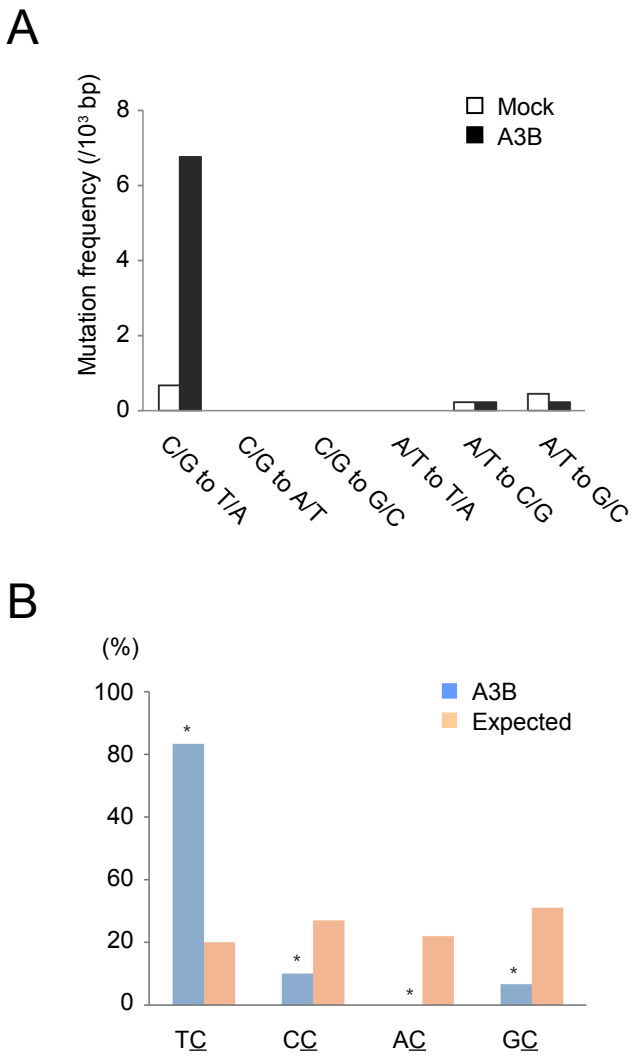
SUPPLEMENTARY MATERIALS

APOBEC3B can impair genomic stability by inducing base substitutions in genomic DNA in human cells

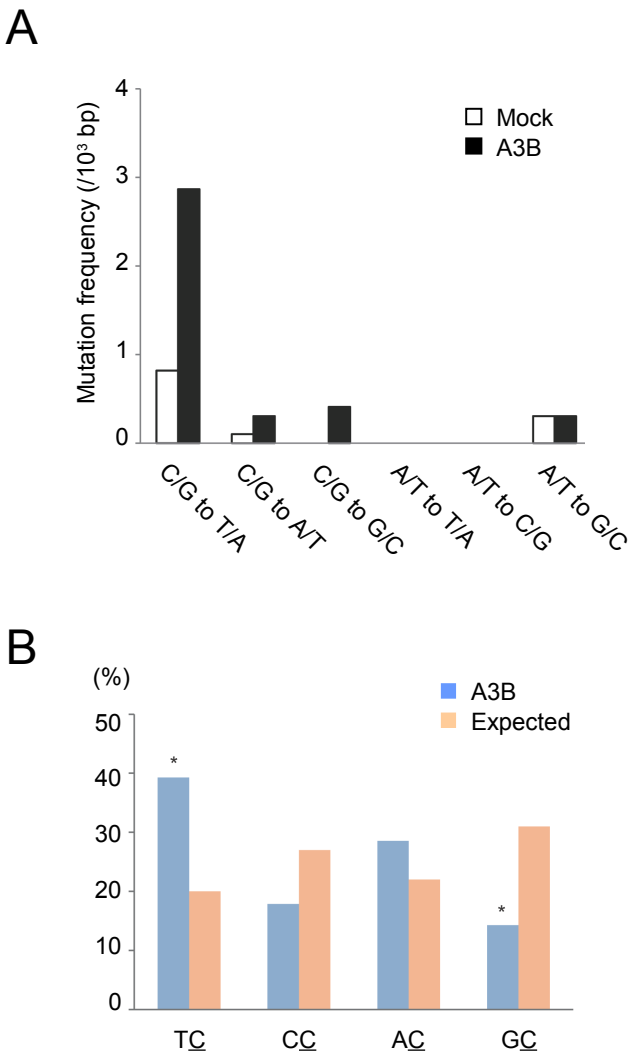
Masanobu Shinohara, Katsuhiro Io, Keisuke Shindo, Masashi Matsui, Takashi Sakamoto, Kohei Tada, Masayuki Kobayashi, Norimitsu Kadowaki & Akifumi Takaori-Kondo

Department of Hematology and Oncology, Graduate school of medicine, Kyoto University

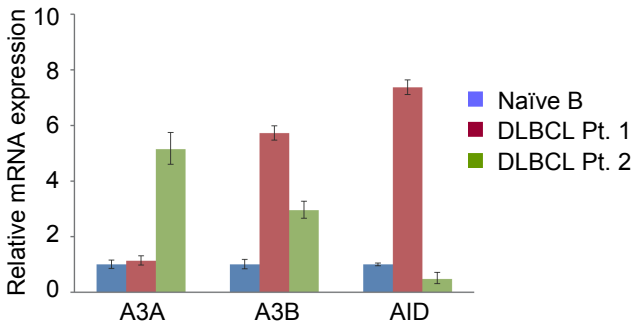
Correspondence should be addressed to K.S. (shind009@kuhp.kyoto-u.ac.jp)



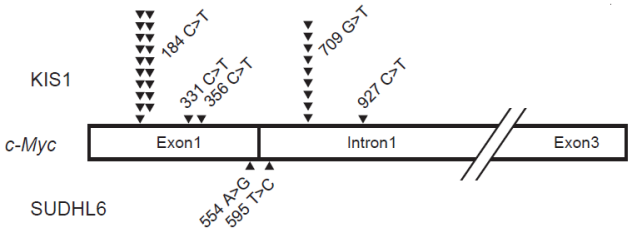
Supplementary Figure S1 Foreign DNA editing by A3B. (A) Frequencies of base substitutions in transfected *EGFP* genes amplified at 94 °C of denaturation temperature (Td). All types of substitutions per 1,000 sequenced base pairs in mock- and A3B WT-expressing HEK293 cells are shown. 8 clones in each group are sequenced (3560 base pairs in total). (B) Dinucleotide contexts in the foreign DNA editing in the amplicons at 94 °C of Td. The rates of indicated dinucleotide sequence at the C to T transitions are shown. Asterisks indicate statistical significance in a $\chi 2$ test ($p < 0.01$).



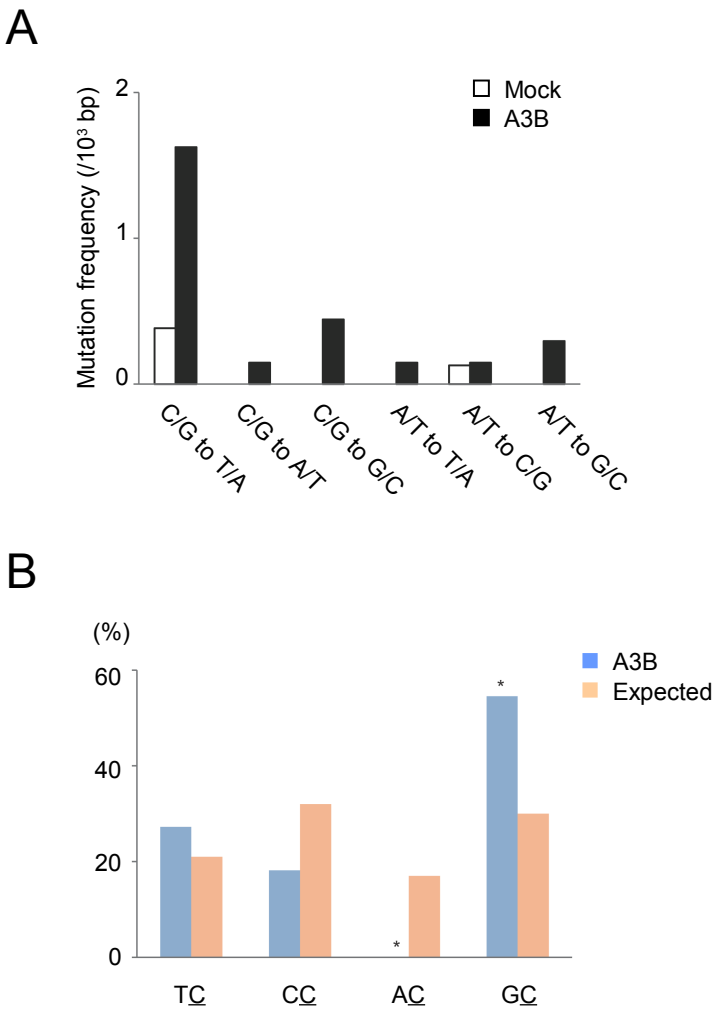
Supplementary Figure S2 Genomic DNA editing by A3B. (A) Frequencies of C/G to T/A transitions in *EGFP* genes integrated in the genomic DNA in HEK293 cells. C/G to T/A transitions per 1,000 sequenced base pairs are shown. (B) Dinucleotide contexts in the genomic DNA editing by A3B in the amplicons at 94 °C of Td. The rates of indicated dinucleotide sequence at the C to T transitions are shown. Asterisks indicate statistical significance in a $\chi 2$ test ($p < 0.01$).



Supplementary Figure S3 Expressions of *A3A*, *A3B*, and *AID* in primary cells from lymphoma patients. Quantitative RT-PCR for *A3A*, *A3B*, and *AID* in primary cells derived from two patients with diffuse large B-cell lymphoma (DLBCL). The levels of target cDNA were normalized to the endogenous glyceraldehyde 3-phosphate dehydrogenase and then compared to those in naïve B-cells isolated from peripheral blood of healthy donor. The analyses were performed in triplicate and standard deviations are shown.



Supplementary Figure S4 Distribution of base alterations in *c-Myc* gene in lymphoma cell lines. A portion of *c-Myc* sequence of 899 base pairs in length containing exon 1 and intron 1 from KIS1 and SUDHL6 cells were amplified and cloned. We sequenced 19 clones (17,081 base pairs in total) in KIS1 cells and 21 clones (18,879 base pairs in total) in SUDHL6 cells. Locations of base alterations are shown above and below the loci with their positions.



Supplementary Figure S5 Genomic DNA editing by A3B in endogenous *cMYC* gene of lymphoma cell. (A) Frequencies of base substitutions in endogenous *c-Myc* genes amplified at 94 °C of Td. All types of substitutions per 1,000 sequenced base pairs in Mock- and A3B WT-expressing SUDHL6 cells are shown. We sequenced 15 clones (6960 base pairs in total) in Mock-transfected cells and 13 clones (6032 base pairs in total) in A3B WT-expressing cells. (B) Dinucleotide contexts in genomic DNA editing in the amplicons at 94 °C of Td. The rates of indicated dinucleotide sequence at the C to T transitions are shown. Asterisks indicate statistical significance in a $\chi 2$ test ($p < 0.01$).

Supplementary Table S1: Variations in *EGFP* sequences from cytidine-deaminase-expressing HEK293/EGFP cells

Type of variant	Frequency of variant (%)				
	mock	A3A	A3B WT	A3B H66/253R	AID
47:T/C	0	0.61	0	0	1.44
51:C/T	0.3	2.7	1.39	0	1.25
52:G/A	0.3	1.71	1.32	0	0.8
55:C/T	0	0.55	1.61	0	0.11
57:G/A	0	0.73	0	0	0.49
58:G/A	0	0.73	0	0	0.34
60:C/T	0	0.49	1.46	0	0.53
61:G/A	1.05	1.84	1.1	0	1.21
62:G/A	0	0	1.24	0	0.53
63:C/T	0.82	3.49	1.61	0.28	1.59
64:G/A	0	1.35	0.81	0	1.71
66:C/T	0.37	0.55	0.44	0	0.42
67:G/A	0	1.71	1.39	0	1.14
72:C/T	0	0.8	0.73	0	0.99
73:G/A	0	0.98	0.81	0	0.61
74:G/A	0.82	0.98	0.37	0	0
75:C/T	0	1.41	0.81	0	1.02
76:C/T	0	0.61	0.66	0	0.8
78:C/T	0	0.49	0.51	0	0
81:G/A	0	1.59	0.37	0	0.08
84:C/T	0.97	1.84	0.37	0	0.61
86:G/A	0	1.04	0.44	0	0.11
87:C/T	0	1.65	1.9	0	1.17
88:G/A	0	1.16	0	0	0.34
90:G/A	0	0.92	0	0	0.34
93:C/A	0	6	0	0	0
93:C/T	0	1.1	1.02	0	0.38
94:G/A	0	0.49	1.46	0.28	1.33
95:G/A	0.3	0.98	0.37	0	0.08
96:C/T	0	2.2	1.02	0.28	0.64
97:G/A	0	0.92	1.02	0	0.53
100:G/A	0	1.04	0	0	0.34
102:C/T	0	1.53	0.66	0.35	0.87
103:G/A	0	1.29	0.66	0	1.17
106:G/A	0	0.49	0.73	0	0
108:C/T	0	1.78	1.32	0	0.76
109:G/A	0	2.33	2.05	0	1.06
112:G/A	0	2.27	1.24	0	0.68
113:C/T	0	0.8	0.29	0	0.38

Type of variant	Frequency of variant (%)				
	mock	A3A	A3B WT	A3B H66/253R	AID
114:C/T	0	0.61	0.29	0	0.42
116:C/T	0	0.12	0.73	0	0
120:C/T	0	7.41	1.76	0.49	0.83
121:G/A	0	2.02	1.17	0	0.76
122:G/A	0	1.29	0.37	0	0
123:C/T	0	1.22	1.39	0.28	0.42
126:G/A	0	1.84	0.51	0	0.53
127:C/T	0	0.8	0.88	0	0
129:G/A	0	1.47	0	0	1.21
132:C/T	0	6.98	0.29	0	0.3
133:C/T	0	0.31	0.59	0	0
144:C/T	0	1.22	1.17	0	0.91
142:C/T	0	0.8	0.15	0	0
146:G/A	0	7.71	1.39	0	0.3
147:C/T	0	1.77	0.29	0.28	0.53
149:C/T	0	0.43	0.44	0	0.64
150:C/T	0	0.31	1.83	0	0.27
152:C/T	0	0.73	0.59	0	0
153:C/T	0	1.59	0.66	0.28	0.57
154:G/A	0	1.1	0.66	0	1.17
155:G/A	0	2.26	0.59	0	0.27
156:C/T	0	2.08	1.02	0	0.42
159:G/A	0	0	1.03	0	0.3
160:C/T	0	0.92	1.25	0.76	0
162:G/A	0	1.71	0.51	0	0
165:C/A	0	0	0.81	0	0.53
165:C/T	0.37	1.77	1.39	0.48	0.8
166:G/A	0	1.59	2.12	0.48	0.95
168:G/A	0	0.86	0.73	0	0.27
168:G/C	0	5.62	0	0	0
170:C/T	0	0.67	0.44	0	0.38
172:T/A	0	0.12	1.1	0	0
173:G/A	0	0.92	0.51	0	0.53
174:G/A	0	0.86	0.51	0	0.3
175:C/T	0	0.79	1.25	0	0.42
176:C/T	0	6.23	0.81	0	0.27
177:C/T	0	6.72	0	0	0.3
179:C/T	0	0	1.03	0.35	0
180:C/T	0	0.79	0.37	0	0.27
183:C/T	0	1.59	1.39	0	0.53
184:G/A	0	2.02	2.42	0	0.42
184:G/T	0	1.04	0	0	0

Type of variant	Frequency of variant (%)				
	mock	A3A	A3B WT	A3B H66/253R	AID
186:G/A	0	0.67	0	0.42	0
188:C/T	0	0.73	0.44	0	0
189:C/T	0	0.73	0.37	0	0
195:G/A	0	1.16	0.37	0	0.53
198:C/T	0	0.61	0	0	0.26
201:C/T	0	1.77	0.73	0	1.1
202:G/A	0	1.95	1.61	0.35	0.83
203:G/A	0	0.49	0.29	0	0.6
204:C/T	1.05	2.86	1.1	0.76	0.57
205:G/A	0	0.61	1.02	0	0.3
207:G/A	0	0.79	1.02	0	0.3
208:C/T	0	1.58	0.88	0	0.49
212:G/A	0	1.89	0.88	0	0.53
213:C/T	0	0.55	0	0	0
215:T/A	0	0.67	0	0	0.57
218:G/A	0	0.61	0.51	0	0.15
220:C/T	0	0.97	0.73	0	0.49
221:G/A	0	1.64	0.8	0.28	0.87
222:C/T	0	1.04	0	0	0.38
226:C/T	0	0	0.73	0	0
228:C/T	0	0	0	0	0.79
229:G/A	0	0.79	0.58	0	0.68
232:C/T	0	1.03	0.73	0	0.08
233:A/G	0	0	0	0	0.64
234:C/T	0	0.97	0	0.42	0.79
237:G/A	0	0.79	0.58	0	0.38
241:C/T	0	0.61	0.37	0	1.1
243:G/A	0	0.67	0.37	0	0.57
244:C/T	0	0.49	1.39	0	0.3
246:C/T	0	0.91	0.73	0	0.79
247:G/A	0	1.82	2.49	0.48	1.29
254:T/G	0	0.42	0	2.37	0.08
255:C/T	0	1.09	0.51	2.37	0.34
261:C/T	0	0.91	0.88	0	0.64
262:G/A	0.37	1.15	0.44	0	1.02
263:C/T	0	0.91	1.24	0	0.3
265:A/T	0	0.79	0	2.37	0.08
267:G/A	0	0.85	0.51	0.2	0.64
270:C/T	0	1.7	0.44	0.47	0.76
271:G/A	0	0.85	0	0	0.49
274:G/A	0	0.73	0	0	0
276:C/T	0	0.48	0	0	0.64

Type of variant	Frequency of variant (%)				
	mock	A3A	A3B WT	A3B H66/253R	AID
279:C/T	0	0.79	0.73	0	0.6
280:G/A	0	0.48	1.17	0	1.02
288:G/T	0	0	0	0	0.72
289:C/T	0	1.09	0.88	0.61	0.94
290:G/A	0	2.18	0.44	0.41	0.98
291:C/T	0	0.42	0.58	0	0.34
293:C/T	1.27	0.55	0.8	0	0.34
294:C/T	0	0.49	0.8	0	0.75
303:C/T	0	0.3	0.51	0	0.38
310:G/A	0	2.18	1.32	0	1.02
312:C/T	0	0.61	0.51	0	0.87
313:G/A	0.3	0.97	1.83	0.54	0.98
314:G/A	0	0.61	0	0	0.49
315:C/T	0	0.79	0.37	0	0.26
318:C/A	0	0	0.88	0	0
328:C/T	0	0.55	1.39	0	0.49
329:G/A	0	1.03	1.68	0	1.21
330:C/T	0	0.36	0	0	0.79
331:G/A	0	0.36	0.51	0	1.32
333:C/T	0	1.09	1.17	0	1.43
334:G/A	0	0.73	0.95	0.41	1.32
339:G/A	0	1.52	0.44	0	0.41
345:C/T	0	0.55	0.58	0	0.64
346:G/A	0	2.55	2.41	0	1.24
351:C/T	0	1.58	1.24	0.41	2.08
352:G/A	0	0.55	0.8	0	0.91
353:A/G	0	0	0	0	0.94
354:C/T	0	0.61	0	0	0
363:G/A	0	0.91	0.37	0	0.49
367:C/T	0	1.03	0.37	0	1.17
368:G/A	0.3	1.03	3	0.34	1.36
369:C/T	0	0.79	0.29	0	0.68
372:C/T	0	0.97	0.29	0	0.75
373:G/A	0.3	2.55	2.7	0	1.51
378:G/A	0.67	0.91	0.37	1.49	0.68
381:G/T	0	0	0	2.37	0
384:C/T	0	0.61	0.58	0	0.72
387:C/A	0	0	1.1	0.27	0
388:G/A	0	3.52	1.31	0.81	1.77
397:G/A	0	1.46	0	0	0.34
402:C/T	0	0.42	1.97	0.41	0.26
403:G/A	0	1.27	0.88	0.34	1.25
404:G/A	0.3	1.03	0	0	0

Type of variant	Frequency of variant (%)				
	mock	A3A	A3B WT	A3B H66/253R	AID
405:C/T	0	1.21	0.73	0.14	0.49
406:A/G	0	0.18	1.83	0	0
408:C/T	0	0.61	0	0	0
411:C/T	0	0.12	0.73	0	0
420:C/T	0	0	0.51	0	0.08
426:G/A	0	0.12	0.15	0	0.72
447:C/T	0	0	1.46	0	0
451:G/A	0	1.16	0.8	0	0.94
451:G/T	0	0	0.66	0	0
459:C/T	0	0.67	0.58	0	0.64
462:G/A	0	0.55	0	0	0.45
463:G/A	0	0.49	0.29	0	0.57
464:C/T	0	0.67	0	0	0
465:C/T	0.98	1.46	1.1	0.27	1.7
466:G/A	0	2.92	1.75	0	0.87
480:C/T	0	0.73	0	0	0.53
481:G/A	0	1.82	1.39	0	1.62
482:G/A	0	0.91	0	0	0.64
483:C/T	0	0.49	0.59	0.27	0.42
492:G/A	0	0.73	0	0.27	1.28
498:C/T	0	0	0.29	0.54	0.3
502:A/C	0	0	0.51	0	0
503:T/C	0.6	0	0	0	0.08
504:C/A	0	0.91	0	0	0.94
504:C/T	0.6	0.73	0.88	0	2.11

Supplementary Table S2: Primers for 3D-PCR, sequencing, and quantitative PCR

Primer pair	Sequence (5'-3')
EGFP 1st (pEGFP-N3)	GGTACCATGGTGAGCAAGGGCGAGGAGCTG
	GCTAGCTTACTTGTACAGCTCGTCCATGCC
EGFP 1st (pDON-EGFP)	GGCTATTCTCGCAGCTCACC
	TGCATACTTCTGCCTGCTGG
EGFP nested	GTGGTGCCCATCCTGGTCGAG
	CCTCGATGTTGTGGCGGATCTT
EGFP for deep seq	ACCGGGGTGGTGCCCATCC
	CGTCCTCGATGTTGTGGCG
A3A for qPCR	GAAGCCAGCCCAGCATCC
	CCCCTGTGCTGGTCCATCTT
A3B for qPCR	GACCCTTTGGTCCTTCGAC
	GCACAGCCCCAGGAGAAG
AID for qPCR	CTGGACTTTGGTTATCTTCG
	AAGGTCATGATGGCTATTTG
HPRT1 for qPCR	TGACCTTGATTTATTTTGCATACC
	CGAGCAAGACGTTCAGTCCT
GAPDH for qPCR	GAAGGTGAAGGTCGGAGTC
	GAAGATGGTGATGGGATTTC
c-Myc for Seq	TCGAGAAGGGCAGGGCTTCTCAGAGGCTTG
	GGCGATATGCGGTCCCTACTCCAAGGAGCT
Pax5 for Seq 1	ACACGCAGCGGGTCCTGAAAGCGGCT
	TCGGACTTGCCCCCTCCGCCTCTCT
Pax5 for Seq 2	CAGCGGTCTCAGTGTTGGAGGC
	CAAGCGCCAGTCTGGATCAGCCCC
A20 for Seq	TGCCTTGACCAGGACTTGGGACTTTGCGA
	CCACTCTGGGCAGACATTTTTGTCTGCAT
c-Myc for qPCR	AGCGACTCTGAGGAGGAAC
	TGTGAGGAGGTTTGCTGTG
Pax5 for qPCR	CTACCCTTTCCTCGCTCC
	CACAGGTCGGAATAATTCAAG
A20 for qPCR	AACATTTTGCTGCTGCCTC
	AGGTGCTTTGTGTGGTTCG
c-Myc nested	AGGGAGATCCGGAGCGAATAGGGGGCTTC
	AACCATTCCCGTTTTCCCTCTGCCTTCTCC