OMTM, Volume *18*

Supplemental Information

Unexpected Mutations by CRISPR-Cas9 CTG Repeat Excision in Myotonic Dystrophy and Use of CRISPR Interference as an Alternative Approach

Miki Ikeda, Mariko Taniguchi-Ikeda, Takema Kato, Yasuko Shinkai, Sonoko Tanaka, Hiroki Hagiwara, Naomichi Sasaki, Toshihiro Masaki, Kiichiro Matsumura, Masahiro Sonoo, Hiroki Kurahashi, and Fumiaki Saito

Figure S1

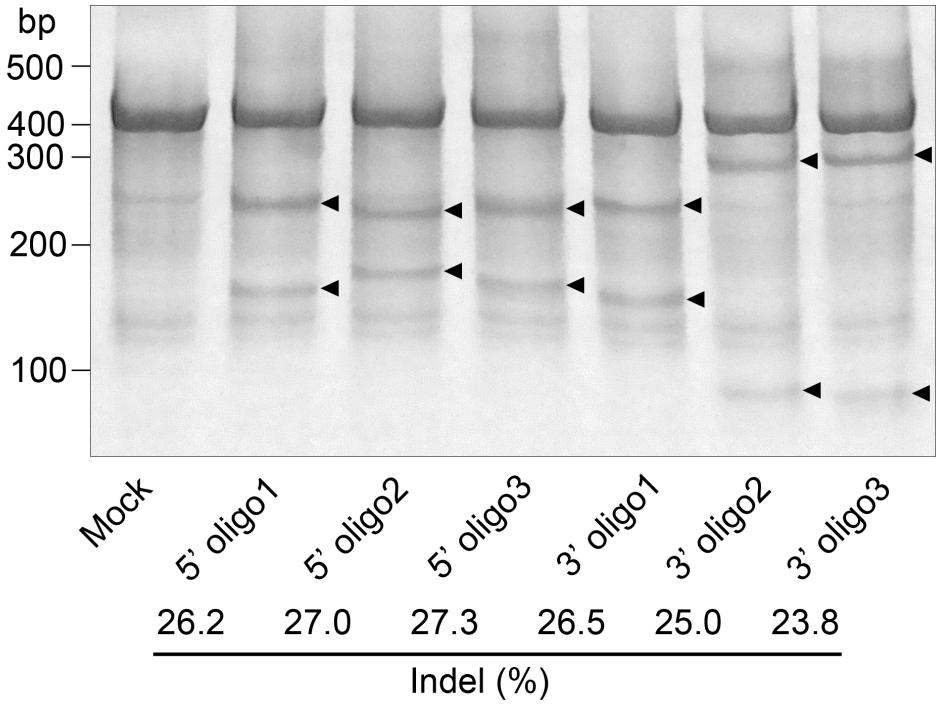
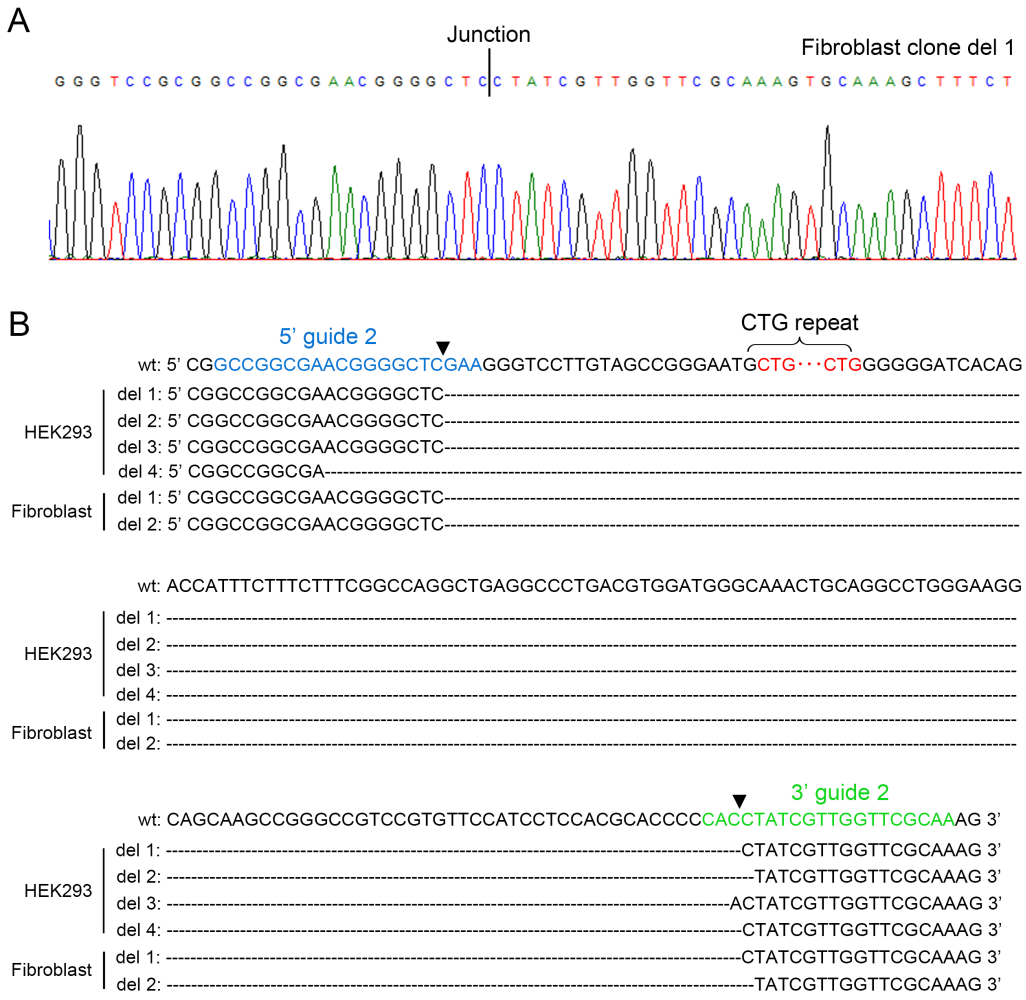


Figure S1. Evaluation of the DSB formation by Cas9 nuclease. The sgRNAs designed at the 3′ UTR of the *DMPK* gene were examined for the efficiency of the formation of DSBs using a T7 endonuclease 1 assay. The indel frequencies of sgRNAs were estimated to be 20-30% using the HEK293 cells. Arrowheads indicate the bands generated by the digestion following the indel formation.

Figure S2



C

cells	clone name	# of clone	% of clone
HEK293	del 1	1	14.3
	del 2	4	57.1
	del 3	1	14.3
	del 4	1	14.3
Fibroblast	del 1	5	83.3
	del 2	1	16.7

Figure S2. Sequence of the 3′ UTR of the *DMPK* gene after the excision of the CTG repeat with Cas9 nuclease. (A) A result of Sanger sequencing, encompassing the junction region of type del 1 in the fibroblast GM03991, is shown. (B) By sequencing the clones obtained from the amplicon of the lower bands, several junctional sequences (del 1 to 4 for HEK293 cells, del 1 and 2 for fibroblasts) were observed. Arrowheads indicate the position of the expected DSBs. (C) The frequencies of each junction were calculated. The junctional sequences were mostly homogeneous.

Figure S3

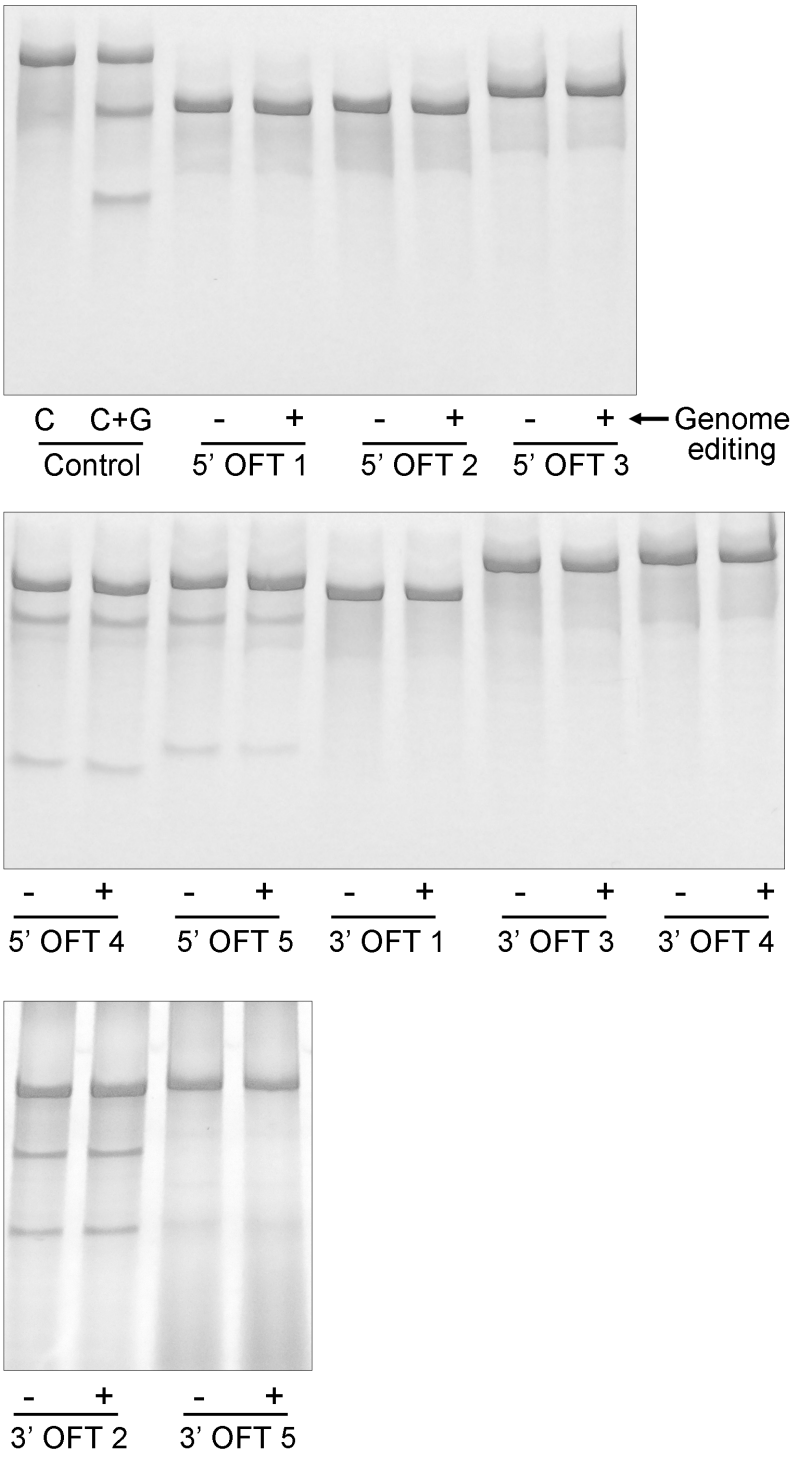


Figure S3. Assessment of the off-target effects using a T7 endonuclease 1 assay. The fibroblast GM03991 was co-transfected with the 5′ guide 3 and 3′ guide 2. The top 5 predicted off-target sites for each sgRNA were amplified by PCR. The T7 endonuclease 1 assay showed no obvious off-target indels at these sites.

Figure S4

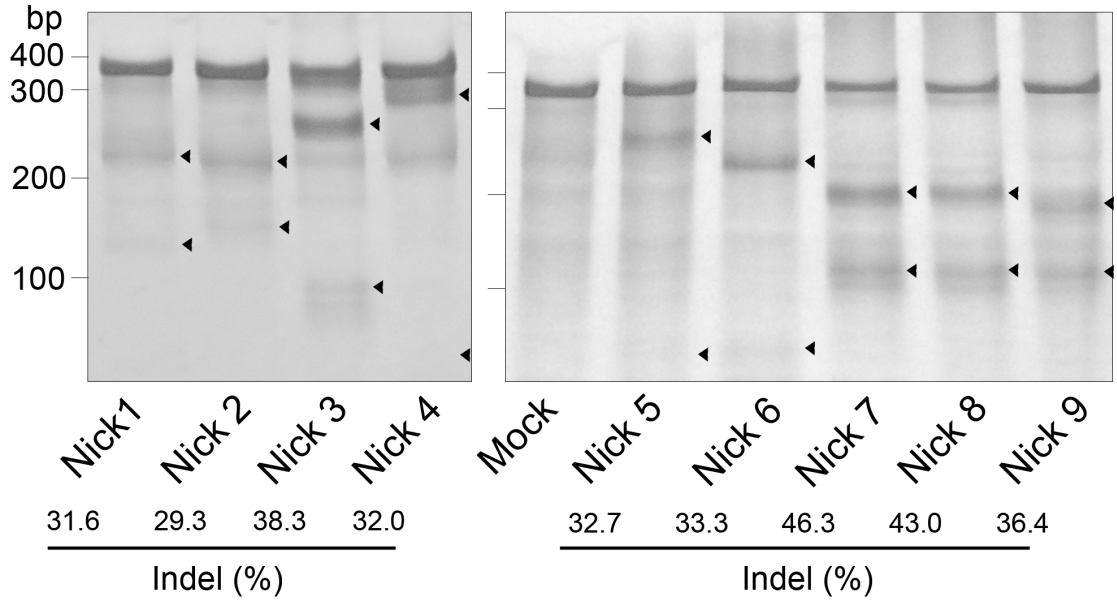
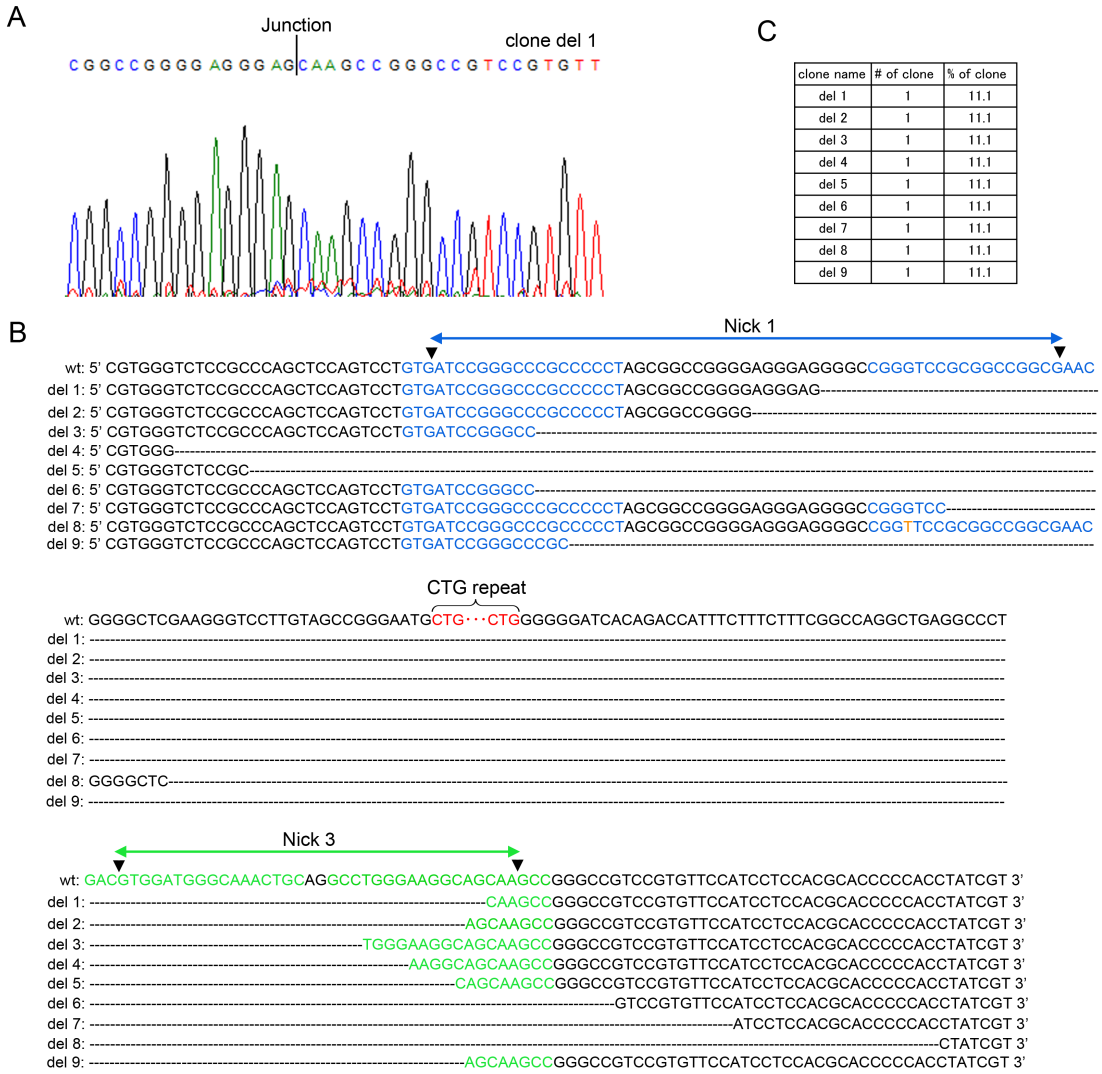


Figure S4. Evaluation of the DSB formation by the double nicking strategy using Cas9 nickase. The nicking pairs of sgRNAs designed at the 3′ UTR of the *DMPK* gene were examined for the efficiency of the formation of DSBs using a T7 endonuclease 1 assay. The indel frequencies of each nicking pairs were estimated to be 30% to over 40% using the HEK293 cells. Arrowheads indicate the bands generated by the digestion following the indel formation.

Figure S5



clone name	# of clone	% of clone
del 1	1	11.1
del 2	1	11.1
del 3	1	11.1
del 4	1	11.1
del 5	1	11.1
del 6	1	11.1
del 7	1	11.1
del 8	1	11.1
del 9	1	11.1

Figure S5. Sequence of the 3′ UTR of the *DMPK* gene after the excision of the CTG repeat with the double nicking strategy. (A) A result of Sanger sequencing, encompassing the junction region of type del 1 in the HEK293 cell is shown. (B) By sequencing the clones obtained from the amplicon of the lower bands, several junctional sequences (del 1 to 9) were observed. Arrowheads indicate the position of the expected nicks. (C) The frequencies of each junction were calculated. The junctional sequences were heterogeneous, and no identical sequence was observed among the clones we tested.

Figure S6

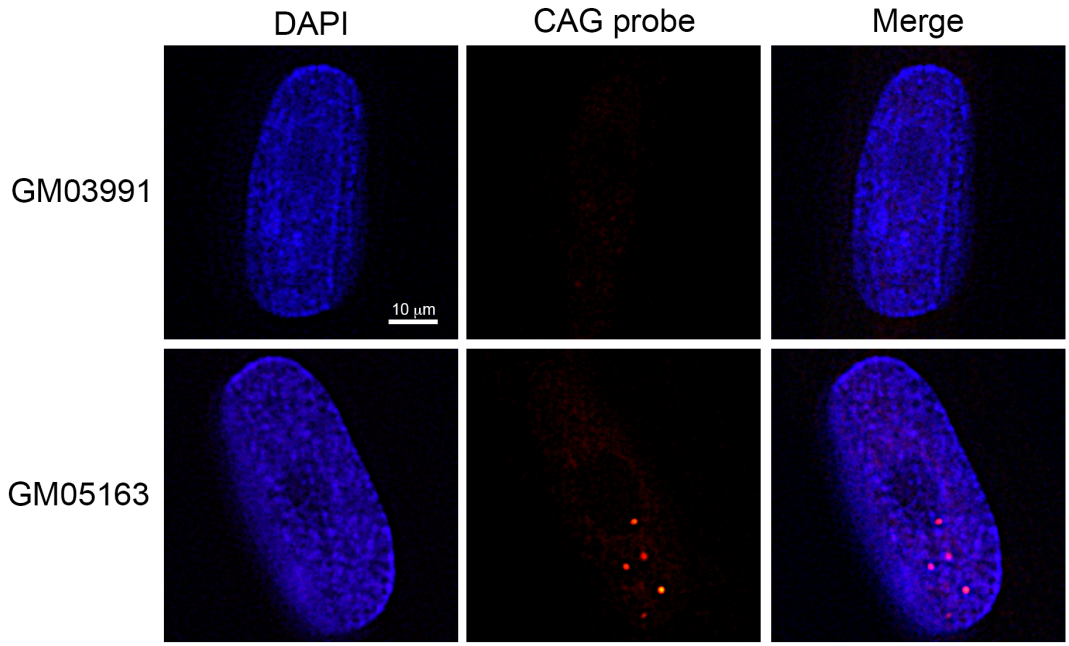


Figure S6. RNA-FISH assay using the fibroblasts derived from DM1 patients. The RNA-FISH assay revealed several intense intranuclear RNA foci in the fibroblast GM05163 but not in GM03991. This result indicates that the length of 50 to 80 CTG repeats are not enough to generate the RNA foci.

Figure S7

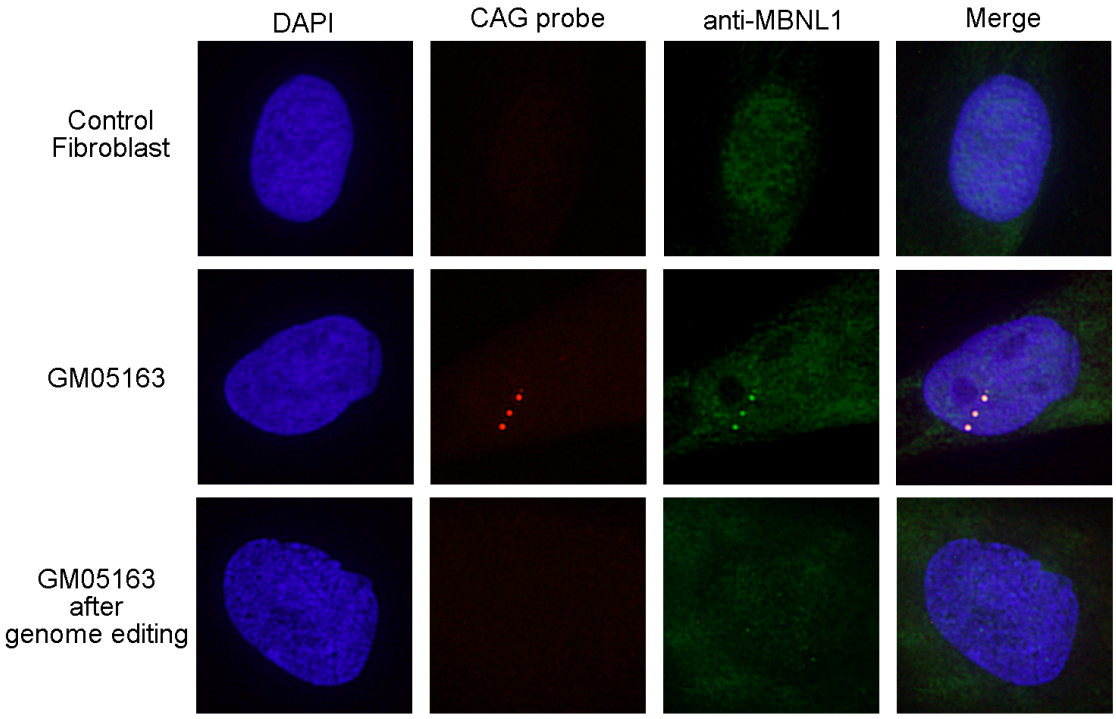


Figure S7. Colocalization of MBNL 1 with RNA foci. The RNA-FISH assay was followed by the immunofluorescent analysis using an anti-MBNL 1 antibody. The MBNL 1 signals were observed to colocalize with the RNA foci in the fibroblast GM05163. Upon genome editing using Cas9 nuclease, both signals were abolished, indicating that the sequestered MBNL 1 was released from the foci.

Figure S8

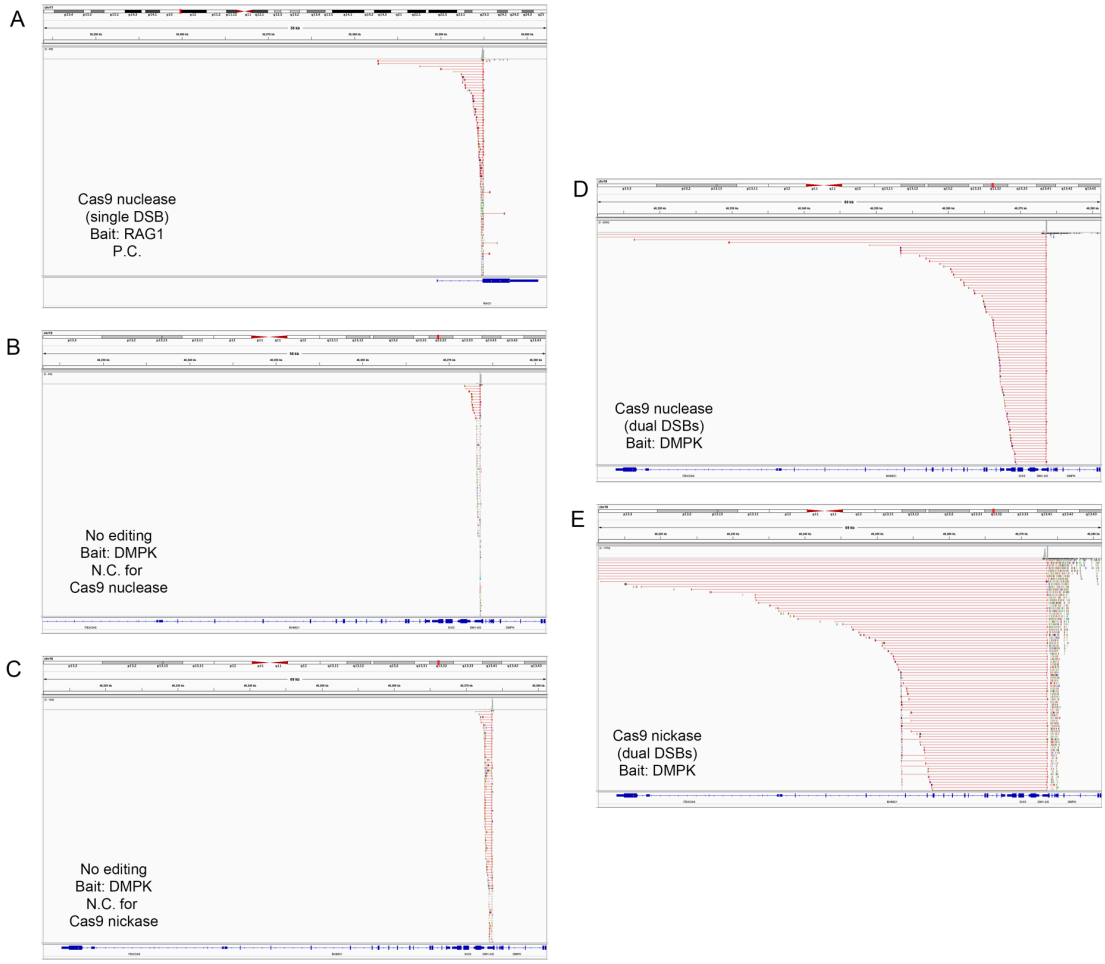


Figure S8. Mapping of the paired reads obtained by LAM-HTGTS using the integrative genomics viewer software. The snapshots show the reads as pairs sorted by their insert size. (A) Upon a generation of a single DSB at the *RAG1* locus, the left-sided reads of some of the paired reads were mapped several kb apart from the breakpoint. The non-edited samples using *DMPK* as a bait, which were used as negative controls for Cas9 nuclease (B) and Cas9 nickase (C), exhibited only minimal gaps between the paired reads. Dual DSBs generated by both Cas9 nuclease (D) and Cas9 nickase (E), especially the latter, resulted in much larger gaps between the paired reads. These large gaps indicate that genomic rearrangements occurred including large deletions. P.C.; positive control, N.C.; negative control.

Table S1

Frequency of translocation

RAG1 Cas9 nuclease

Translocation	Bait		Prey		Count of	Frequency in total	Frequency in total
hot-spot	chromosome	position	chromosome	position	Reads	split reads (1137)	mapped reads (2476303)
#1	11	36594900	7	155931000	3	0.263852%	0.000121%
	11	36594900	7	155931100	3	0.263852%	0.000121%
#2	11	36594900	12	47002900	8	0.703606%	0.000323%
	11	36594900	12	47003000	9	0.791557%	0.000363%
#3	11	36594900	15	75746500	10	0.879507%	0.000404%
	11	36594900	15	75746600	4	0.351803%	0.000162%
	11	36594900	15	75746700	5	0.439754%	0.000202%
#4	11	36594900	19	1417400	12	1.055409%	0.000485%
	11	36594900	19	1417500	8	0.703606%	0.000323%
	11	36594900	19	1417600	3	0.263852%	0.000121%
Total	N/A	N/A	N/A	N/A	65	5.716799%	0.002625%

DMPK Cas9 nuclease

Translocation	Bait		Prey		Count of	Frequency in total	Frequency in total
hot-spot	chromosome	position	chromosome	position	Reads	split reads (26069)	mapped reads (6596133)
#1	19	46273300	1	227059400	3	0.0115079%	0.000045%
#2	19	46273600	14	62534300	3	0.0115079%	0.000045%
#3	19	46273600	15	68132400	25	0.0958993%	0.000379%
	19	46273600	15	68132500	44	0.1687828%	0.000667%
	19	46273600	15	68132600	21	0.0805554%	0.000318%
#4	19	46273600	17	79579500	3	0.0115079%	0.000045%
#5	19	46273600	19	2136700	3	0.0115079%	0.000045%
#6	19	46273600	X	15717100	3	0.0115079%	0.000045%
Total	N/A	N/A	N/A	N/A	105	0.4027772%	0.001592%

DMPK Cas9 nickase

Translocation	Bait		Prey		Count of	Frequency in total	Frequency in total
hot-spot	chromosome	position	chromosome	position	Reads	split reads (30619)	mapped reads (5485316)
#1	19	46273600	2	27615800	3	0.009798%	0.000055%
#2	19	46273600	2	33141300	3	0.009798%	0.000055%
#3	19	46273600	3	3075400	3	0.009798%	0.000055%
#4	19	46273600	5	78782800	4	0.013064%	0.000073%
#5	19	46273600	5	179558400	3	0.009798%	0.000055%
#6	19	46273600	7	134156100	3	0.009798%	0.000055%
#7	19	46273600	8	1754600	3	0.009798%	0.000055%
#8	19	46273600	9	134700900	3	0.009798%	0.000055%
#9	19	46273600	9	139685700	3	0.009798%	0.000055%
#10	19	46273600	11	33757700	3	0.009798%	0.000055%
#11	19	46273600	11	74950100	3	0.009798%	0.000055%
#12	19	46273600	12	74399000	3	0.009798%	0.000055%
#13	19	46273600	15	56252500	3	0.009798%	0.000055%
#14	19	46273600	15	68132500	20	0.065319%	0.000365%
	19	46273600	15	68132600	15	0.048989%	0.000273%
#15	19	46273600	16	1481600	3	0.009798%	0.000055%
#16	19	46273600	16	2296500	4	0.013064%	0.000073%
#17	19	46273600	16	22328400	3	0.009798%	0.000055%
#18	19	46273600	17	36041800	3	0.009798%	0.000055%
#19	19	46273600	17	39620700	3	0.009798%	0.000055%
#20	19	46273600	17	46178700	13	0.042457%	0.000237%
#21	19	46273600	17	80023600	3	0.009798%	0.000055%
#22	19	46273600	19	2312500	3	0.009798%	0.000055%
#23	19	46273600	19	17531800	3	0.009798%	0.000055%
#24	19	46273600	20	34892400	3	0.009798%	0.000055%
#25	19	46273600	20	39292200	3	0.009798%	0.000055%
Total	N/A	N/A	N/A	N/A	119	0.388648%	0.002169%

Table S2

Possible on-target large deletions

RAG1	Chromosome 11 reference span				Estimated deletion
Read pair #	Left alignment		Right alignment		size (bp)
1	36587558	36587665	36594893	36595029	7227
2	36589939	36590121	36594878	36595029	4756
3	36591473	36591523	36594888	36595029	3364
4	36592346	36592438	36594884	36595029	2445
5	36592453	36592667	36594880	36595029	2212
6	36592498	36592666	36594897	36595029	2230
7	36592760	36592868	36594880	36595029	2011
8	36592766	36592902	36594882	36595029	1979
9	36592803	36592967	36594884	36595029	1916
10	36593334	36593418	36594886	36595029	1467
11	36593520	36593663	36594880	36595029	1216
12	36593683	36593793	36594880	36595029	1086
13	36593729	36593829	36594883	36595029	1053
14	36593691	36593857	36594879	36595029	1021
15	36593735	36593852	36594880	36595029	1027
16	36593697	36593880	36594878	36595029	997
17	36593903	36594002	36594883	36595029	880
18	36593933	36594119	36594878	36595029	758
19	36593993	36594125	36594878	36595029	752
20	36594085	36594162	36594883	36595029	720

DMPK nuclease	Chromosome 19 reference span				Estimated deletion
Read pair #	Left alignment		Right alignment		size (bp)
1	46253337	46253481	46273561	46273724	20079
2	46253371	46253536	46273576	46273724	20039
3	46253407	46253536	46273557	46273724	20020
4	46258843	46258880	46273567	46273724	14686
5	46259353	46259504	46273568	46273724	14063
6	46260312	46260463	46273557	46273724	13093
7	46260705	46260837	46273576	46273724	12738
8	46260851	46260918	46273557	47273724	12638
9	46261728	46261835	46273572	46273724	11736
10	46262176	46262323	46273574	46273724	11250
11	46262817	46262964	46273567	46273723	10602
12	46263637	46263763	46273557	46273724	9793
13	46264698	46264794	46273574	46273724	8779
14	46264843	46264947	46273557	46273724	8609
15	46264883	46264975	46273563	46273724	8587
16	46264962	46265059	46273567	46273723	8507
17	46265040	46265226	46273571	46273724	8344
18	46265197	46265232	46273561	46273724	8328
19	46265244	46265422	46273567	46273724	8144
20	46266027	46266137	46273571	46273724	7433

DMPK nickase	Chromosome 19 reference span				Estimated deletion
Read pair #	Left alignment		Right alignment		size (bp)
1	46209259	46209340	46273599	46273599	64258
2	46210852	46210917	46273602	46273602	62684
3	46211747	46211796	46273558	46273558	61761
4	46215181	46215320	46273611	46273611	58290
5	46217317	46217406	46273611	46273611	56204
6	46227014	46227072	46273590	46273590	46517
7	46233099	46233157	46273624	46273624	40466
8	46235123	46235235	46273605	46273605	38369
9	46237678	46237802	46273614	46273614	35811
10	46244211	46244330	46273612	46273612	29281
11	46245267	46245398	46273612	46273612	28213
12	46246492	46246541	46273627	46273627	27085
13	46246629	46246754	46273608	46273608	26853
14	46248913	46248984	46273617	46273617	24632
15	46249727	46249849	46273599	46273599	23749
16	46250937	46251086	46273601	46273601	22514
17	46251222	46251334	46273611	46273611	22276
18	46251995	46252114	46273606	46273606	21491
19	46252333	46252435	46273592	46273592	21156
20	46252490	46252613	46273613	46273613	20999

Table S3

Primers and sgRNAs used in this paper

DMPK PCR primers

Fowered	CGACTCCGGGGCCCCGTTGGAAGACT	Reverse	TGCACAAGAAAGCTTTGCAC

Primers for T7 endonuclease I assay

	Locus	Name	Foward	Reverse
on-target	chr19:-46273705	DMPK	CGACTCCGGGGCCCCGTTGGAAGACT	TTCCCGAGTAAGCAGGCAGAG
off-target for 5' guide 3	chr1:+241682887	5' OFT 1	CCGCCCAGAAATTCTACCCAAG	GTTACCTCAAAACGCCCCGG
	chr15:-32274042	5' OFT 2	GACAGGTGCCAGTGGATGTAAAC	CCTGATGGCACACTTAGACTGAC
	chr13:-43724123	5' OFT 3	CTTTACCATCTGTGTGTGCCTCTC	GAAACCAGAAGGGGCTGGTTAAG
	chr20:+2682376	5' OFT 4	ATTTGGCCTGAGCACTTGCAGGG	GCAGTCCTTTCAAGTTGAGGCC
	chr9:-139929364	5' OFT 5	TGCTCACACACCACGGAGCT	TCAGCCTCACACCACCCCAT
off-target for 3' guide 2	chr17:+76322345	3' OFT 1	CTCGGCTTTCACGTGGCCTA	TGCGATTATTCAGTTGGCTCAGGC
	chr10:+7791015	3' OFT 2	CCGTTCCAAACACTAGATCCGTTC	ACTCCTGGCCTCAAGTGGTC
	chr7:+120984995	3' OFT 3	CCCATGATCATGGCCCACAC	GGCACAGTTACAGGAATTGTGGC
	chr16:+63838238	3' OFT 4	CTGGAGAAGCAACAGAGATTCAAGAAAGAC	GGCCATAGTAGAAGTCAGAGGTG
	chr15:+40318690	3' OFT 5	GAGGTGGGAGGATTGCTTGAG	CCTACCCATATGGTTGATACTCCC

sgRNAs for Cas9 nuclease

Name	Protospacer sequence (5' to 3')	PAM	Strand
5' guide 1	CGAGCCCCGTTCGCCGGCCG	CGG	-
5' guide 2	GCCGGCGAACGGGGCTCGAA	GGG	+
5' guide 3	ACCCTTCGAGCCCCGTTCGC	CGG	-
3' guide 1	GCTGAGGCCCTGACGTGGAT	GGG	+
3' guide 2	TTGCGAACCAACGATAGGTG	GGG	-
3' guide 3	GCACTTTGCGAACCAACGAT	AGG	-
RAG1A	GCCTCTTTCCCACCCACCTT	GGG	+

sgRNA pairs for Cas9 nickase (double nicking)

	Upstream			Downstream		
Name	Protospacer sequence (5' to 3')	PAM	Strand	Protospacer sequence (5' to 3')	PAM	Strand
Nick 1	AGGGGGCGGGCCCGGATCAC	AGG	-	CGGGTCCGCGGCCGGCGAAC	GGG	+
Nick 2	CTCCCCGGCCGCTAGGGGGC	GGG	-	GCCGGCGAACGGGGCTCGAA	GGG	+
Nick 3	GCAGTTTGCCCATCCACGTC	AGG	-	GCCTGGGAAGGCAGCAAGCC	GGG	+
Nick 4	AGGATGGAACACGGACGGCC	CGG	-	CACGCACCCCCACCTATCGT	TGG	+
Nick 5	GCGGCTTCTGTGCCGTGCCC	CGG	-	GTTCACAACCGCTCCGAGCG	TGG	+
Nick 6	CGCTCGGAGCGGTTGTGAAC	TGG	-	GATCCGGGCCCGCCCCCTAG	CGG	+
Nick 7	CTCCCCGGCCGCTAGGGGGC	GGG	-	CGGGTCCGCGGCCGGCGAAC	GGG	+
Nick 8	CCTCCCTCCCCGGCCGCTAG	GGG	-	CGGGTCCGCGGCCGGCGAAC	GGG	+
Nick 9	CCTCCCTCCCCGGCCGCTAG	GGG	-	GCCGGCGAACGGGGCTCGAA	GGG	+

sgRNAs for CRISPR interference

Name	Protospacer sequence (5' to 3')	PAM	Strand
i guide 1	AGGAGGCCTCGGCCGGCCGCA	GAGAG	+
i guide 2	GGGGCTCCAGCCCCAGGAAGC	CCGGGT	-
i guide 3	TACGTGGCCGACTTCTTGCAG	TGGGGT	+

Table S4

Primers and adaptors used for LAM-HTGTS	
Bio-LAM-PCR primers	
DMPK	Biotin-GCCAACTCACCGCAGTCTGG
RAG1	Biotin-AGGACTGCTGGAGATTGCTC
Oligos for bridge adaptor	
Upper oligo	GCGACTATAGGGCACGCGTGGNNNNNN-NH_2
Lower oligo	Phosphorylation-CCACGCGTGCCCTATAGTCGC-NH_2
	"N" means random nucleotide.
Nested PCR primers	
DMPK forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCGGGGCCCCGTTGGAAGACT
DMPK reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCGACTATAGGGCACGCGTGG
RAG1 forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAGAGGGTTTCCCCTCAAAG
RAG1 reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCGACTATAGGGCACGCGTGG
Dual-index primers	
517	AATGATACGGCGACCACCGAGATCTACACGCGTAAGATCGTCGGCAGCGTC
504	AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTCGGCAGCGTC
507	AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCGGCAGCGTC
508	AATGATACGGCGACCACCGAGATCTACACCTAAGCCTTCGTCGGCAGCGTC
701	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGG
703	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGG
707	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGG
708	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGG
	Red caracters indicate *DMPK* locus specific sequence.