

In the format provided by the authors and unedited.

Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis

Changyang Zhou^{1,2,9}, Yidi Sun^{2,3,4,9}, Rui Yan^{5,9}, Yajing Liu^{2,6,9}, Erwei Zuo^{1,7,9}, Chan Gu⁵, Linxiao Han¹, Yu Wei¹, Xinde Hu^{1,2}, Rong Zeng^{3,6}, Yixue Li^{5,6,8*}, Haibo Zhou^{1*}, Fan Guo^{5*} & Hui Yang^{1*}

¹Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. ²College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China. ³CAS Key Laboratory of Systems Biology, CAS Center for Excellence in Molecular Cell Science, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. ⁴Bio-Med Big Data Center, Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China. ⁵Center for Translational Medicine, Ministry of Education Key Laboratory of Birth Defects and Related Diseases of Women and Children, Department of Obstetrics and Gynecology, West China Second University Hospital, College of Life Sciences, Sichuan University, Chengdu, China. ⁶School of Life Science and Technology, Shanghai Tech University, Shanghai, China. ⁷Center for Animal Genomics, Agricultural Genome Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China. ⁸Shanghai Jiao Tong University, Fudan University, Shanghai Academy of Science & Technology, Shanghai, China. ⁹These authors contributed equally: Changyang Zhou, Yidi Sun, Rui Yan, Yajing Liu, Erwei Zuo. *e-mail: yxli@sibs.ac.cn; hbzhou@ion.ac.cn; guofan@scu.edu.cn; huiyang@ion.ac.cn

Supplementary Information

Off-target RNA mutation induced by DNA base editing and its elimination by deaminase engineering

Changyang Zhou^{1,2,9}, Yidi Sun^{2,3,5,9}, Rui Yan^{4,9}, Yajing Liu^{2,7,9}, Erwei Zuo^{1,6,9}, Chan Gu⁴,
Linxiao Han¹, Yu Wei¹, Xinde Hu^{1,2}, Rong Zeng^{3,7}, Yixue Li^{5,7,8,10}, Haibo Zhou^{1,10}, Fan Guo^{4,10},
Hui Yang^{1,10}

Correspondence: yxli@sibs.ac.cn, hbzhou@ion.ac.cn, guofan@scu.edu.cn,
huiyang@ion.ac.cn,

Supplementary Table 1. Summary of the RNA-seq data.

Sample ID	Group	Raw Reads (M)	Raw Q30(%)
QP115-1	ABE7.10	36.19	93.73
QP115-2	ABE7.10	34.11	93.81
QP115-3	ABE7.10	29.77	93.49
6759-1	ABE7.10-site 1	37.84	93.65
6759-2	ABE7.10-site 1	38.17	93.77
6759-3	ABE7.10-site 1	39.72	93.46
6759-36-1	ABE7.10-site 1	49.43	94.78
6759-36-2	ABE7.10-site 1	39.61	94.67
6759-36-3	ABE7.10-site 1	46.64	94.45
6761-1	ABE7.10-site 2	35.89	93.8
6761-2	ABE7.10-site 2	44.58	93.7
6761-3	ABE7.10-site 2	38.28	93.92
1123-1	ABEmax-site 1	56.29	93.57
1123-2	ABEmax-site 1	64.76	93.01
1123-3	ABEmax-site 1	63.05	93.86
1116-1	ABE7.10 ^{D53E} -site 1	55	93.75
1116-2	ABE7.10 ^{D53E} -site 1	45.83	93.79
1116-3	ABE7.10 ^{D53E} -site 1	42.21	93.88
1117-1	ABE7.10 ^{F148A} -site 1	46.03	93.8
1117-2	ABE7.10 ^{F148A} -site 1	65.48	93.19
1117-3	ABE7.10 ^{F148A} -site 1	41.29	93.4
1149-1	ABE7.10 ^{F148A} -site 2	44.59	95.09
1149-2	ABE7.10 ^{F148A} -site 2	28.59	95.26
1149-3	ABE7.10 ^{F148A} -site 2	40.63	95.54
ZP-1112-1	APOBEC1	29.92	94.24
ZP-1112-2	APOBEC1	42.48	93.74
ZP-1112-3	APOBEC1	35	93.62
LP105-1	BE3	39.44	93.76
LP105-2	BE3	36.53	93.77
LP105-3	BE3	44.62	93.14
6757-1	BE3-RNF2	47.45	93.8
6757-2	BE3-RNF2	39.74	93.58
6757-3	BE3-RNF2	39.62	93.5
6755-1	BE3-site 3	50.31	93.49
6755-2	BE3-site 3	40.72	93.53
6755-36-1	BE3-site 3 (36h)	58.75	94.28
6755-36-2	BE3-site 3 (36h)	48.75	94.36
6755-36-3	BE3-site 3 (36h)	50.18	94.28
1124sg-1	BE3 (FNLS)-site 3	55.69	93.43
1124sg-2	BE3 (FNLS)-site 3	57.47	93.46
1124sg-3	BE3 (FNLS)-site 3	61.58	93.07
ZP-1115-1	BE3 (hA3A)-site 3	42.68	93.53
ZP-1115-2	BE3 (hA3A)-site 3	39.25	93.48
ZP-1115-3	BE3 (hA3A)-site 3	40.22	93.1
ZP-1114-1	BE3 ^{W90A} -site 3	50.88	93.82
ZP-1114-2	BE3 ^{W90A} -site 3	31.12	93.99
ZP-1114-3	BE3 ^{W90A} -site 3	34.86	93.82

WP1-1	BE3 ^{W90Y+R126E} -site 3	39.31	95.4
WP1-2	BE3 ^{W90Y+R126E} -site 3	41.69	95.22
1162-1	BE3 (hA3A ^{R128A})-site 3	44.58	94.33
1162-2	BE3 (hA3A ^{R128A})-site 3	54.74	94.14
1162-3	BE3 (hA3A ^{R128A})-site 3	54.23	94.06
WP2-1	BE3 (hA3A ^{Y130F})-site 3	56.29	93.57
WP2-2	BE3 (hA3A ^{Y130F})-site 3	64.76	93.01
WP2-3	BE3 (hA3A ^{Y130F})-site 3	63.05	93.86
XXP31-1	GFP	47.13	93.54
XXP31-2	GFP	40.91	93.86
XXP31-3	GFP	35.74	94.09
XXP31-new-1	GFP	48.35	93.49
XXP31-new-2	GFP	51.28	94.35
XXP31-new-3	GFP	34.12	93.47
ZP-1113-1	TadA-TadA*	36.29	93.64
ZP-1113-2	TadA-TadA*	47.11	93.86
ZP-1113-3	TadA-TadA*	42.22	94.15
WT-1	WT	57.23	93.18
WT-2	WT	40.34	93.32
WT-3	WT	33.53	93.76

Supplementary Table 2. The number of off-target RNA SNVs.

Group	No. of SNVs	Group	No. of SNVs
GFP-#1	592	BE3-RNF2-#1	6460
GFP-#2	594	BE3-RNF2-#2	4017
GFP-#3	410	BE3-RNF2-#3	5536
GFP-#4	961	TadA-TadA*-#1	6566
GFP-#5	1034	TadA-TadA*-#2	8273
GFP-#6	967	TadA-TadA*-#3	7444
APOBEC1-#1	24407	ABE7.10-#1	4035
APOBEC1-#2	35165	ABE7.10-#2	3035
APOBEC1-#3	27125	ABE7.10-#3	2836
BE3-#1	13260	ABE7.10-site 1-#1	3575
BE3-#2	12543	ABE7.10-site 1-#2	3499
BE3-#3	9525	ABE7.10-site 2-#1	5345
BE3-site 3-#1	3499	ABE7.10-site 2-#2	5733
BE3-site 3-#2	3856	ABE7.10-site 2-#3	4942

Supplementary Table 3. Precise p values between any two groups in Fig. 1, f, g.

Fig. 1f	GFP	APOBEC1	BE3	BE3-RNF2	BE3-site 3
APOBEC1	3.22E-06				
BE3	1.84E-06	0.02358958			
BE3-RNF2	3.50E-05	0.01484277	0.01353352		
BE3-site 3	1.23E-05	0.01577894	0.01750179	0.13839447	

Fig. 1g	GFP	TadA-TadA*	ABE7.10	ABE7.10-site 1	ABE7.10-site 2
TadA-TadA*	3.13E-07				
ABE7.10	5.53E-05	0.00336727			
ABE7.10-site 1	0.000182969	0.01523444	0.59182854		
ABE7.10-site 2	1.55E-07	0.03463096	0.01461902	0.01379675	

Two-sided unpaired t-test. $n = 3$ biologically independent samples for each group except for BE3-site 3 at site 3 locus ($n = 2$ biologically independent samples) and GFP ($n = 6$ biologically independent samples).

Supplementary Table 8. The number of off-target RNA SNVs for samples transfected with engineered BE3 and ABE7.10.

Group	No. of SNVs
BE3 (hA3A)-site 3-#1	2579
BE3 (hA3A)-site 3-#2	2479
BE3 (hA3A)-site 3-#3	2155
BE3 (hA3A ^{R128A})-site 3-#1	777
BE3 (hA3A ^{R128A})-site 3-#2	1094
BE3 (hA3A ^{R128A})-site 3-#3	943
BE3 (hA3A ^{Y130F})-site 3-#1	711
BE3 (hA3A ^{Y130F})-site 3-#2	747
BE3 (hA3A ^{Y130F})-site 3-#3	774
BE3 ^{W90A} -site 3-#1	602
BE3 ^{W90A} -site 3-#2	314
BE3 ^{W90A} -site 3-#3	295
BE3 ^{W90Y+R126E} -site 3-#1	1185
BE3 ^{W90Y+R126E} -site 3-#2	1121
ABE7.10 ^{D53E} -site 1-#1	3126
ABE7.10 ^{D53E} -site 1-#2	2897
ABE7.10 ^{D53E} -site 1-#3	2092
ABE7.10 ^{F148A} -site 1-#1	749
ABE7.10 ^{F148A} -site 1-#2	1434
ABE7.10 ^{F148A} -site 1-#3	733
ABE7.10 ^{F148A} -site 2-#1	888
ABE7.10 ^{F148A} -site 2-#2	536
ABE7.10 ^{F148A} -site 2-#3	640

Supplementary Table 9. P values between any two groups in Fig. 4, c, e.

Fig. 4c	GFP	BE3-site 3	BE3^{W90A} - site 3	BE3^{W90Y+} R126E -site 3	BE3 (hA3A)- site 3	BE3 (hA3A^{R128A}) -site 3	BE3 (hA3A^{Y130F}) -site 3
BE3-site 3	1.23E-05						
BE3^{W90A}-site 3	0.34421424	0.008576701					
BE3^{W90Y+R126E} - site 3	0.003645	0.01136	0.01145				
BE3 (hA3A)- site 3	4.64E-05	0.027699783	0.0003473 34	0.007478			
BE3 (hA3A^{R128A})- site 3	0.22	0.01393	0.01446	0.1334	0.001183		
BE3 (hA3A^{Y130F})- site 3	0.99	0.03728	0.06128	0.01468	0.005187	0.1638	

Fig. 4e	GFP	ABE7.10- site 1	ABE7.10^D 53E -site 1	ABE7.10^F 148A -site 1	ABE7.10 F148A -site 2
ABE7.10-site 1	0.00018297				
ABE7.10^{D53E}- site 1	0.00015104	0.115487084			
ABE7.10^{F148A}- site 1	0.42391495	0.006839321	0.0135897 37		
ABE7.10^{F148A}- site 2	0.2818	0.0004544	0.01568	0.3499	

Two-sided unpaired t-test. $n = 3$ biologically independent samples for each group except for GFP ($n = 6$ biologically independent samples).

Supplementary Table 10. Primers used to generate PCR products for Sanger sequencing.

fwd_BE3-site 3	gggagcttggcatgagaaacc
rev_BE3-site 3	gttgagctcgaccctgaaggg
fwd_BE3-RNF2	ttcaacgtaggaatttgggtgggac
rev_BE3-RNF2	tcaccactgttcacccagtacc
fwd_ABE7.10-site 1	gccattctaccaacaatagaggccc
rev_ABE7.10-site 1	tcccctcagcattcagccactaata
fwd_ABE7.10-site 2	agttcatgggtattacgaggaatacagg
rev_ABE7.10-site 2	tgaaccgggaggaggaggttc
fwd_ABE7.10-site 3	ccgacagccagtgggtaagt
rev_ABE7.10-site 3	gcttttcaccgactgcacag
fwd_ABE7.10-site 7	gatgccctccatcttctccg
rev_ABE7.10-site 7	aggtttgcatagacctgcc
fwd_ABE7.10-site 12	acccatgtgctgacatagg
rev_ABE7.10-site 12	tggtgattatggttacacagcg
fwd_BE3 OF-1	gtgttttcacagagatagactctttgc
rev_BE3 OF-1	cacacacacacagacacaagcctt
fwd_BE3 OF-2	cagaacagacaggcccagattg
rev_BE3 OF-2	gccgcatctgtcgaacaatg
fwd_ABE TOPRS	cagtggctacattaatatccagctct
rev_ABE TOPRS	gcccgatctcctttaacatggca
fwd_ABE CSDE1	agtacgcttatagatcacatgggtgg
rev_ABE CSDE1	gtgcgctgttctcacaactt
fwd_ABE OF-1	gcatgatttagatcggatacggc
rev_ABE OF-1	gctgccagggtcaaggctcgt

Supplementary Table 11. sgRNA sequences.

BE3-site 3	GGCCCAGACTGAGCACGTGA
BE3-RNF2	GTCATCTTAGTCATTACCTG
ABE7.10-site 1	GAACACAAAGCATAGACTGC
ABE7.10-site 2	GAGTATGAGGCATAGACTGC
ABE7.10-site 3	GTCAAGAAAGCAGAGACTGC
ABE7.10-site 7	GAATACTAAGCATAGACTCC
ABE7.10-site 12	GTAGAAAAAGTATAGACTGC

Supplementary Table 12. The number of detected genes in each single cell.

Cell	No. of detected genes	Cell	No. of detected genes	Cell	No. of detected genes
WT_SC 1	11249	BE3_SC 1	11267	ABE7.10_SC 1	10572
WT_SC 2	11730	BE3_SC 2	11729	ABE7.10_SC 2	10841
WT_SC 3	11612	BE3_SC 3	11573	ABE7.10_SC 3	11723
WT_SC 4	11036	BE3_SC 4	11647	ABE7.10_SC 4	11231
WT_SC 5	11045	BE3_SC 5	11316	ABE7.10_SC 5	11967
WT_SC 6	12034	BE3_SC 6	11296	ABE7.10_SC 6	10739
WT_SC 7	11588	BE3_SC 7	11517	ABE7.10_SC 7	9890
WT_SC 8	11541	BE3_SC 8	10207	ABE7.10_SC 8	10645
WT_SC 9	11337	BE3_SC 9	10653	ABE7.10_SC 9	9607
WT_SC 10	11130	BE3_SC 10	11020	ABE7.10_SC 10	11105
WT_SC 11	11068	BE3_SC 11	11099	ABE7.10_SC 11	10729
WT_SC 12	11226	BE3_SC 12	11735	ABE7.10_SC 12	12349
WT_SC 13	10909	BE3_SC 13	10904	ABE7.10_SC 13	10541
WT_SC 14	11634	BE3_SC 14	10973	ABE7.10_SC 14	10579
WT_SC 15	11151	BE3_SC 15	11702	ABE7.10_SC 15	9044
WT_SC 16	11314	BE3_SC 16	11451	ABE7.10_SC 16	10488
GFP_SC 1	13704	BE3_SC 17	11729	ABE7.10_SC 17	9458
GFP_SC 2	11328	BE3_SC 18	10811	ABE7.10_SC 18	10184
GFP_SC 3	11242	BE3_SC 19	11736	ABE7.10_SC 19	10508
GFP_SC 4	10476	BE3_SC 20	10137	ABE7.10_SC 20	9702
GFP_SC 5	10558	BE3_SC 21	11072	ABE7.10_SC 21	10580
GFP_SC 6	11368	BE3_SC 22	11302	ABE7.10_SC 22	9210
GFP_SC 7	11339	BE3_SC 23	11269	ABE7.10_SC 23	9419
GFP_SC 8	11087	BE3_SC 24	11070	ABE7.10_SC 24	8484
GFP_SC 9	10871	BE3_SC 25	10530	ABE7.10_SC 25	9442
GFP_SC 10	11063	BE3_SC 26	11863	ABE7.10_SC 26	10179
GFP_SC 11	10805	BE3_SC 27	11126	ABE7.10_SC 27	9578
GFP_SC 12	10405	BE3_SC 28	11091	ABE7.10_SC 28	11188
GFP_SC 13	10794	BE3_SC 29	11179		
GFP_SC 14	11114	BE3_SC 30	11161		
GFP_SC 15	10063	BE3_SC 31	10164		
GFP_SC 16	10724				

Supplementary Table 13. Precise p values in Fig. 2.

Fig. 2a	APOBEC1 vs GFP	BE3 vs GFP	BE3-site 3 vs GFP	BE3-RNF2 vs GFP
	1.42E-09	2.03E-10	0.01673	5.90E-10
Fig. 2b	TadA-TadA* vs GFP	ABE7.10 vs GFP	ABE7.10-site 1 vs GFP	ABE7.10-site 2 vs GFP
	0.00011	9.95E-05	0.0001048	9.66E-05

Two-sided unpaired t-test. $n = 3$ biologically independent samples for each group except for GFP ($n = 6$ biologically independent samples) and BE3-site 3 ($n = 2$ biologically independent samples).

Supplementary Table 14. Precise p values in Fig. 3.

Fig. 3b	BE3-site 3 vs GFP	ABE7.10-site 1 vs GFP
	0.0285	0.0034
Fig. 3c	BE3-site 3 vs GFP	ABE7.10-site 1 vs GFP
	0.04024	9.85E-05
Fig. 3d	BE3-site 3 vs GFP	ABE7.10-site 1 vs GFP
	0.03577	3.01E-05

Two-sided unpaired t-test. GFP ($n = 15$ cells), BE3-site 3 ($n = 5$ cells) and ABE7.10-site 1 ($n = 9$ cells).

Supplementary Table 15. Precise p values in Extended Data Figure 3.

	ABE7.10-site 1 (36h) vs ABE7.10-site 1 (72h)	ABEmax-site 1 (72h) vs ABE7.10-site 1 (72h)
Extended Data Figure 3c	0.002865	0.003094
Extended Data Figure 3e	ABE7.10-site 1 (36h) vs ABE7.10-site 1 (72h)	ABEmax-site 1 (72h) vs ABE7.10-site 1 (72h)
	0.0034	0.0017

Two-sided unpaired t-test. $n = 3$ biologically independent samples for each group.

Supplementary Table 16. Precise p values in Extended Data Figure 3.

Extended Data Figure 5a	BE3 vs GFP	ABE7.10 vs GFP	ABE7.10 vs BE3
3'UTR	0.00549451	0.00549451	2.84E-06
5'UTR	1	0.02564103	2.84E-06
Exonic	0.01098901	0.00549451	2.84E-06
Downstream	0.76923077	0.00549451	2.84E-06
Intronic	0.00549451	0.00549451	0.36530807
Non_coding RNA	0.00549451	0.00549451	2.84E-06
Upstream	0.00549451	0.29120879	2.84E-06

Two-sided unpaired t-test. $n = 3$ biologically independent samples for GFP, $n = 8$ biologically independent samples for BE3 group and $n = 9$ biologically independent samples for ABE7.10 group.

Supplementary Table 17. Precise p values in Extended Data Figure 10.

	ABE7.10 (TadA-TadA*- Cas9n) vs GFP	ABE7.10 (TadA^{F148A}- TadA*-Cas9n) vs GFP	ABE7.10 (TadA^{F148A}- TadA*^{F148A}-Cas9n) vs GFP
Extended Data Figure 10c	0.0003	0.0001	0.3596

Two-sided unpaired t-test. $n = 3$ biologically independent samples for each group.