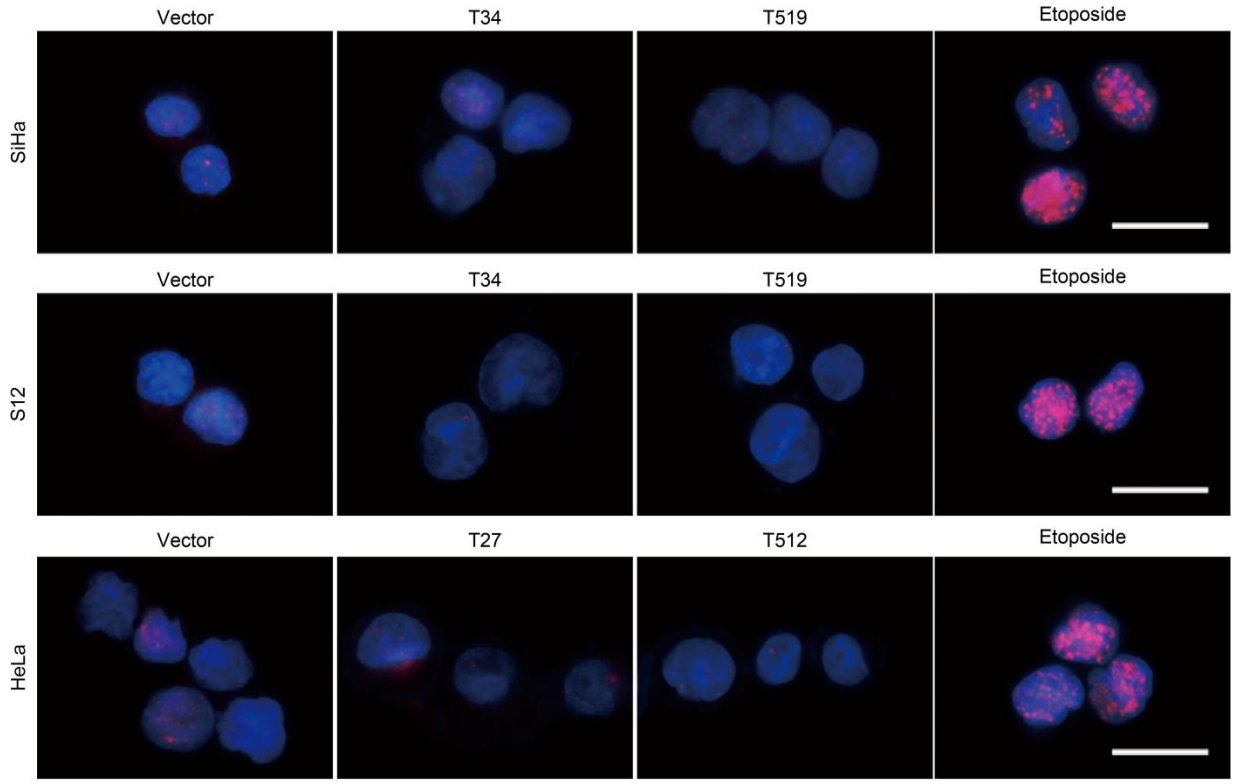
Supplemental Figure 5 The average γ-H2AX foci per nuclei in cells treated with TALENs detected by FISH. Averageγ-H2AX foci per nuclei SiHa (**A**) and S12 (**B**) after treatment with T27 and T512 and in HeLa (**C**) after treated with T34 and T519 were presented as box plot. *, $p < 0.01$.

Supplemental Figure 6



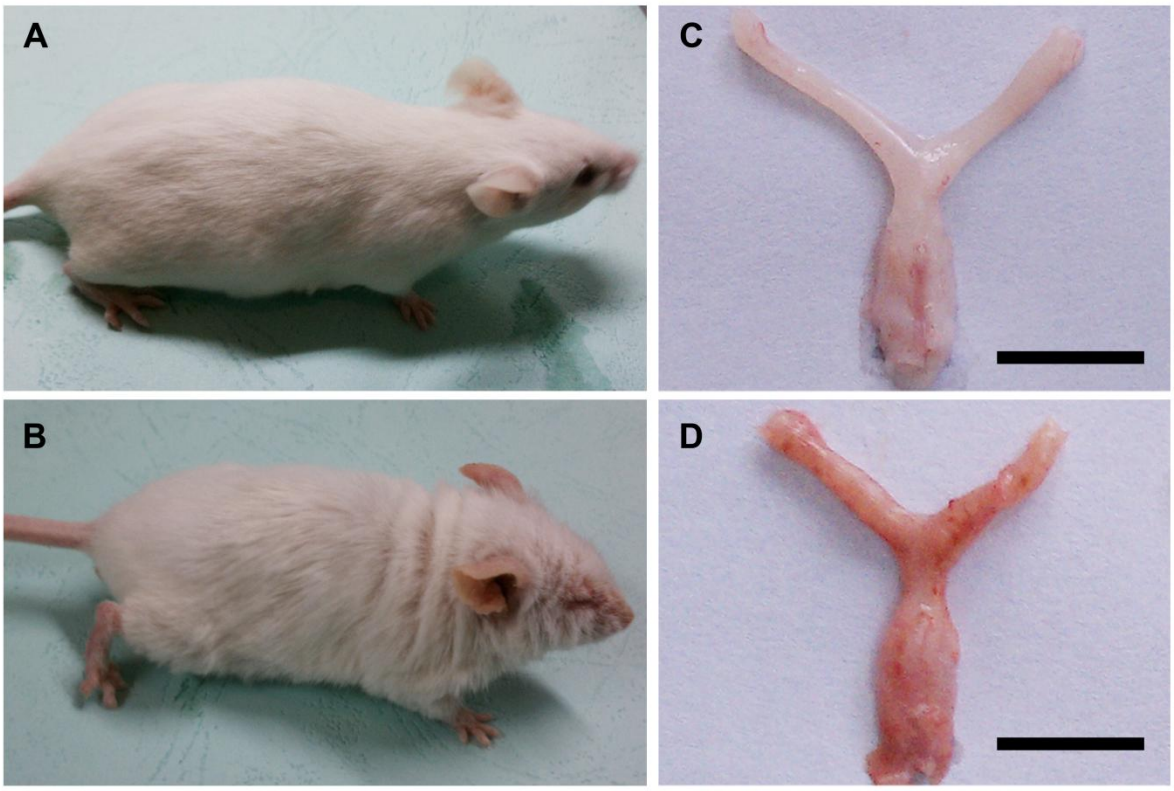
Supplemental Figure 6 Expression of γ-H2AX in HEK293cells treated with TALENs. The representative images of γ-H2AX (red) in Vector-, T27-, T512, T34- and T519-treated HEK293 cells were shown. Cell nuclei were indicated by DAPI staining (blue). These experiments were performed triplicate. HEK293 cells treated with Etoposide were used as positive control. Scale bars, 20 μm.

Supplemental Figure 7



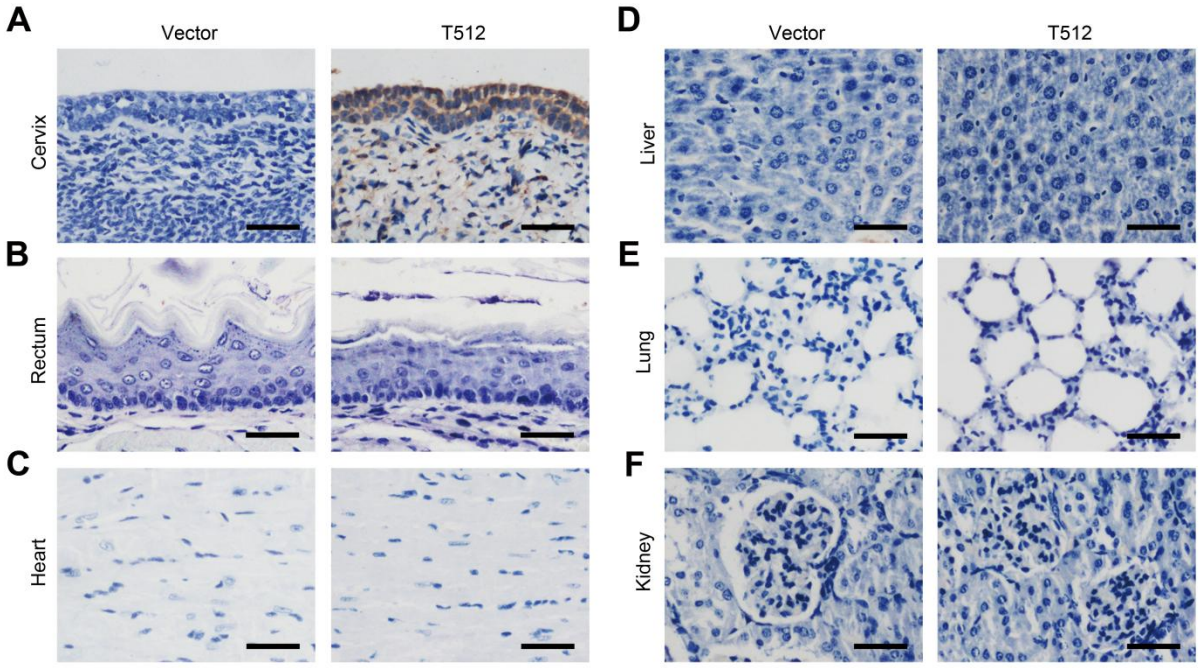
Supplemental Figure 7 Expression of γ-H2AX in cells treated with TALENs. The representative images of γ-H2AX (red) in Vector-, T34- and T519-treated SiHa and S12 cells and Vector-, T27- and T512-treated HeLa cells were shown. Cell nuclei were indicated by DAPI staining (blue). These experiments were performed triplicate. Cells treated with Etoposide were used as positive control. Scale bars, 20 μm.

Supplemental Figure 8



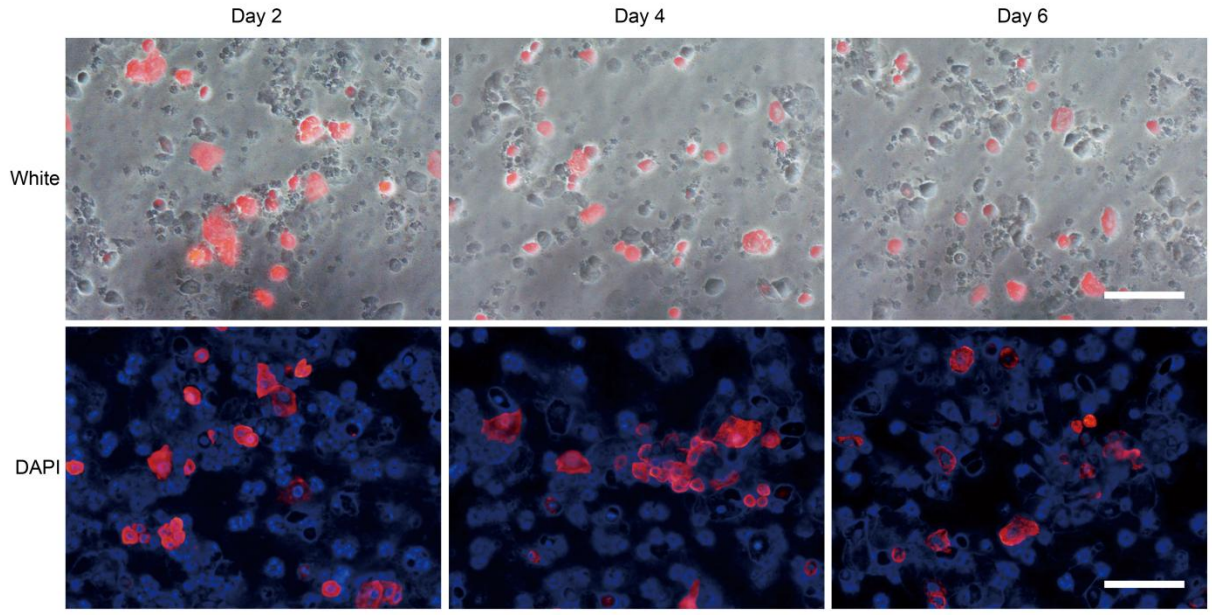
Supplemental Figure 8 Representative photographs of wild-type and K14-HPV16 transgenic mice and their cervicovaginal tissues. (**A**) wild-type mouse; (**B**) the cervicovaginal tissue of wild-type mice; (**C**) K14-HPV16 transgenic mouse; (**D**) the representative cervicovaginal tissue of K14-HPV16 transgenic mice. Scale bars, 1 cm.

Supplemental Figure 9



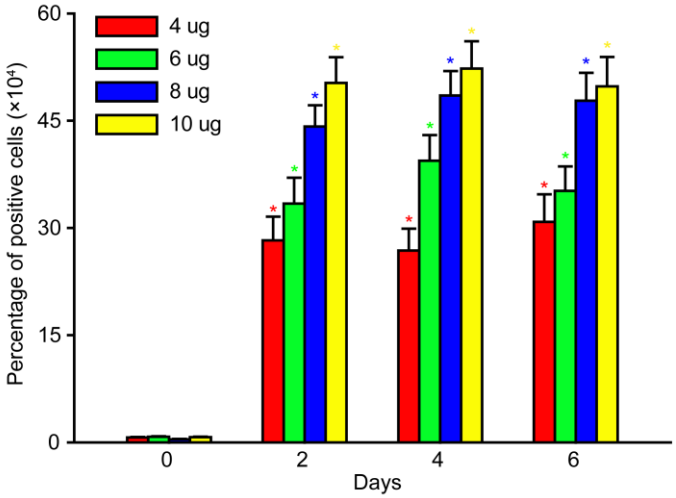
Supplemental Figure 9 IHC staining of FLAG-tagged T512 in cervixs and distal organs. (**A**) cervix, (**B**) rectum, (**C**) heart, (**D**) liver, (**E**) lung and (**F**) kidney. Scale bars, 20 μm.

Supplemental Figure 10



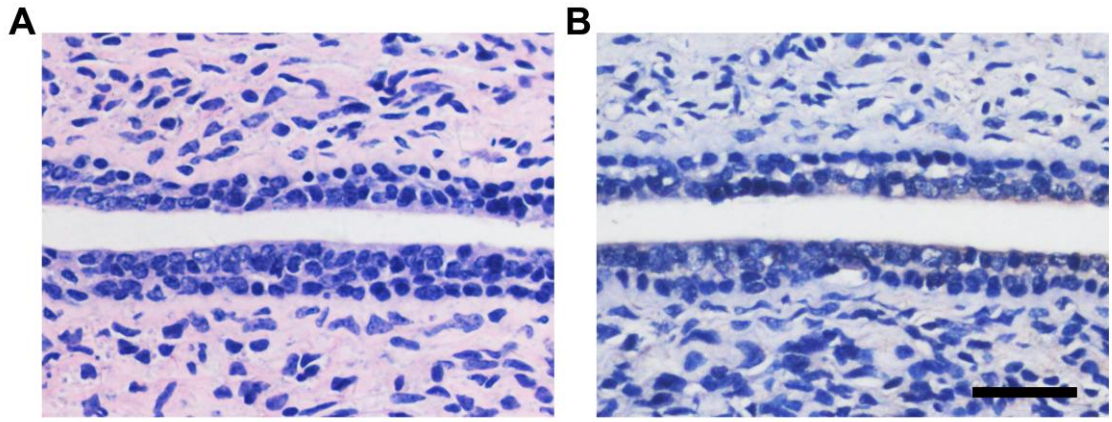
Supplemental Figure 10 The representative images of mRFP fluorescence in exfoliated cervical cells at days 2, 4 and 6 after local transfection . Ten micrograms of the mRFP expression plasmids that were incubated with the corresponding volumes of polymer were transfected intravaginally. The exfoliated cervical cells from the treated mice were gathered and smeared. The DAPI staining was done after fixed by 4% Paraformaldehyde at room temperature. Scale bars, 100 μm.

Supplemental Figure 11



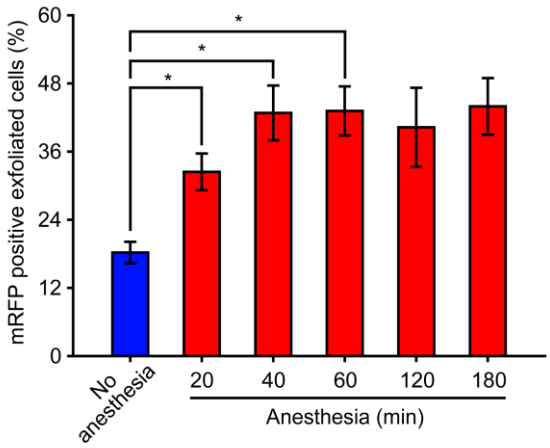
Supplemental Figure 11 Optimization of the transfection efficiency through the use of a range of DNA-to-polymer ratios. The indicated micrograms of the mRFP expression plasmids that were incubated with 1.2 µL of polymer were transfected intravaginally. At days 2, 4 and 6, exfoliated cervical cells from the treated mice (similar to a Pap smear test) were gathered, and the mRFP-positive cells were counted. $n = 3$ for each group of treated mice. *, $p < 0.01$ compared to control (day 0).

Supplemental Figure 12



Supplemental Figure 12 H&E and IHC staining of mice which were treated with T512 at 3-day intervals for a total of 24 days and kept for 2 months. (**A**) H&E; (**B**) IHC staining. Scale bar, 20 μm.

Supplemental Figure 13



Supplemental Figure 13 Improvement of intravaginal transfection efficiency under anesthesia. Ten micrograms of the mRFP expression plasmids were incubated with 1.2 μL of polymer and transfected intravaginally without anesthesia or under anesthesia for the indicated times, mRFP-positive exfoliated cervical cells were counted 48 h later (n = 3 for each group of treated mice).

Supplemental Note 1. Amino acid sequence of *Fok*I variants：

WT:

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNSTQ DRILEMKVME FFMKVYGYRG
KHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMQRYVE ENQTRNKHIN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHITN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

Sharkey:

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNPTQ DRILEMKVME FFMKVYGYRG
EHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMQRYVE ENQTRNKHIN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHITN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

KK:

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNSTQ DRILEMKVME FFMKVYGYRG
KHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMQRYVK ENQTRNKHIN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHKTN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

EL:

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNSTQ DRILEMKVME FFMKVYGYRG
KHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMERYVE ENQTRNKHLN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHITN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

S+KK (S is short for Sharkey):

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNPTQ DRILEMKVME FFMKVYGYRG
EHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMQRYVK ENQTRNKHIN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHKTN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

S+EL (S is short for Sharkey):

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNPTQ DRILEMKVME FFMKVYGYRG
EHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMERYVE ENQTRNKHLN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHITN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

KKR:

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNSTQ DRILEMKVME FFMKVYGYRG
KHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMQRYVK ENQTRNKHIN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNRKTN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

ELD:

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNSTQ DRILEMKVME FFMKVYGYRG
KHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMERYVE ENQTRDKHLN

PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHITN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

S+KKR (S is short for Sharkey):

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNPTQ DRILEMKVME FFMKVYGYRG
EHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMQRYVK ENQTRNKHIN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNRKTN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF

S+ELD (S is short for Sharkey):

QLVKSEL EEKKSELRHK LKYVPHEYIE LIEIARNPTQ DRILEMKVME FFMKVYGYRG
EHLGGSRKPD GAIYTVGSPI DYGVIVDTKA YSGGYNLPIG QADEMERYVE ENQTRDKHLN
PNEWWKVYPS SVTEFKFLFV SGHFKGNYKA QLTRLNHITN CNGAVLSVEE LLIGGEMIKA
GTLTLEEVRR KFNNGEINF