YMTHE, Volume *31*

Supplemental Information

Engineering of efficiency-enhanced Cas9 and base editors with improved gene therapy efficacies

Shuming Yin, Mei Zhang, Yang Liu, Xiaoyue Sun, Yuting Guan, Xi Chen, Lei Yang, Yanan Huo, Jing Yang, Xiaohui Zhang, Honghui Han, Jiqin Zhang, Min-Min Xiao, Mingyao Liu, Jiazhi Hu, Liren Wang, and Dali Li

A

Name	Length (aa)	Sources
HMGB1	215	*Homo sapiens*
HMGB2	209	*Homo sapiens*
HMGB3	200	*Homo sapiens*
HMGN1	100	*Homo sapiens*
HMGN2	90	*Homo sapiens*
HMGI	107	*Homo sapiens*
HMGI-C	118	*Homo sapiens*
HMGY	96	*Homo sapiens*
Sso7d	64	*Sulfolobus solfataricus P2*
Sac7d	66	*Sulfolobus acidocaldarius*
HMG-D	112	*Drosophila melanogaster*

B

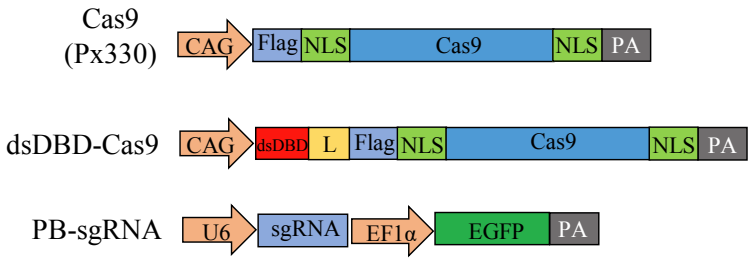


Figure S1. The information of double-strand DNA binding domains and plasmid designs. (A) The list of non-sequence-specific double-strand DNA binding domains (dsDBDs); **(B)** The architectures of Cas9, dsDBD-Cas9 and PB-sgRNA. NLS, nuclear localization signal. PA, polyadenylation signal. L, linker.

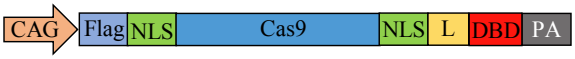
A

Linker	Amino acid sequence
L1	GGSGGSGGS
L2	GGSGGSGGSGGSGGSGGS
L3	SGSETPGTSESATPES
L4	SGGSSGGSSGSETPGTSESATPESSGGSSGGS
L5	SGGSSGGSSGSETPGTSESATPESSGSETPGT SESATPESSGGSSGGS

B

dsDBD-Cas9 fusions:

Cas9-dsDBD fusions:



C

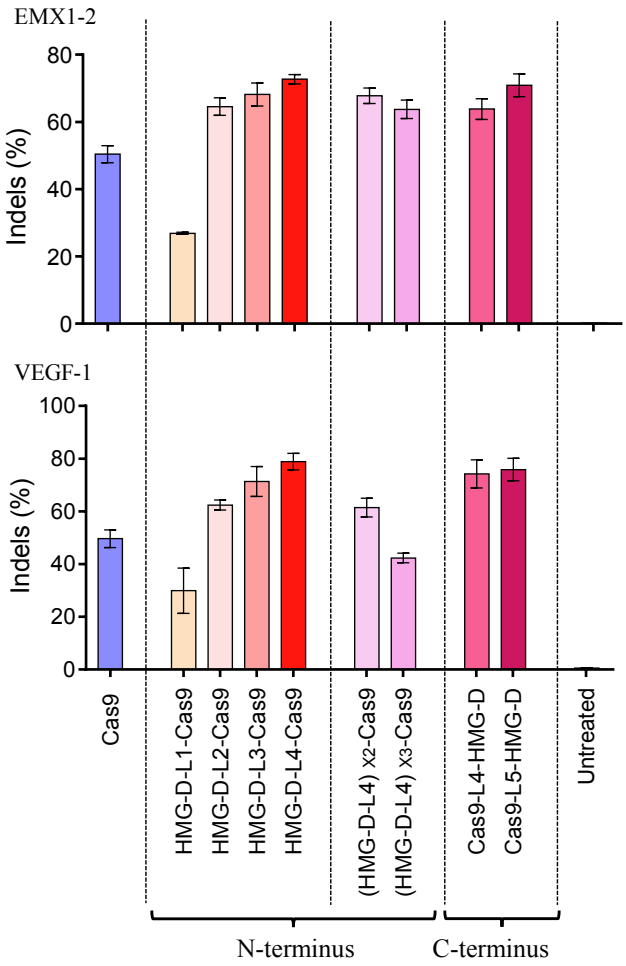
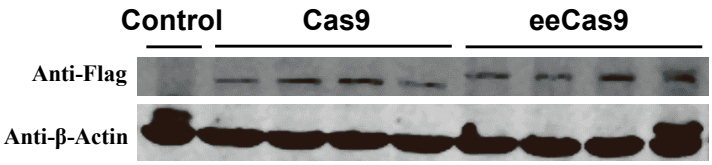


Figure S2. Optimization of the fusion of Cas9 and HMG-D architecture. (A) Sequences of linkers L1~L5. **(B)** The architectures of N-terminal or C-terminal fusion of HMG-D. NLS, nuclear localization signal. PA, polyadenylation signal. L, linker. **(C)** Quantification of the editing efficiency at EMX1 and VEGF-1 in HEK293T cells. Data represent means±s.d. (*n*=3 independent experiments).

A



Relative β-Actin ratio	
Control	0
Cas9	1.00
eeCas9	0.92

B

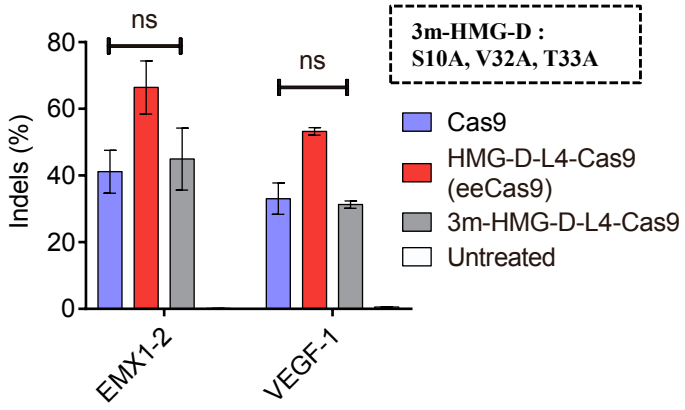


Figure S3. DNA binding capability of HMG-D is critical to its enhancement effect. (A) Top, protein expression analysis of Cas9 and eeCas9 by Western Blot. Full length Cas9 and eeCas9 protein were probed with anti-Flag antibody. Bottom, quantification of relative expression levels of Cas9 and eeCas9. The protein levels of Cas9 and eeCas9 was normalized to that of β-Actin in each sample and the average values were calculated. **(B)** Performance comparison of Cas9, eeCas9, 3m-HMG-D-L4-Cas9 at EMX1-2 and VEGF-1 by HTS analysis. Data represent means ± s.d. (*n*=3 independent experiments). ns, not significant.

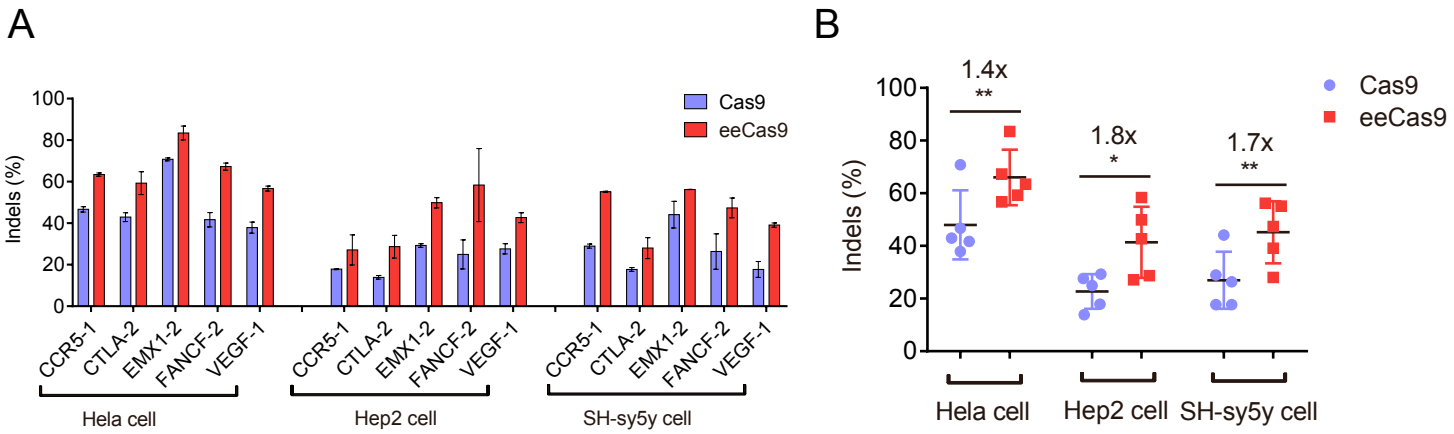


Figure S4. Evaluation of the gene editing efficiency of Cas9 and eeCas9 in other cell type by HTS analysis. (A) Five endogenous targets (CCR5-1, CTLA-2, EMX1-2, FANCF-2 and VEGF-1) were tested in Hela, Hep2, and SH-sy5y cells, respectively. **(B)** Summary of the average gene editing efficiency of Cas9 and eeCas9 at five endogenous targets in fig. A. In all graphs, data represent means ± s.d. (n=3 independent experiments). *$p < 0.05$. *$p < 0.01$.

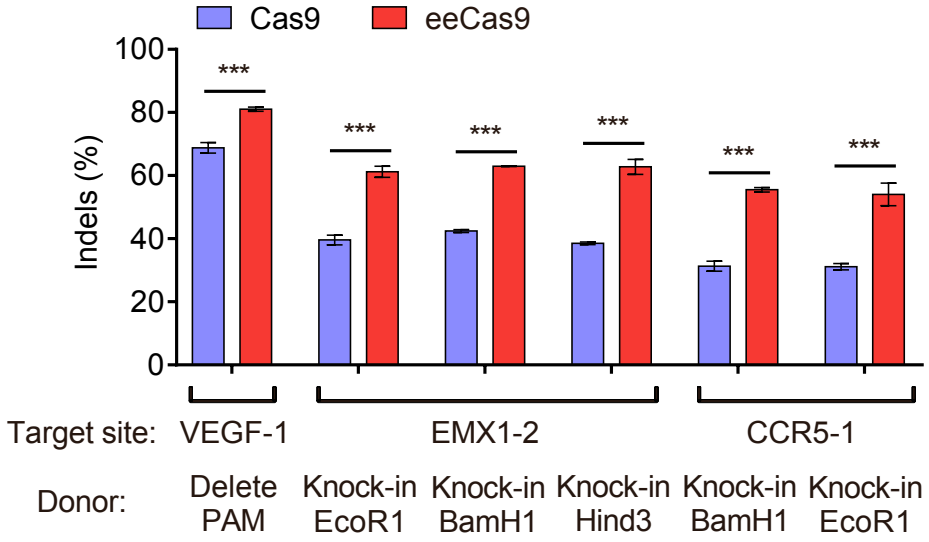


Figure S5. The indels frequency caused by Cas9 and eeCas9 when evaluation of HDR efficiency. Data represent means ± s.d. (n=3 independent experiments). ***$p < 0.001$.

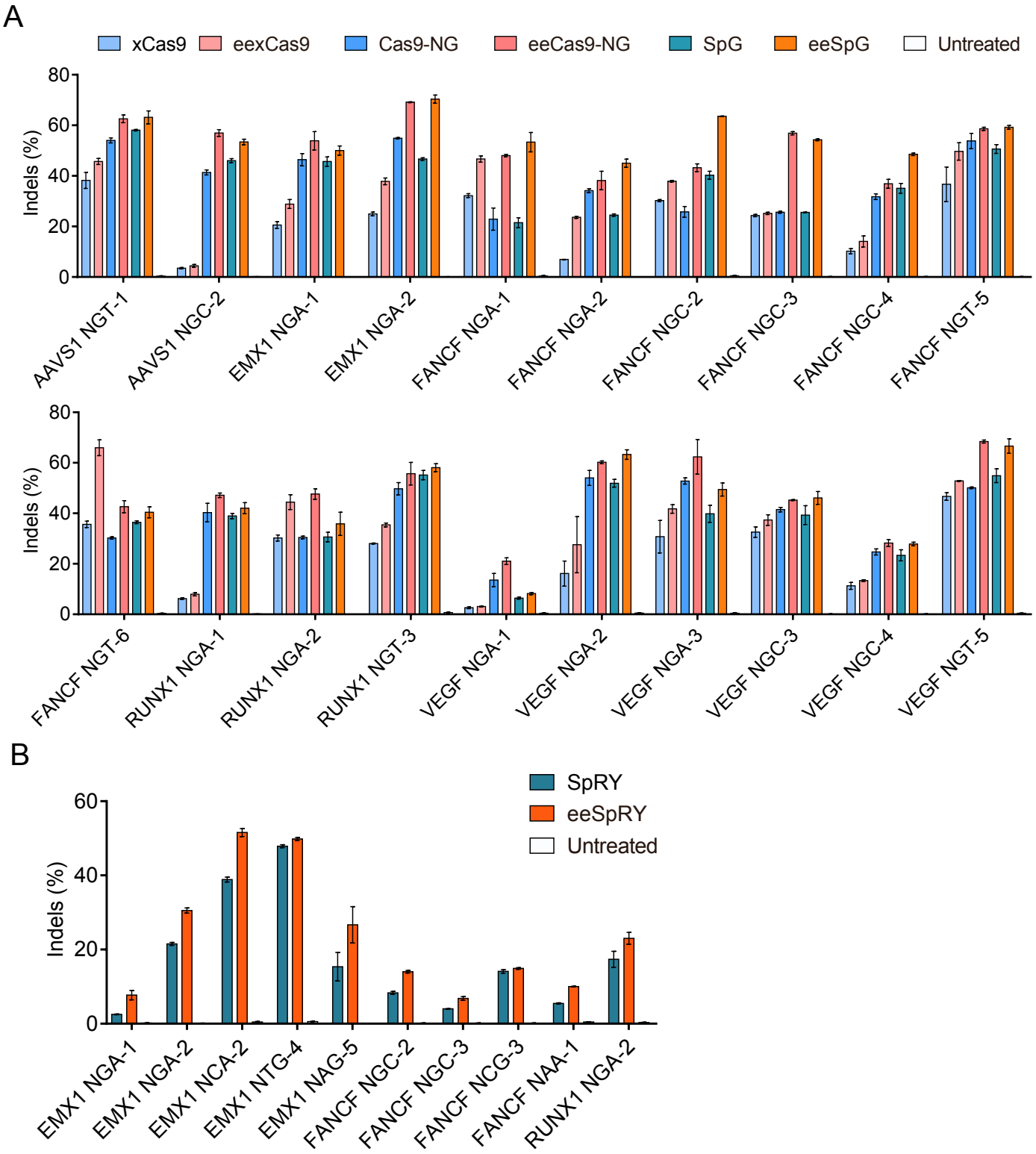


Figure S6. Evaluation of the gene editing efficiency of various Cas9 variants and eeCas9 variants at endogenous targets for non-NGG PAM. (A) Evaluation of the gene editing efficiency of xCas9, Cas9-NG, SpG and eexCas9, eeCas9-NG and eeSpG at 20 endogenous targets for non-NGG PAM in HEK293T cells. **(B)** Evaluation of the gene editing efficiency of SpRY and eeSpRY at 10 endogenous targets in HEK293T cells. In all graphs, data represent means ± s.d. (*n*=3 independent experiments).

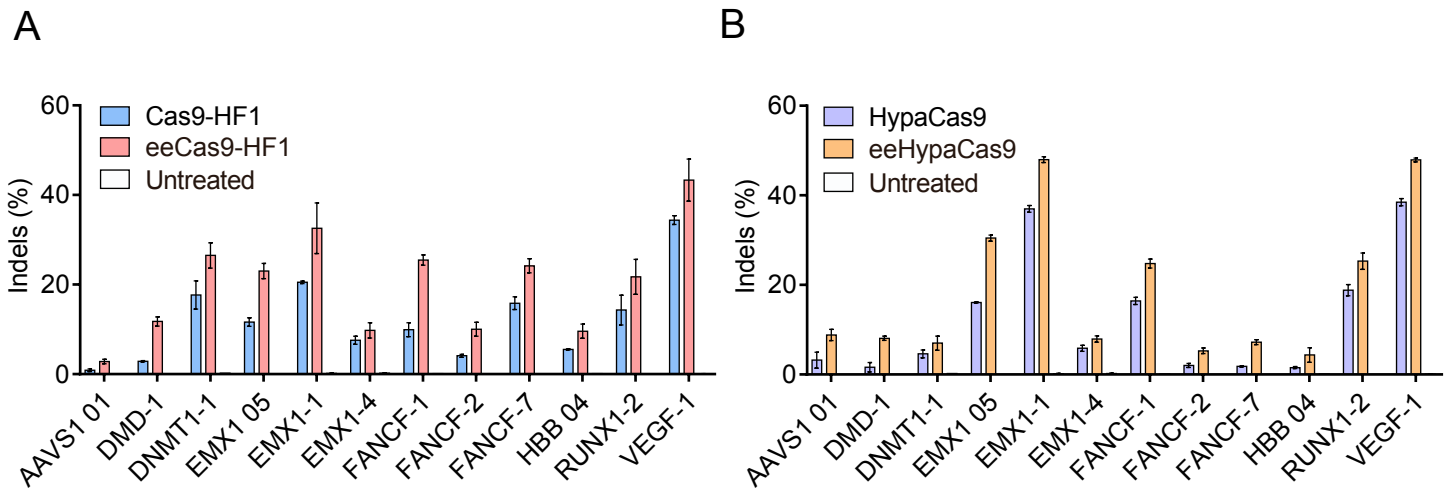


Figure S7. Evaluation of the gene editing efficiency of two high-fidelity Cas9s with or without HMG-D in HEK293T cells. (A) Evaluation of the gene editing efficiency of Cas9-HF1 and eeCas9-HF1 at 12 endogenous targets in HEK293T cells. (**B**) Evaluation of the gene editing efficiency of HypaCas9 and eeHypaCas9 at 12 endogenous targets in HEK293T cells. In all graphs, data represent means ± s.d. (*n*=3 independent experiments).

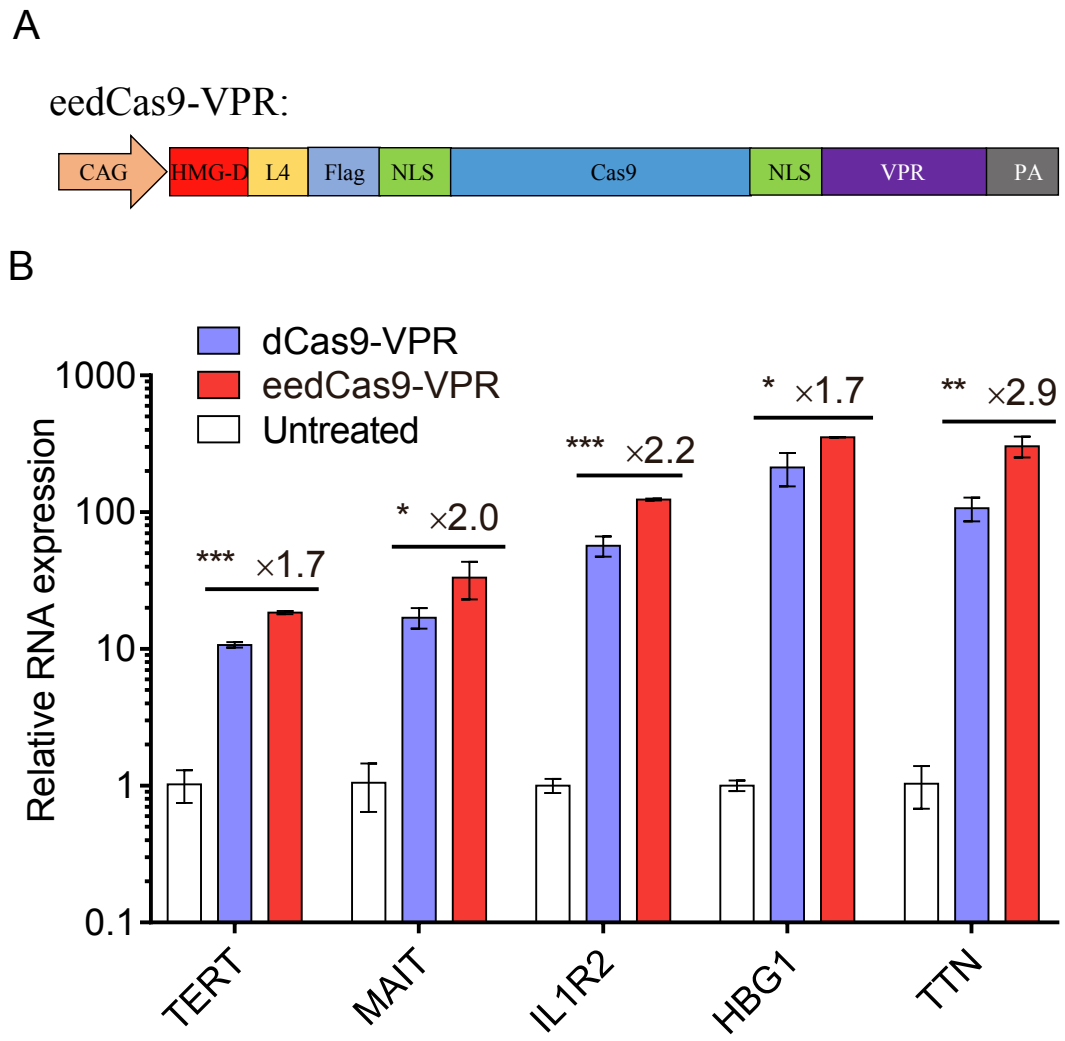


Figure S8. Development of eedCas9-VPR and evaluation of its gene activation capability in HEK293T cells. (A) Schematic view of eedCas9-VPR. VPR contains VP64, p65 and RTA. **(B)** Comparison of dCas9-VPR and eedCas9-VPR at five endogenous targets in 293T cells. The number above the bars represent ratio of eedCas9-VPR vs. dCas9-VPR mediated gene activation. All values were normalized to β-actin mRNA expression level. Data represent means ± s.d. (n=3 independent experiments). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

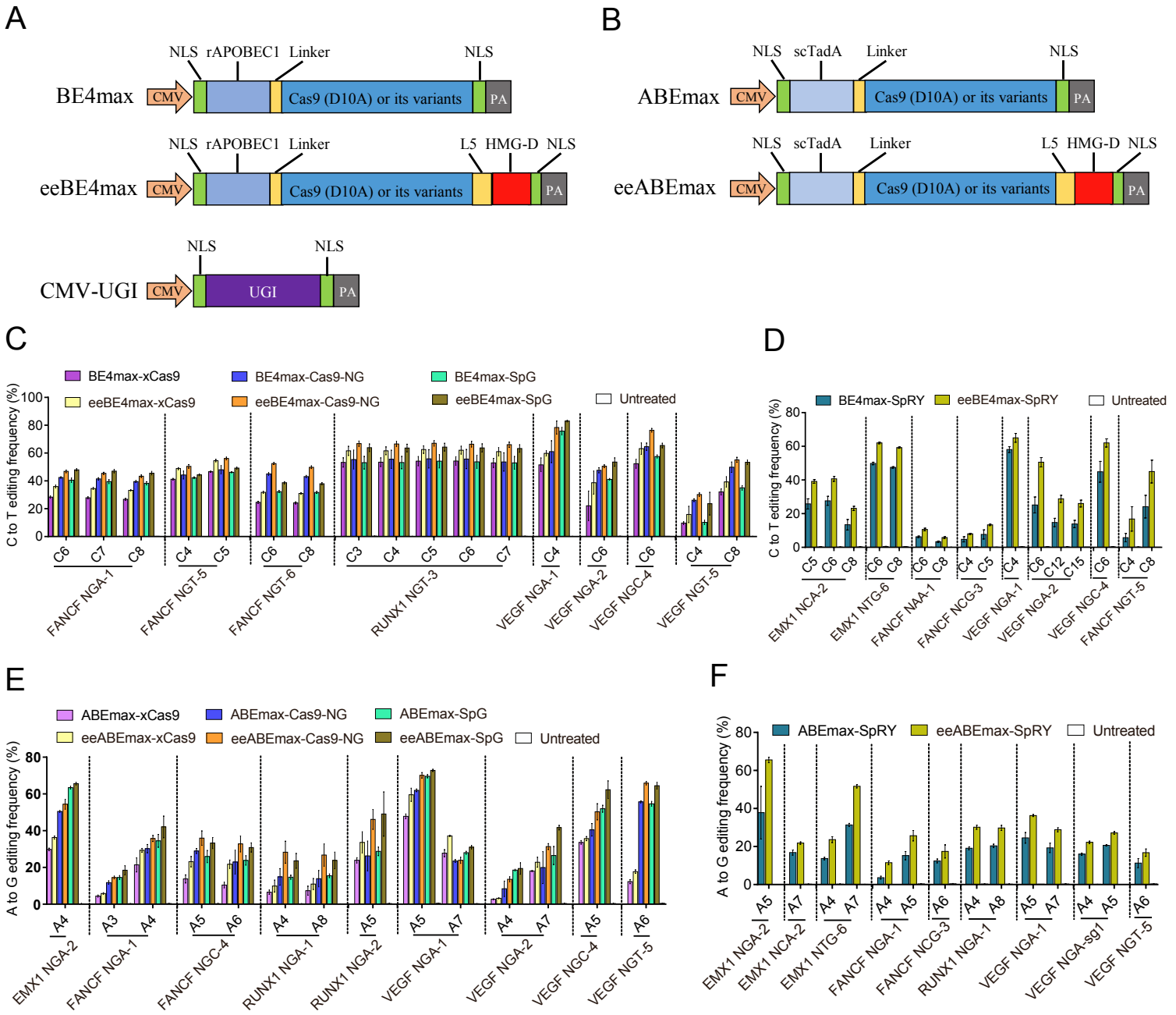


Figure S9. Evaluation of the base editing efficiency of eeBE4max variants and eeABEmax variants at multiple endogenous target sites. (A) Architectures of BE4max, eeBE4max and CMV-UGI. **(B)** Architectures of ABEmax and eeABEmax. **(C-D)** Comparison of C-to-T base editing efficiency between BE4max variants and eeBE4max variants at endogenous targets for non-NGG PAM. (**E-F**) Comparison of A-to-G base editing efficiency between ABEmax variants and eeABEmax variants at endogenous targets for non-NGG PAM. Data in graphs C-F represent means ± s.d. (*n*=3 independent experiments).

Fig. S10

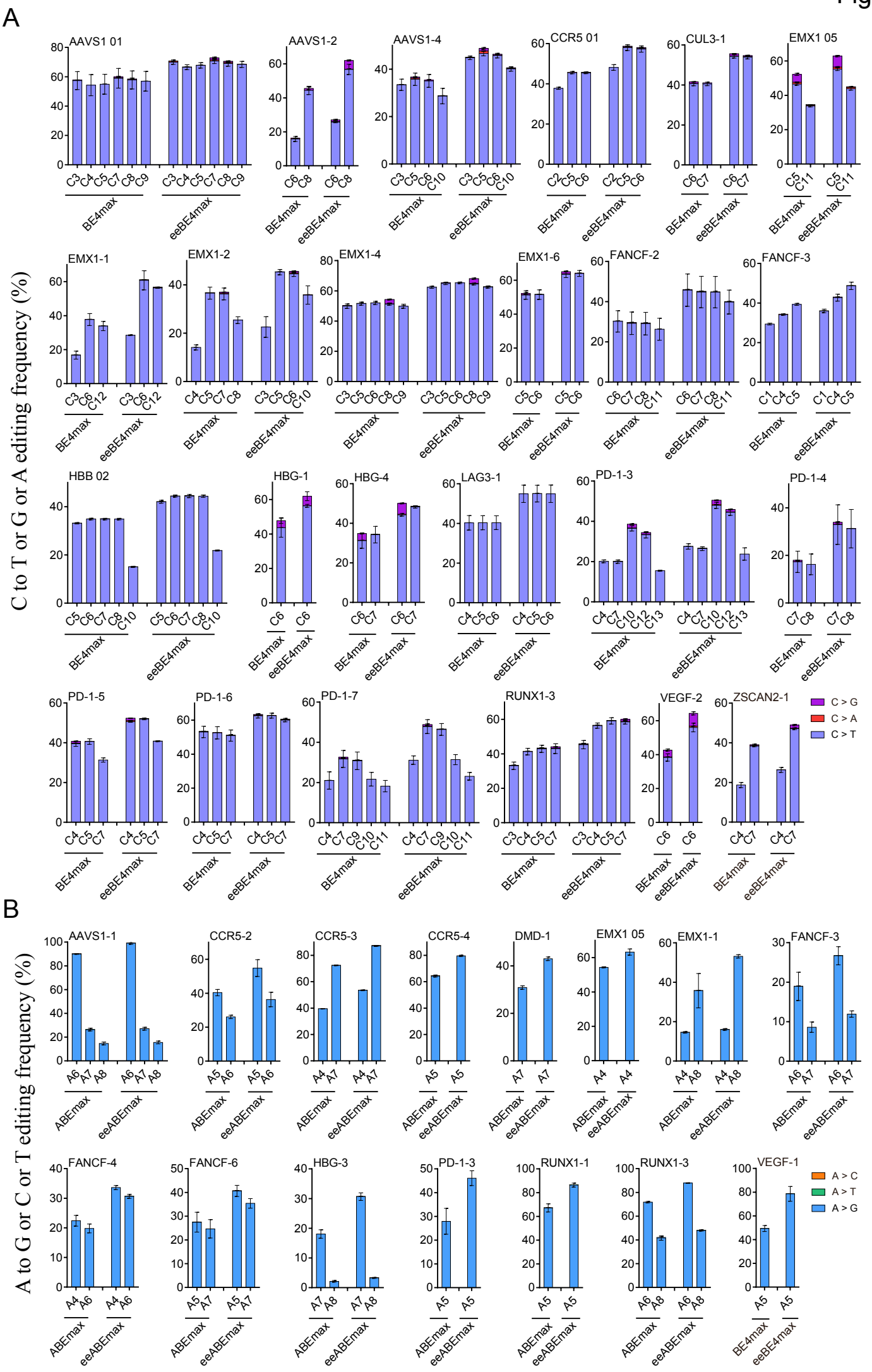

A
C to T or G or A editing frequency (%)
AAVS1 01
AAVS1-2
AAVS1-4
CCR5 01
CUL3-1
EMX1 05
EMX1-1
EMX1-2
EMX1-4
EMX1-6
FANCF-2
FANCF-3
HBB 02
HBG-1
HBG-4
LAG3-1
PD-1-3
PD-1-4
PD-1-5
PD-1-6
PD-1-7
RUNX1-3
VEGF-2
ZSCAN2-1
C > G
C > A
C > T
BE4max
eeBE4max
B
A to G or C or T editing frequency (%)
AAVS1-1
CCR5-2
CCR5-3
CCR5-4
DMD-1
EMX1 05
EMX1-1
FANCF-3
FANCF-4
FANCF-6
HBG-3
PD-1-3
RUNX1-1
RUNX1-3
VEGF-1
A > C
A > T
A > G
ABEmax
eeABEmax
BE4max
eeBE4max

Figure S10. Comparison of the product purity of multiple base editors. (A) Comparison of the base editing product purity of BE4max and eeBE4max at endogenous targets in 293T cells. C-to-T, C-to-G and C-to-A conversion frequency were analyzed by HTS, respectively. **(B)** Comparison of the base editing product purity of ABEmax and eeABEmax at endogenous targets in 293T cells. A-to-G, A-to-C and A-to-T conversion frequency were analyzed by HTS, respectively. In all graphs, data represent means±s.d. (*n*=3 independent experiments).

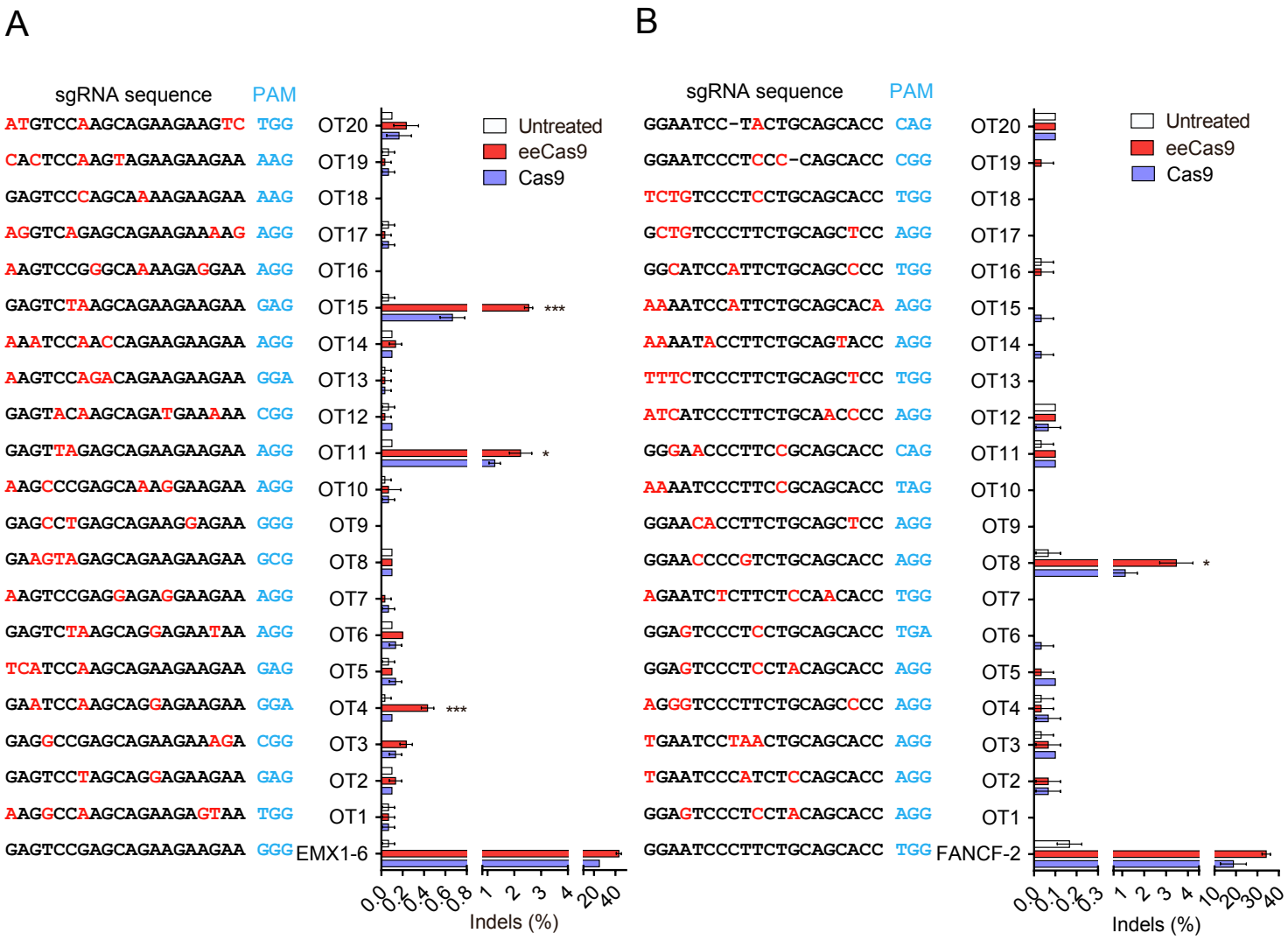


Figure S11. Evaluation of Cas9 and eeCas9 specificity at predicted off-target sites in 293T cells. On- and off-targets analysis by Cas9 or eeCas9 at EMX1-6 site **(A)** or FANCF-2 site **(B).** Mismatched nucleotides are shown in read. PAM sequences are shown in blue. *$p < 0.05$, **$p < 0.01$. In all graphs, data represent means ± s.d. (n=3 independent experiments).

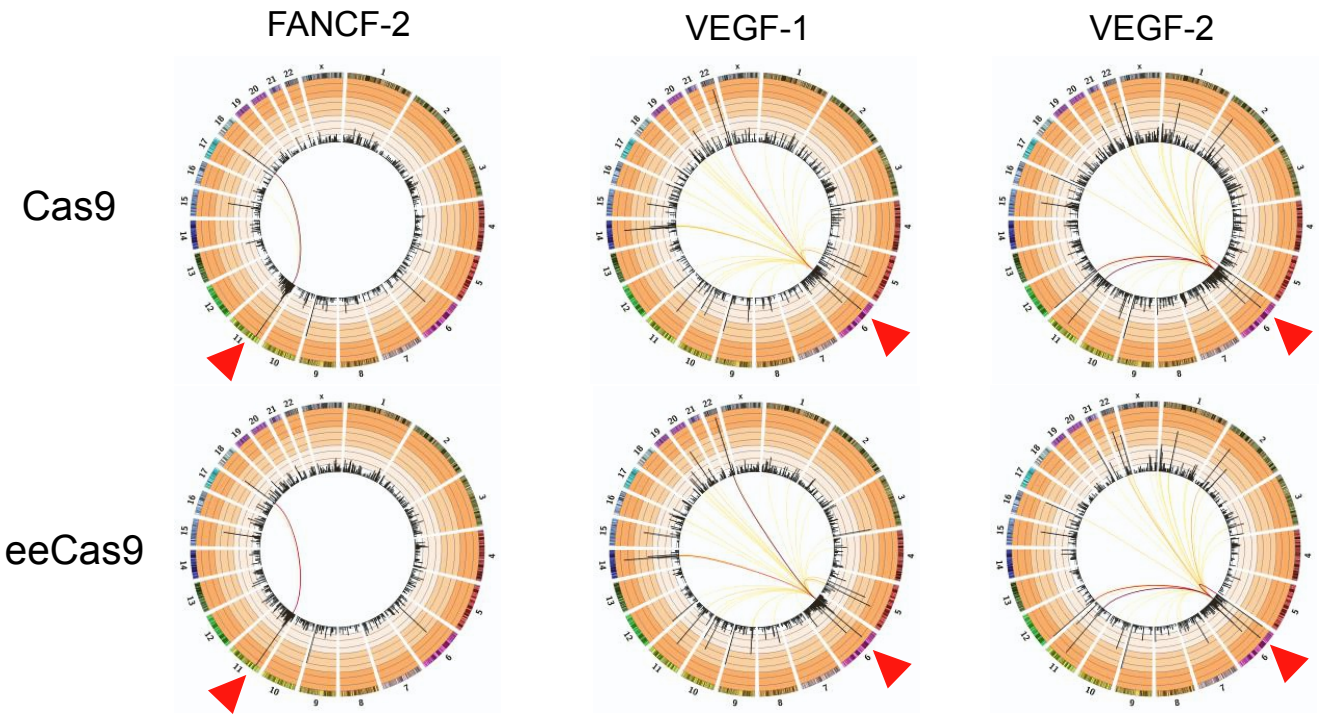


Figure S12. Evaluation of Cas9 and eeCas9 specificity by PEM-seq analysis. Circos plots show the off-targets sites of Cas9 and eeCas9 for FANCF-2, VEGF-1 and VEGF-2 target. Red arrows indicated the cleavage site of the indicated targets. Colored lines connected the on-target site to off-target hotspots.

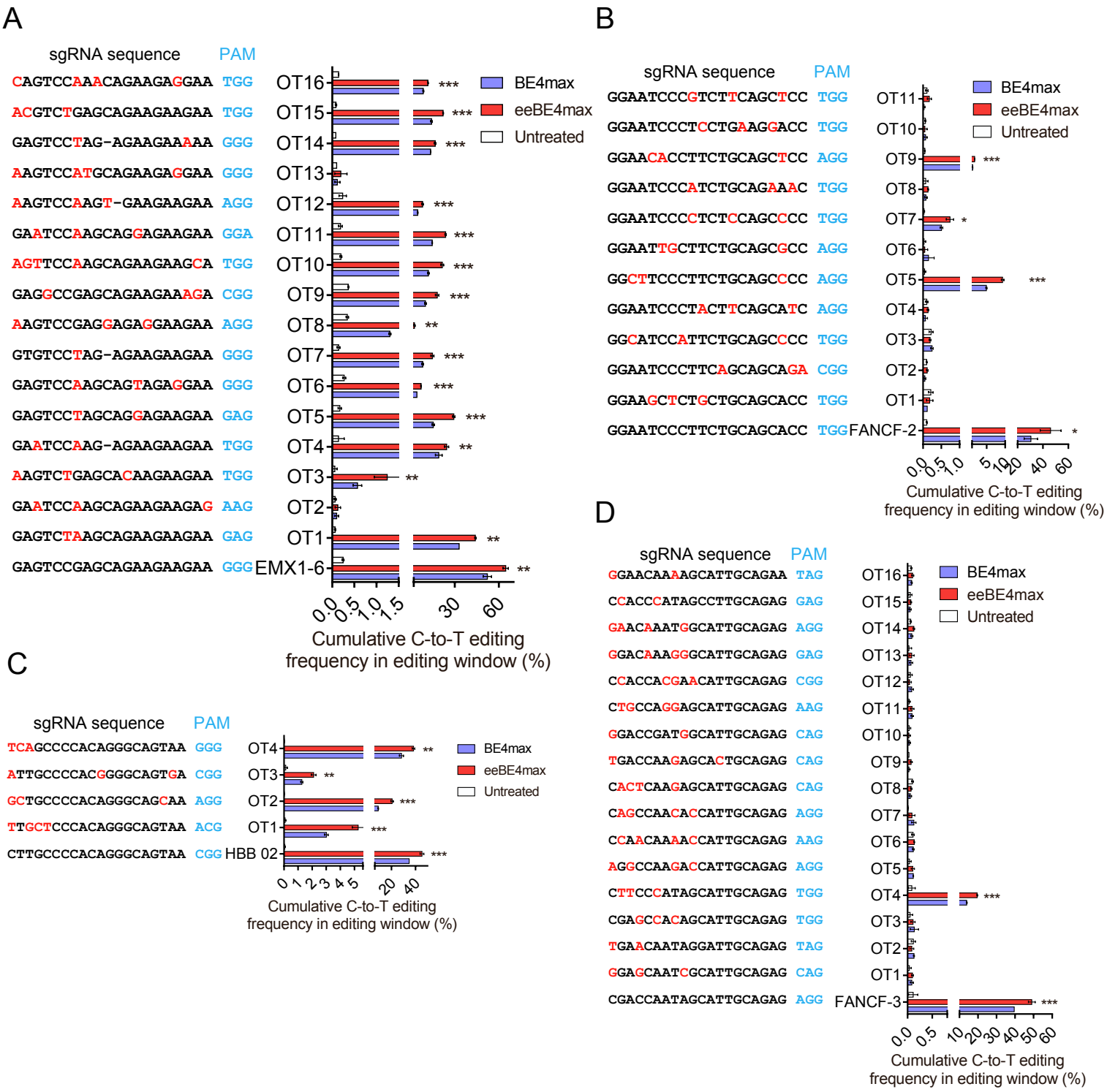


Figure S13. Evaluation of the off-target efficiency of eeBE4max and BE4max by HTS analysis in 293T cells. Comparing the off-target efficiency of eeBE4max and BE4max at EMX1-6 **(A)**, FANCF-2 **(B)**, HBB 02 **(C)** and FANCF-3 **(D)** target sites. FANCF-3 was web-predicted (https://www.benchling.com/) off-target site, the others were previously reported off-target sites. Mismatched nucleotides in off-targets sequence are shown in read. PAM sequences are shown in blue. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. In all graphs, data represent means±s.d. (n=3 independent experiments).

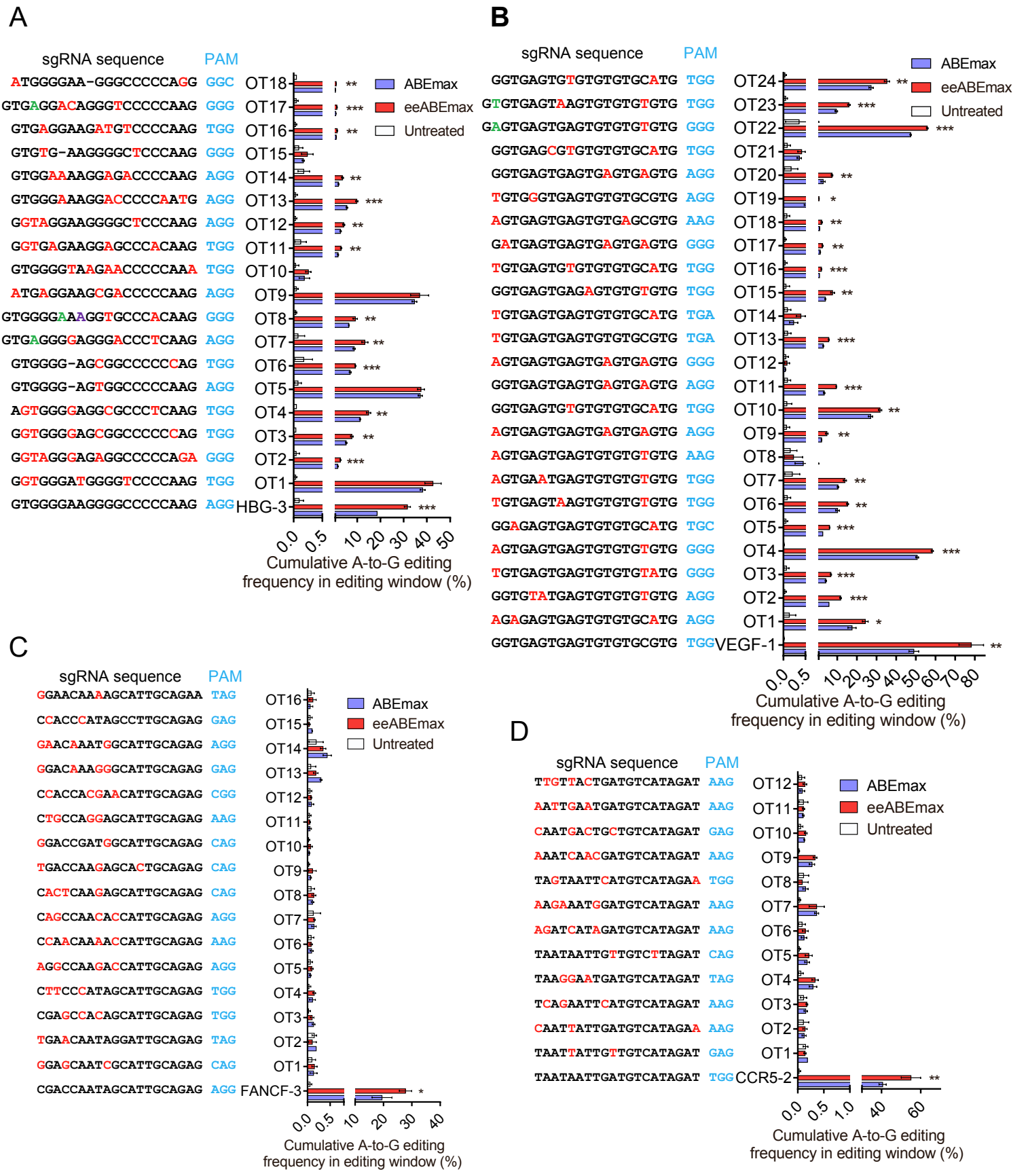


Figure S14. Evaluation of the off-target efficiency of eeABEmax and ABEmax by HTS analysis in 293T cells. Comparing the off-target efficiency of eeABEmax and ABEmax at HBG-3 **(A)**, VEGF-1 **(B)**, FANCF-3 **(C)** and CCR5-2 **(D)** target sites. CCR5-2 and FANCF-3 were web-predicted (https://www.benchling.com/) off-target site, HBG-3 and VEGF-1 were previously reported off-target sites. Mismatched nucleotides in off-targets sequence are shown in red. PAM sequence was shown in blue. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. In all graphs, data represent means±s.d. (n=3 independent experiments).

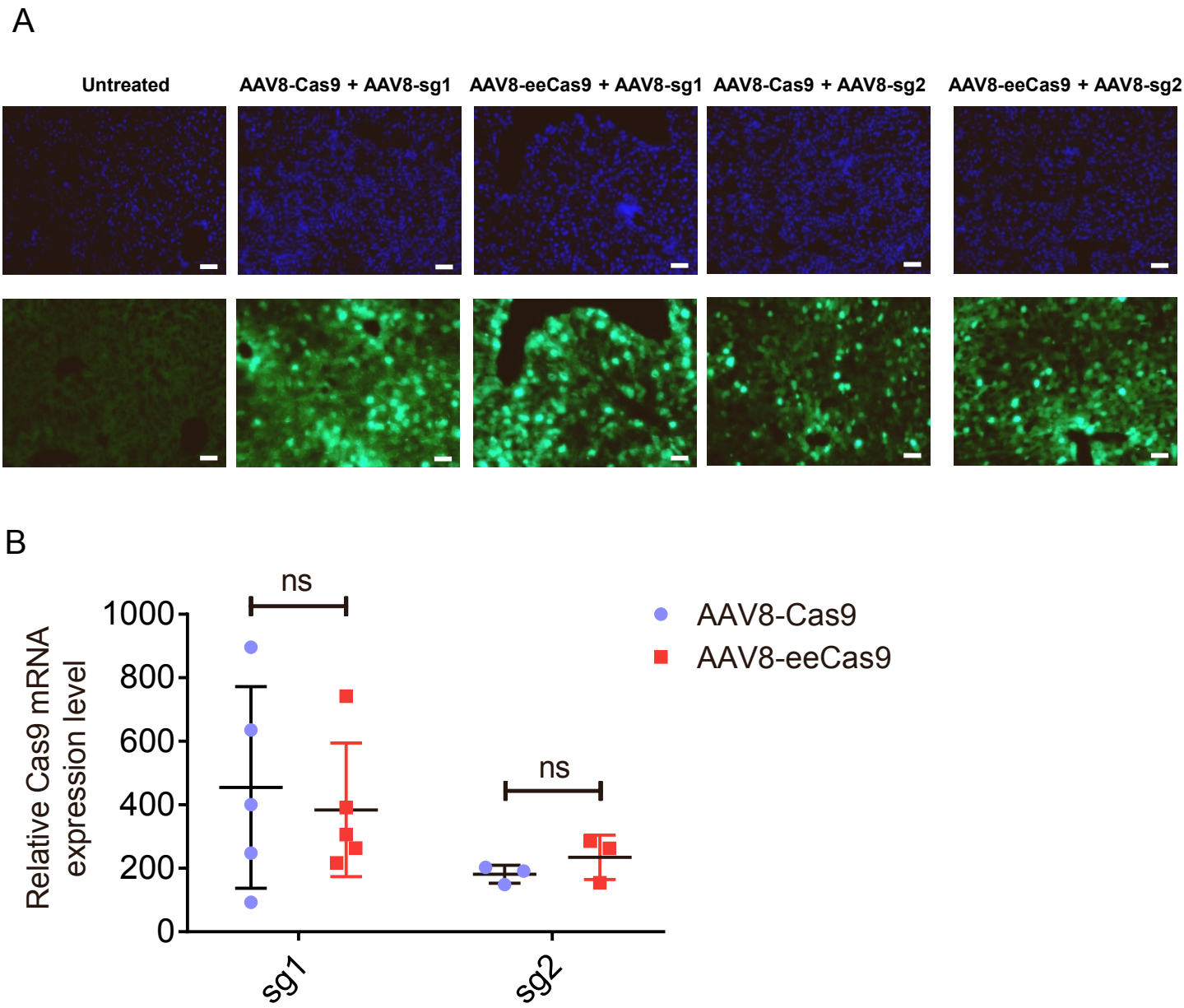


Figure S15. Evaluation of the viral transduction efficiency and Cas9 expression level of mice liver. (A) Representative immunofluorescence images from mice liver transduced with AAV. Scale bar: 100μm. **(B)** Comparing the expression levels of eeCas9 and Cas9 in liver by qPCR. Data represent means ± s.d. (*n*=3 independent experiments). ns, not significant.

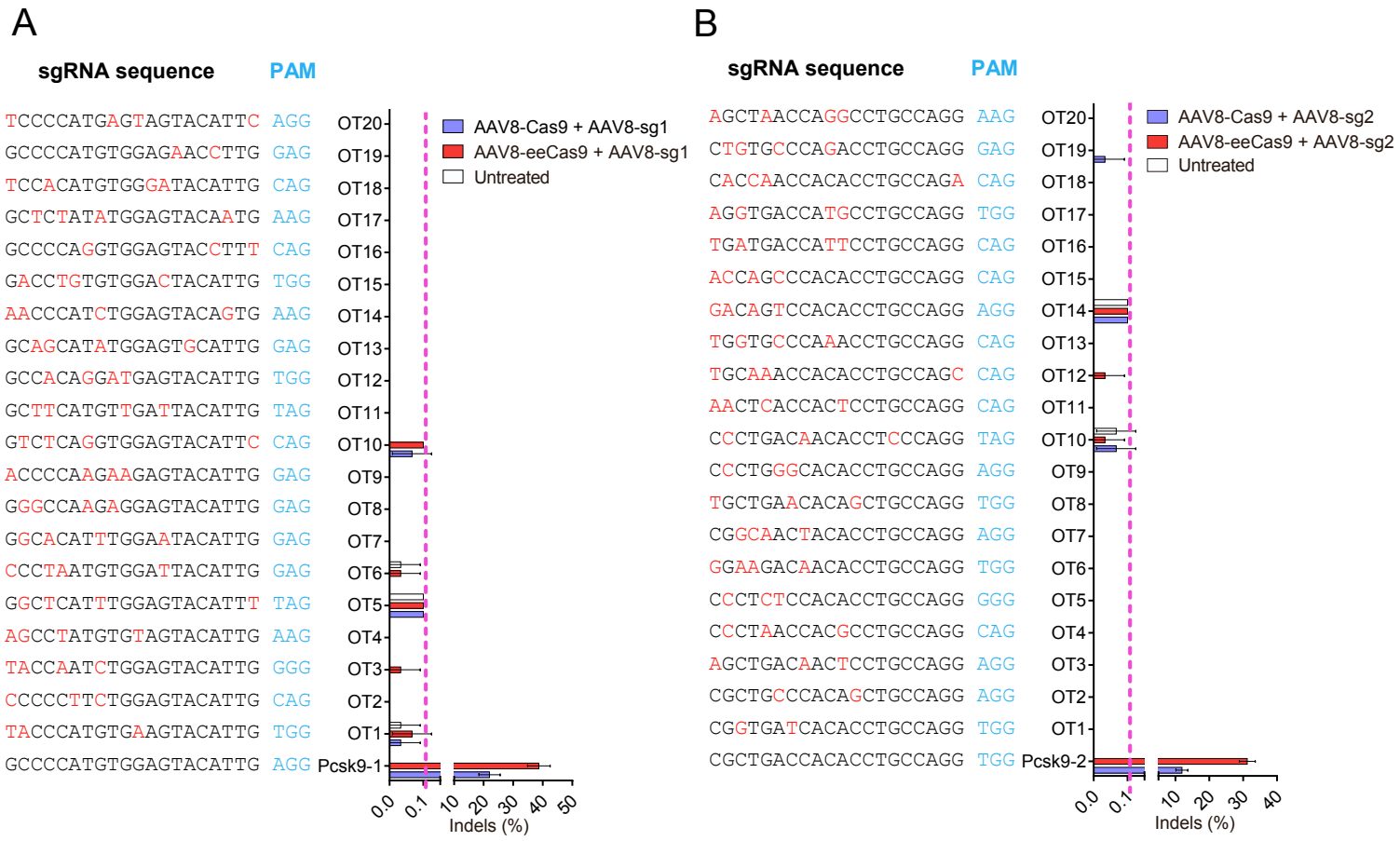
A

Mice (ID)	Treatment	Serum cholesterol(mg dL^{-1})		Indels (%)
		0 d	35 d	
1-5#	AAV8-Cas9 + AAV8-sg1	280.2	147.9	21.5%
1-4#		282.9	185.0	19.8%
3-6#		260.6	176.8	18.7%
2-11#		219.1	153.4	24.2%
3-5#		345.4	161.6	25.8%
1-1#		231.2	136.9	21.4%
1-2#	AAV8-eeCas9 + AAV8-sg1	249.9	95.9	41.3%
1-3#		261.0	125.1	38.8%
1-6%		216.4	126.0	34.3%
3-7#		258.2	90.4	40.5%
2-8#		166.6	115.0	35.7%
2-15#		338.8	120.5	39.8%
3-1#	Untreated	331.8	299.2	0.1%
3-2#		336.9	282.1	0.0%
3-3#		268.4	323.2	0.0%
2-2#		213.6	249.2	0.2%
2-3#		261.1	350.9	0.1%

B

Mice (ID)	Treatment	Serum cholesterol(mg dL^{-1})		Indels (%)
		0 d	35 d	
1#	AAV8-Cas9 + AAV8-sg2	387.1	208.3	13.1%
2#		259.6	195.2	13.5%
8#		256.0	221.5	10.2%
7#	AAV8-eeCas9 + AAV8-sg2	356.1	97.0	29.2%
1-2#		245.3	64.6	33.8%
2-4#		289.5	142.5	31.6%
WT-1#	Untreated	259.0	255.0	0.0%
WT-2#		365.3	309.0	0.1%
WT-3#		324.1	259.8	0.0%

Figure S16. Evaluation of the gene editing efficiency and therapeutic efficacy for targeting Pcsk9-1 or Pcsk9-2 site of *Pcsk9* gene in adult mice. (**A**) Summary of the serum cholesterol levels and the indels frequencies of each mice for Pcsk9-1 site. (**B**) Summary of the serum cholesterol levels and the indels frequencies of each mice for Pcsk9-2 site. Indel frequencies in graphs A and B were determined by HTS.



C

	Control	Pcsk9 1-1#	Pcsk9 1-2#
Raw SNVs	28656	26488	26711
SNVs filter	7891	7280	7436
Sample-specific SNVs (Remove control SNVs)	0	1551	1047
Candidate SNVs (off-targets)	0	0	0

D

	Contro l	Pcsk9 1-1#	Pcsk9 1-2#
Raw Indels	1005	600	647
Indels filter	835	564	614
Sample-specific Indels (Remove control Indels)	0	117	129
Candidate Indels (off-targets)	0	0	0

Figure S17. Analysis of the off-target effects for mice treated with Cas9 or eeCas9 virus. (A-B) DNA on-target and off-target efficiency of Pcsk9-1 **(A)** and Pcsk9-2 **(B)** target site was determined by HTS. The PAM sequences are shown in blue. The mismatched nucleotides are shown in red. In all graphs, data represent means±s.d. (*n*=3 independent experiments). **(C-D)** Summary of the number of indels and SNVs detected by PEM-seq.

Supplementary Table 1. Amino acid sequences of double-stranded DNA binding domains (dsDBDs).

Name	Amino acid sequence	Origin
HMGB1	MGKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAK EKGKFEDMAKADKARYEREMKTYIPPKGETKKKFKDPNAPKRPPSAFFLFCSEYRP KIKGEHPGLSIGDVAKKLGEMWNNTAADDKQPYEKKAAKLKEKYEKDIAAYRAKG KPDAAKKGVVKAEKSKKKKEEEEDEEDEEDEEEEEDEEDEDEEEDDDDE	*Homo sapiens*
HMGB2	MGKGDPNKPRGKMSSYAFFVQTCREEHKKKHPDSSVNFAEFSKKCSERWKTMSAK EKSKFEDMAKSDKARYDREMKNYVPPKGDKKGKKKDPNAPKRPPSAFFLFCSEHR PKIKSEHPGLSIGDTAKKLGEMWSEQSAKDKQPYEQKAAKLKEKYEKDIAAYRAKG KSEAGKKGPGRPTGSKKKNEPEDEEEEEEEEDEDEEEEDEDEE	*Homo sapiens*
HMGB3	MAKGDPKKPKGKMSAYAFFVQTCREEHKKKNPEVPVNFAEFSKKCSERWKTMSGK EKSKFDEMAKADKVRYDREMKDYGPAKGGKKKKDPNAPKRPPSGFFLFCSEFRPKI KSTNPGISIGDVAKKLGEMWNNLNDSEKQPYITKAAKLKEKYEKDVADYKSKGKFD GAKGPAKVARKKVEEEDEEEEEEEEEEEEEDE	*Homo sapiens*
HMGN1	MPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD	*Homo sapiens*
HMGN2	MPKRKAEGDAKGDKAKVKDEPQRRSARLSAKPAPPKPEPKPKKAPAKKGEKVPKG KKGKADAGKEGNNPAENGDAKTDQAQKAEGAGDAK	*Homo sapiens*
HMGI	MSESSSKSSQPLASKQEKDGTEKRGRGRPRKQPPVSPGTALVGSQKEPSEVPTPKRPR GRPKGSKNKGAAKTRKTTTTPGRKPRGRPKKLEKEEEEGISQESSEEEQ	*Homo sapiens*
HMGI-C	MSARGEGAGQPSTSAQGQPAAPAPQKRGRGRPRKQQQEPTGEPSPKRPRGRPKGSK NKSPSKAAQKKAEATGEKRPRGRPRKWPQQVVQKKPAQVNVALPGKDHPGNLIYL LFSKNAT	*Homo sapiens*
HMGY	MSESSSKSSQPLASKQEKDGTEKRGRGRPRKQPPKEPSEVPTPKRPRGRPKGSKNKG AAKTRKTTTTPGRKPRGRPKKLEKEEEEGISQESSEEEQ	*Homo sapiens*
Sso7d	MATVKFKYKGEEKEVDISKIKKVWRVGKMISFTYDEGGGKTGRGAVSEKDAPKELL QMLEKQKK	*Sulfolobus sp.*
Sac7d	MVKVKFKYKGEEKEVDTSKIKKVWRVGKMVSFTYDDNGKTGRGAVSEKDAPKEL LDMLARAEREKK	*Sulfolobus sp.*
HMG-D	MSDKPKRPLSAYMLWLNSARESIKRENPGIKVTEVAKRGGELWRAMKDKSEWEAK AAKAKDDYDRAVKEFEANGGSSAANGGGAKKRAKPAKKVAKKSKKEESDEDDDD ESE	*Drosophila melanogaster*

Supplementary Table 2. List of on-target sites and HTS primer sequences.

See in Supplemental Spreadsheet.

Supplementary Table 3. List of synthetic ssODN sequences.

Name	Sequence (5'-3')
VEGF-1 delete PAM	G*T*GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGGGTTGAGGGCGTTGGAGCGGGGAGAAGGCC AGGGGTCA*C*T
EMX1-2 knock-in EcoR1	G*A*AGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCGAATTCACATCAACCGGTGGCGCATTGCCA CGAAGCAGGCCAAT*G*G
EMX1-2 knock-in BamH1	G*A*AGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCGGATCCACATCAACCGGTGGCGCATTGCCA CGAAGCAGGCCAAT*G*G
EMX1-2 knock-in Hind3	G*A*AGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCAAGCTTACATCAACCGGTGGCGCATTGCCA CGAAGCAGGCCAAT*G*G
CCR5-1 knock-in BamH1	T*A*TCAAGTGTCAAGTCCAATCTATGACATCAATTATTATAGGATCCCATCGGAGCCCTGCCAAAAAATCAAT GTGAAGCAAATC*G*C
CCR5-1 knock-in EcoR1	T*A*TCAAGTGTCAAGTCCAATCTATGACATCAATTATTATAGAATTCCATCGGAGCCCTGCCAAAAAATCAAT GTGAAGCAAATC*G*C

Supplementary Table 4. List of off-target sites and HTS primers.

See in Supplemental Spreadsheet.

Supplementary Table 5. List of qPCR primer sequences.

Name	Forward primer	Reverse primer
HBG1	AGATGCTGGAGGAGAAACCC	AGGTGCCCTTGAGATCATCC
IL1R2	AGCTTCTCTGGGGTCAAGACT	TCTCAACAGAAGACCCTGGC
MAIT	TGGCTGGGGTTTGAACCTTT	AGGAAGCTGTTCCAGACTGC
TERT	CAGAGCCAGTCTCACCTTCA	ACATGCGTGAAACCTGTACG
TTN	TGTTGCCACTGGTGCTAAAG	ACAGCAGTCTTCTCCGCTTC
Cas9	ACAATCTGACCAAGGCCGAG	TTCCGGAAATCGGACACCAG
Actin	CGTCATACTCCTGCTTGCTG	GTACGCCAACACAGTTGCTG
HBG	GGTTATCAATAAGCTCCTAGTCC	ACAACCAGGAGCCTTCCCA
HBA	GCCCTGGAGAGGATGTTC	TTCTTGCCGTGGCCCTTA

Supplementary Table 6. List of AAV vectors and sequences.

ITR **miniCMV** **Cas9 or eeCas9** **short PolyA**

AAV8-Cas9:

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAG

TGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGCGGCCGCTAGCGACTCACGGGGATTTCCAAGT
CTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGAC
GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCCGCGGCCACCGGTGCCA
CCATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAG
CGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCC
GTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATC
GGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAAC
CGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGG
TGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTA
CCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGG
GGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGC
TGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAA
ATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAA
GAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGAT
CGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAG
ATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGC
AGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGG
AAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCT
GCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGA
TTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGG
GAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTT
CCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGA
GTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAA
AAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCT
TCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAA
GGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAG
GAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTG
AGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAA
ACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCA
CGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGT
GATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGA
GAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA
CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGAT
GTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAG
AGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGA
AAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACC
CGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGA
AAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCAC
GCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACT
ACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACAT
CATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGA
GATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTG
CAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGA
AGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGA

GTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAA

AGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCC

GGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGG

GCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTT

CTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAG

GCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAG

GTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAG

CTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGTAAGTTTAAACGAATTCAATAAAAGAT

CTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTGCGGCCGTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTAGAGCTCGCGG

CCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGA

CGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

AAV8-eeCas9:

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAG

TGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGCGGCCGCTAGCGACTCACGGGGATTTCCAAGT

CTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGAC

GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCCGCGGCCACCGGTGCCA

CCATGAGCGACAAACCGAAACGGCCGCTTTCTGCTTACATGCTCTGGCTGAACTCAGCACGAGAGAGCATCAAGAGAGAAAATCCCGG

CATTAAAGTCACTGAGGTGGCTAAACGCGGTGGCGAACTTTGGAGGGCAATGAAAGATAAGTCCGAGTGGGAGGCCAAGGCAGCGAA

AGCGAAGGATGACTATGACCGCGCCGTAAAAGAGTTTGAGGCAAATGGGGGTAGCAGCGCGGCAAATGGAGGCGGCGCGAAGAAAAG

GGCAAAGCCTGCTAAGAAAGTTGCTAAGAAGTCTAAAAAAGAAGAATCAGATGAGGATGATGACGACGAGTCTGAGGCTAGCTCTGGC

GGATCTAGCGGTGGATCTAGCGGCTCTGAGACCCCTGGAACATCCGAATCCGCCACTCCAGAGAGCAGCGGAGGCTCTTCTGGAGGAT

CACTCGAGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAAAGACGATGACGATAAGATGGCCCCAAAGAAG

AAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGG

GCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTG

ATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAG

AACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCC

TGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCAT

CTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTC

CGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAAC

CAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTG

GAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACT

TCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCC

AGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC

CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTG

CGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGC

CAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGAC

CTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAG

GAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGC

CAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGG

CGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTG

TACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGC

AGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGA

GTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGG

ACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGAT

CGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAG

GCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAAC

AGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCC

TGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGA

AAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC

GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGC

AGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTA

CGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGC

AAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAG

AGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAA

ACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAG

TGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCAC

CACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCG

ACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAA

CATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGG

GAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGG

TGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAA

GAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAA

GAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTAC

AAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG

CCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAA

GGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAG

TTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGC

AGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAG

AGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTC

AGCTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGTAAGTTTAAACGAATTCAATAAAAG

ATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTGCGGCCGTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTAGAGCTCGC

GGCCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCC

GACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG