

Supplementary Material

Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme

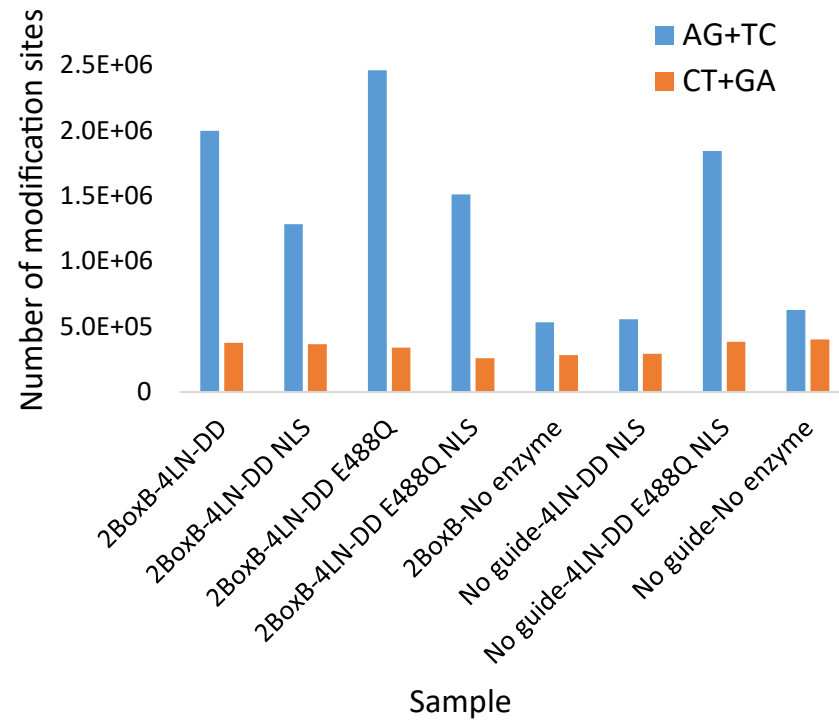
Isabel C. Vallecillo-Viejo^{1,2}, Noa Liscovitch-Brauer³, Maria Fernanda Montiel-Gonzalez¹, Eli Eisenberg^{3*} and Joshua J.C. Rosenthal^{1*}

¹ Eugene Bell Center for Regenerative Biology and Tissue Engineering, The Marine Biological Laboratory, Woods Hole, MA 02543

² Department of Medicine, University of Puerto Rico, Medical Sciences Campus, San Juan, PR 00921

³ Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel

* To whom correspondence should be addressed: jrosenthal@mbl.edu (JJCR) and elieis@post.tau.ac.il (EE)



Supplementary Figure S1. Transcriptome-wide off-targets produced by SDRE. Number of RNA-DNA mismatches identified within the complete genome in each of the transfected cells and controls.

Sample name	IVR2	IVR3	IVR4	IVR5	IVR6	IVR7	IVR8	IVR1
Treatment	2BoxB-4LN-DD	2BoxB-4LN-DD NLS	2BoxB-4LN-DD E488Q	2BoxB-4LN-DD E488Q NLS	2BoxB-No enzyme	No guide-4LN-DD NLS	No guide-4LN-DD E488Q NLS	No guide-No enzyme

RNA-DNA differences (all genome)

AC	44705	45691	39060	33275	33824	35073	44321	51157
AG	1E+06	647385	1245263	765303	269183	280097	929450	314342
AT	68197	65483	57268	35962	39988	41742	48403	81432
CA	77273	72659	68928	66784	55559	55874	73823	85752
CG	47230	46812	42148	32814	33890	34475	44860	49852
CT	186892	182119	169539	129215	140993	146870	191535	201065
GA	189784	183959	171279	130911	142614	148105	192934	203505
GC	47666	47090	42263	32835	34349	35328	45826	49924
GT	77977	73460	69549	66432	56457	56051	73482	86463
TA	69831	66452	59211	37757	40948	42116	49193	82776
TC	989713	636540	1215121	746397	264912	277761	913872	313051
TG	45094	46444	39271	34116	33974	35240	44759	52397

RNA-DNA differences (CoDing Sequence only):

All CDS differences:

AC	10895	13027	10189	9307	9549	9213	9253	12796
AG	621945	305989	924168	543385	36076	37470	366173	41707
AT	18312	19358	16179	11636	9688	9343	8562	20542
CA	15004	16231	15522	23691	15975	14103	16940	13761
CG	9140	9612	9342	7722	5885	5554	6019	8490
CT	56505	59090	58387	48275	40890	40352	44309	52237
GA	49622	51338	51023	43918	37792	36873	39678	47894
GC	20318	21076	20187	16414	14099	13809	14821	19228
GT	22556	21888	22415	22112	12907	12488	12810	21819
TA	9821	10758	10013	8786	7246	6900	7284	9855
TC	92014	90163	91188	74752	66111	63717	74921	75710
TG	11600	12574	12371	12180	6765	6827	7283	11308

Sample name	IVR2	IVR3	IVR4	IVR5	IVR6	IVR7	IVR8	IVR1
Treatment	2BoxB-4LN-DD	2BoxB-4LN-DD NLS	2BoxB-4LN-DD E488Q	2BoxB-4LN-DD E488Q NLS	2BoxB-No enzyme	No guide-4LN-DD NLS	No guide-4LN-DD E488Q NLS	No guide-No enzyme

CDS differences with ≥ 3 supporting reads and editing level ≥ 0.1 :

AC	270	271	214	177	326	343	290	309
AG	20092	9673	78197	32264	990	1041	6613	954
AT	199	202	148	148	232	252	205	246
CA	342	309	264	247	433	435	355	390
CG	424	350	333	271	432	464	409	426
CT	888	844	822	701	1069	1113	930	1003
GA	905	875	753	670	1038	1100	929	983
GC	432	386	308	243	475	496	437	451
GT	238	229	234	178	304	309	266	290
TA	217	199	172	159	252	256	205	232
TC	1467	1299	1467	1108	1081	1167	1216	1079
TG	220	165	166	125	237	252	228	222

Supplementary Table S1. Summary of transcriptome-wide off-targets by SDRE. DNA-RNA mismatches identified in the complete genome and in coding sequences are shown.

Chr	Position	Gene	Region	Strand	IVR2 #A	IVR2 #G	IVR3 #A	IVR3 #G	IVR4 #A	IVR4 #G	IVR5 #A	IVR5 #G	IVR6 #A	IVR6 #G	IVR7 #A	IVR7 #G	IVR8 #A	IVR8 #G	IVR1 #A	IVR1 #G
chr1	160302244	COPA	CDS	-	292	56	257	40	170	43	198	25	890	61	1077	99	613	89	441	33
chr11	105804694	GRIA4	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chr11	105815132	GRIA4	intron	+	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
chr11	105816106	GRIA4	intron	+	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
chr11	105816129	GRIA4	intron	+	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
chr11	105816145	GRIA4	intron	+	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
chr11	105816160	GRIA4	intron	+	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
chr12	5021742	KCNA1	CDS	+	0	0	1	0	0	2	2	0	0	0	1	0	0	0	0	0
chr13	46090371	COG3	CDS	+	107	67	87	42	37	27	63	23	170	68	202	101	176	89	116	45
chr14	26917530	NOVA1	CDS	-	31	6	20	2	14	4	7	4	23	0	35	0	60	7	12	0
chr14	101506074	mir376C	miRNA	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chr17	43045220	C1QL1	CDS	-	3	0	6	2	2	0	4	0	1	0	0	0	2	0	3	0
chr19	47152854	DACT3	CDS	-	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
chr2	20450819	PUM2	3'UTR	-	125	2	94	1	53	8	66	2	233	3	351	7	183	6	93	2
chr2	21233202	APOB	CDS	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chr2	210835613	UNC80	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chr20	36147533	BLCAP	CDS	-	278	20	169	16	140	21	101	9	347	15	452	7	352	11	274	16
chr20	36147563	BLCAP	CDS	-	248	40	140	23	121	26	105	7	306	17	397	32	296	47	243	23
chr20	36147572	BLCAP	CDS	-	208	78	111	57	86	60	71	41	266	52	354	68	246	95	208	61
chr20	36148080	BLCAP	intron	-	4	12	2	4	4	6	2	2	0	1	0	0	7	2	2	6
chr20	52104918	TSHZ2	3'UTR	+	7	4	14	0	7	0	4	1	26	0	30	0	8	2	16	0
chr21	30953750	GRIK1	CDS	-	2	0	3	0	0	0	0	0	0	0	0	0	1	0	2	0
chr21	34922801	SON	CDS	+	773	256	548	123	302	135	269	129	1309	3	1458	6	1277	217	504	3
chr21	34923319	SON	CDS	+	720	494	485	282	304	201	263	170	1267	22	1568	45	1329	391	749	13
chr21	46595620	ADARB1	intron	+	0	0	2	0	1	0	9	0	0	0	0	0	11	0	9	0
chr3	53820892	CACNA1D	CDS	+	2	0	1	0	7	0	7	0	8	0	8	0	9	0	7	0
chr3	58141801	FLNB	CDS	+	152	17	147	6	75	9	107	11	270	0	250	0	270	8	287	0
chr3	62423807	CADPS	CDS	-	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
chr4	57976234	IGFBP7	CDS	-	10	1	1	0	3	1	6	4	2	0	0	0	1	0	20	0
chr4	57976286	IGFBP7	CDS	-	8	0	6	0	4	0	12	0	3	0	2	0	0	0	11	0
chr4	158257875	GRIA2	CDS	+	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
chr4	158257879	GRIA2	CDS	+	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
chr4	158258136	GRIA2	intron	+	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

Chr	Position	Gene	Region	Strand	IVR2 #A	IVR2 #G	IVR3 #A	IVR3 #G	IVR4 #A	IVR4 #G	IVR5 #A	IVR5 #G	IVR6 #A	IVR6 #G	IVR7 #A	IVR7 #G	IVR8 #A	IVR8 #G	IVR1 #A	IVR1 #G
chr4	158258137	GRIA2	intron	+	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
chr4	158281294	GRIA2	CDS	+	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
chr5	156736808	CYFIP2	CDS	+	80	11	90	2	67	11	50	3	190	0	232	0	177	2	72	0
chr6	34100903	GRM4	CDS	-	4	1	1	2	0	0	0	0	0	0	17	0	9	1	7	0
chr6	44120349	TMEM63B	CDS	+	195	4	140	1	93	18	79	10	147	0	143	0	307	12	195	0
chr6	102337689	GRIK2	CDS	+	5	0	1	1	0	0	0	0	3	0	2	0	1	0	2	0
chr6	102337702	GRIK2	CDS	+	4	0	1	1	0	0	0	0	2	0	1	0	2	0	2	0
chr6	102372589	GRIK2	CDS	+	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0
chr6	102372630	GRIK2	intron	+	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
chr6	102374616	GRIK2	intron	+	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0
chr6	102374643	GRIK2	intron	+	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0
chr6	150093334	PCMT1	intron	+	6	23	5	1	1	14	5	7	18	0	20	1	15	12	15	0
chr8	103841636	AZIN1	CDS	-	468	16	362	5	301	25	218	3	1842	7	2290	6	970	13	767	1
chr8	103841637	AZIN1	CDS	-	481	3	366	0	324	1	220	1	1852	0	2281	3	984	0	772	0
chr9	97847739	mir23B	miRNA	+	2	0	1	0	0	0	1	0	1	0	3	0	5	0	3	0
chrX	114082682	HTR2C	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chrX	114082684	HTR2C	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chrX	114082689	HTR2C	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chrX	114082694	HTR2C	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chrX	122598962	GRIA3	CDS	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chrX	122598998	GRIA3	intron	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chrX	135111055	SLC9A6	intron	+	2	1	0	0	0	0	1	0	3	0	1	0	0	0	0	0
chrX	135111070	SLC9A6	intron	+	4	0	0	0	0	0	1	0	3	0	2	0	1	0	0	0
chrX	151358319	GABRA3	CDS	-	1	1	0	0	3	0	1	1	4	0	0	0	1	0	0	0
chrX	153579737	FLNA	intron	-	240	65	207	25	209	68	107	13	45	0	66	3	345	17	213	0
chrX	153579950	FLNA	CDS	-	956	278	722	180	543	258	371	135	987	4	1147	5	1744	235	1331	5

Supplementary Table S2. Off-targets produced by SDRE in conserved mammalian sites. Chr = chromosome

OLIGO NAME	SEQUENCE	TARGET SEQUENCE	DESCRIPTION
3XNLS gene block	<u>GACAACTAGTAGCTCCAAA</u> <u>AAAGAAGAGAAAGGTAGATC</u> <u>CAAAAAGAAGAGAAAGGTA</u> <u>GATCCAAAAAGAAGAGAAA</u> <u>GGTAGGATCCACCTCCGGAG</u> <i>TAAATGCCCGTACG</i>	N/A	Geneblock with the SV40 Large T Antigen NLS and linker for cloning in 4λN-DD and 4λN-DD E488Q. 12nt overlaps with inverse PCR primers are shown in bold. 3X NLS is underlined. The GSTSGV linker is shown in italics.
1st_λN_λNDDpc DNA_Fwd	AATGCCCGTACGCGCCGGCG CGAACG	970-995 nt of 4λN-DD and 4λN-DD E488Q in pcDNA 3.1 (+)	Forward primer for inverse PCR to clone the 3XNLS gene block in 4λNDD and 4λNDD E488Q in pcDNA
Inv pcDNA3 λN-DD R	GTTTGCAGCTTTCATTGAGC TTGTTTCTCAGCGC	1322-1356 nt of 4λN-DD and 4λN-DD E488Q in pcDNA 3.1 (+)	Reverse primer for inverse PCR to clone the 3XNLS gene block in 4λNDD and 4λNDD E488Q in pcDNA
Guide2BB-Y122X S	<u>CACCGAAGGCATAAGGCCCT</u> <u>GAAAAGGGCCTATGCCTAG</u> <u>CCAAATCGCGATAGAGCGTT</u> <u>GGCCCTGAAAAGGGCCCT</u> TGTTATC	Guide is complementary to nt's 334-343, 347-375, and 379-388 of CFTR	Sense primer to synthesize 2boxB RNA Guide A for targeting CFTR Y122X <i>in cellula</i> . RNA guide has two A:C mismatches under the target A's. Extra nt's required by the BLOCK-iT U6 RNAi ENTRY VECTOR kit are in bold. boxB hairpins are underlined.
Guide2BB-Y122X AS	<u>AAAAGATAACAAGGGGCCCT</u> <u>TTTTCAGGGCCAACGCTCTAT</u> <u>CGCGATTTGGCTAGGCATAG</u> <u>GCCCTTTTTCAGGGCCTTATG</u> CCTTC	Guide is complementary to nt's 334-343, 347-375, and 379-388 of CFTR	Antisense primer to synthesize 2boxB RNA Guide A for targeting CFTR Y122X <i>in cellula</i> . RNA guide has two A:C mismatches under the target A's. Extra nt's required by the BLOCK-iT U6 RNAi ENTRY VECTOR kit are in bold. boxB hairpins are underlined.

OpGuide2BB-Y122X S	CACCAGGCATAAGCGGCCCT <u>GAAAAGGGCCTGCCTAGCC</u> AAATCGCGATAGAGCGTTGG <u>CCCTGAAAAGGGCCCTTG</u> TTATC	Guide is complementary to nt's 334-343, 347-373, and 377-386	Sense primer to synthesize 2boxB RNA Guide B for targeting CFTR Y122X <i>in cellula</i> . RNA guide has two A:C mismatches under the target A's. Extra nt's required by the BLOCK-iT U6 RNAi ENTRY VECTOR kit are in bold. boxB hairpins are underlined.
OpGuide2BB-Y122X AS	AAAAGATAACAAGGGGCCCT <u>TTTTCAGGGCCAACGCTCTAT</u> CGCGATTTGGCTAGGCAGGC <u>CCTTTTTCAGGGCCGCTTATG</u> CCT	Guide is complementary to nt's 334-343, 347-373, and 377-386	Antisense primer to synthesize 2boxB RNA Guide A for targeting CFTR Y122X <i>in cellula</i> . RNA guide has two A:C mismatches under the target A's. Extra nt's required by the BLOCK-iT U6 RNAi ENTRY VECTOR kit are in bold. boxB hairpins are underlined.
CFTR123RT	CGGACAATTTCCCTCTCAGC AGTGGGCTATTGAAGTATCTC ACATA	901-921 nt of CFTR	Reverse RT primer for cDNA synthesis of CFTR Y122X after <i>in cellula</i> editing assays. Bold – ActinV.
CFTR123F	TGCCCTTCGGCGATGTTTTTT CTGGAGATT	213-242 nt of CFTR	Forward primer for PCR amplification with CFTR Y122X after <i>in cellula</i> editing assays.
CFTRY122X_REV1	GGAGCGATCCACACGAAATG TGCC	591-614 nt of CFTR	Reverse primer for PCR amplification of CFTR Y122X after <i>in cellula</i> editing assays

CFTRY122X_ SEQ	GCTATTCTCATCTGCATTCC	445-464 nt of CFTR	Reverse primer for sequencing CFTR after <i>in cellula</i> editing assays.
CFTR Fwd Begin	ATGCAGAGGTGCCTCTGGA AAAGG	1-25 nt of CFTR	Forward primer for PCR amplification of CFTR Y122X after <i>in cellula</i> editing assays.
CFTR 18 RT	CGGACAATTTCCCTCTCAGC AGTGGGGACACAGCCCCCAT CCAC	4018-4036 nt of CFTR	Reverse RT primer for cDNA synthesis of CFTR Y122X after <i>in cellula</i> editing assays. Bold=ActinV.
CFTR 22 RT	CGGACAATTTCCCTCTCAGC AGTGGCCAGGAGACAGGAGC ATCTCC	2017-2037 nt of CFTR	Reverse RT primer for cDNA synthesis of CFTR Y122X after <i>in cellula</i> editing assays. Bold=ActinV.
RT Actin V IVS 2corr	CGGACAATTTCCCTCTCAGC AGTGGTTTTTTTTTTTTTTTTT ATTAGG	N/A	Reverse RT primer for cDNA synthesis of CFTR Y122X after <i>in cellula</i> editing assays. Bold=ActinV.
CFTR 16	CTGGCAGAGGTGGCTGCTTC	2611-2630 nt of CFTR	Forward primer for PCR amplification and sequencing CFTR after <i>in cellula</i> editing assays.
CFTR 17	GCGATCTGTGAGCCGAGTCT	3471-3490 nt of CFTR	Forward primer for sequencing CFTR after <i>in cellula</i> editing assays.
CFTR 19	GAAAGAGGACAGTTGTTGGC GG	1345-1366 nt of CFTR	Forward primer for sequencing CFTR after <i>in cellula</i> editing assays.
CFTR 123SEQ	CCGAGTCAGTTTCAGTTCTGT TTGTC	866-891 nt of CFTR	Reverse primer for sequencing CFTR after <i>in cellula</i> editing assays.
CFTR 9	GAGAAACGGTGTAAGGTCTC AG	1988 – 2009 nt of CFTR	Reverse primer for sequencing CFTR after <i>in cellula</i> editing assays.

CFTR-4 SEQ	GGCTGTCATCACCATTAGAAG TTTTTC	1256-1282 nt of CFTR	Reverse primer for sequencing CFTR after <i>in cellula</i> editing assays.
CFTR 24	ATGGGACATTTTCAGAACTCC A	1880-1901 nt of CFTR	Forward primer for PCR amplification and sequencing CFTR after <i>in cellula</i> editing assays.
CFTR SEQ END	GTCCTTGTGGATGGGGGCTG TGTC	4012-4035 nt of CFTR	Forward primer for sequencing CFTR after <i>in cellula</i> editing assays.
Y122XFwdMiseq	CACGACGCTCTTCCGATCTG <u>GGAAGTCACCAAAGCAGTAC</u> AGCC	213-242 nt of CFTR	Forward primer for amplifying CFTR Y122X for Miseq. Sequence complementary to CFTR (same sequence as CFTR 123 primer) is underlined. The rest encodes templates for Illumina MiSeq sequencing primers.
Y122XRevMiseq	GACGTGTGCTCTTCCGATCT <u>GGAGCGATCCACACGAAATG</u> TGCC	591-614 nt of CFTR	Reverse primer for amplifying CFTR Y122X for Miseq. Sequence complementary to CFTR (same sequence as CFTRY122X_REV1) is underlined. The rest encodes templates for Illumina MiSeq sequencing primers.
Illumina Fwd	AATGATACGGCGACCACCGA GATCTACACTCTTCCCTACA CGACGCTCTTCCGATCT	N/A	Forward primer for second round PCR amplification for Illumina MiSeq sequencing

Illumina Rev	CAAGCAGAAGACGGCATACG AGAT*6NTIndex*GTGACTGGA GTTCAGACGTGTGCTCTTCC GATCT	N/A	Reverse primer for second round PCR amplification for Illumina MiSeq sequencing
--------------	--	-----	---

Supplementary Table S3. List of oligonucleotides used in this study

SAMPLE	RNA GUIDE	EDITING ENZYME	# READS
IVR1	None	None	105,608,417
IVR2	2boxB	4λN-DD	88,930,808
IVR3	2boxB	NLS-4λN-DD	74,629,491
IVR4	2boxB	4λN-DD E488Q	59,027,782
IVR5	2boxB	NLS-4λN-DD E488Q	45,700,559
IVR6	2boxB	None	120,838,138
IVR7	None	NLS-4λN-DD	141,347,218
IVR8	None	NLS-4λN-DD E488Q	127,471,855

Supplementary Table S4. Samples used for assessing transcriptomic off-target RNA editing