OMTN, Volume *25*

Supplemental information

MultiEditR: The first tool for the detection and quantification of RNA editing from Sanger sequencing demonstrates comparable fidelity to RNA-seq

Mitchell G. Kluesner, Rafail Nikolaos Tasakis, Taga Lerner, Annette Arnold, Sandra Wüst, Marco Binder, Beau R. Webber, Branden S. Moriarity, and Riccardo Pecori

SUPPLEMENTARY MATERIALS

pJET-CmAG-WT and pJET-CmAG-6x

CTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCC
ACTCCACCGGCGGCATGGACGAGCTGTACAAGaagcttaAACAAGTAGCTGGTGCCAAG
GAAAAAATAACTTCTTTCATGGAAAATTATAGAATTACAGATAATGATGTACTAATT
GCCATAGATAGTGCCAAAATCAACTTCAATGAAAAA**C**TCTCTCAA**C**TTGAGACATA
CGCGATA**C**AATTTGATCAGTATATTAAAGATAATTATGATCCACATGA**C**TTAAAAAG
AA**C**TATTGCTGAGATTATTGAT**C**GAATCATTGAAAAGTTAAAAATTCTTGATGAACA
GTATCATATCCGTGTAAATCTAGCAAAATCAATCCATAATCTCTATTTATTTGTTGAA
AACGTTGATCTTAACCAAGTCAGTAGTAGTAACACCTCTTGGATCCAAAATGTGGAT
TCCAATTATCAAGTCAGAATCCAAATTCAAGAAAAACTACAGCAGCTCAGGACACA
AATTCAGAATATAGACATTCAGCAGCTTGCTGCAGAGGTAAAACGACAG

In **yellow** are highlighted the 6 Cs which are unedited (C) in pJET-CmAG-WT and edited (T) in the pJET-CmAG-6x.

mCherry-mApob-eGFP (CmAG) coding region:

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTT
CAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGG
GCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGT
GGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAG
GCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAG
GGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGAC
CCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCA
CCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCC
TCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAG
GCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGG
CCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATC
ACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCG
CCACTCCACCGGCGGCATGGACGAGCTGTACAAGaagcttaAACAAGTAGCTGGTGCCA
AGGAAAAAATAACTTCTTTCATGGAAAATTATAGAATTACAGATAATGATGTACTAA
TTGCCATAGATAGTGCCAAAATCAACTTCAATGAAAAACTCTCTCAACTTGAGACAT
ACGCGATACAATTTGATCAGTATATTAAAGATAATTATGATCCACATGACTTAAAAA
GAACTATTGCTGAGATTATTGATCGAATCATTGAAAAGTTAAAAATTCTTGATGAAC
AGTATCATATCCGTGTAAATCTAGCAAAATCAATCCATAATCTCTATTTATTTGTTGA
AAACGTTGATCTTAACCAAGTCAGTAGTAGTAACACCTCTTGGATCCAAAATGTGGA

TTCCAATTATCAAGTCAGAATCCAAATTCAAGAAAAACTACAGCAGCTCAGGACAC
AAATTCAGAATATAGACATTCAGCAGCTTGCTGCAGAGGTAAAACGACAGACCCGG
GATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT
GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCG
GCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACC
ACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTG
CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCC
ATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC
TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC
AACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAA
GGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACC
ACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
TACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACAT
GGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTA
CAAGTAA

In red mCherry, in yellow mouse Apob and in green eGFP.

mCherry-Apob-eGFP W58X (CAGX) coding region:

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTT
CAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGG
GCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGT
GGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAG
GCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAG
GGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGAC
CCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCA
CCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCC
TCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAG
GCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGG
CCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATC
ACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCG
CCACTCCACCGGCGGCATGGACGAGCTGTACAAGAAGCTTACCATGGCCAAGGAGA
AACTGACTGCTCTCACAAAAAAGTATAGAATTACAGAAAATGATATACAAATTGCA
TTAGATGATGCCAAAATCAACTTTAATGAAAAACTATCTCAACTGCAGACATATATG
ATACAATTTGATCAGTATATTAAAGATAGTTATGATTTACATGATTTGAAAATAGCT
ATTGCTAATATTATTGATGAAATCATTGAAAAATTAAAAAGTCTTGATGAGCACTAT
CATATCCGTGTAAATTTAGTAAAAACAATCCATGATCTACATTTGTTTATTGAAAAT
ATTGATTTTAACAAAAGTGGAAGTAGTACTGCATCCTGGATTCAAAATGTGGATACT

AAGTACCAAATCAGAATCCAGATACAAGAAAAACTGCAGCAGCTTAAGAGACACAT
ACAGAATATAGACATCCAGCACCTAGCTGGAATTCTGCAGTCGACGGTACCGCGGG
CCCGGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGG
GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGT
GTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCT
GCACCACCGGCAAGCTGCCCGTGCCCTAGCCCACCCTCGTGACCACCCTGACCTACG
GCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGT
CCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC
ATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGC
CGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACA
ACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGAT
CACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAG
CTGTACAAGTAA

In red mCherry, in yellow human Apob and in green eGFP. In light blue is highlighted W58X mutation and in red the A to edit to G to re-activate the eGFP.

SUPPLEMENTARY TABLES

Supplementary Table 1 - List of all the regions within transcripts analysed with REDItools v1 for NGS comparison with MultiEditR. For each transcript strandness, reference genome and chromosomal coordinates of the amplicon are indicated.

Gene (strand)	Reference Genome	Chromosomal coordinates of the amplicon
ARSD (reg.1) (-)	hg19	chrX:2,824,741-2,825,282
ARSD (reg.2) (-)	hg19	chrX:2,824,114-2,824,494
CTSS (-)	hg19	chr1:150,704,237-150,704,783
DDX58 (-)	hg19	chr9:32,455,300-32,502,734
MAVS (+)	hg19	chr20:3,827,446-3,856,770
SSR3 (-)	hg19	chr3:156,257,929-156,272,973
Slc39a10 (-)	mm10	chr1:46,807,710-46,808,199
Aldoc (+)	mm10	chr11:78,327,206-78,327,625
Atp6ap2 (+)	mm10	chrX:12,615,885-12,616,656
Serinc1 (-)	mm10	chr10:57,515,845-57,516,141
B2m (+)	mm10	chr2:122,147,686-122,153,083
Rab7 (-)	mm10	chr6:87,999,398-87,999,968
Reep5 (-)	mm10	chr18:34,346,596-34,347,035
Casp6 (+)	mm10	chr3:129,913,560-129,914,011

Supplementary Table 2 - List of all the regions within transcripts analysed by Amplicon-seq. For each transcript strandness, reference genome and chromosomal coordinates of the amplicon are indicated.

Gene (strand)	Reference Genome	Chromosomal coordinates of the amplicon
ARSD (-)	hg19	chrX:2824741-2825282
CTSS (-)	hg19	chr1:150704237-150704783
DDX58 (-)	hg19	chr9:32456205-32456648
MAVS (+)	hg19	chr20:3851441-3851945
Slc39a10 (-)	mm10	chr1:46807710-46808199
Aldoc (+)	mm10	chr11:78327206-78327625
Atp6ap2 (+)	mm10	chrX:12616190-12616656
Serinc1 (-)	mm10	chr10:57515845-57516325
B2m (+)	mm10	chr2:122152650-122153068
Rab7 (-)	mm10	chr6:87999398-87999968

Supplementary Table 3 - List of all the primers and oligos used in this study. For all the primers and oligos listed above lowercase bases represent the homology part needed during the cloning (#1-8, 17) or the adaptor sequence needed for Amplicon-seq (#57-76). For oligos #13-14 the base in bold is the one modified during the site-directed mutagenesis. In oligo #17 gRNA the bases highlighted in green codify the boxB hairpin. Oligos #27-56 were used to generate the PCR fragments used in the comparison of MultiEditR and NGS methods (Figure 2 and S4).

1	mApob_F	ggcggcatggacgagctgtacaagAAGCTTAAACAAGTAGCTGGTGCCAAGGAA
2	mApob_R	accatggtggcgaccggtggatcccgggtCTGTCGTTTTACCTCTGCAGCAAG
3	RBM47_F	ttagtgaaccgtcagatccgctagcATGACTGCTGAAGATTCCGCC
4	RBM47_R	agagtcgcggccgctttacttgtacaTCAGTAAGTCTGGTAGACGTCG
5	CRISPR sgRNA ADAR1ko_ exon4_F	caccGGACAGGAGACGGAATTCGC
6	CRISPR sgRNA ADAR1ko_ exon4_R	aaacGCGAATTCCGTCTCCTGTCC
7	non-targeting sgRNA_F	caccGTATTACTGATATTGGT
8	non-targeting sgRNA_R	aaacACCAATATCAGTAATAC
9	A1KO_11_fw	CACCGAGCAAGATGAGTTCCGAGAC

10	A1KO_11_rv	AAACGTCTCGGAACTCATCTTGCTC
11	A1KO_39_fw	CACCGTAGCTGTTGATCCCACTCTG
12	A1KO_39_rv	AAACCAGAGTGGGATCAACAGCTAC
13	W58X_uAg_F	CTGCCCGTGCCCTAGCCCACCCT
14	W58X_uAg_R	AGGGTGGGCTAGGGCACGGGCAG
15	pENTR_plasm_F	TTTTTTCTAGACCCAGCTTTCTTGTA
16	pENTR_plasm_R	GGTGTTTCGTCCTTTCCACA
17	Oligo_gRNA_W58X	tggaaaggacgaaacaccTCAGGGTGGTGGCCCTGAAAAAGGGCCGAGGGTGGGCCAGGGCACGGGCAGCTTGCGGCCCTGAAAAAGGGCCTGGTGCAGATttttttctagacccagctt
18	Apob_F	TGCATCCTGGATTCAAAATGTGG
19	eGFP_R	TTGAAGTCGATGCCCTTCAG
20	A1_KO1A_fw	CATTGATGGCTCTGTGGGTGTTC
21	A1_KO1A_rv	GCTGAAAAGCACCCAGGGAC
22	A1_KO2A_fw	GTACCTCTCAGATCCTTTGAGAAGTC

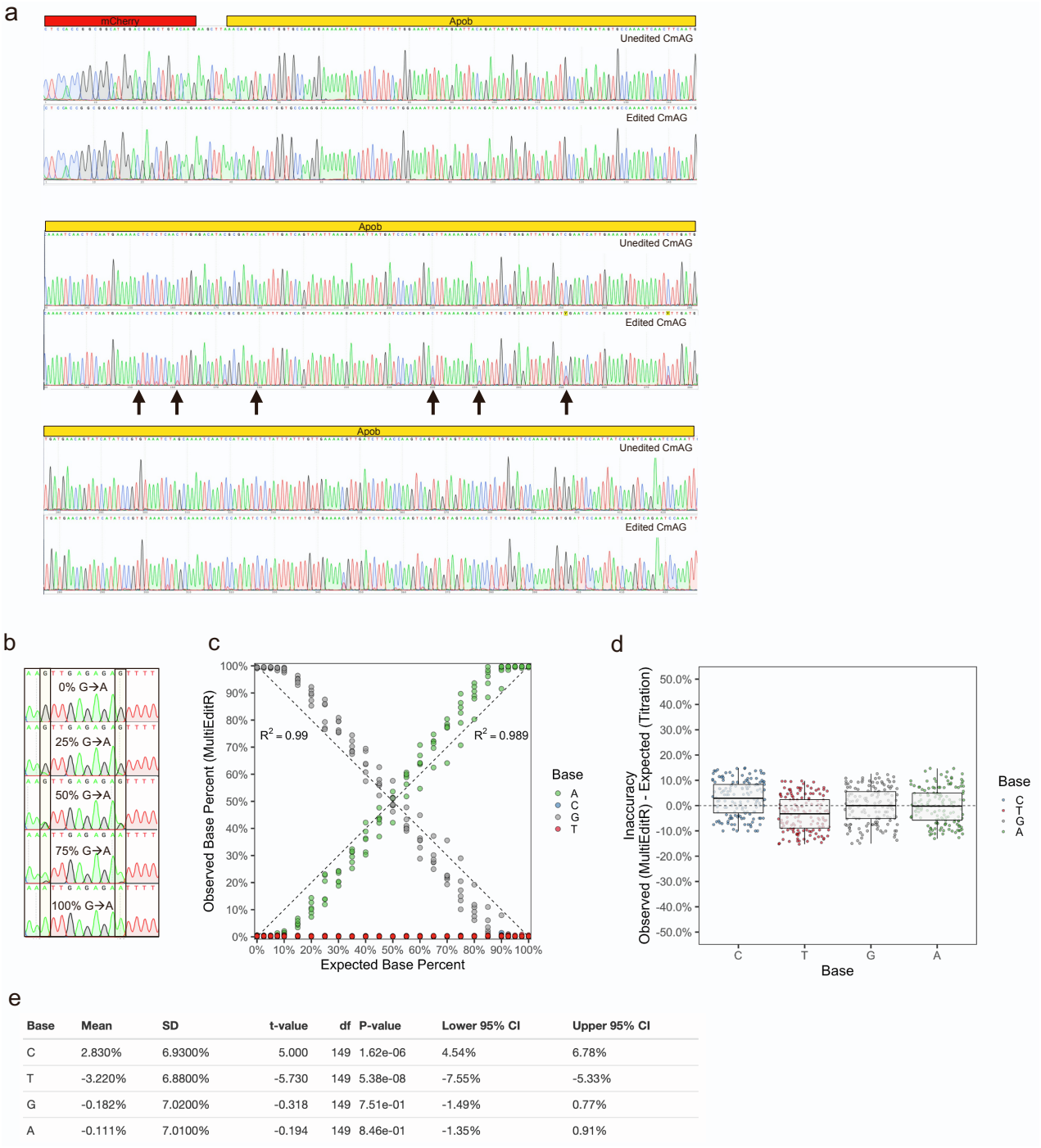
23	A1_KO2A_rv	GCATGCTGTAACCCTGTAGTTC
24	B2m_fw	CAAGCATCATGATGCTCTGAAG
25	B2m_rv	GTAAAAGTAACAAAAGCAGAAGTAGCC
26	mCherry-mApob-eGFP_SEQ	CTCCCACAACGAGGACTACACC
27	MAVS_F	TACCCTGCCTGGCCTCAAACTATTA
28	MAVS_R	ACTTCATGCTGTCTGGGAGCAA
29	DDX58_F	ATTTGGCCCTGTTGAGCACTCT
30	DDX58_R	ACGTCCAGGAAACCGCAAACTA
31	SSR3_F	TGGGCTCCATGCAAGAACTTGGAA
32	SSR3_R	GCACTGCCATCTAGTGGCAAGTTT
33	ARSD_reg1_F	GAGCCTGACTGCGTTGCAAACAAA
34	ARSD_reg1_R	AAGACGGAGGGGTGAAAACATCG
35	ARSD_reg2_F	CACCCCTCCGTCTTAAAGCATTGT

36	ARSD_reg2_R	CCCTTGATCGGCTCATTACTTGGA
37	CTSS_F	CTGGTGATGGGGTTTAGCAACT
38	CTSS_R	TGCCAAATAAAGGCCCTGGTCA
39	Aldoc_F	AAGGGCTATGACCCACTTCCATGT
40	Aldoc_R	CGATTCCAATTCGAGCGATTGAGG
41	Atp6ap2_F	GCTAGACTTAGACAACAGGTTTGG
42	Atp6ap2_R	GTCTACAGATTGAAGCCATACCAC
43	B2m_F	CAAGCATCATGATGCTCTGAAG
44	B2m_R	GTAAAAGTAACAAAAGCAGAAGTAGCC
45	Casp6_F	AACCTAGCAAGTAGGGCCATCTGT
46	Casp6_R	CATGACCAAGTCAAATAGGCCCAC
47	Cd36_F	CTGGCTACATCTTTGGTAAAGCCG
48	Cd36_R	GGGCCACCCCAGTCATGATAG
49	Rab7_F	TCCGTTCTGAGCAGGCTGTTTTGT

50	Rab7_R	TCTGGTGGGTTCTCCTTTCCTCTT
51	Reep5_F	GCCTTGGAAACTTCCCGCTGTATT
52	Reep5_R	GTCTCCCTTCTCAGGTCACAGGT
53	Serinc1_F	TGTATCGCTGCTGTCCAGCAT
54	Serinc1_R	GGCTGGAACATGAAGATGAACTGC
55	Slc39a10_F	ACTGCGGGAATATACTCCCACTCT
56	Slc39a10_R	GCAAGTATCACCTTGCAGGACCAT
57	ARSD_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagGAGCCTGACTGCGTTGCAAACAAA
58	ARSD_NGS_R	gtctcgtgggctcggagatgtgtataagagacagAAGACGGAGGGGTGAAAACATCG
59	CTSS_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagCTGGTGATGGGGTTTAGCAACT
60	CTSS_NGS_R	gtctcgtgggctcggagatgtgtataagagacagTGCCAAATAAAGGCCCTGGTCA
61	DDX58_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagATTTGGCCCTGTTGAGCACTCT
62	DDX58_NGS_R	gtctcgtgggctcggagatgtgtataagagacagGAGAACAATGGCACACGTTAAGAG
63	MAVS_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagCCATGAGCCCATCACCCAAC

64	MAVS_NGS_R	gtctcgtgggctcggagatgtgtataagagacagACTTCATGCTGTCTGGGAGCAA
65	Slc39a10_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagACTGCGGGAATATACTCCCACTCT
66	Slc39a10_NGS_R	gtctcgtgggctcggagatgtgtataagagacagGCAAGTATCACCTTGCAGGACCAT
67	Aldoc_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagAAGGGCTATGACCCACTTCCATGT
68	Aldoc_NGS_R	gtctcgtgggctcggagatgtgtataagagacagCGATTCCAATTCGAGCGATTGAGG
69	Atp6ap2_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagGGAGAACGCACTGGGTTTCTTA
70	Atp6ap2_NGS_R	gtctcgtgggctcggagatgtgtataagagacagGTCTACAGATTGAAGCCATACCAC
71	Serinc1_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagAGGCTCGGGTTAGGCACTAAGATA
72	Serinc1_NGS_R	gtctcgtgggctcggagatgtgtataagagacagGGCTGGAACATGAAGATGAACTGC
73	B2m_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagcaagcatcatgatGCTCTGAAG
74	B2m_NGS_R	gtctcgtgggctcggagatgtgtataagagacagGTAAAAGTAACAAAAGCAGAAGTAGCC
75	Rab7_NGS_F	tcgtcggcagcgtcagatgtgtataagagacagTCCGTTCTGAGCAGGCTGTTTTGT
76	Rab7_NGS_R	gtctcgtgggctcggagatgtgtataagagacagTCTGGTGGGTTCTCCTTTCCTCTT

SUPPLEMENTARY FIGURES



e

Base	Mean	SD	t-value	df	P-value	Lower 95% CI	Upper 95% CI
C	2.830%	6.9300%	5.000	149	1.62e-06	4.54%	6.78%
T	-3.220%	6.8800%	-5.730	149	5.38e-08	-7.55%	-5.33%
G	-0.182%	7.0200%	-0.318	149	7.51e-01	-1.49%	0.77%
A	-0.111%	7.0100%	-0.194	149	8.46e-01	-1.35%	0.91%

Fig. S1 | C-to-U editing of Apob by APOBEC1 within CmAG plasmid.

a, Sanger sequencing chromatograms from HEK293T cells transfected only with CmAG or with CmAG, APOBEC1 and RBM47. Amplification of Apob region from CmAG transcript and

following Sanger sequencing show abundant C-to-T editing in several sites along the sequenced area. Black arrows point to the 6 Cs which are unedited (C) in pJET-CmAG-WT and edited (T) in the pJET-CmAG-6x (Supplementary materials and Fig. 1a). **b**, Chromatograms from G-to-A titration showing a change in peak height at two sites. **c**, Titration of pJET-CmAG-WT with pJET-CmAG-6x sequenced with the forward primer, coefficient of determination was calculated relative to the identity line y = x, N = 6 sites per chromatogram. **d**, Inaccuracy of MultiEditR relative to expected titration values, significance was determined using Student's one-sample t-test relative to an inaccuracy of 0%. **e**, Table of results from Student's one-sample t-test of inaccuracies.

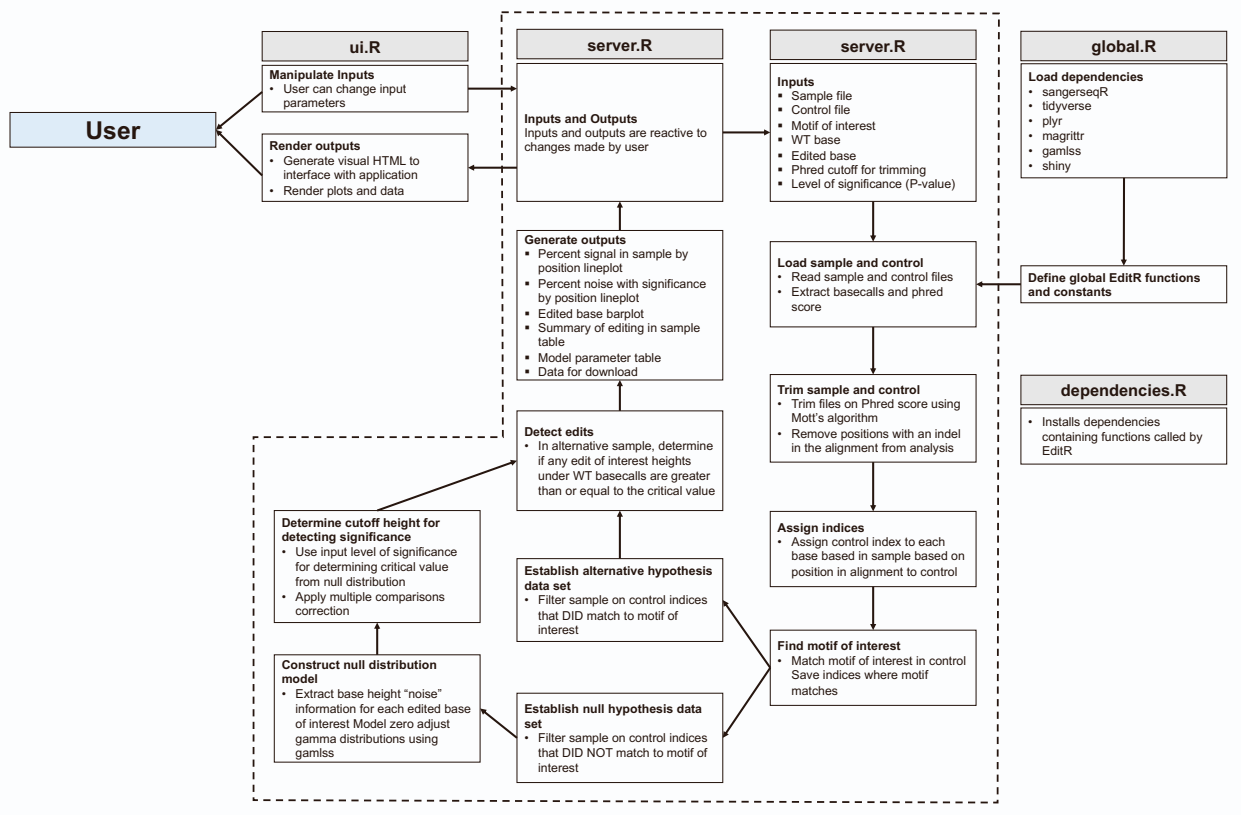


Fig. S2 | MultiEditR algorithm.

MultiEditR algorithm flow chart in the context of the MultiEditR web application made using R Shiny (z.umn.edu/multieditr). Source code available at:

https://github.com/MoriarityLab/MultiEditR

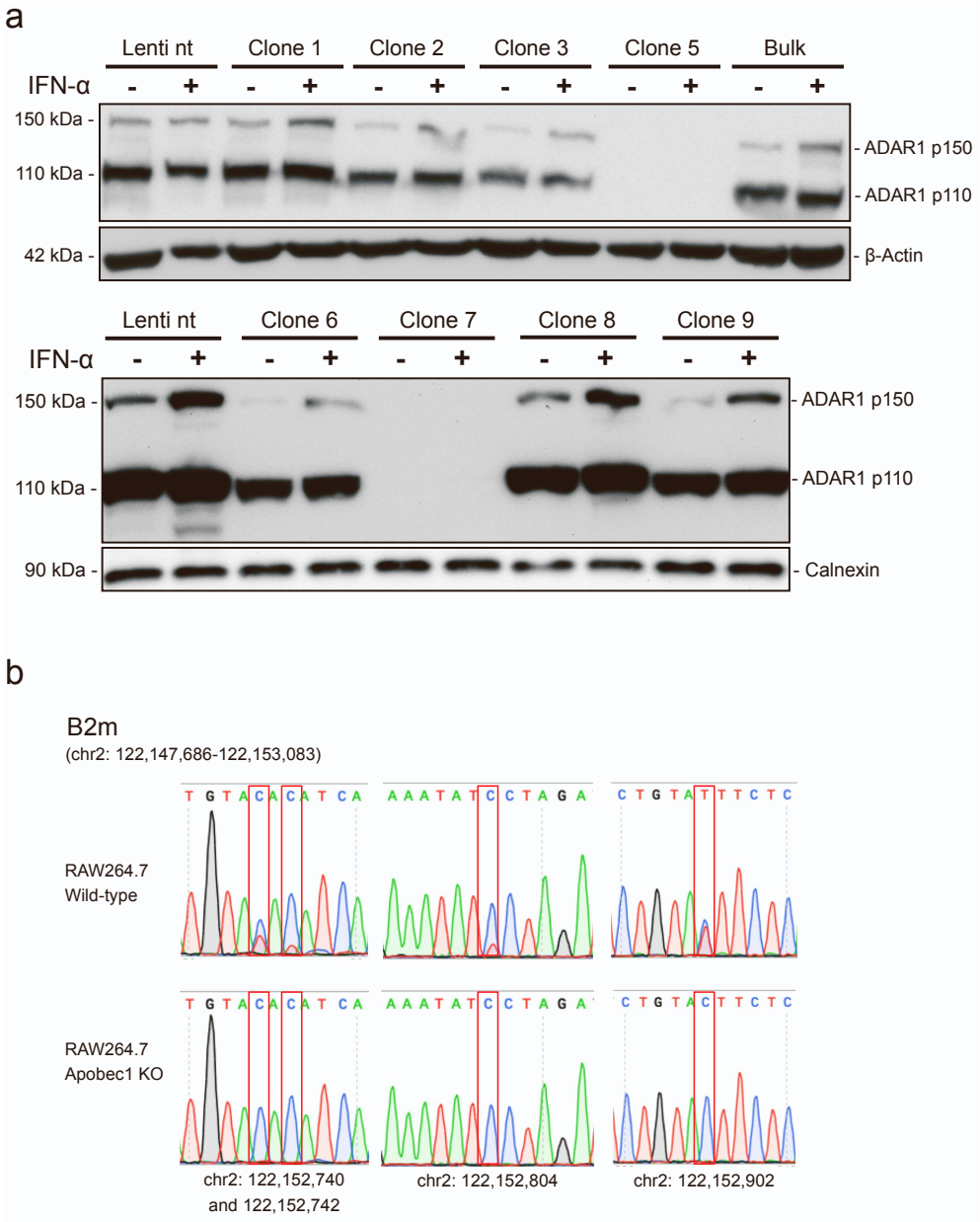
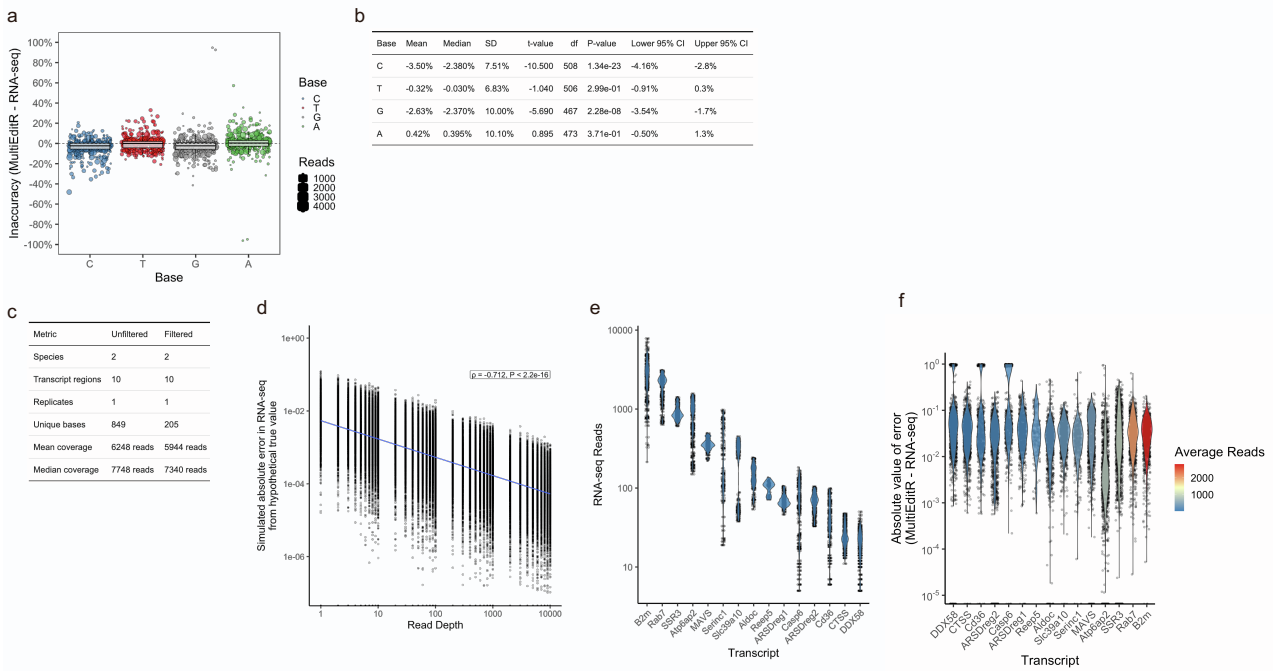


Fig. S3 | Generation of human ADAR1 and murine APOBEC1 KO cell lines.

a, Western blot of A549 KO clones and bulk, after transduction with lenti-CRISPR-ADAR1 exon 4. A549 cells were stimulated with 200 U/ml of IFN-α (pbl assay science, cat. no.#11100-1) for 16 h to evaluate the presence of ADAR1 p150. "Lenti nt" is the non-targeting control and was generated by transducing cells with the same vector (LentiCRISPRv2) but containing a non-targeting sgRNA (see Supplementary Table 3). β-Actin and Calnexin were used as loading controls

(Mouse β-Actin monoclonal antibody, Sigma-Aldrich, cat. no.#A5441; Rabbit Calnexin polyclonal antibody, Enzo Life science, cat. no.#ADI-SPA-865-F). **b**, Sanger sequencing chromatograms for RAW 264.7 wild-type and APOBEC1 knock-out after RT-PCR amplification of B2m 3'UTR region (oligos #24-25). The complete absence of C-to-U editing along the sequenced region confirms the deficiency of APOBEC1 editing activity within the knock-out clone.



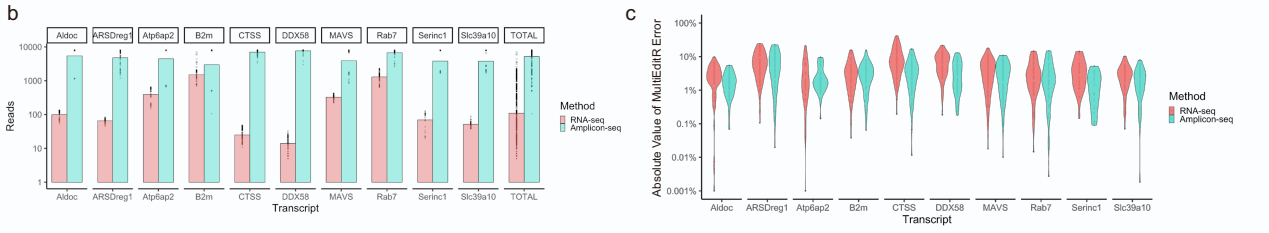
Base	Mean	Median	SD	t-value	df	P-value	Lower 95% CI	Upper 95% CI
C	-3.50%	-2.380%	7.51%	-10.500	508	1.34e-23	-4.16%	-2.8%
T	-0.32%	-0.030%	6.83%	-1.040	506	2.99e-01	-0.91%	0.3%
G	-2.63%	-2.370%	10.00%	-5.690	467	2.28e-08	-3.54%	-1.7%
A	0.42%	0.395%	10.10%	0.895	473	3.71e-01	-0.50%	1.3%

Fig. S4 | Descriptive statistics, coverage and error of samples in RNA-seq data set.

a-b, Comparison of MultiEditR to RNA-seq measurements. All data is filtered on between 1% and 99% editing measured by either method. Coefficients of determination represent regression to the identity line. Dot size is proportional to read coverage at the base of interest in RNA-seq. Tables represent results of Student's pairwise t-test between measurement methods. **c**, Descriptive statistics of samples in RNA-seq data set. **d**, Simulation of expected absolute error in RNA-seq editing value from true editing value as a function of read depth. Absolute error of editing measurement is negatively correlated with read depth (Spearman's correlation coefficient, $\rho = -0.712$, $P < 2.2e\text{-}16$). **e**, Violin plots of read depth at each base for each transcript analyzed. **f**, Violin plots of absolute value of error in MultiEditR editing measurement minus REDItools RNA-seq editing measurement by each transcript. Plots are colored by average read depth.

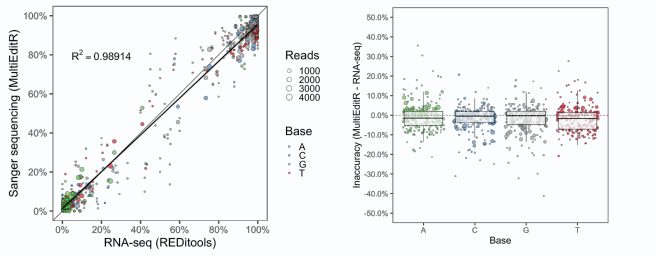
a

Metric	Unfiltered	Filtered
Species	2	2
Transcript regions	10	10
Replicates	1	1
Unique bases	849	205
Mean coverage	6248 reads	5944 reads
Median coverage	7748 reads	7340 reads



d

Transcript	P-value	Mean difference	Mean RNA-seq Reads	SD RNA-seq Reads	Mean Amplicon-seq Reads	SD Amplicon-seq Reads
Aldoc	1.60e-06	0.922%	103.0	25.40	6610	2750
ARSDreg1	8.29e-02	0.592%	65.6	7.46	5700	2480
Atp6ap2	1.61e-02	0.828%	437.0	171.00	6220	3120
B2m	1.32e-01	0.291%	1710.0	1060.00	5360	3590
CTSS	0.00e+00	5.480%	26.9	10.40	7110	1060
DDX58	0.00e+00	3.340%	15.5	6.31	7670	739
MAVS	3.97e-04	0.566%	331.0	48.10	5160	3070
Rab7	6.02e-04	0.244%	1380.0	474.00	6900	1290
Serinc1	5.00e-07	2.740%	80.2	34.30	4880	3060
Slc39a10	2.52e-05	0.785%	52.5	9.35	4610	2780



g

Base	Mean	Median	SD	t-value	df	P-value	Lower 95% CI	Upper 95% CI
C	-1.79%	-0.49%	6.57%	-3.90	205	1.28e-04	-2.69%	-0.885%
T	-1.92%	-1.72%	6.89%	-4.06	211	7.01e-05	-2.85%	-0.986%
G	-2.08%	-0.31%	7.69%	-3.86	202	1.54e-04	-3.15%	-1.020%
A	-0.73%	-1.48%	7.46%	-1.41	209	1.59e-01	-1.74%	0.287%

h

Base	Mean	Median	SD	t-value	df	P-value	Lower 95% CI	Upper 95% CI
C	-2.08%	-1.280%	4.350%	-6.850	205	8.37e-11	-2.67%	-1.48%
T	-1.66%	-1.270%	4.720%	-5.110	211	7.08e-07	-2.30%	-1.02%
G	-2.59%	-1.740%	5.110%	-7.230	202	9.96e-12	-3.30%	-1.89%
A	-0.25%	-0.831%	5.070%	-0.715	209	4.76e-01	-0.94%	0.44%

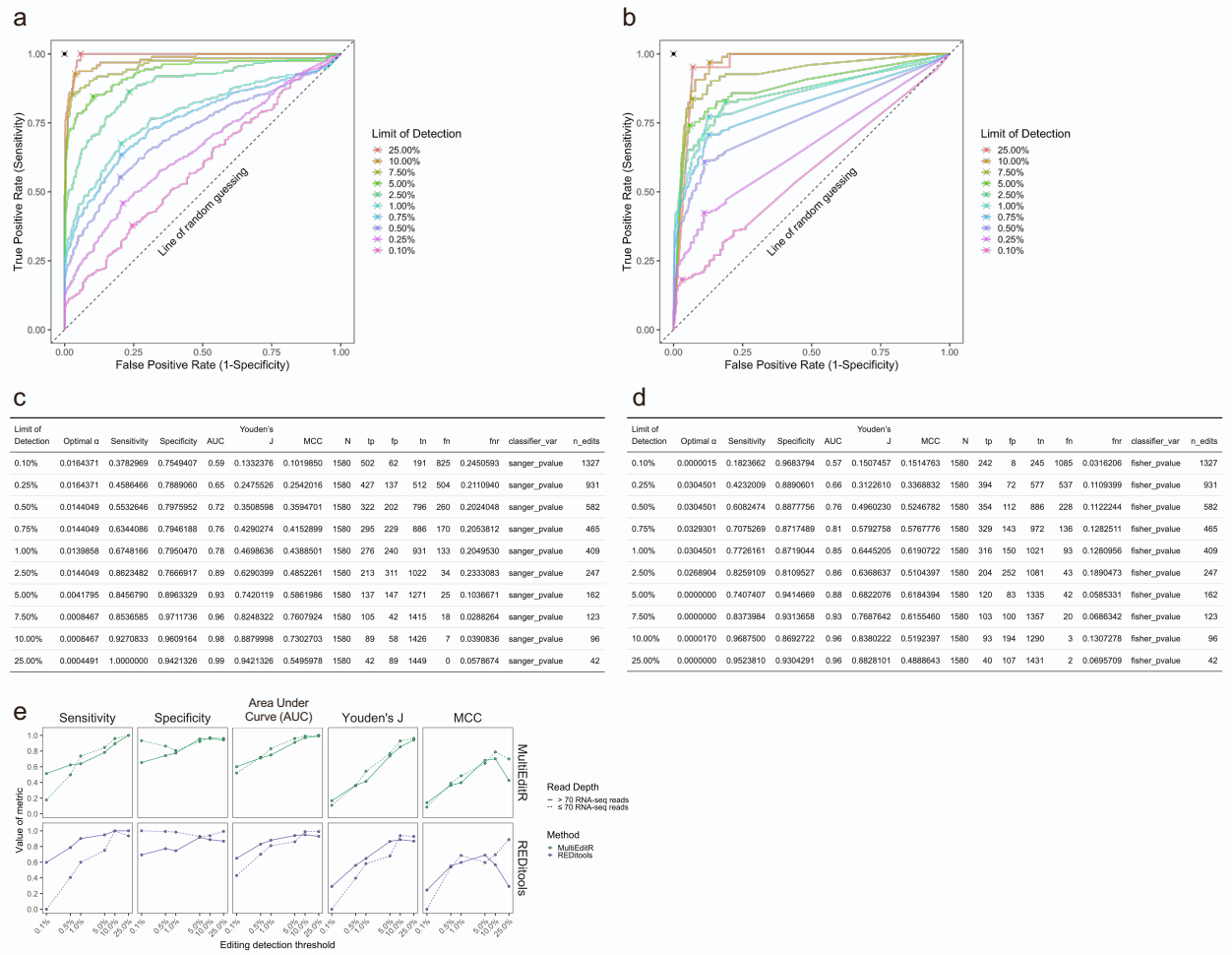
i

Base	Mean	Median	SD	t-value	df	P-value	Lower 95% CI	Upper 95% CI
C	-0.29%	-0.32%	5.370%	-0.774	205	4.40e-01	-1.03%	0.45%
T	0.26%	0.50%	5.410%	0.700	211	4.85e-01	-0.47%	0.99%
G	-0.51%	-0.56%	5.550%	-1.310	202	1.92e-01	-1.28%	0.26%
A	0.48%	0.60%	5.570%	1.240	209	2.15e-01	-0.28%	1.24%

Fig. S5 | Descriptive statistics, coverage and error of samples in the combined RNA-seq and Amplicon-seq data set.

a, Descriptive statistics of samples in the combined data set. **b**, Barplots of read depth in each sample for RNA-seq and Amplicon-seq data for each transcript. Amplicon-seq read depth is higher for every transcript. **c**, Violin plots of absolute value of error in MultiEditR editing measurements minus REDItools RNA-seq or Amplicon-seq editing measurements. **d**, Table comparison of RNA-seq to Amplicon-seq editing values for each transcript. **e-f**, Comparison of MultiEditR to RNA-seq measurements. Data is filtered on between 1% and 99% editing measured by Amplicon-seq. Coefficients of determination represent regression to the identity line. Dot size is proportional to read coverage at the base of interest in RNA-seq. **g**, Tables of Student's pairwise t-test results

between MultiEditR and RNA-seq. **h**, Tables of Student's pairwise t-test results between MultiEditR and Amplicon-seq. **i**, Tables of Student's pairwise t-test results between RNA-seq and Amplicon-seq.



c

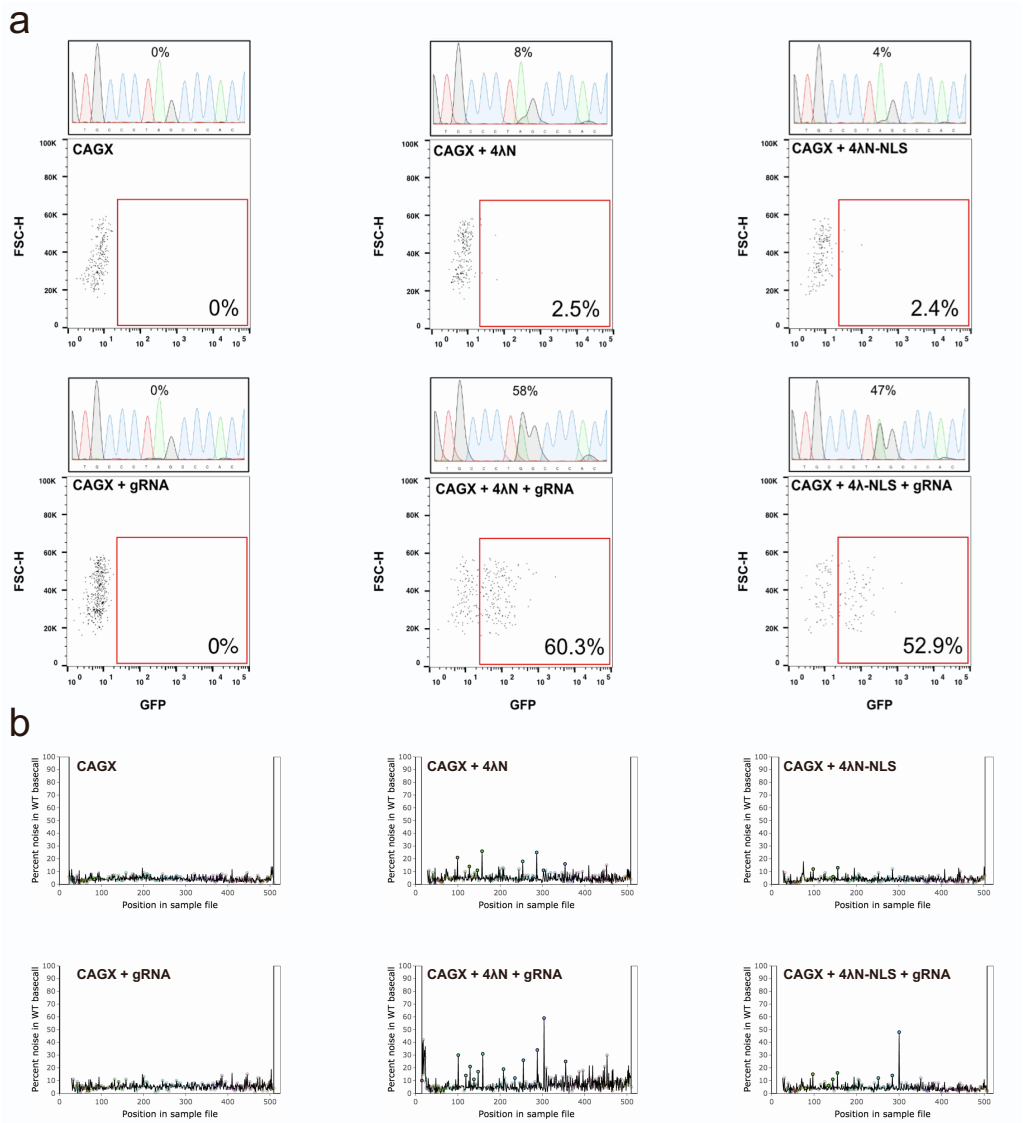
Limit of Detection	Optimal α	Sensitivity	Specificity	AUC	Youden's J	MCC	N	tp	fp	tn	fn	fnr	classifier_var	n_edits
0.10%	0.0164371	0.3782969	0.7549407	0.59	0.1332376	0.1019850	1580	502	62	191	825	0.2450593	sanger_pvalue	1327
0.25%	0.0164371	0.4586466	0.7889060	0.65	0.2475526	0.2542016	1580	427	137	512	504	0.2110940	sanger_pvalue	931
0.50%	0.0144049	0.5532646	0.7975952	0.72	0.3508598	0.3594701	1580	322	202	796	260	0.2024048	sanger_pvalue	582
0.75%	0.0144049	0.6344086	0.7946188	0.76	0.4290274	0.4152899	1580	295	229	886	170	0.2053812	sanger_pvalue	465
1.00%	0.0139858	0.6748166	0.7950470	0.78	0.4698636	0.4388501	1580	276	240	931	133	0.2049530	sanger_pvalue	409
2.50%	0.0144049	0.8623482	0.7666917	0.89	0.6290399	0.4852261	1580	213	311	1022	34	0.2333083	sanger_pvalue	247
5.00%	0.0041795	0.8456790	0.8963329	0.93	0.7420119	0.5861986	1580	137	147	1271	25	0.1036671	sanger_pvalue	162
7.50%	0.0008467	0.8536585	0.9711736	0.96	0.8248322	0.7607924	1580	105	42	1415	18	0.0288264	sanger_pvalue	123
10.00%	0.0008467	0.9270833	0.9609164	0.98	0.8879998	0.7302703	1580	89	58	1426	7	0.0390836	sanger_pvalue	96
25.00%	0.0004491	1.0000000	0.9421326	0.99	0.9421326	0.5495978	1580	42	89	1449	0	0.0578674	sanger_pvalue	42

d

Limit of Detection	Optimal α	Sensitivity	Specificity	AUC	Youden's J	MCC	N	tp	fp	tn	fn	fnr	classifier_var	n_edits
0.10%	0.0000015	0.1823662	0.9683794	0.57	0.1507457	0.1514763	1580	242	8	245	1085	0.0316206	fisher_pvalue	1327
0.25%	0.0304501	0.4232009	0.8890601	0.66	0.3122610	0.3368832	1580	394	72	577	537	0.1109399	fisher_pvalue	931
0.50%	0.0304501	0.6082474	0.8877756	0.76	0.4960230	0.5246782	1580	354	112	886	228	0.1122244	fisher_pvalue	582
0.75%	0.0329301	0.7075269	0.8717489	0.81	0.5792758	0.5767776	1580	329	143	972	136	0.1282511	fisher_pvalue	465
1.00%	0.0304501	0.7726161	0.8719044	0.85	0.6445205	0.6190722	1580	316	150	1021	93	0.1280956	fisher_pvalue	409
2.50%	0.0268904	0.8259109	0.8109527	0.86	0.6368637	0.5104397	1580	204	252	1081	43	0.1890473	fisher_pvalue	247
5.00%	0.0000000	0.7407407	0.9414669	0.88	0.6822076	0.6184394	1580	120	83	1335	42	0.0585331	fisher_pvalue	162
7.50%	0.0000000	0.8373984	0.9313658	0.93	0.7687642	0.6155460	1580	103	100	1357	20	0.0686342	fisher_pvalue	123
10.00%	0.0000170	0.9687500	0.8692722	0.96	0.8380222	0.5192397	1580	93	194	1290	3	0.1307278	fisher_pvalue	96
25.00%	0.0000000	0.9523810	0.9304291	0.96	0.8828101	0.4888643	1580	40	107	1431	2	0.0695709	fisher_pvalue	42

Fig. S6 | Receiver Operating Characteristics (ROC) analysis of MultiEditR and REDItools RNA-seq, compared to REDItools Amplicon-seq.

a, MultiEditR ROC across different limits of detection. Hashed points represent optimal cutpoint. **b**, REDItools RNA-seq ROC across different limits of detection. Hashed points represent optimal cutpoint. **c**, MultiEditR ROC analysis parameter table across levels of detection. **d**, REDItools RNA-seq ROC analysis parameter table across levels of detection. **e**, Lineplots for ROC parameters for MultiEditR and REDItools across limits of detection, broken up by regions that had high RNA seq coverage (>70 reads), and low RNA-seq coverage (≤ 70 reads). MultiEditR is more robust in regions that have poor coverage in RNA-seq.



c

enzyme	guide	target_edit	gfp	mei	n_sig_edits	ave_sig_edits	ave_nontarget_edits	phred	p_value	target_MEI_ratio	target_nontarget_ratio	norm_target_MEI_ratio
CAGX	- gRNA	0.01	0.000	0.0151	0	0.0000	0.0000	1e-04	1e-04	0.662	0.000	1.00
CAGX + 4λN	- gRNA	0.08	0.025	0.0317	5	0.1720	0.1560	1e-03	1e-04	2.520	0.514	3.81
CAGX + 4λN-NLS	- gRNA	0.04	0.024	0.0179	3	0.0775	0.0642	1e-04	1e-04	2.230	0.623	3.37
CAGX	+ gRNA	0.03	0.000	0.0309	0	0.0000	0.0000	1e-04	1e-04	0.971	0.000	1.47
CAGX + 4λN	+ gRNA	0.58	0.603	0.0484	8	0.2570	0.1740	3e-03	1e-04	12.000	3.330	18.10
CAGX + 4λN-NLS	+ gRNA	0.47	0.529	0.0260	4	0.1990	0.0819	3e-03	1e-04	18.100	5.740	27.30

Fig. S7 | Reactivation of eGFP by RNA targeted editing.

a, FACS analysis for GFP positive cells after transfection with 4λN and 4λN-NLS in presence and absence of the gRNA to re-activate eGFP by RNA base-editing. The percentage in the plot represents the GFP+ cells of the mCherry+ cells detected by FACS. Above each FACS plot, it is shown A-to-G editing at position W58X for that sample; the percentage of editing is calculated by MultiEditR. **b**, Plots generated in MultiEditR showing the distribution of A-to-G editing along the sequenced region. From this plot appears evident that 4λN has a higher editing activity which leads to a more prominent RNA off-target effect both in presence and absence of the gRNA. **c**, MultiEditR parameters used in this analysis.